

Genetic predictors of GLP1 receptor agonist weight loss and side effects

<https://doi.org/10.1038/s41586-026-10330-z>

Received: 2 October 2025

Accepted: 26 February 2026

Published online: 8 April 2026

Open access

 Check for updates

Qiaojuan Jane Su¹, James R. Ashenhurst¹, Wanwan Xu¹, Vinh Tran¹, R. RYANNE WU¹, Catherine H. Weldon¹, Jingchunzi Shi¹, Barry Hicks¹, 23andMe Research Team*, Noura S. Abul-Husn¹, Stella Aslibekyan¹, Michael V. Holmes¹, Bertram L. Koelsch¹ & Adam Auton^{1✉}

The development of glucagon-like peptide 1 (GLP1) receptor agonists, including semaglutide and tirzepatide, has transformed the clinical management of overweight and obesity. However, substantial inter-person variability exists in both weight loss efficacy and the incidence of side effects¹. To investigate the genetic basis of this variability, here we conduct a genome-wide association study of self-reported weight loss and treatment-related side effects in 27,885 people following GLP1 receptor agonist therapy. We identify a missense variant in *GLP1R* that is associated significantly with increased efficacy of GLP1 medications ($P = 2.9 \times 10^{-10}$), with an additional -0.76 kg of weight loss expected per copy of the effect allele. Furthermore, we identify associations linking variation in both *GLP1R* and *GIPR* to GLP1 medication-related nausea or vomiting, with the *GIPR* association being restricted to people using tirzepatide. We incorporate these findings into a broader model of GLP1 medication response, and demonstrate the ability to stratify patients by efficacy and side effect risk. These findings provide direct genetic evidence that variation in the drug target genes contributes to inter-person variability in response and lay the foundation for precision medicine approaches in the treatment of obesity.

Approximately 40% of adults in the USA are classified as living with obesity—a chronic condition that affects the risk of several health conditions including type 2 diabetes (T2D) and cardiovascular disease. Until recently, therapeutic strategies for people with obesity were of limited efficacy², with the main strategies for weight loss focused on lifestyle interventions, including diet and exercise³.

GLP1 and gastric inhibitory polypeptide (GIP) are two primary gastrointestinal hormones (incretins) secreted from the gastrointestinal tract after food ingestion. Incretins stimulate pancreatic β -cell proliferation and insulin secretion, delay gastric emptying and act centrally to suppress appetite. Recently, potent incretin mimetics, including semaglutide (a GLP1 receptor agonist) and tirzepatide (a dual GLP1 and GIP receptor agonist), have become widely available for patients with obesity or who have overweight and have a cardiovascular risk factor. The clinical efficacy of semaglutide and tirzepatide has led to these drugs being among the most commonly prescribed in the USA, with an estimated one in eight people having used a GLP1 receptor agonist medication⁴.

For people using GLP1 receptor agonists, there is large variation in weight loss¹. In a study of semaglutide efficacy, the average reduction in weight from baseline was 10.2%, but 4.9% of patients achieved over 25% reduction from baseline, and 32.2% achieved less than a 5% reduction from baseline or even weight gain⁵. Identifying factors that predict a person's response to GLP1 medications may help guide treatment strategy, including choice of drug, dose and speed of dose escalation.

In other contexts, genetic variation is known to have an important role in treatment response, both for intended and adverse events^{6–8}. We hypothesized that some of the variation in GLP1 medication efficacy could be attributed to genetics, and surveyed 23andMe research participants regarding their use of GLP1 medications. Using these data, we conducted a large genome-wide association study (GWAS), and identified robust evidence that variants in the *GLP1R* locus are associated with both differential weight loss and side effects. Further analysis revealed an additional nausea and vomiting association in the *GIPR* locus specifically within the tirzepatide-treated population. We incorporated our genetic findings into broader models that combine demographic, clinical and genetic features to predict efficacy and side effects, and demonstrated the ability of this model to stratify patients by weight loss in a held out electronic health record (EHR) dataset. Our results highlight the opportunity for pharmacogenetics and precision medicine approaches applied to GLP1 medication.

In August 2024, we surveyed 23andMe research participants about their GLP1 medication usage, focusing on those taking weight loss medication. As of August 2025, we had collected over 27,885 survey respondents who had reported using at least one of Ozempic, Wegovy, compounded semaglutide, Mounjaro, Zepbound or compounded tirzepatide (Extended Data Table 1). Study participants were mostly female (82.4%) with a median age of 52 years. Most respondents were of European ancestry (78.3%), although the study also included substantial representation from Latino (12.9%) and African American (4.2%) ancestries.

¹23andMe Research Institute, Palo Alto, CA, USA. *A list of authors and their affiliations appears at the end of the paper. ✉e-mail: aauton@23andme.com

Study participants reported a median body mass index (BMI) of 35.1 kg m⁻² before initiating GLP1 treatment, with 96.8% of participants having a BMI of at least 25 at baseline. The median time that study participants reported taking GLP1 medications was 8.3 months. Following GLP1 treatment, participants reported having lost a median of 4.1 kg m⁻² units of BMI (or 11.3 kg of weight), equivalent to a median of 11.7% of pre-treatment weight. Participants reported greater BMI loss using tirzepatide compared with semaglutide (median 4.75 versus 3.71 kg m⁻²; $P = 9.7 \times 10^{-29}$, median test), despite similar treatment times (8.1 versus 8.4 months), consistent with previous reports⁹.

Non-genetic predictors of BMI loss

We examined non-genetic factors that were predictive of GLP1 medication efficacy, as measured by the percentage change in BMI from baseline ($\Delta\text{BMI}_{\%}$). As noted in other studies, medication seemed more effective in women¹⁰ ($\Delta\text{BMI}_{\%} = -12.2\%$ in women versus -10.0% in men; $P = 5.0 \times 10^{-31}$, Mood's median test; Supplementary Table 1), and in people of European ancestry¹¹ ($\Delta\text{BMI}_{\%} = -12.1\%$ in Europeans versus -10.6% in non-Europeans; $P = 4.7 \times 10^{-16}$, Mood's median test; Supplementary Table 2) before correcting for other factors.

To investigate the relative contribution of non-genetic factors that predict GLP1 medication efficacy, we fit a linear model to assess the combined effect of age, sex, pre-treatment BMI, drug type, drug dose and time on drug on total BMI loss (Supplementary Table 3). In this combined analysis, we observe that the degree of BMI loss depends strongly on sex, drug type, time on treatment and drug dosage, consistent with previous reports¹¹, and has a relatively weak dependence on pre-treatment BMI (with higher starting BMI associated with greater BMI loss). Efficacy also showed a modest reduction with age: each additional 10 years of age corresponded to a 0.5% reduction in BMI change (around 0.45 kg reduced weight loss efficacy), which is directionally consistent with recent clinical trial meta-analyses^{12,13}. Collectively, this model explained about 21.4% of variance in percentage BMI loss (adjusted R^2).

Previous literature has highlighted the role of ancestry¹¹ and T2D status¹⁴ in determining GLP1 medication efficacy for weight loss. We therefore investigated the predictive value of these variables in our cohort, after adjusting for the non-genetic factors described above (Supplementary Information). Consistent with GLP1 receptor agonists being developed initially to treat T2D, 23.1% of survey participants reported using GLP1 medication to improve blood sugar. T2D status was found to be a highly significant predictor of weight loss efficacy ($P = 2.0 \times 10^{-73}$; Supplementary Table 4), with an average predicted reduction of 2.87 percentage points in BMI loss for people with a T2D diagnosis. Similarly, our data indicate differential efficacy by ancestry (Supplementary Table 4), with GLP1 medications being most effective in people of European ancestry, less effective in people of Latino ancestry and least effective in people of African American ancestry.

Comparison of self-report data with EHRs

Research participants were able to share EHR data with 23andMe through Apple HealthKit (Methods). Using this mechanism, we obtained EHR data from 909 participants with a recorded GLP1 medication prescription (Supplementary Table 5). Of these, 195 had also completed the GLP1 survey, allowing us an opportunity to investigate the relationship between EHR and self-report data in the context of GLP1 medication and weight loss.

Qualitatively, both self-report and EHR data showed similar distributions for weight loss efficacy (Extended Data Fig. 1a). However, EHR data showed smaller changes in BMI compared with baseline, with a median $\Delta\text{BMI}_{\%}$ of -5.79% , which is significantly less than that reported in the self-report surveys (-11.8% , $P = 2.4 \times 10^{-55}$). The EHR cohort was younger and had a larger proportion of men than the survey cohort (median age

44 years versus 52 years, and 37% male versus 18%), and both of these factors may contribute to some of the apparent difference in weight loss efficacy between self-report and EHR data. The discrepancy was reduced when restricting to the 195 participants with both survey and EHR data, but remained significant (self-report = -14.14% versus -8.43% in EHR, $P = 1.1 \times 10^{-9}$ paired t -test; Extended Data Fig. 1b), despite the fact that the EHR and self-reported $\Delta\text{BMI}_{\%}$ measurements were reasonably well correlated (Pearson $r = 0.57$).

Of the participants with both EHR and survey data, 125 self-reported using semaglutide, of whom 97.6% had a corroborating EHR record. Likewise, 70 participants self-reported using tirzepatide, of whom 85.7% had a corroborating EHR record. These data indicate that survey participants generally accurately self-reported which drug type they were using, although EHR data reported a higher fraction of users relying on compounded versions of the medications (18.9% in EHR versus 13.5% in self-report for semaglutide, and 6.1% versus 4.8% for tirzepatide). For dosage information, we found that 58.4% of semaglutide users and 64.3% of tirzepatide users self-reported a dosage equal to that recorded in the EHR. In general, the self-report data indicate that medication users often report higher doses than those recorded in their EHRs, with 88.8% of semaglutide and 84.3% of tirzepatide users self-reporting a dosage greater or equal to the EHR records. Furthermore, we found that treatment intervals reported in EHRs were reasonably consistent with self-report data (7.8 months versus 8.3 months for EHR and self-report respectively; Extended Data Fig. 1c), and were closely matched in participants with both survey and EHR data (median 9.45 months in EHR compared with 9.58 months in survey; Extended Data Fig. 1d).

GWAS of GLP1 medication efficacy

We performed a GWAS of the percentage BMI loss phenotype ($n = 15,237$), and identified a strong genome-wide significant association on chromosome (chr.) 6 (Fig. 1a,b and Supplementary Table 6). The index single-nucleotide polymorphism (SNP) of the main association was rs10305420 ($P = 2.9 \times 10^{-10}$, reference allele C, effect allele T), which conferred an additional 0.641% BMI loss per T allele, approximately equivalent to an additional 0.76 kg weight loss per allele (95% CI, $[-1.27, -0.34]$ kg). We observed no evidence of dominance effects ($P = 0.80$), indicating that the allelic dosage of the T allele contributes to efficacy in an additive manner. In addition, we did not observe compelling evidence for additional independent associations within the locus (Extended Data Fig. 2a). We also confirmed that the association of rs10305420 with GLP1 medication efficacy cannot be accounted for by T2D or smoking status (Supplementary Information and Supplementary Tables 7 and 8).

The associated SNP, rs10305420, is a missense variant in the signal peptide of the *GLPIR* gene, changing the seventh amino acid from proline to leucine (p.Pro7Leu). To determine whether this variant is likely to be causal, we considered all 42 variants within the 99% credible set of the association, and annotated by predicted functional consequence (Supplementary Table 9). Within the credible set, rs10305420 is the only coding variant, and has the maximum posterior probability (35%). In addition, we tested for co-localization with expression quantitative trait loci in the locus, and identified none that colocalized with the GWAS signal (Supplementary Information). As such, we conclude that rs10305420 is probably the causal variant.

In the gnomAD database¹⁵, the rs10305420 T allele is most common in the European (40%) and Middle Eastern (38%) ancestry groups, followed by the Admixed American (28%), South Asian (20%) and East Asian (16%) ancestry groups, and least common in the African (7%) ancestry group. We retested rs10305420 for association in each of five additional populations; Latino, African American, East Asian, South Asian and Middle Eastern. Although the SNP did not achieve significance ($P < 0.001$) in any non-European population, the estimated effects were

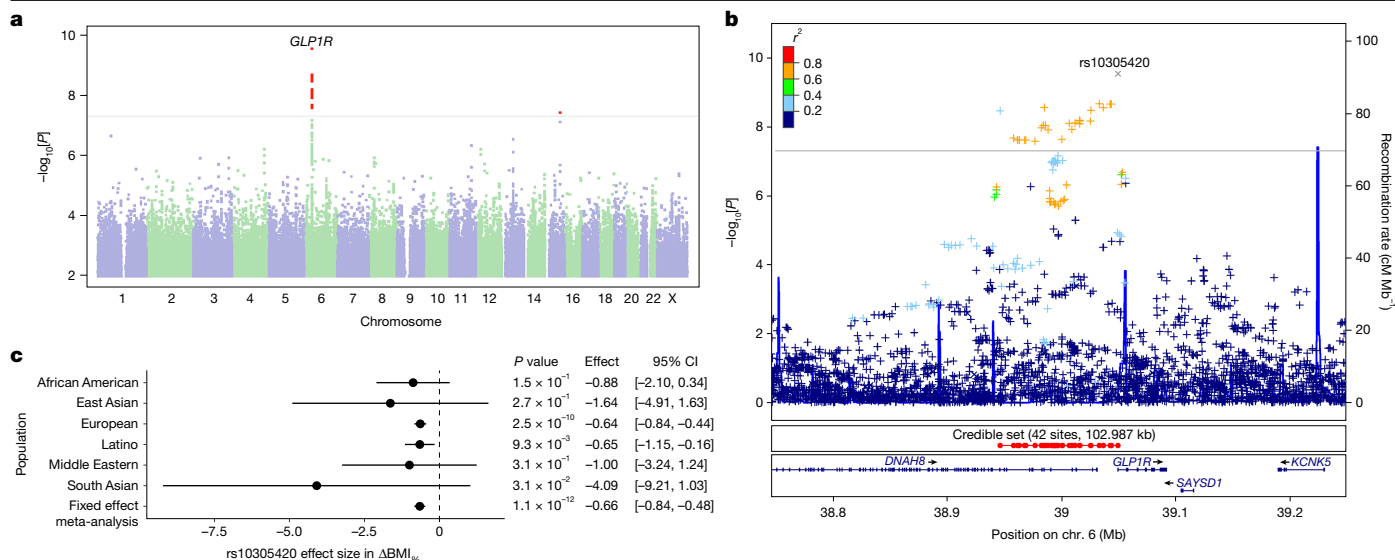


Fig. 1 | Genetic associations with GLP1 medication efficacy. **a**, Manhattan plot of percentage BMI change ($\Delta\text{BMI}_{\%}$) GWAS. SNPs achieving genome-wide significance ($P < 5 \times 10^{-8}$) are highlighted in red. **b**, Regional plot around the *GLP1R* locus. Colours indicate the strength of linkage disequilibrium (r^2) relative

all directionally consistent (Fig. 1c) and a fixed effect meta-analysis increased the significance of the observed association ($P = 1.1 \times 10^{-12}$).

We compared the SNP effect among people treated with semaglutide versus tirzepatide (Extended Data Fig. 3a,b). When testing each drug type separately, we estimate the effect of the SNP on $\Delta\text{BMI}_{\%}$ to be larger in tirzepatide (effect = -0.95%) than semaglutide (effect = -0.51%), although the 95% confidence intervals were overlapping (Extended Data Fig. 3c). In a joint regression model including a *snp:drug type* interaction term, we determined the difference to be weakly significant ($P = 0.02$).

We further replicated the SNP association in the All of Us cohort, using data derived from 4,855 people with EHR data. The association between percentage BMI change and rs10305420 was replication significant, with directionality consistent with that observed in the 23andMe cohort ($P = 0.001$; effect = -0.47% ; Supplementary Table 10). The association did not replicate in UK Biobank, although the expected power for replication in this cohort was low (Supplementary Table 11 and Supplementary Information).

GWAS of GLP1 medication side effects

We performed a GWAS of 11 side effect phenotypes that contrast patients with moderate or severe side effects to those who experienced mild or no side effects. We identified associations within each of the nausea and vomiting side effect GWAS, both of which occurred within the vicinity of *GLP1R* locus (Extended Data Fig. 4). The index SNP for the vomiting signal was rs11760106 ($P = 2.5 \times 10^{-27}$; T allele odds ratio = 1.57), whereas the index SNP for the nausea signal was rs9357296 ($P = 2.6 \times 10^{-28}$; G allele odds ratio = 1.36). We did not observe any difference in the SNP effect on side effects for semaglutide versus tirzepatide (Extended Data Fig. 3d,e). Conditional analysis did not identify independent associations within the locus for either trait (Extended Data Fig. 2b,c).

Although both of these index SNPs were in moderate linkage disequilibrium with the BMI-loss coding variant (rs11760106 versus rs10305420 $r^2 = 0.57$; rs9357296 versus rs10305420 $r^2 = 0.75$), neither signal included the coding variant within the 99% credible set¹⁶. To determine whether the signals share the same causal variant, we applied co-localization analysis¹⁷ and found that the signals co-localize with high probability ($\Delta\text{BMI}_{\%}$ versus nausea, $H_4 = 96.6\%$; $\Delta\text{BMI}_{\%}$ versus

to the index SNP (rs10305420). Non-coding variants are indicated with plus symbols; coding variants are indicated with multiplication symbols. **c**, Estimated effect sizes of the index variant in $\Delta\text{BMI}_{\%}$ for each ancestral population. Circles, point estimates; horizontal bars, 95% confidence intervals (CI).

vomiting, $H_4 = 88.5\%$; nausea versus vomiting, $H_4 = 92.1\%$). Multi-trait co-localization supports the conclusion that these signals probably represent the same signal, with a 72.6% posterior probability of co-localization (Supplementary Information). From the co-localization analysis, we infer that increased nausea or vomiting is associated with greater BMI loss efficacy (Extended Data Fig. 5).

Unlike semaglutide, tirzepatide is a dual receptor agonist targeting both *GLP1R* and *GIPR*. By performing a GWAS in the tirzepatide-treated population alone (Supplementary Table 12), we identified an association between the vomiting side effect and *GIPR* (rs71338792, $P = 4.2 \times 10^{-9}$, odds ratio = 1.84; Fig. 2a). The index variant is in near perfect linkage disequilibrium ($r^2 = 0.99$) with a missense variant within *GIPR* (rs1800437, $P = 5.1 \times 10^{-9}$; Fig. 2b), which we conclude is the causal variant for this association (Supplementary Table 13). This variant alters the 354th amino acid of the protein sequence from glutamic acid (G allele) to glutamine (C allele; p.Glu354Gln). The G allele (our reported effect allele, odds ratio = 0.546) is protective, whereas the C allele is associated with a higher risk of the vomiting side effect (odds ratio = $1/0.546 = 1.83$).

