nature portfolio

Corresponding author(s): Jung-Im Shin

Last updated by author(s): Jan 11, 2026

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	NA. Data was not collected using software.
Data analysis	All analyses were conducted in Stata/SE 16.1 (StataCorp, College Station, TX, USA).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Data were obtained from Optum Labs and consist of de-identified electronic health record data. These third-party proprietary data are not publicly available under the Optum Labs data use agreement. The authors do not have permission to share the data directly. Qualified researchers may request access from Optum Labs, subject to approval of a data use agreement and compliance with applicable requirements (https://www.optumlabs.com).

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Sex was collected directly as recorded in electronic health records and reflected what the patient or healthcare provider has recorded. Gender was not separately captured in the dataset. Sex was included as a baseline covariate in all analyses. As reported in Extended Table 1, there were 6,092,537 person-trials, of which 2,895,070 (47.5%) were female and 3,197,467 (52.5%) were male.
Reporting on race, ethnicity, or other socially relevant groupings	Race and ethnicity were collected directly as recorded in electronic health records and reflected what the patient or healthcare provider has recorded. Categories included White, Black, Hispanic, Asian, and other/unknown. Race and ethnicity were included as covariates in the propensity score model to balance these characteristics between treatment groups. Data are provided in Extended Table 1.
Population characteristics	The study population included 174,678 individuals identified with type 1 diabetes by a previously validated algorithm in the Optum Labs Data Warehouse. A detailed table of demographic, social and clinical characteristics is provided in Extended Table 1.
Recruitment	This was a retrospective analysis of electronic health records. No active recruitment occurred. The study included all individuals meeting eligibility criteria within the Optum Labs Data Warehouse from January 2013 to March 2024.
Ethics oversight	The study was approved by the Johns Hopkins Bloomberg School of Public Health Institutional Review Board. Informed consent was waived as this study used de-identified data and presented no more than minimal risk of harm to participants.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We conducted a retrospective cohort study using target trial emulation methodology to analyze de-identified electronic health record data from the Optum Labs Data Warehouse. Data were quantitative.
Research sample	Youth and adults with type 1 diabetes identified using a previously validated algorithm who received care within health systems enrolled in the Optum Labs Data Warehouse (OLDW) between January 2013 and March 2024. OLDW has longitudinal de-identified electronic health record data including demographics, vital signs, diagnoses, laboratory test results, and prescriptions from 300 million enrollees and patients, representing a mixture of ages and geographical regions across the United States. The study population included 174,678 individuals with type 1 diabetes who contributed 6,092,537 person-trials (mean age 43 years, 47% female).
Sampling strategy	All individuals meeting predefined eligibility criteria within the database were enrolled. No sample size calculation was performed a priori. The large sample size provides adequate power to detect clinically meaningful differences in cardiovascular and kidney outcomes.
Data collection	Data were from Optum Labs Data Warehouse (OLDW), which collects de-identified electronic health record data including demographic information, vital signs, prescription medication orders, laboratory test results, and health care encounters. Data are updated quarterly from participating health systems and harmonized into a standardized format. Researchers were not present during clinical encounters when data were originally recorded. Because this study involved secondary analysis of de-identified observational data, there was no experimental condition; therefore, researcher blinding to interventions or study hypotheses was not applicable.
Timing	The Optum Labs Data Warehouse has de-identified electronic health record data from 2007 onwards with ongoing updates. For this study, we analyzed data from January 1, 2013 through June 30, 2024.
Data exclusions	As shown in Extended Figure 1, we excluded person-trials with contraindications to GLP-1RA (n=33,097), with end-stage of kidney disease (n=251,529), with use of GLP-1RA within 1-year prior (n=69,224), with no follow-up or death (n=891,551), with no body mass index measurement at baseline (n=462,756), and with no hemoglobin A1c at baseline (n=1,570,228). These restrictions yielded an analytic sample of 6,092,537 person-trials contributed by 174,678 unique individuals.
Non-participation	NA - this is a secondary data analysis. Follow-up ended at the occurrence of outcomes, death, or last encounter (through June 30, 2024).

Randomization	NA - this is an observational cohort study, so participants were not randomized. We used propensity score overlap weighting to balance baseline covariates between treatment groups.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging