Real-world cardiovascular effects of liraglutide: transportability analysis of the LEADER trial

Kevin Josey, PhD[1,2], Wenhui Liu, MS[1], Theodore Warsavage, MS[1], Morten Medici, PhD[3], Kajsa Kvist, PhD[3], Catherine G. Derington, PharmD, MS[4,5], Jane E.B. Reusch, MD[1,6], Debashis Ghosh, PhD[2], Sridharan Raghavan, MD, PhD[1,7]

1 US Department of Veterans Affairs Eastern Colorado Health Care System, Aurora, CO.
2 Department of Biostatistics and Informatics, Colorado School of Public Health, Aurora, CO.
3 Novo Nordisk A/S, Copenhagen, Denmark.
4 Division of Cardiology, Department of Medicine, University of Colorado School of Medicine, Aurora, CO.
5 Adult & Child Center for Outcomes Research & Delivery Science, University of Colorado Anschutz Medical Campus, Aurora, CO.
6 Division of Diabetes, Endocrinology, and Metabolism, Department of Medicine, University of Colorado School of Medicine, Aurora, CO.
7 Division of General Internal Medicine, Department of Medicine, University of Colorado School of Medicine, Aurora, CO.

Key Points

Question: Are the benefits of liraglutide observed in the LEADER randomized trial externally valid to US veterans with diabetes - a high-risk population underrepresented in the trial?

Findings: When the LEADER study results were weighted to provide treatment effect estimates relevant to trial-eligible veterans, the effects of liraglutide on cardiovascular events and all-cause mortality were similar to that observed in the LEADER trial.

Meaning: The cardiovascular and mortality benefits of liraglutide generalized to a population of US adults under-studied in trials of glucagon-like peptide-1 receptor agonists - evidence with implications for treatment access policies in the Veterans Affairs Healthcare System.

Abstract

Importance: Whether randomized clinical trial (RCT) results pertain to populations not included in trials remains a question of substantial interest.

Objective: To evaluate generalizability of the cardiovascular benefits of liraglutide - a glucagon-like peptide-1 receptor agonist (GLP-1RA) - observed in the Liraglutide Effect and Action in Diabetes: Evaluation of Cardiovascular Outcome Results (LEADER) trial in US Veterans Affairs Healthcare System (VA) patients with diabetes, a population at high cardiovascular disease risk and underrepresented in LEADER.

Design: We used transportability methods to integrate real-world and RCT data and estimate the average treatment effect of liraglutide versus placebo had LEADER enrolled VA diabetes patients. We balanced baseline characteristics between LEADER participants and trial-eligible veterans using approximate balancing weights and estimated transported treatment effects with doubly-robust augmented inverse probability weighting. Sensitivity analyses varied VA cohort composition and balancing variables.

Setting: Multi-national RCT and US VA

Participants: LEADER participants and VA diabetes patients from 2015-2023

Interventions: Liraglutide versus placebo

Main Outcomes and Measures: Risk differences (RD) in survival probabilities and hazard ratios (HR) of trial-adjudicated major adverse cardiovascular events (MACE) (composite of non-fatal myocardial infarction, non-fatal stroke, and cardiovascular mortality) and all-cause mortality.

Results: Data from 9,336 LEADER participants were transported onto 357,075 trial-eligible VA diabetes patients. Transported effects of liraglutide compared to placebo on MACE and all-cause mortality in veterans ("VA-weighted LEADER") consistently overlapped the treatment effects observed in LEADER: RD of 3-year MACE of 2.0% [95% CI 0.8, 3.2] in VA-weighted LEADER versus 1.6% [0.3, 2.9] in LEADER; MACE HR 0.74 [0.61, 0.90] in VA-weighted LEADER versus 0.87 [0.78, 0.97] in LEADER; RD of 3-year all-cause mortality of 1.5% [0.6, 2.4] in VA-weighted LEADER versus 0.9% [-0.09, 1.9] in LEADER; all-cause mortality HR 0.71 [0.57, 0.89] in VA-weighted LEADER versus 0.85 [0.74, 0.97] in LEADER. Results were robust to all sensitivity analyses, including in veterans currently ineligible for GLP-1RA treatment in the VA.

Conclusions and Relevance: The benefits of liraglutide observed in LEADER generalized to veterans with diabetes - real-world evidence that can guide diabetes treatment for a high-risk population underrepresented in RCTs and inform formulary decisions for an integrated healthcare system.

Introduction

Randomized clinical trials (RCTs) provide gold-standard evidence of medication efficacy but suffer from uncertain external validity, particularly in populations underrepresented in RCTs. The most direct method for assessing medication effects in a specific target population – conducting an RCT in that population – is often infeasible due to RCTs being costly and time-consuming. Observational studies with real-world data performed after a medication is in routine use are an alternative to RCTs but can be limited by bias due to confounding and selection. An accepted approach to solving this problem is transportability analysis,[1,2] a study design that integrates RCT and real-world data to estimate what the treatment effect observed in an RCT would have been if individuals from the real-world target population of interest had been enrolled in the trial.[3] Critically, these methods preserve the benefits of randomization (i.e., minimizing confounding) and can be performed with data on potential users of a medication without requiring real-world uptake, limiting selection bias and delays in real-world evidence generation. Transportability analyses can help address at least two important clinical knowledge gaps: 1. Evidence of efficacy of new medications populations poorly represented in RCTs highlighted by recent US Food and Drug Administration guidance;[4,5] and 2. Evidence to guide formulary policies for healthcare payors servicing potential users of a medication who differ from RCT participants.

Glucagon-like peptide-1 receptor agonists (GLP-1RAs) are potent agents for the treatment of type 2 diabetes and obesity[6,7] that have demonstrated cardiovascular benefits observed in large RCTs.[6] The Liraglutide Effect and Action in Diabetes: Evaluation of Cardiovascular Outcome Results (LEADER) study provided the first evidence of cardiovascular benefit of a GLP-1RA – liraglutide – in patients with type 2 diabetes.[8] However, the LEADER study and other RCTs supporting the cardiovascular benefits of GLP-1RAs have consistently under-enrolled key constituencies of potential users of the medications, leading to evidence gaps in certain populations.[9-11]

Individuals with type 2 diabetes receiving care in the US Department of Veterans Affairs (VA) represent one such population of potential GLP-1RA users for whom there is

little direct RCT evidence of cardiovascular benefit. Veterans with diabetes comprised a small minority (<5% of the LEADER study population) of RCTs that established cardiovascular benefit of GLP-1RA,[6,8,12,13] but are at high cardiovascular disease risk owing to greater prevalence of cardiovascular risk factors than non-veteran US adults with diabetes.[14,15] Moreover, evidence-based VA formulary policies for GLP-1RA are limited by a lack of RCT evidence in veterans. The objective of this study was to transport the LEADER trial to veterans with type 2 diabetes and estimate the cardiovascular effects of liraglutide in this real-world population. We demonstrate how transportability methods can be used to estimate treatment effects in populations under-represented in RCTs by integrating RCT and real-world data.

Methods

Data sources

Individual-participant data from the LEADER study was shared with the study team through a data use agreement with Novo Nordisk A/S. All VA data were derived from the Corporate Data Warehouse, which records diagnoses, labs, and medication use for VA users. Veterans with type 2 diabetes were identified using a validated algorithm for the VA based on the presence of two or more ICD-9 or ICD-10 diagnosis codes for diabetes within a two-year period starting from January 1, 2002.[16,17] To ensure the study population demonstrated evidence of active VA care, we required participants to have at least two outpatient encounters including at least one VA primary care provider visit and at least one prescription medication filled through the VA within two years of the first occurrence of a diabetes diagnosis code. The VA Eastern Colorado Health Care System Research & Development Committee and the Colorado Multiple Institutional Review Board provided human subjects review and approval of the study.

Target populations

We constructed target populations of veterans with type 2 diabetes receiving care in the VA between 2015 and 2023, corresponding to years after the initial results of the

LEADER trial were first publicly reported. The approach to mapping the major inclusion and exclusion criteria of the LEADER trial to VA data to create a LEADER-eligible VA cohort are shown in **eTable 1**. As in LEADER, VA patients could be untreated for diabetes or treated with oral hypoglycemic medications, long-acting insulin, or a combination of these agents. Unlike the LEADER study, we excluded individuals with estimated glomerular filtration rate (eGFR) calculated with the Modification of Diet in Renal Disease[18] equation <15 ml/min/1.73m^2. In this way, the primary VA target population ("Cohort A") most closely matches the LEADER study sample, requiring individuals to have a history of atherosclerotic cardiovascular disease (ASCVD), age ≥50 years, and HbA1c ≥ 7% at baseline. LEADER included individuals with age ≥60 years with at least one cardiovascular risk factor, but this criterion was not mapped to the VA due to uncertainty about the data quality for the cardiovascular risk factors.

To examine the potential benefit of liraglutide use in pragmatic populations that were not trial-eligible, additional VA target populations were created in sensitivity analyses, relaxing three key inclusion criteria of LEADER: Cohort B relaxed the HbA1c requirement; Cohort C removed the age restriction; Cohort D did not require a history of ASCVD; and Cohort E imposed none of these three criteria.

Intervention

The real-world treatment in the target population is not used in the analysis. The intervention of interest was the LEADER RCT intervention: liraglutide (1.8 mg daily or maximum tolerated dose) or placebo.[8]

Outcomes

All outcomes are derived from the LEADER data; real-world outcomes in the target population are not used in the analysis. The primary endpoint was time to first major adverse cardiovascular event (MACE), a composite of cardiovascular death, non-fatal myocardial infarction (MI), or non-fatal stroke.[8] Time to non-fatal MI, non-fatal stroke, and all-cause mortality were secondary outcomes. We also evaluated the incidence of

selected adverse events reported in LEADER: all-cause medication discontinuation, gastrointestinal events (treatment-interrupting severe nausea/vomiting/diarrhea), pancreatitis, a gallbladder-related composite (acute gallstone disease, cholelithiasis, or cholecystitis), and any incident malignant neoplasm.

Covariates

We identified potential effect modifiers and confounders of trial participation – variables that might differ between the trial and VA populations and modify treatment effects: age, self-reported race, measures of diabetes severity (HbA1c, diabetes duration, number of diabetes medications), comorbidities (history of cardiovascular conditions; chronic kidney disease stage; atrial fibrillation; chronic obstructive pulmonary disease [COPD]; cancer; liver disease; dementia), and risk factors (body mass index [BMI], blood pressure, lipids, kidney function, smoking status).[17,19] We did not include sex as a covariate in the primary analysis due to the VA sample being predominantly men but performed a sensitivity analysis including sex as a balancing variable.

Analytical approach

The conceptualization and biostatistical methods underlying the transportability analysis are included in the **Online Supplemental Methods**. We computed pseudo-observations[20] for individual survival probabilities of MACE at 6-month intervals from 6 to 54 months. By using pseudo-observations at fixed time points, we can treat the survival pseudo-observations as quasi-continuous outcomes and apply standard causal inference methods without having to fit complex censoring-adjusted time-to-event models.[21]

We used a doubly-robust estimation framework to transport treatment effects from the RCT to the target population.[1] To adjust for differences between the RCT and VA target population, we constructed weights for each trial participant that reflect how "representative" they were of the VA population using approximate balancing weights, which utilize a penalty term to ensure that weighted covariate means in the RCT match the target population within a small tolerance.[22,23] We verified covariate balance between the

weighted trial sample and the target VA sample by examining standardized mean differences (SMDs) for all covariates.[24] In addition to weighting, we fit an outcome regression model on the trial data to predict survival probability as a function of treatment and covariates using an ensemble machine learning approach to flexibly model this relationship.[25] The final transported treatment effect estimates for the VA target population were obtained by combining the information from the weighting and outcome models in an augmented inverse probability weighted estimator.[1,3,26] If either the outcome or the weighting model is correctly specified, the treatment effect estimate will be consistent.[27]

For each target population and outcome, we computed the risk difference (RD) in event-free survival (liraglutide minus placebo) and its 95% CI estimated using the efficient influence function of the transported treatment effect with plug-in estimates of the requisite nuisance parameters. We also calculated hazard ratios (HR) by fitting weighted Cox proportional hazards models on the LEADER trial, the weights being the same as those used to estimate survival differences.[28]

All analyses were completed in R v4.4.1 (R Core Team 2025, Vienna, Austria).

Sensitivity Analyses

First, we included sex as a balancing covariate. Second, we conducted a leave-one-out analysis to see if any single covariate's omission from the set of balancing variables substantially changed the transported estimates. Third, we repeated the analyses in Cohorts B through E (described above) to assess the sensitivity of results to the target population definition, in particular target populations inclusive of veterans currently ineligible to receive GLP-1RA based on VA formulary policies.

Results

Baseline Characteristics

Cohort A (LEADER-eligible VA cohort) included 357,075 veterans meeting LEADER's inclusion criteria, compared to 9,336 participants in the LEADER trial. **Table 1** summarizes baseline characteristics of these two groups. Veterans were older on average (mean age of

70.0 years versus 64.3 years in LEADER), predominantly male (97.5% vs. 64.3% in LEADER), and were more likely to be Black or African American (14.1% vs. 8.3%). Baseline BMI (approximately 32.5 kg/m^2) and HbA1c (8.8% vs. 8.7%) were similar between VA and LEADER samples. Most comorbidities were more prevalent in veterans than in LEADER, baseline kidney function was worse on average in veterans, and current or former smoking was more common in veterans (**Table 1** and **eTable 2**). After applying sampling weights, SMDs between the LEADER and VA samples were reduced to <0.05 for all balancing variables (**eFigure 1**).

Outcomes

The transported survival probabilities for the primary MACE outcome and secondary outcomes in the primary VA target population were lower for both the liraglutide and placebo arms than in the original LEADER study (**Figure 1A**). Randomization to liraglutide was associated with a risk reduction of the primary MACE composite outcome and each secondary outcome compared to placebo. At 36 months of follow-up, the RD in outcome-free survival between the liraglutide and placebo arms was: 2.01% (95% CI: 0.80%, 3.22%), favoring liraglutide, for the composite MACE outcome in VA-weighted LEADER and 1.59% (0.30%, 3.04%) in LEADER; 0.17% (-0.05%, 0.85%) for stroke in VA-weighted LEADER and 0.20% (-0.49%, 0.90%) in LEADER; 1.00% (0.09%, 1.91%) for MI in VA-weighted LEADER and 0.97% (0.04%, 1.90%) in LEADER; and 1.50% (0.62%, 2.39%) for all-cause mortality in VA-weighted LEADER and 0.90% (-0.09%, 1.89%) in LEADER (**Figure 1B**). For all outcomes and at all follow-up time points, the transported RD in the VA target population fell within the 95% confidence interval of the corresponding treatment effect in LEADER (**Figure 1B**). The RDs of the composite MACE outcome and all-cause mortality, but not for stroke and MI, were qualitatively greater when LEADER was transported to the VA target population.

Correspondingly, the transported HR for MACE in VA-weighted LEADER was 0.76 (0.62, 0.94) for liraglutide versus placebo, which was a somewhat larger relative effect than observed in LEADER (HR 0.87 [0.78–0.97]) (**Figure 2**). A similar pattern was observed for

each of the secondary outcomes: the CIs overlap between the relative effects estimated in LEADER and in the VA target population (**Figure 2**).

Adverse Events

To assess external validity in the VA target population of safety outcomes assessed in LEADER, we estimated transported effects of liraglutide versus placebo on selected adverse events. As with the primary and secondary efficacy outcomes, the transported RDs for all adverse events examined fell within the 95% confidence intervals estimated in LEADER (**Figure 3**). For several of the adverse events examined, we observed a qualitatively greater risk associated with randomization to liraglutide versus placebo in VA-weighted LEADER than in LEADER (**Figure 3**).

Sensitivity Analyses

The transported RDs estimated for the VA target population were nearly identical without and with sex as a balancing variable (**eFigure 2**). In the leave-one-out-analysis, baseline age, HbA1c, and number of diabetes medications qualitatively impacted the transported effect estimates on the absolute (RD) and relative (HR) scales for the primary and secondary outcomes (**eFigure 3**). However, regardless of the balancing variable omitted, the transported effects in the VA target population for all outcomes on the RD and HR scales remained within the 95% confidence intervals of the LEADER study treatment effect estimates. Finally, the absolute RD and HR of the primary and secondary efficacy outcomes was similar across all VA target populations – within the 95% confidence intervals of the LEADER study with qualitatively greater absolute and relative risk reduction for the composite MACE and all-cause mortality (**Figure 4** and **eFigure 4**).

Discussion

Transporting treatment effects across populations, we found that the LEADER trial examining cardiovascular outcomes associated with liraglutide generalized to a real-world cohort of US veterans with type 2 diabetes. Liraglutide was associated with reduced risk of

the composite MACE outcome and all-cause mortality in a largely male, elderly VA target population. While the transported effect estimates for all outcomes fell within the 95% confidence intervals estimated in the LEADER study, the modestly greater RD for the primary MACE outcome and all-cause mortality implies the number needed to treat (NNT) to prevent one MACE event or one death over three years may be smaller in veterans than in the LEADER population (NNT of 50 versus 62.5 for MACE and 67 versus 111 for all-cause mortality). Moreover, the sensitivity analysis in broader VA target populations suggests liraglutide may confer benefit outside the trial inclusion/exclusion parameters and outside the current criteria for use of GLP-1RA in the VA. This reinforces the cardioprotective role of liraglutide as previously demonstrated in RCTs like LEADER, and importantly, extends that evidence to a broader population of potential users.

Our findings extend prior studies that replicate or generalize GLP-1RA cardiovascular outcome trials with real-world data in a variety of settings. The sensitivity analysis results in Cohorts D and E demonstrating cardiovascular benefit of liraglutide in veterans irrespective of pre-existing ASCVD mirror real-world studies evaluating GLP-1RA for ASCVD primary prevention and subgroup analyses in GLP-1RA RCTs that did not demonstrate heterogeneity by ASCVD history.[6,29] Our results also support a recent systematic review that found concordance between RCTs and real-world studies for GLP-1RAs[30] and prior work examining overlap between real-world patients and RCT participants for type 2 diabetes treatments.[31] Importantly, transportability analyses explicitly combine RCT data with real-world data – a hybrid design – unlike prior secondary analyses of RCTs or studies exclusively using real-world data. This approach has the advantage of preserving the internal validity of the RCT while extending inferences to the target population, addressing the gap between efficacy (does the drug work in a highly selected population?) and effectiveness/safety (does it work and is it safe in typical patients?).

In the VA, formulary decisions and criteria for use are often based on demonstrated outcomes in veterans. Our study provides evidence that can inform those decisions in the absence of VA-based RCTs.[32,33] The primary analyses support the use of GLP-1RAs for patients with type 2 diabetes and established ASCVD or high risk of ASCVD. The sensitivity

analyses in broader VA target populations suggest that several of the key LEADER inclusion criteria were inessential for liraglutide to be beneficial in veterans. As several of the RCT inclusion criteria – e.g., prior ASCVD and elevated A1c – remain among the criteria for use of GLP-1RA in the VA, the results with more broadly defined target populations argue for consideration of broader use balanced against cost, clinical characteristics, patient preferences, and drug availability.

Our study has important limitations. First, transportability assumes that any censoring is non-informative or that the censoring mechanisms would be similar in the target population as in the RCT. Second, while we included many covariates for adjustment, unmeasured differences or effect modifiers between the trial and VA populations could exist. For example, psychosocial factors, diet, exercise, adherence behaviors, or quality of background care might differ and were not captured. However, leave-one-out sensitivity analysis demonstrated that no single effect modifier included in the analysis meaningfully impacted the transported effect estimate, suggesting that the impact of an unmeasured effect modifier would have to substantially exceed the impact of any of the measured factors examined in the study. Third, the VA population is predominantly older men with extensive comorbidities. While excess disease burden in veterans helps motivate our study, it also means that the external validity of our conclusions is strong for VA and similar target populations, but caution is needed in extrapolating to populations with different characteristics. Finally, even with a very large VA sample, the statistical power for transported estimates is governed by the trial size and the variation in weights. We found that the effective sample size of LEADER that provides information relevant to the primary VA target population was only 2477 out of the 9336 participants in the LEADER study and declined further across broader target populations in sensitivity analyses (**eFigure 5**). The effective sample size estimates highlight the limited statistical power to detect significant differences between LEADER and transported effects, though that was not the goal of the study. The resulting wide confidence intervals reflect that transporting an effect is not as precise as conducting an RCT in the target population.

Conclusions

By transporting the LEADER trial results to a real-world veteran population, we found that liraglutide's cardiovascular and mortality benefits generalize to US veterans with type 2 diabetes, supporting the use of liraglutide to mitigate cardiovascular disease risk in a real-world population that was poorly represented in the RCT. The study also illustrates the value of causal transportability analyses in bridging the gap between clinical trial efficacy and real-world effectiveness. As healthcare systems increasingly rely on real-world evidence to inform practice and policy, the approach applied here – leveraging robust RCT results and tailoring them to a target population – can serve as a model for evaluating external validity of RCTs of other new therapies.

Acknowledgements

Author contributions

K.J., S.R., W.G.L, and D.G. designed the study. M.M. and K.K. assisted with access to and cleaning of LEADER data for analyses. K.J., W.G.L., and T.W. performed all analyses. K.J. and S.R. drafted the manuscript. All authors provided critical revision and input on the final manuscript.

Funding statement

The funders had no role in the conduct of the study. This publication does not represent the views of the Department of Veteran Affairs or the United States Government. S.R. was supported by VA award IK2-CX001907 and VA award I01-BX006417. J.E.B.R was supported by 1P30DK116073.

Data and code availability statement

Code for all statistical analysis is available upon request. A deidentified, anonymized limited data set derived from the datasets used for the analysis can be made available

upon reasonable request from researchers with necessary human subjects research oversight and in accordance with VA data sharing policies.

Conflicts of interest

M.M. and K.K. are employees of Novo Nordisk and provided data access and assistance with reproducing analyses of the original LEADER study. They provided critical feedback on the manuscript but did not shape or modify the reported results, interpretation, or conclusions. The remaining authors report no conflicts of interest.

References

1. Dahabreh IJ, Robertson SE, Steingrimsson JA, Stuart EA, Hernan MA. Extending inferences from a randomized trial to a new target population. *Stat Med*. 2020;39(14):1999-2014.

2. Degtiar I, Rose S. A Review of Generalizability and Transportability. *Annual Review of Statistics and Its Application*. 2023;10(1):null.

3. Josey KP, Berkowitz SA, Ghosh D, Raghavan S. Transporting experimental results with entropy balancing. *Stat Med*. 2021;40(19):4310-4326.

4. Califf RM. Now is the time to fix the evidence generation system. *Clin Trials*. 2023;20(1):3-12.

5. US Food and Drug Administration. Diversity plans to improve enrollment of participants from underrepresented racial and ethnic populations in clinical trials guidance for industry. In: Services UDoHaH, ed. Silver Spring, MD2022.

6. Badve SV, Bilal A, Lee MMY, et al. Effects of GLP-1 receptor agonists on kidney and cardiovascular disease outcomes: a meta-analysis of randomised controlled trials. *Lancet Diabetes Endocrinol*. 2025;13(1):15-28.

7. Moiz A, Filion KB, Toutounchi H, et al. Efficacy and Safety of Glucagon-Like Peptide-1 Receptor Agonists for Weight Loss Among Adults Without Diabetes : A Systematic Review of Randomized Controlled Trials. *Ann Intern Med*. 2025;178(2):199-217.

8. Marso SP, Daniels GH, Brown-Frandsen K, et al. Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes. *N Engl J Med*. 2016;375(4):311-322.

9. Knudsen JS, Thomsen RW, Pottegard A, Knop FK, Sorensen HT. Differences Between Randomized Clinical Trial Patients and Real-World Initiators of the Glucagon-Like Peptide 1 Receptor Agonist Liraglutide. *Diabetes Care*. 2018;41(9):e133-e135.

10. Weeda ER, Burnette B, Rowe S, Haas S, Zecopoulos A, Alnewais ME. Disparity in the rural-urban distribution of clinical trial sites used in studies cited by diabetes treatment guidelines. *Diabetes Obes Metab*. 2022;24(3):561-563.