In the gnomAD database¹⁵, the rs1800437 C allele is most common in the East Asian (21%) and European (20%) ancestry groups, followed by the Middle Eastern (18%) and South Asian (14%) ancestry groups, and least common in the African or African American (11%) and Admixed American (11%) ancestry groups. We retested rs1800437 for association with the vomiting side effect in the Latino population, which was the only non-European population with sufficient sample size, and found it to be directionally consistent ($P = 0.03$, odds ratio = 0.49). A fixed effect meta-analysis increased the significance of the observed association ($P = 1.1 \times 10^{-10}$, odds ratio = 0.54). No evidence of effect was observed within the semaglutide population (Fig. 2c).

We observed that 11.8% of tirzepatide-treated rs1800437-CC carrier individuals experienced vomiting, compared with 9.4% for semaglutide-treated people carrying the same genotype, although the difference was not significant ($P = 0.12$, two-sample z-test; Supplementary Table 14). We observed no evidence of dominance at rs1800437 ($P = 0.16$). However, when testing for an interaction between the *GLP1R* variant (rs10305420) and the *GIPR* variant (rs1800437), we identified weak evidence for interaction ($P = 0.018$; Supplementary Table 15). We estimate that people who are homozygous for the risk alleles at both the *GLP1R* and *GIPR* loci have a 14.8-fold increased odds (95% confidence

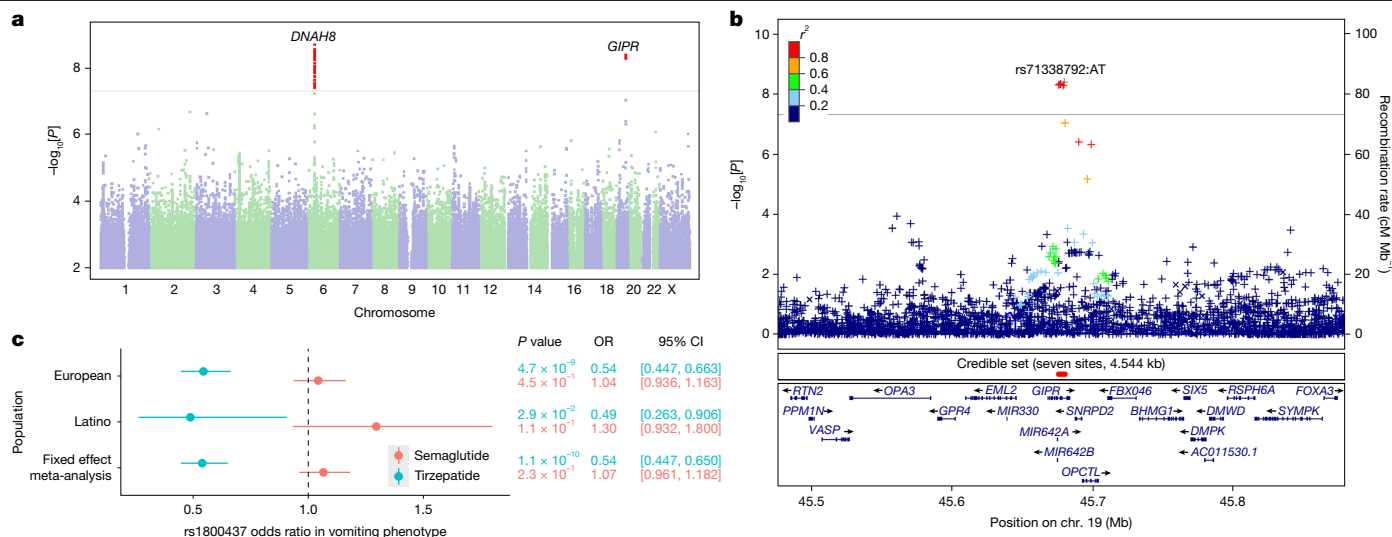


Fig. 2 | Genetic associations with GLP1 medication side effects. **a**, Manhattan plot of experiencing moderate-to-severe vomiting while on tirzepatide treatment. SNPs achieving genome-wide significance ($P < 5 \times 10^{-8}$) are highlighted in red. **b**, Regional plot around the *GIPR* locus. Colours indicate strength of linkage disequilibrium (r^2) relative to the index SNP (rs71338792). Note that the probable causal missense variant, rs1800437, is located within

interval [6.2, 35.8] of tirzepatide-mediated vomiting, compared with those homozygous for the non-risk alleles.

Across phenotypes, our data indicate an association between rs1800437 and tirzepatide treatment for both the vomiting and nausea phenotypes, and potentially constipation, but the variant does not seem to influence efficacy ($P = 0.73$; Extended Data Fig. 6).

Phenome-wide associations

In the GWAS catalogue¹⁸, the index BMI-loss SNP in *GLPIR*, rs10305420, has been reported as associated with T2D^{19,20}, smoking initiation²¹, fasting glucose²⁰ and haemoglobin A1c (HbA1c) measurements²⁰. These data are consistent with the T allele decreasing risk for T2D, as well as decreasing glucose and HbA1c levels, but slightly increasing the risk of smoking initiation (Supplementary Table 16). Although the SNP is not associated with any traits in the UK Biobank with $P < 1 \times 10^{-4}$, rs10305420 is associated with obesity related-traits at weak significance (P value around 1×10^{-5}) in FinnGen. In the MVP cohort²⁰, rs10305420 is associated significantly with BMI-associated traits (P value around 2×10^{-15}), although it is in linkage with an upstream association (Extended Data Fig. 7).

In the 23andMe database, we observed rs10305420 to be associated with a number of traits (Supplementary Table 17), including T2D and traits linked to glucose metabolism. The T allele was associated with reduced risk of T2D, and increased likelihood of having been a smoker—observations that are consistent with public data. In addition, we observed association with diet-related phenotypes, including preference of sugary foods, red meat consumption and number of cavities. Finally, the T allele was associated with reduced risk of morning sickness during pregnancy. Notably, the 23andMe data did not reveal an association with BMI or weight, indicating that the association of the SNP may be identifiable more strongly in the context of GLP1 treatment.

The *GIPR* missense variant, rs1800437, has been associated with a large number of traits in the public domain, including BMI and body mass phenotypes, T2D, blood glucose and a number of metabolic or haematologic traits including blood pressure, reticulocyte and/or erythrocyte volume and urinary calculus (Supplementary Table 18). This broad pleiotropy was supported by associations observed in the 23andMe database (Supplementary Table 19).

the cluster of significant SNPs and is in very high linkage disequilibrium with the index SNP ($r^2 = 0.99$). Non-coding variants are indicated with plus symbols; coding variants are indicated with multiplication symbols. **c**, Estimated effect sizes of rs1800437 in the vomiting side effect, partitioned by drug type, for Europeans, Latinos and a fixed effect meta-analysis. Circles, point estimates; horizontal bars, 95% confidence intervals (CI). OR, odds ratio.

Modelling of treatment response

We aimed to build models of treatment response that combined both genetic and non-genetic predictors. As such, we integrated treatment, clinical, demographic, disease diagnosis and genetic variables as predictors in models of both efficacy and side effects.

The model of BMI-loss efficacy demonstrated good performance in the self-report data, explaining 25% of the variance (R^2) in both the training set and held out test set, with most of the variance explained by non-genetic factors (Supplementary Table 20). T2D, non-alcoholic fatty liver disease (also known as metabolic dysfunction-associated steatotic liver disease) and hypertension diagnoses were all associated with lower weight loss efficacy, consistent with previous reports^{22–24}. Calibration plots confirmed the model was well calibrated in the test set (Extended Data Fig. 8).

To assess the utility of the model in a real world setting, we applied the model to a set of EHR data not used during model construction, and predicted efficacy at 6 months, assuming that the drug type and final dose was unknown at time point zero. We observe that people predicted to achieve higher weight loss did indeed achieve higher weight loss in the longitudinal EHR data (Fig. 3a).

For the side effect models, we focused on the two phenotypes for which GWAS SNPs were identified (Supplementary Table 21). The nausea model achieved a receiver operating characteristic area under the curve of 65.4% in the test set, whereas the vomiting model achieved a receiver operating characteristic area under the curve of 68.0% (Fig. 3b,c), although, in both cases, the contribution from genetics was relatively modest compared with that from non-genetic factors (Supplementary Table 22).

Discussion

We report a large-scale GWAS of treatment response to GLP1 medications and identify a missense variant in *GLPIR*, which encodes the therapeutic target of semaglutide and tirzepatide that associates with BMI loss. In addition, variants linked to both *GLPIR* and *GIPR* are associated with GLP1 medication-related nausea and vomiting.

The fact that the loci we identified map to the genes encoding pharmacodynamic targets of GLP1 medications provides compelling

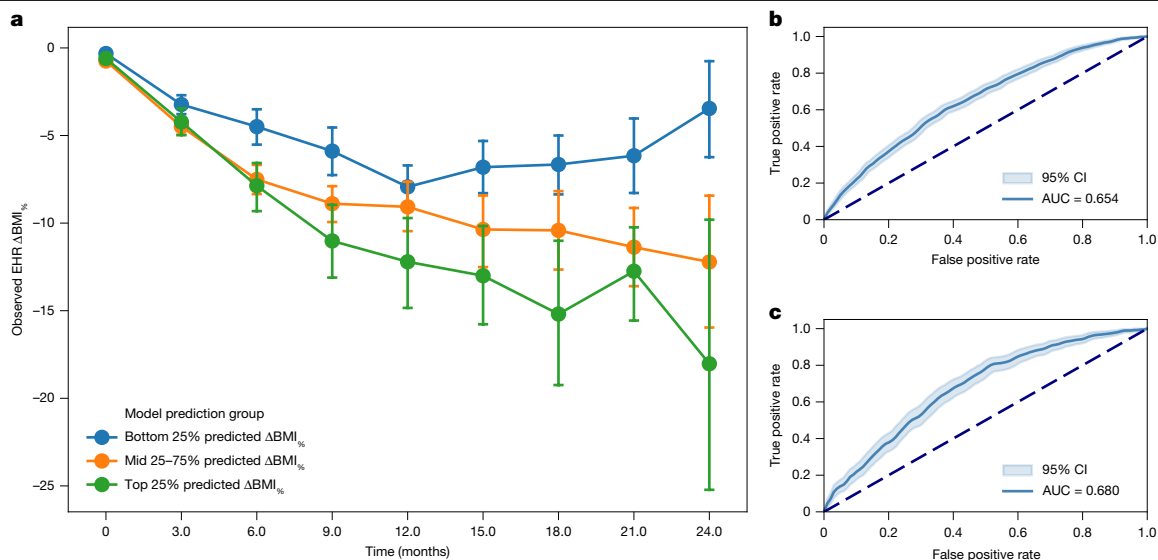


Fig. 3 | Combined genetic and non-genetic model performance in test set. **a**, Validation of $\Delta\text{BMI}_{\%}$ model in longitudinal HealthKit EHR data. Circles, point estimates; horizontal bars, 95% confidence intervals. **b**, Receiver operating

characteristic curve for nausea side effect, as assessed in held out test data.

c, Receiver operating characteristic curve for vomiting side effect, as assessed in held out test data. AUC, area under the curve; CI, confidence interval.

biological plausibility for their role in treatment response. Genetic variants that alter the function of a therapeutic receptor are known to influence drug response through several mechanisms.

The *GLP1R* variant rs10305420 is a missense variant in the signal peptide region, and may influence stability of the peptide secondary structure (Extended Data Fig. 9). Specifically, the effect allele encodes for leucine, which is more hydrophobic than the reference proline allele and may therefore increase stability of the hydrophobic region of the signal peptide (Supplementary Information). As such, the association with efficacy may be the result of enhanced trafficking of the protein within the cell that ultimately modulates the cell surface abundance of the receptor, rather than the result of alterations in the receptor protein structure that may influence ligand affinity.

Previous literature has also identified rs10305420 as influencing GLP1 receptor agonist efficacy^{25,26}, although with opposite directionality to our study. It is unclear why directionality would vary between studies, although we note these studies were performed in specific disease contexts, and are in substantially smaller samples. Power calculations suggest that previous studies would have had limited power to detect an effect size similar to that identified in our study (Supplementary Table 23), and also differed in terms of drug type, disease cohort and analytical choices. Given the independent replication of our observed directionality in an external cohort using EHR data, the directionality inferred by our study is unlikely to be the result of a data quality issue.

Similarly, the *GIPR* variant rs1800437 corresponds to the missense substitution p.Glu354Gln. The Gln variant has been well characterized as a partial loss-of-function mutation^{27,28}. Our data indicate an increased risk of vomiting and nausea side effects associated with the Gln allele in tirzepatide-treated patients, which may result from a diminished ability of the GIP component of the drug to buffer the nausea-inducing effects of the GLP1 component. Recent preclinical studies have demonstrated that central GIP receptor activation can attenuate the aversive effects, such as nausea and vomiting, typically induced by GLP1 receptor agonism²⁹.

Understanding the genetic basis of GLP1 medication treatment response is of substantial clinical importance. There is large inter-person variability in response to GLP1 medications both in terms of weight loss and side effects, with the two potentially being related (that is, people who experience more nausea may lose more weight³⁰). The discovery of robust predictors of treatment response could enable the

prediction, from the outset, of a person's probable treatment journey, and thereby pave the way for precision medicine through the clinical implementation of combined phenotypic and genotypic prediction models.

We have shown that self-report data is a complementary instrument for collecting data regarding treatment response that is qualitatively consistent with medical records, while also enabling collection of information regarding side effects that may not be readily available in EHRs. Although EHR-recorded weight loss was generally less than the self-reported values, several factors may explain this discrepancy. Previous work has noted that self-report data remain a valid and useful measure in many contexts³¹, but downward biases in self-reported weight are common³² and may be influenced by factors including health awareness, culture and social norms³³. Conversely, EHR data may offer an incomplete picture of a person's weight variation. If patients receive GLP1 medications outside their primary health system (for example, from telehealth) or change providers during the treatment, the EHR data collected in our study may not capture the full course of treatment, and this may explain our observation of generally shorter treatment durations, and hence less weight loss, recorded in EHRs.

Our study detected a robust genetic association with GLP1 medication weight loss efficacy and associated side effects. Although the genetic effect sizes we detected are modest, it is likely that additional data will reveal further associations and increase the predictive utility of genetics in this context. Future research should develop longitudinal datasets to enable analysis of weight loss and side effects in the context of dosage escalations, while exploring genetic factors, which could then inform how genetics might be leveraged in clinical decision-making at treatment initiation and beyond.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-026-10330-z>.

- Wilding, J. P. H. et al. Once-weekly semaglutide in adults with overweight or obesity. *New Engl. J. Med.* **384**, 989–1002 (2021).
- Bessesen, D. H. & Gaal, L. F. V. Progress and challenges in anti-obesity pharmacotherapy. *Lancet Diabetes Endocrinol.* **6**, 237–248 (2018).

3. Wing, R. R. et al. Cardiovascular effects of intensive lifestyle intervention in type 2 diabetes. *N. Engl. J. Med.* **369**, 145–154 (2013).
4. Montero, A., Sparks, G., Presiado, M. & Hamel, L. KFF Health Tracking Poll May 2024: the public's use and views of GLP-1 drugs. *KFF* <https://www.kff.org/health-costs/kff-health-tracking-poll-may-2024-the-publics-use-and-views-of-glp-1-drugs/> (2024).
5. Ryan, D. H. et al. Long-term weight loss effects of semaglutide in obesity without diabetes in the SELECT trial. *Nat. Med.* **30**, 2049–2057 (2024).
6. Aithal, G. P., Day, C. P., Kesteven, P. J. & Daly, A. K. Association of polymorphisms in the cytochrome P450 CYP2C9 with warfarin dose requirement and risk of bleeding complications. *Lancet* **353**, 717–719 (1999).
7. Chasman, D. I. et al. Genetic determinants of statin-induced low-density lipoprotein cholesterol reduction. *Circ. Cardiovasc. Genet.* **5**, 257–264 (2012).
8. Mallal, S. et al. HLA-B*5701 screening for hypersensitivity to abacavir. *N. Engl. J. Med.* **358**, 568–579 (2008).
9. Rodriguez, P. J. et al. Semaglutide vs tirzepatide for weight loss in adults with overweight or obesity. *JAMA Intern. Med.* **184**, 1056–1064 (2024).
10. Yang, Y. et al. Sex differences in the efficacy of glucagon-like peptide-1 receptor agonists for weight reduction: a systematic review and meta-analysis. *J. Diabetes* **17**, e70063 (2025).
11. German, J. et al. Association between plausible genetic factors and weight loss from GLP1-RA and bariatric surgery. *Nat. Med.* **31**, 2269–2276 (2025).
12. Guo, H. et al. Comparative efficacy and safety of GLP-1 receptor agonists for weight reduction: a model-based meta-analysis of placebo-controlled trials. *Obes. Pillars* **13**, 100162 (2025).
13. Wong, H. J. et al. Efficacy of GLP-1 receptor agonists on weight loss, BMI, and waist circumference for patients with obesity or overweight: a systematic review, meta-analysis, and meta-regression of 47 randomized controlled trials. *Diabetes Care* **48**, 292–300 (2025).
14. Jensterle, M., Rizzo, M., Haluzik, M. & Janež, A. Efficacy of GLP-1 RA approved for weight management in patients with or without diabetes: a narrative review. *Adv. Ther.* **39**, 2452–2467 (2022).
15. Chen, S. et al. A genomic mutational constraint map using variation in 76,156 human genomes. *Nature* **625**, 92–100 (2024).
16. The Wellcome Trust Case Control Consortium et al. Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nat. Genet.* **44**, 1294–1301 (2012).
17. Giambartolomei, C. et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).
18. Cerezo, M. et al. The NHGRI-EBI GWAS Catalog: standards for reusability, sustainability and diversity. *Nucleic Acids Res.* **53**, D998–D1005 (2024).
19. Vujkovic, M. et al. Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. *Nat. Genet.* **52**, 680–691 (2020).
20. Verma, A. et al. Diversity and scale: genetic architecture of 2068 traits in the VA Million Veteran Program. *Science* **385**, eadj1182 (2024).
21. Saunders, G. R. B. et al. Genetic diversity fuels gene discovery for tobacco and alcohol use. *Nature* **612**, 720–724 (2022).
22. Lau, D. C. W. et al. Once-weekly cagrilintide for weight management in people with overweight and obesity: a multicentre, randomised, double-blind, placebo-controlled and active-controlled, dose-finding phase 2 trial. *Lancet* **398**, 2160–2172 (2021).
23. Davies, M. et al. Semaglutide 2.4 mg once a week in adults with overweight or obesity, and type 2 diabetes (STEP 2): a randomised, double-blind, double-dummy, placebo-controlled, phase 3 trial. *Lancet* **397**, 971–984 (2021).
24. Levy, M. E. et al. Influence of BMI-associated genetic variants and metabolic risk factors on weight loss with semaglutide: a longitudinal clinico-genomic cohort study. Preprint at *medRxiv* <https://doi.org/10.1101/2024.10.31.24316494> (2024).
25. Jensterle, M., Pirš, B., Goričar, K., Dolžan, V. & Janež, A. Genetic variability in GLP-1 receptor is associated with inter-individual differences in weight lowering potential of liraglutide in obese women with PCOS: a pilot study. *Eur. J. Clin. Pharmacol.* **71**, 817–824 (2015).
26. Yu, M., Wang, K., Liu, H. & Cao, R. GLP1R variant is associated with response to exenatide in overweight Chinese Type 2 diabetes patients. *Pharmacogenomics* **20**, 273–277 (2019).
27. DeBoever, C. et al. Medical relevance of common protein-altering variants in GPCR genes across 337,205 individuals in the UK Biobank. Preprint at *bioRxiv* <https://doi.org/10.1101/2019.12.13.876250> (2020).
28. Gabe, M. B. N. et al. Enhanced agonist residence time, internalization rate and signalling of the GIP receptor variant [E354Q] facilitate receptor desensitization and long-term impairment of the GIP system. *Basic Clin. Pharmacol. Toxicol.* **126**, 122–132 (2020).
29. Borner, T. et al. GIP receptor agonism attenuates GLP-1 receptor agonist-induced nausea and emesis in preclinical models. *Diabetes* **70**, 2545–2553 (2021).
30. Vosoughi, K., Roghani, R. S. & Camilleri, M. Effects of GLP-1 agonists on proportion of weight loss in obesity with or without diabetes: systematic review and meta-analysis. *Obes. Med.* **35**, 100456 (2022).
31. Thomas, J., Paulet, M. & Rajpura, J. R. Consistency between self-reported and recorded values for clinical measures. *Cardiol. Res. Pract.* **2016**, 4364761 (2016).
32. Hattori, A. & Sturm, R. The obesity epidemic and changes in self-report biases in BMI. *Obesity* **21**, 856–860 (2013).
33. Chia, Y. C. et al. Measurement accuracy and reliability of self-reported versus measured weight and height among adults in Malaysia: findings from a nationwide blood pressure screening programme. *PLoS ONE* **18**, e0280483 (2023).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026