11. Kelley AT, Mizokami-Stout K, O'Brien MJ, Bowen ME, Sussman J. Hispanic representation in diabetes cardiovascular outcomes trials. *BMJ Open Diabetes Res Care*. 2019;7(1):e000656.

12. Gerstein HC, Colhoun HM, Dagenais GR, et al. Dulaglutide and cardiovascular outcomes in type 2 diabetes (REWIND): a double-blind, randomised placebo-controlled trial. *Lancet*. 2019;394(10193):121-130.

13. Marso SP, Bain SC, Consoli A, et al. Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes. *N Engl J Med*. 2016;375(19):1834-1844.

14. Florez HJ, Ghosh A, Pop-Busui R, et al. Differences in complications, cardiovascular risk factor, and diabetes management among participants enrolled at veterans affairs (VA) and non-VA medical centers in the glycemia reduction approaches in diabetes: A comparative effectiveness study (GRADE). *Diabetes Res Clin Pract*. 2022;184:109188.

15. Fryar CD, Herrick K, Afful J, Ogden CL. Cardiovascular Disease Risk Factors Among Male Veterans, U.S., 2009-2012. *Am J Prev Med*. 2016;50(1):101-105.

16. Miller DR, Safford MM, Pogach LM. Who has diabetes? Best estimates of diabetes prevalence in the Department of Veterans Affairs based on computerized patient data. *Diabetes Care*. 2004;27 Suppl 2:B10-21.

17. Raghavan S, Vassy JL, Ho YL, et al. Diabetes Mellitus-Related All-Cause and Cardiovascular Mortality in a National Cohort of Adults. *J Am Heart Assoc*. 2019;8(4):e011295.

18. Levey AS, Coresh J, Greene T, et al. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. *Ann Intern Med*. 2006;145(4):247-254.

19. Burr RD, Tyler WB, Miller OF, 3rd. Burning discomfort and swelling of the hands and face of a young child. Erythropoietic protoporphyria (EPP). *Arch Dermatol*. 1993;129(5):635, 637-638.

20. Andersen PK, Perme MP. Pseudo-observations in survival analysis. *Stat Methods Med Res*. 2010;19(1):71-99.

21. Wang J. A simple, doubly robust, efficient estimator for survival functions using pseudo observations. *Pharm Stat*. 2018;17(1):38-48.

22. Wang Y, Zubizarreta JR. Minimal dispersion approximately balancing weights: asymptotic properties and practical considerations. *Biometrika*. 2019;107(1):93-105.

23. Keele LJ, Ben-Michael E, Feller A, Kelz R, Miratrix L. Hospital quality risk standardization via approximate balancing weights. *The Annals of Applied Statistics*. 2023;17(2):901-928, 928.

24. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med*. 2009;28(25):3083-3107.

25. van der Laan MJ, Polley EC, Hubbard AE. Super Learner. *Statistical Applications in Genetics and Molecular Biology*. 2007;6(1).

26. Lee D, Yang S, Dong L, Wang X, Zeng D, Cai J. Improving trial generalizability using observational studies. *Biometrics*. 2023;79(2):1213-1225.

27. Kang JDY, Schafer JL. Demystifying Double Robustness: A Comparison of Alternative Strategies for Estimating a Population Mean from Incomplete Data. *Statistical Science*. 2007;22(4):523-539, 517.

28. Shu D, Young JG, Toh S, Wang R. Variance estimation in inverse probability weighted Cox models. *Biometrics*. 2021;77(3):1101-1117.

29. Richardson TL, Jr., Halvorson AE, Hackstadt AJ, et al. Primary Occurrence of Cardiovascular Events After Adding Sodium-Glucose Cotransporter-2 Inhibitors or Glucagon-like Peptide-1 Receptor Agonists Compared With Dipeptidyl Peptidase-4 Inhibitors: A Cohort Study in Veterans With Diabetes. *Ann Intern Med*. 2023;176(6):751-760.

30. Wang W, Tang W-C, Webster-Clark M, Yu OH, Filion KB. Replicating cardiovascular outcome trials of medications used to treat type 2 diabetes using real-world data: A systematic review of observational studies. *medRxiv*. 2024:2024.2011.2028.24317099.

31. Wittbrodt ET, Eudicone JM, Bell KF, Enhoffer DM, Latham K, Green JB. Generalizability of glucagon-like peptide-1 receptor agonist cardiovascular outcome trials enrollment criteria to the US type 2 diabetes population. *Am J Manag Care*. 2018;24(8 Suppl):S146-S155.

32. Tummalapalli SL, Montealegre JL, Warnock N, Green M, Ibrahim SA, Estrella MM. Coverage, Formulary Restrictions, and Affordability of Sodium-Glucose Cotransporter 2 Inhibitors by US Insurance Plan Types. *JAMA Health Forum*. 2021;2(12):e214205.

33. Zerzan JT, Gibson M, Libby AM. Improving state Medicaid policies with comparative effectiveness research: a key role for academic health centers. *Acad Med*. 2011;86(6):695-700.