23andMe Research Team

Stella Aslibekyan¹, Adam Auton¹, Robert K. Bell¹, Katelyn Kukar Bond¹, Zayn Cochinwala¹, Sayantan Das¹, Kahsaia de Brito¹, Devika Dhamija¹, Payambr Dibaeinia¹, Emily DelloRusso¹, Chris Eijssbouts¹, Sarah L. Elson¹, Shirin Fuller¹, Chris German¹, Julie M. Granka¹, Larry Hengl¹, Barry Hicks¹, David A. Hinds¹, Michael V. Holmes¹, Reza Jabal¹, Aly Khan¹, Matthew J. Kmiecik¹, Alan Kwong¹, Yanyu Liang¹, Keng-Han Lin¹, Matthew H. McIntyre¹, Alex Moran¹, Carrie Northover¹, Shubham Saini¹, Anjali J. Shastri¹, Jingchunzi Shi¹, Suyash Shringarpure¹, Teague Sterling¹, Qiaojuan Jane Su¹, Vinh Tran¹, Joyce Y. Tung¹, Catherine H. Weldon¹ & Wanwan Xu¹

Methods

Overview of study recruitment

Participants in this study were recruited from the customer base of 23andMe. Participants provided informed consent and volunteered to participate in the research online, under a protocol approved by the external Association for the Accreditation of Human Research Protection Programs-accredited Salus Institutional Review Board (<https://www.versiticlinicaltrials.org/salusirb>). Participants were included in the analysis on the basis of consent status verified at the time data analyses were initiated.

The 23andMe GLP1 survey was launched to research participants in August 2024. The survey aimed to capture participants' experiences with GLP1 receptor agonist medication, and was targeted to 23andMe participants who had previously responded in the affirmative to the question, 'Have you ever taken prescription medications to help you lose weight?'. The survey included questions regarding drug brand, dosing regimen, time on treatment, efficacy (including pre-treatment weight and weight on treatment), and side effects, as well as reasons for pursuing or stopping GLP1 treatment. We focused the survey and subsequent analysis on primarily six drug varieties; Ozempic, Wegovy, compounded semaglutide, Mounjaro, Zepbound and compounded tirzepatide, the first three of which represent variations of semaglutide, and the last three represent variations in tirzepatide. A full list of survey questions can be found in Supplementary Table 24.

Phenotype definitions

Using the information derived from the surveys, we defined phenotypes that aimed to capture aspects of drug efficacy and side effects. We defined our efficacy phenotype as the contrast between pre-treatment BMI to post-treatment BMI (or current BMI, if treatment is on-going). In general, for study participants who reported taking more than one GLP1 medication, we selected the GLP1 medication that they reported taking for the longest period of time. Specifically, we defined a percentage BMI change phenotype as:

$$\Delta\text{BMI}_{\%} = 100(\text{BMI}_2 - \text{BMI}_1)/\text{BMI}_1$$

where BMI_1 and BMI_2 represent pre-treatment and post-treatment BMI, respectively, measured in weight in kilograms per height in metres squared. We applied quality control filters to people with weight less than 36 kg or greater than 181 kg, height less than 1.39 m or greater than 2.06 m, BMI less than 14 kg m^{-2} or greater than 70 kg m^{-2} , or age less than 18 years. In aggregate, these initial filters removed 80 people (0.29%). Inspection of the $\Delta\text{BMI}_{\%}$ phenotype revealed a heavy tailed distribution, so we further quality controlled the $\Delta\text{BMI}_{\%}$ phenotype to remove outlier participants with BMI changes above 20% or below -45% (Extended Data Fig. 10). The $\Delta\text{BMI}_{\%}$ estimates were set to missing for participants who did not pass quality control.

To enable genetic associations to be interpreted in units of weight rather than $\Delta\text{BMI}_{\%}$, we also defined a corresponding Δweight phenotype, defined as the change in weight from baseline in kilograms. We note that, because adult height is treated as constant during the treatment window, the percentage change in BMI ($\Delta\text{BMI}_{\%}$) is mathematically identical to the percentage change in weight ($\Delta\text{weight}_{\%}$).

For the side effect phenotypes, we defined separate case-control phenotypes for each side effect recorded in the survey, contrasting those who self-rated their side effects as moderate or severe (cases) to those who self-rated their side effects as mild or non-existent (controls). As before, for study participants who reported taking more than one GLP1 medication, we selected the GLP1 medication that they reported taking for the longest period of time.

We further defined phenotypes to represent covariates, specifically for drug type (semaglutide = 1 versus tirzepatide = 0), dosage and days on treatment. For the dosage phenotype, we used the reported most

recent weekly dosage in milligrams; this was either the final dose or the current dose for people still taking medication.

Comparison of self-report and EHR data

As part of the 23andMe experience, research participants are offered the opportunity to share EHR information collected on their Apple iPhone devices. Specifically, the Apple Health application enables connection to healthcare providers for the purposes of sharing EHR information with third parties through Apple HealthKit (<https://developer.apple.com/documentation/healthkit>). 23andMe research participants can elect to share their EHR information for research purposes. We used these data to perform comparisons with the self-report survey data. Full details of comparison analyses are provided in Supplementary Information.

Non-genetic predictors of BMI loss

To analyse the dependence of achieved BMI loss on non-genetic factors such as drug type, dosage and time on treatment, we fit the following model:

$$\begin{aligned} \Delta\text{BMI}_{\%} \sim & \text{age} + \text{sex} + \text{BMI}_1 + \text{drugType} + \text{dose} + \text{days}_{\text{treat}} \\ & + \text{drugType}:\text{dose} + \text{drugType}:\text{days}_{\text{treat}} \\ & + \text{dose}:\text{days}_{\text{treat}} + \text{dose}:\text{days}_{\text{treat}}:\text{drugType} \end{aligned} \quad (1)$$

where 'drugType' is an indicator variable that equals 1 for individuals using semaglutide and 0 for tirzepatide, 'dose' represents the dose in milligrams, $\text{days}_{\text{treat}}$ represents the total days on the relevant drug and ':' represents an interaction term between two or more variables. Note that semaglutide and tirzepatide typically have different standard dosing levels, which is handled in the regression model by the 'drugType:dose' interaction term.

Genotyping and SNP imputation

DNA extraction and genotyping were performed on saliva samples by Clinical Laboratory Improvement Amendments-certified and College of American Pathologists-accredited clinical laboratories of Laboratory Corporation of America. Samples were genotyped on one of five genotyping platforms. The V1 and V2 platforms were variants of the Illumina HumanHap550 BeadChip and contained a total of about 560,000 SNPs, including about 25,000 custom SNPs selected by 23andMe. The V3 platform was based on the Illumina OmniExpress BeadChip and contained a total of about 950,000 SNPs and custom content to improve the overlap with our V2 array. The V4 platform was a fully custom array of about 950,000 SNPs and included a lower redundancy subset of V2 and V3 SNPs with additional coverage of lower-frequency coding variation. The V5 platform was based on the Illumina Global Screening Array, consisting of approximately 654,000 preselected SNPs and approximately 50,000 custom content variants. Participant genotype data were imputed against a reference panel composed of data from the Haplotype Reference Consortium³⁴ and augmented with additional sequences to boost imputation performance (Supplementary Information).