	VA	LEADER	p-value
Total N	357,075	9,336	
Male sex, n (%)	348,101 (97.5)	6,001 (64.3)	<0.001
Race, n (%)			<0.001
Black or African-American	50,312 (14.1)	776 (8.3)	
White	279,341 (78.2)	7,237 (77.5)	
Other	27,422 (7.7)	1,323 (14.2)	
Age (years), mean (SD)	70.0 (8.7)	64.3 (7.2)	<0.001
Hemoglobin A1c (%), mean (SD)	8.8 (1.7)	8.7 (1.5)	<0.001
# of diabetes medications, n (%)			<0.001
≤1	173,632 (48.6)	1,313 (14.1)	
2	87,554 (24.5)	3,294 (35.3)	
≥3	95,889 (26.9)	4,729 (50.7)	
BMI (kg/m^2), mean (SD)	32.6 (6.4)	32.5 (6.3)	0.466
eGFR (ml/min/1.73m^2), mean (SD)	66.4 (21.4)	79.1 (22.1)	<0.001
CAD, n (%)	196,030 (54.9)	887 (9.5)	<0.001
HF, n (%)	75,334 (21.1)	1,304 (14.0)	<0.001
Prior stroke, n (%)	85,302 (23.9)	1,506 (16.1)	<0.001
Prior MI, n (%)	44,406 (12.4)	2,862 (30.7)	<0.001
Prior revascularization, n (%)	81,584 (22.8)	3,481 (37.3)	<0.001
CKD stage, n (%)			<0.001
Stage 1	57,932 (16.2)	3,700 (39.6)	
Stage 2	142,760 (40.0)	3,656 (39.2)	
Stage 3 or 4	156,383 (43.8)	1,980 (21.2)	
Any cardiovascular disease, n (%)	303,221 (84.9)	7,731 (82.8)	<0.001
Hypertension, n (%)	306,579 (85.9)	8,508 (91.1)	<0.001
Hyperlipidemia, n (%)	246,848 (69.1)	7,067 (75.7)	<0.001
Atrial fibrillation, n (%)	58,036 (16.3)	179 (1.9)	<0.001
Dementia, n (%)	19,146 (5.4)	16 (0.2)	<0.001
COPD, n (%)	79,735 (22.3)	135 (1.4)	<0.001
Cancer, n (%)	131,325 (36.8)	551 (5.9)	<0.001
Liver disease, n (%)	26,929 (7.5)	594 (6.4)	<0.001
Smoking status, n (%)			<0.001
Current	77,955 (21.8)	1,130 (12.1)	
Former	199,577 (55.9)	4,338 (46.5)	
Never	79,543 (22.3)	3,868 (41.4)	

Table 1. Study participant characteristics at baseline.Abbreviations: BMI, Body mass index; eGFR, estimated glomerular filtration rate; CAD, coronary artery disease; HF, heart failure; MI, myocardial infarction; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease

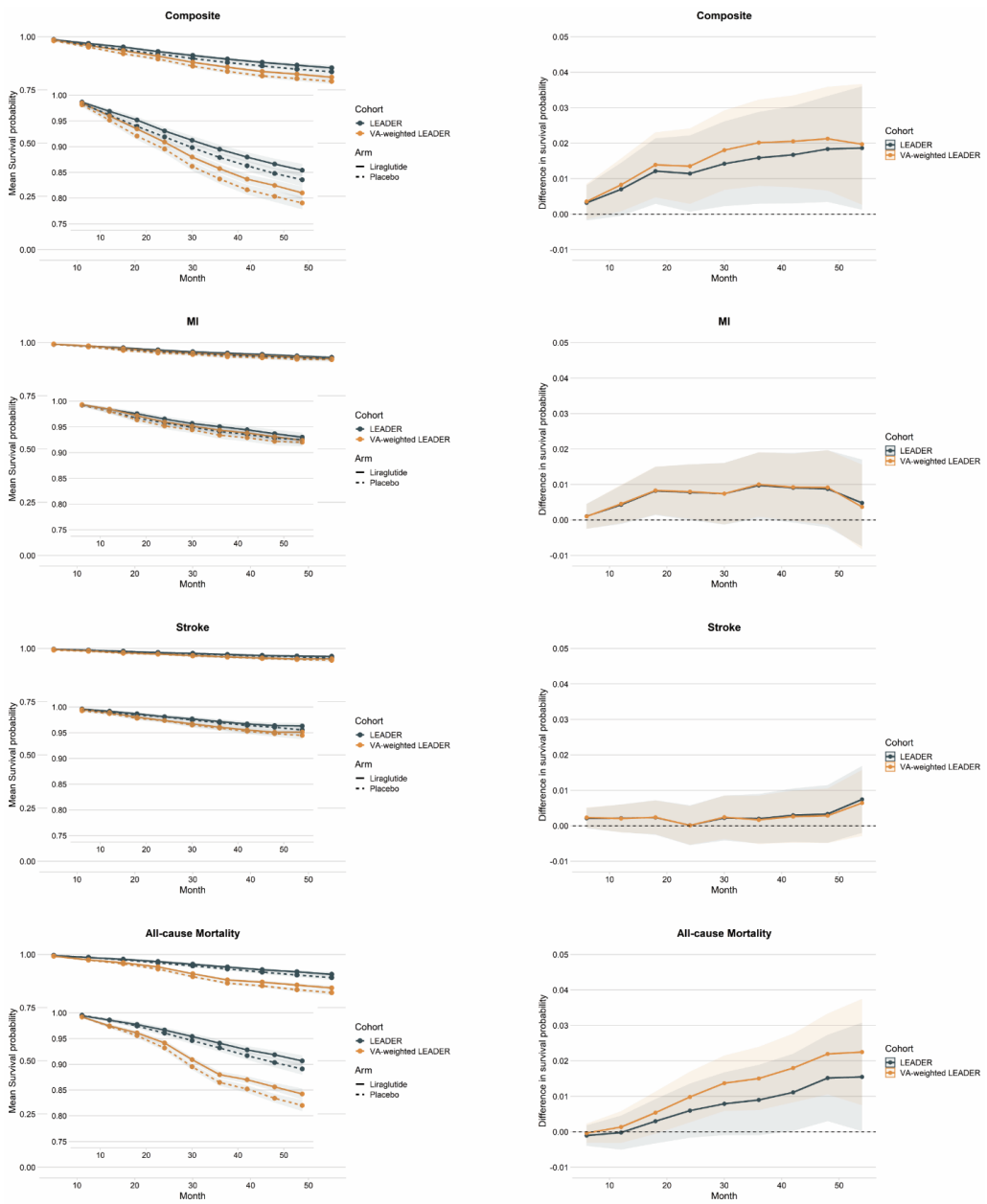

Figure 1. (**Left panel**) Predicted survival probabilities in liraglutide (solid lines) and placebo (dashed lines) arms of LEADER trial (blue) and transported to VA target population (yellow) for composite major adverse cardiovascular events (Composite), non-fatal myocardial infarction (MI), non-fatal stroke, and all-cause mortality. (**Right panel**) Absolute risk differences in event-free survival (liraglutide minus placebo) for composite major adverse cardiovascular events (Composite), non-fatal myocardial infarction (MI), non-fatal stroke, and all-cause mortality in LEADER (blue) and transported to VA target population (yellow); 95% confidence intervals are shown with colored shading.