Association testing

We performed a GWAS of $\Delta\text{BMI}_{\%}$ in people of European ancestry using methods that have been described previously³⁵. In brief, unrelated participants were included in the GWAS analyses on the basis of European ancestry as determined by a genetic ancestry classification algorithm³⁶. The GWAS was performed including covariates as described in equation 1 above, with the addition of five genetic principal components to account for fine-scale genetic ancestry, and indicator variables to account for variation in the genotyping platform. Among 21,822 people of European ancestry, we required participants to have complete data needed to construct the target phenotype and GWAS covariates (that

is, data available for pre-treatment weight, post-treatment weight, drug type, dosage, time on treatment and factors such as age, sex and height), resulting in 18,488 participants. Finally, participants were filtered on relatedness such that no two people shared more than 700 cM identity by descent³⁷, which corresponds approximately to the minimal expected sharing between first cousins in an outbred population, resulting in a final GWAS sample size of 15,237. An equivalent procedure was used for a GWAS of side effect phenotypes. For the purposes of testing drug-specific associations, we repeated the GWAS procedure for the semaglutide and tirzepatide-treated populations separately, removing the drug-type covariate and interaction terms as appropriate. All GWASs were adjusted for inflation using genomic control, with the inflation factor being no more than 1.035 in all phenotypes.

Given the smaller sample sizes available in non-European populations, we did not perform genome-wide association testing in these populations, and instead focused analyses on variants discovered as associated in the European GWAS. For these variants, we tested for association in non-European populations following a similar approach to that described above.

Replication

We performed replication of the identified efficacy association in the All of Us cohort³⁸, using Controlled Tier Dataset v.8. We extracted genomic data, EHR data and a drug code referring to either semaglutide or tirzepatide from 9,579 participants. After filtering to retain participants with information regarding pre-treatment and post-treatment BMI and genotype data passing quality control, we obtained 4,889 participants, of which 3,948 had complete data when incorporating covariates akin to those used in the GWAS. For the replication analysis, we tested for association between the EHR-derived $\Delta\text{BMI}_{\%}$ and the genotype, including covariates. We repeated the replication analysis having performed mean-imputation of missing drug dose data, allowing a larger sample size of 4,855 to be analysed.

We also attempted replication analysis in the UK Biobank cohort, although the available data predate the availability of semaglutide or tirzepatide, and hence relied on earlier variants of GLP1 receptor agonists. Full details of the replication analysis methodology is provided in Supplementary Information.

Genetic and non-genetic risk modelling

To construct combined genetic and non-genetic models of $\Delta\text{BMI}_{\%}$ and risk of treatment-related side effects, we selected treatment, clinical, demographic, disease diagnosis and genetic variables as predictors. In addition to the covariates included in the GWAS, we also included years of education as a proxy for socio-economic status, and binary indicators of previous disease diagnosis for T2D, hypertension and non-alcoholic fatty liver disease. All continuous predictor variables were standardized before modelling to allow for the comparison of effect sizes.

We used a linear multi-variable model to fit $\Delta\text{BMI}_{\%}$. Given the binary nature of side effect phenotype definitions, we fitted multi-variable logistic regression models (equivalent to a generalized linear model with a binomial family and logit link function). The dataset was partitioned randomly into training (70% of the sample) and held out test (30%) sets, with the test set being used to assess model performance. Further details are outlined in Supplementary Information.

Model performance of efficacy was further assessed by applying the model derived from our self-report data in a sample of 642 people who had provided HealthKit EHR data but had not completed the GLP1 survey, and hence were not used in the construction of the model. To replicate the situation where efficacy predictions are made before treatment, we assumed the dose, treatment duration and drug type variables were unknown, and imputed these values in the model to an arbitrary constant value for all participants.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The full set of de-identified summary statistics can be made available to qualified investigators who enter into an agreement with 23andMe that protects participant confidentiality. Interested investigators should visit the 23andMe Dataset Access Program at <https://research.23andme.com/dataset-access/>. Summary statistics for all SNPs with $P < 1 \times 10^{-4}$ from all primary GWASs are provided in the Supplementary Data. Individual-level data from 23andMe participants used in the analyses are not publicly available due to participant confidentiality and in accordance with the Institutional Review Board-approved protocol under which the study was conducted.

34. McCarthy, S. et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).
35. Shelton, J. F. et al. Trans-ancestry analysis reveals genetic and nongenetic associations with COVID-19 susceptibility and severity. *Nat. Genet.* **53**, 801–808 (2021).
36. Durand, E. Y. et al. A scalable pipeline for local ancestry inference using tens of thousands of reference haplotypes. Preprint at *bioRxiv* <https://doi.org/10.1101/2021.01.19.427308> (2021).
37. Henn, B. M. et al. Cryptic distant relatives are common in both isolated and cosmopolitan genetic samples. *PLoS ONE* **7**, e34267 (2012).
38. The All of Us Research Program Genomics Investigators et al. Genomic data in the All of Us Research Program. *Nature* **627**, 340–346 (2024).

Acknowledgements We thank the 23andMe research participants and employees who made this study possible. We thank A. Guan and J. Wen for analytical validation of the identified genetic variants. 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 further thank the UK Biobank Resource, with the replication analysis in this study having been conducted under application number 95801.

Author contributions N.S.A.-H., S.A., M.V.H., B.L.K. and A.A. designed the study. V.T., R.R.W. and N.S.A.-H. developed the recruitment and participant engagement strategy and acquired and processed the data. Q.J.S., J.R.A., W.X., C.W., J.S., B.H. and A.A. analysed the data. Q.J.S., J.R.A., W.X., R.R.W., N.S.A.-H., S.A., M.V.H., B.L.K. and A.A. interpreted the data. Q.J.S., M.V.H. and A.A. wrote the manuscript. All authors participated in the preparation of the manuscript by reading and commenting on the drafts before submission.

Competing interests All authors are current or former employees of 23andMe. N.S.A.-H. holds an academic appointment at the Icahn School of Medicine at Mount Sinai. M.V.H. is currently affiliated with Genentech.

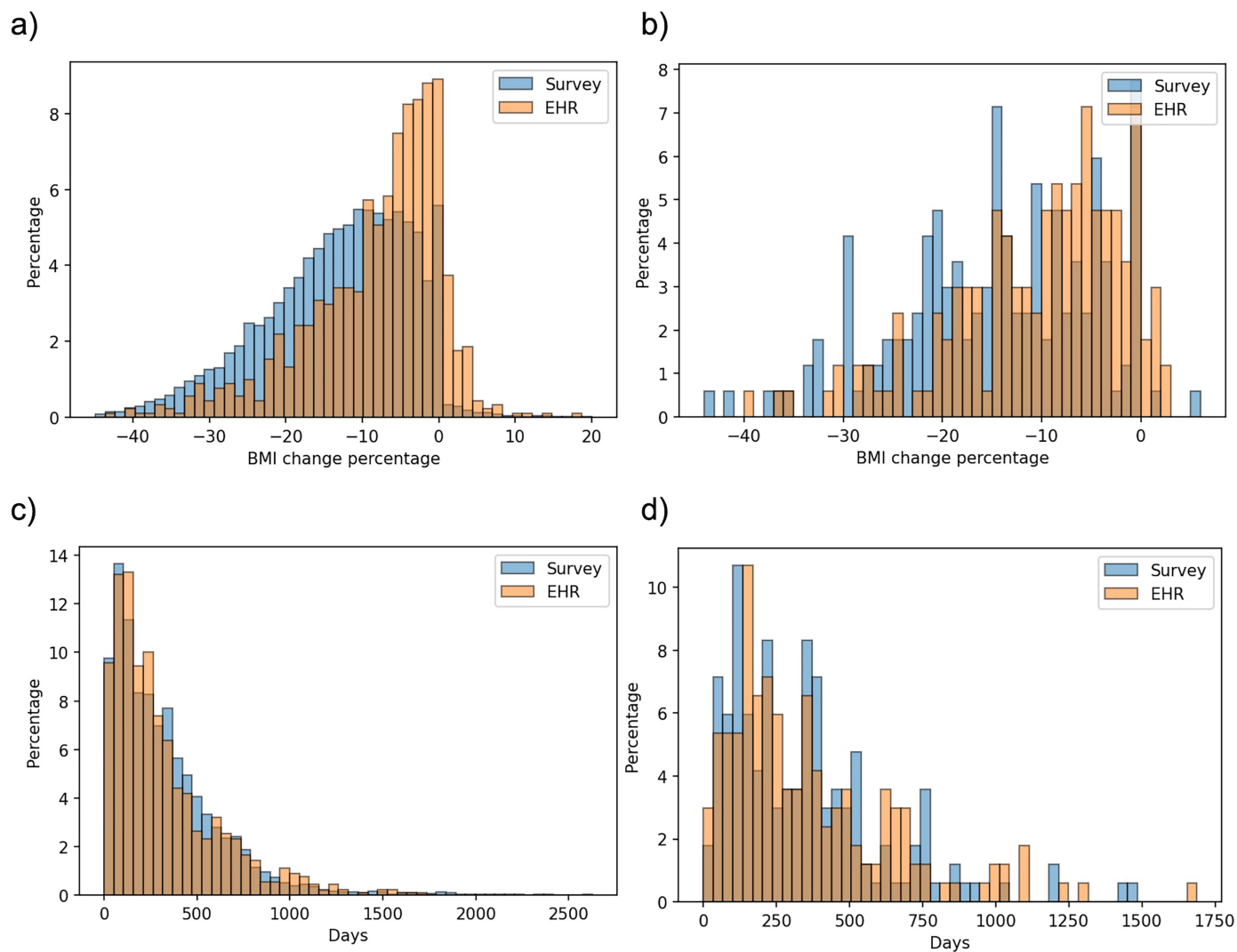
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-026-10330-z>.

Correspondence and requests for materials should be addressed to Adam Auton.