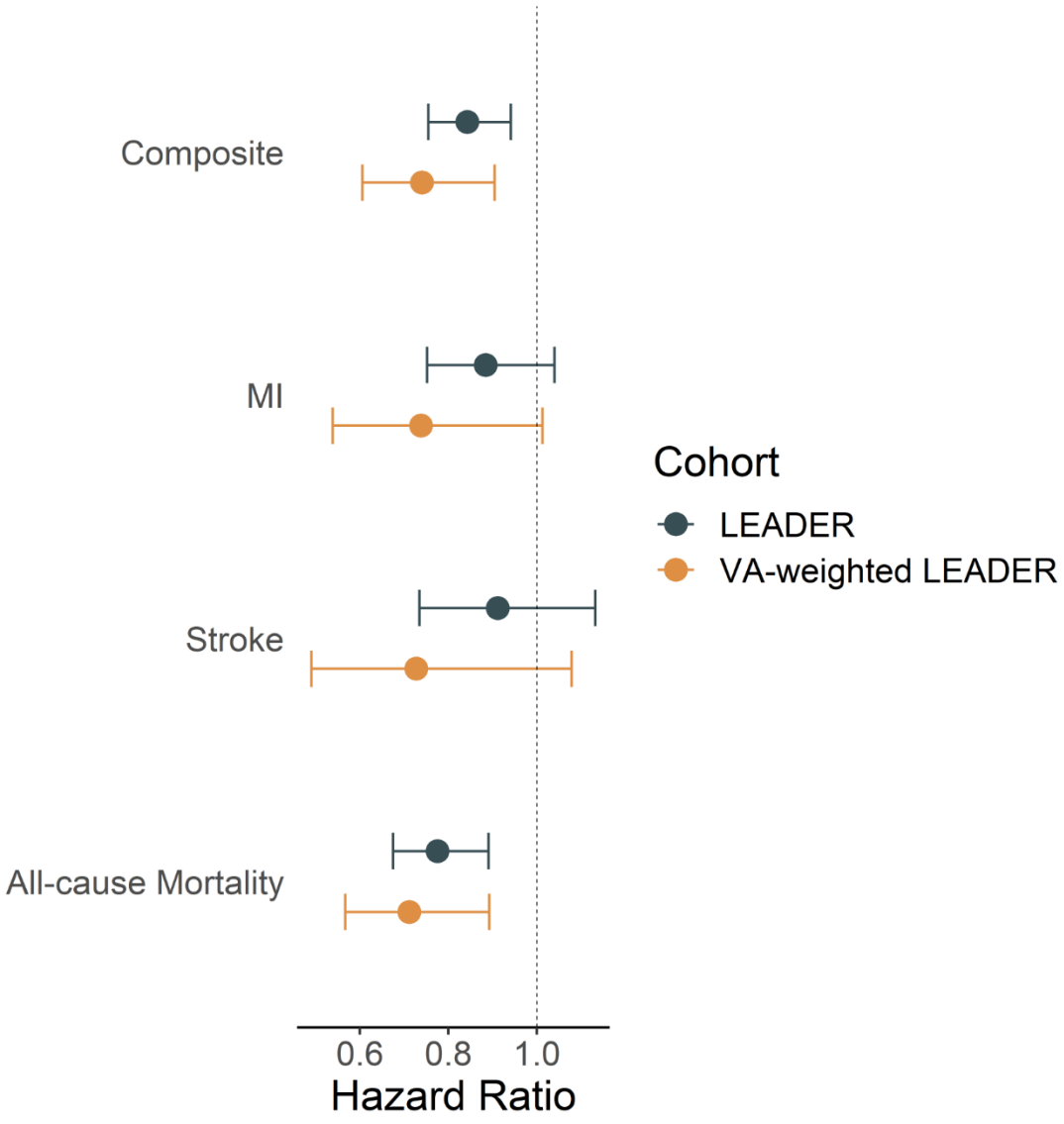

Figure 2. Hazard ratios and 95% confidence intervals estimated in LEADER (blue) and transported to VA target population (yellow) comparing randomized treatment to liraglutide relative to placebo for composite major adverse cardiovascular events (Composite), non-fatal myocardial infarction (MI), non-fatal stroke, and all-cause mortality. Hazard ratios less than 1 (vertical dotted line) favor liraglutide.