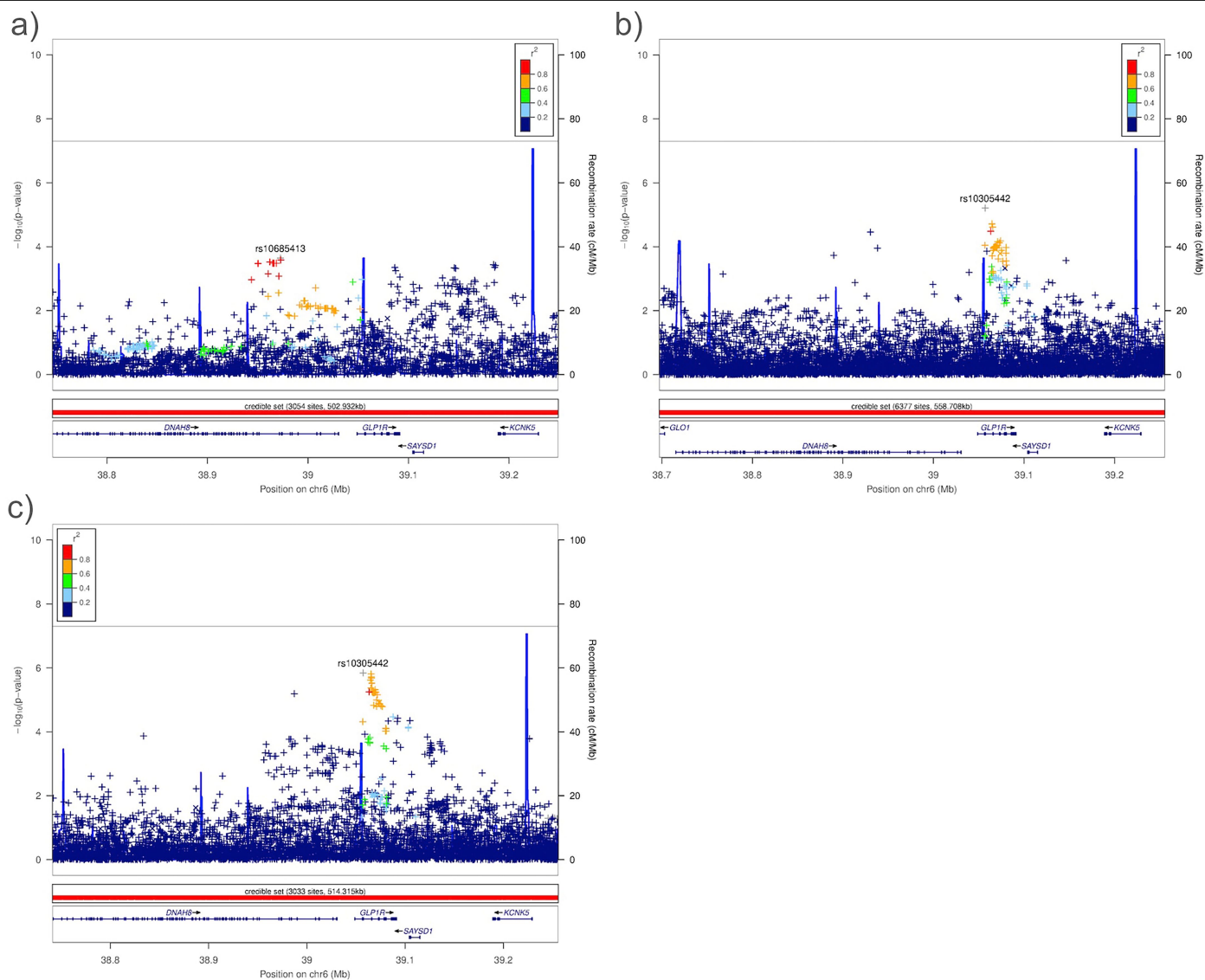
Peer review information *Nature* thanks Dipender Gill, Ruth Loos and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



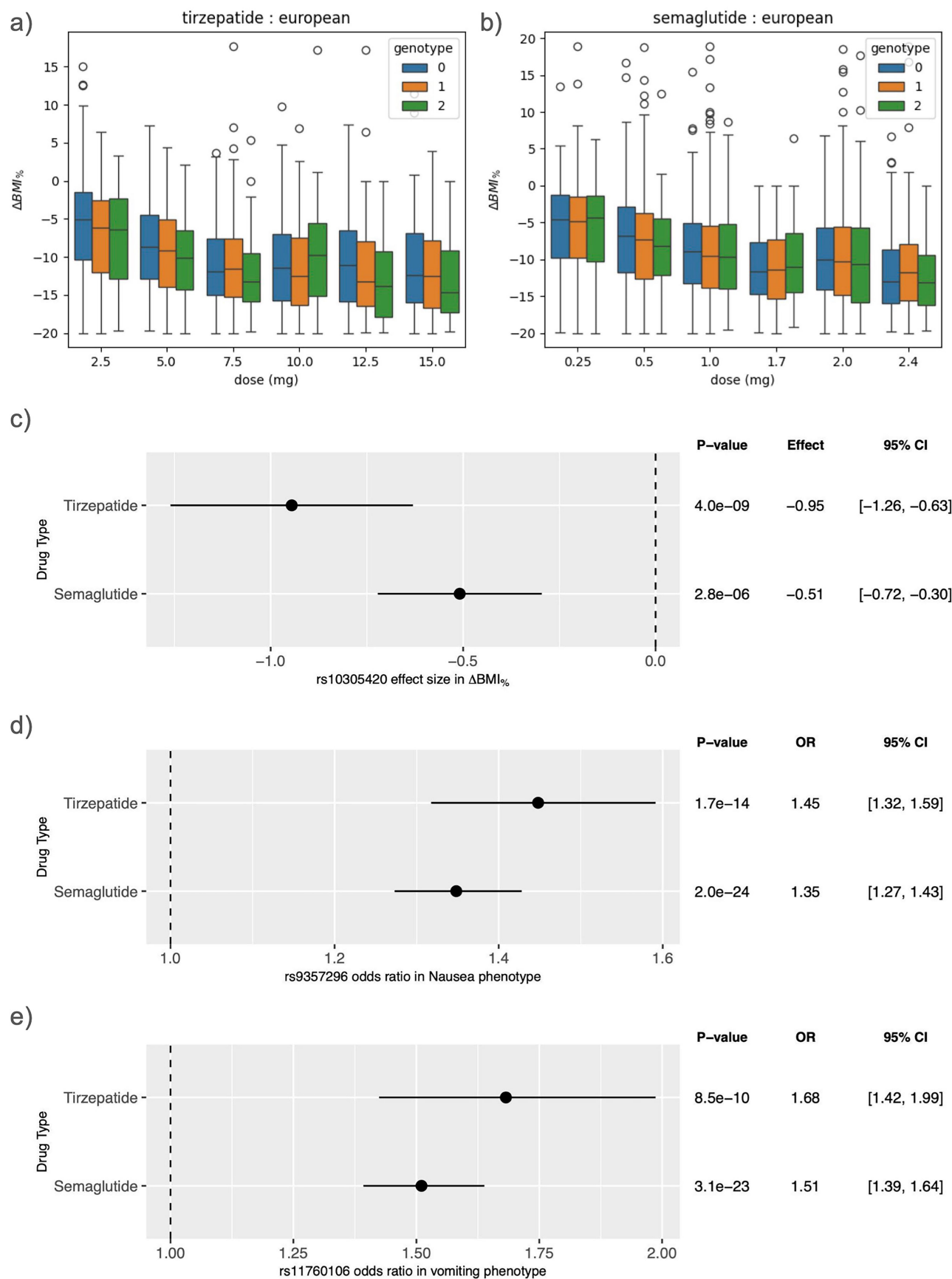
Extended Data Fig. 1 | Comparison of self-report and EHR data. a) Distribution of $\Delta BMI_{\%}$ as recorded in all individuals with survey data and to 909 individuals with EHR data. b) Distribution of $\Delta BMI_{\%}$ as recorded in individuals with measurements from both survey and EHR. c) Distribution of total treatment

time, measured in days, as recorded in all individuals with survey data and to 909 individuals with EHR data. d) Distribution of total treatment time, measured in days, as recorded in individuals with measurements from both survey and EHR.



Extended Data Fig. 2 | Conditional analysis of *GLP1R* locus. Each figure here shows the residual association signal having conditioned on the index SNP of the relevant GWAS. a) $\Delta\text{BMI}_{\%}$ conditioned on rs10305420. b) GLP-1 medication