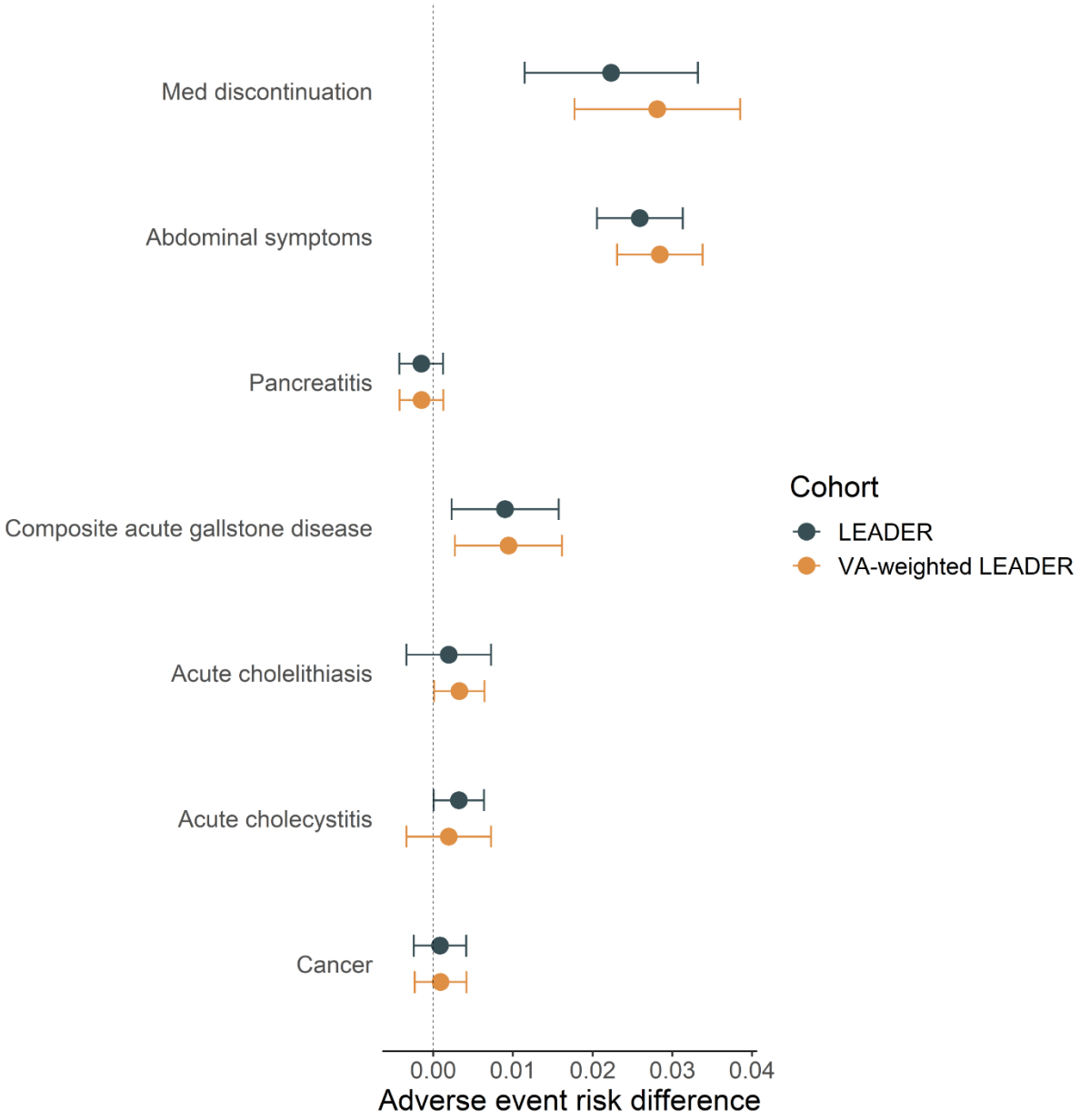

Figure 3. Absolute risk differences at the end of follow-up and 95% confidence intervals for several adverse events recorded in the LEADER trial (blue) and transported to the VA target population (yellow). Risk differences greater than 0 (vertical dotted line) reflect greater risk of adverse events in individuals randomized to liraglutide compared to placebo.

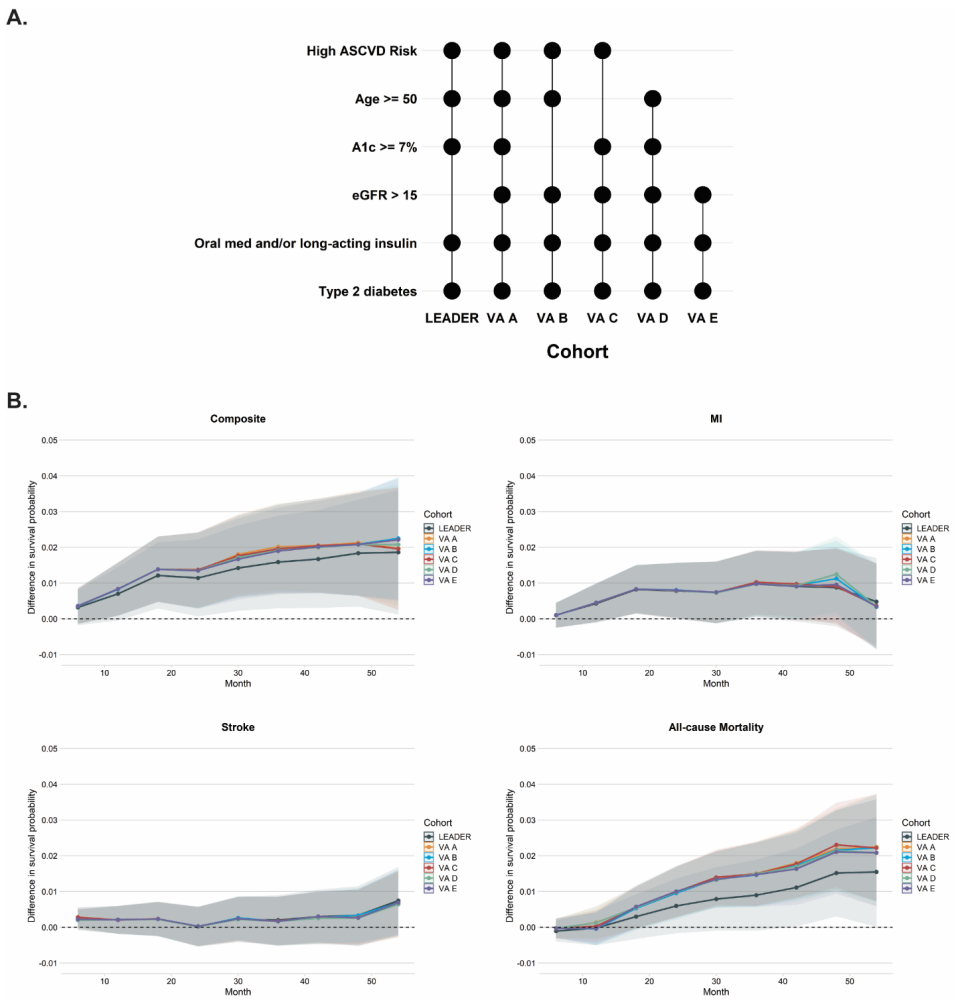

Figure 4. Comparison of treatment effects in LEADER and LEADER transported to a series of VA target populations with varying inclusion/exclusion criteria. (**A**) Key inclusion criteria for LEADER and for each VA target population included in the analysis. (**B**) Transported treatment effect estimated as survival probability differences for composite major adverse cardiovascular events (Composite), non-fatal myocardial infarction (MI), non-fatal stroke, and all-cause mortality; 95% confidence intervals are shown with colored shading for each transported treatment effect estimate. VA A through VA E indicate results after weighting LEADER to the corresponding target population. Abbreviations: ASCVD, atherosclerotic cardiovascular disease; A1c, hemoglobin A1c; eGFR, estimated glomerular filtration rate calculated using the Modification of Diet in Renal Disease equation (units of ml/min/1.73m^2).