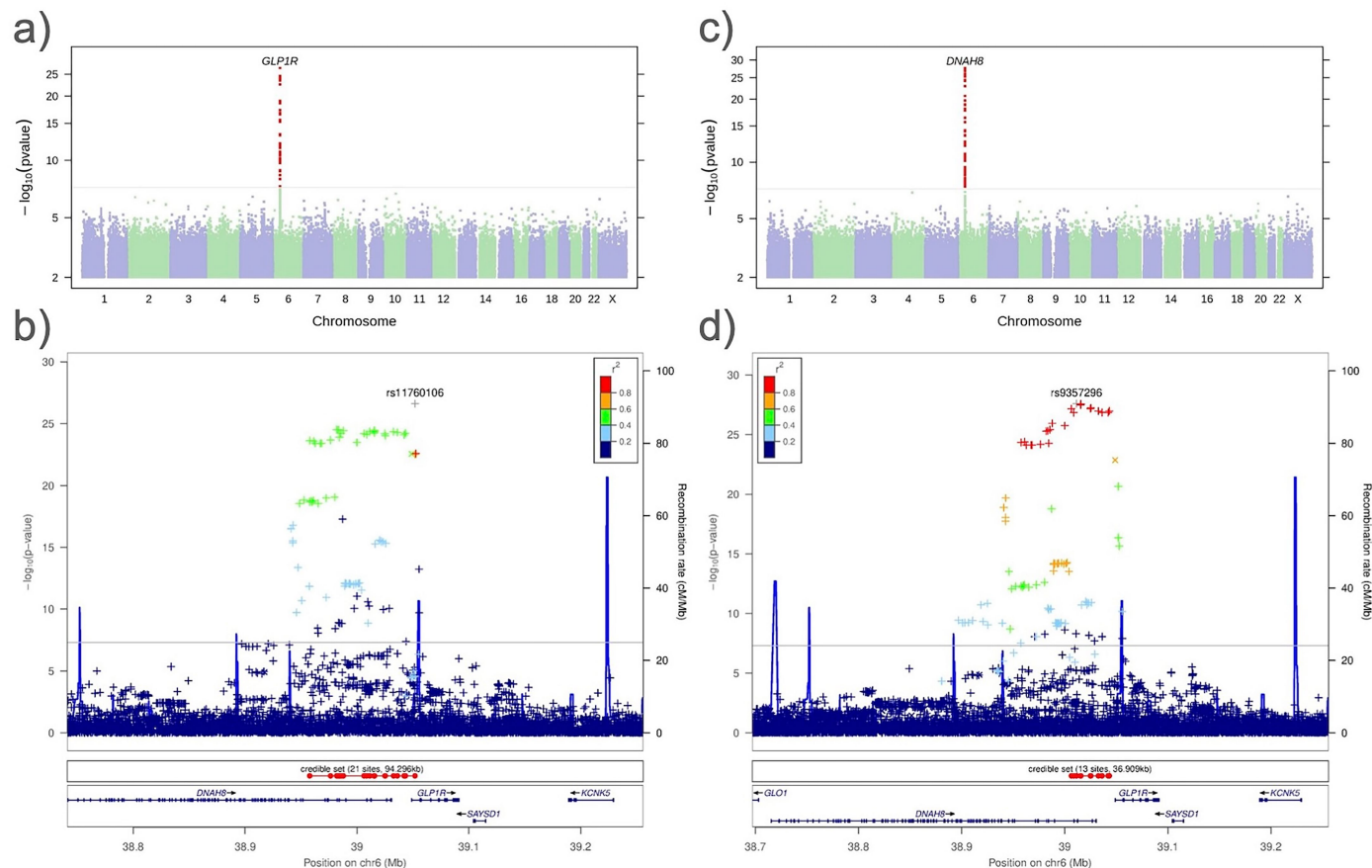
induced nausea conditioned on rs9357296. c) GLP-1 medication induced vomiting conditioned on rs11760106.



Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Effects of GWAS index SNPs in the vicinity of *GLPIR* locus on efficacy and side effects, partitioned by drug type. a) $\Delta\text{BMI}_{\%}$ stratified by dose and rs10305420 genotype for tirzepatide, as measured in individuals of European ancestry. The boxes represent the quartiles of the data while the whiskers represent 1.5 times the interquartile range. Outliers are

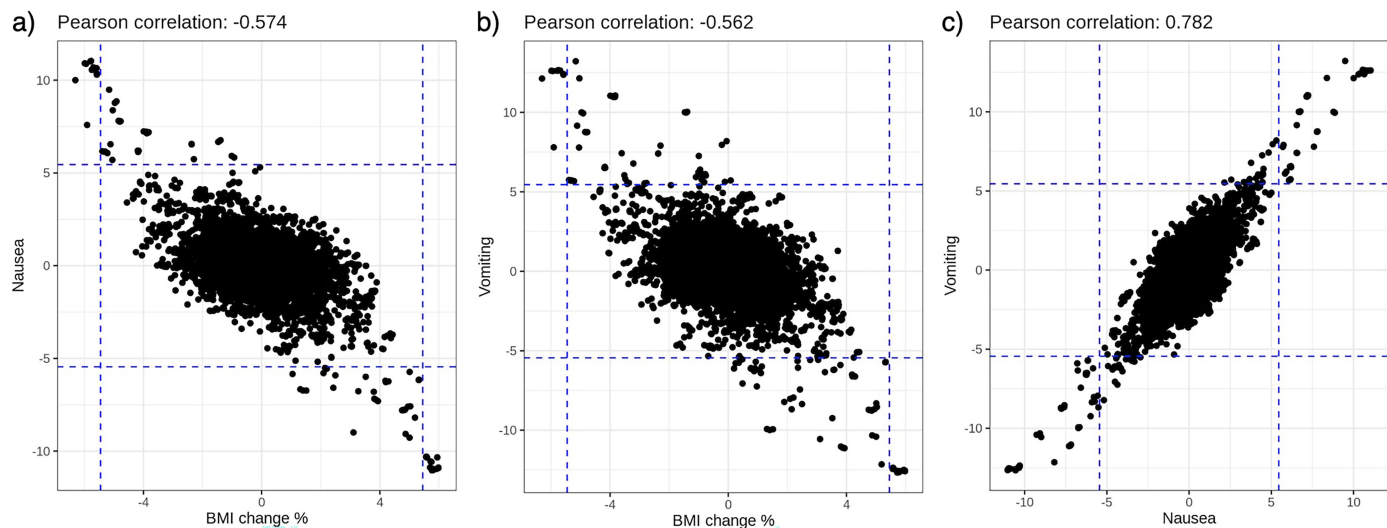
indicated by circles. b) Same plot for semaglutide. c) Effect size estimates for the rs10305420 association with $\Delta\text{BMI}_{\%}$. d) Effect size estimates for the rs9357296 association with nausea. e) Effect size estimates for the rs11760106 association with vomiting. For plots c, d, and e, point estimates are represented as circles, and horizontal bars indicate 95% confidence intervals.



Extended Data Fig. 4 | GWAS of side effects while on GLP-1 medication.

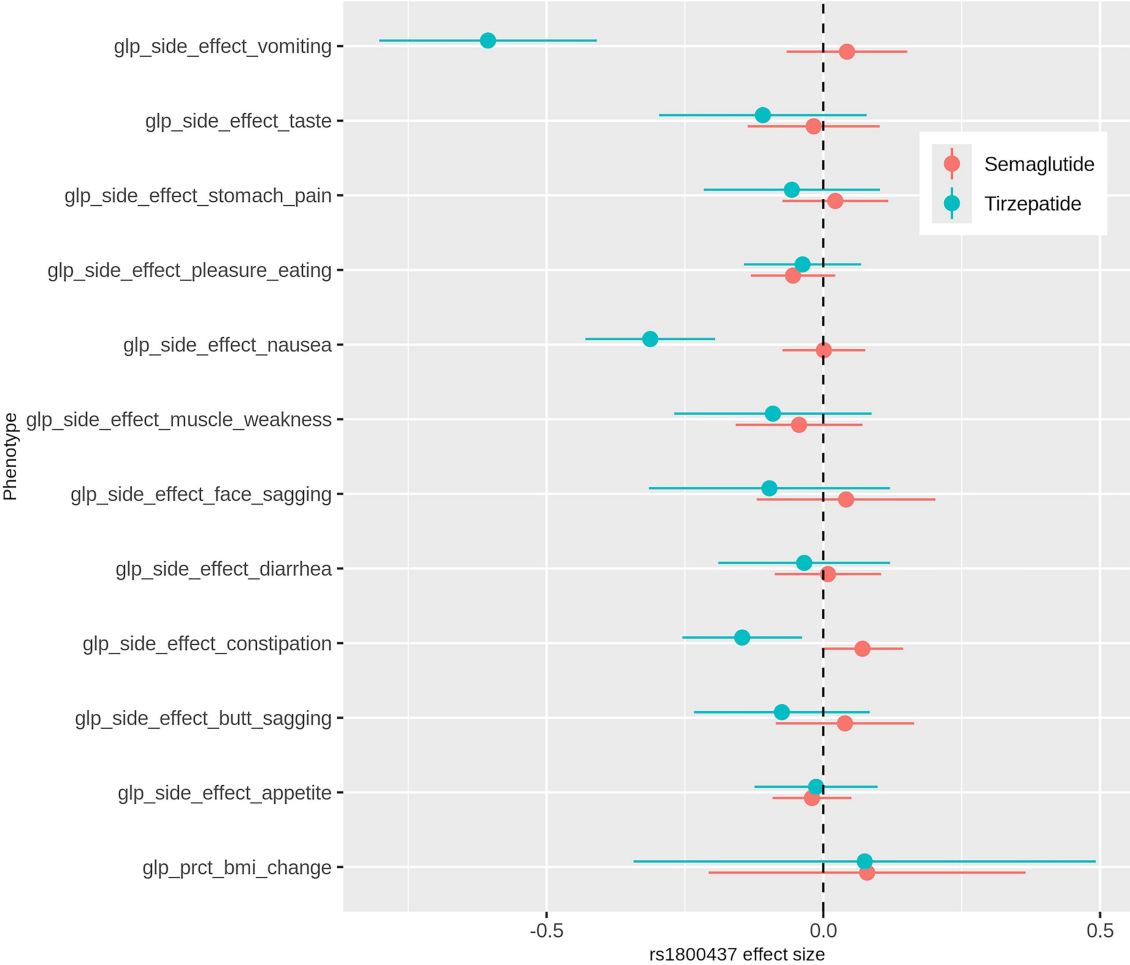
a) Manhattan plot of experiencing vomiting while on GLP-1 treatment. SNPs achieving genome-wide significance ($p < 5e-8$) are highlighted in red. b) Regional plot around vomiting association in the *GLP1R* locus. c) Manhattan plot of experiencing nausea while on GLP-1 treatment. d) Regional plot around

nausea association in the *GLP1R* locus. For regional plots, colors indicate the strength of linkage disequilibrium (r^2) relative to the index SNP. Non-coding variants are indicated with '+' symbols; coding variants are indicated with 'x' symbols.



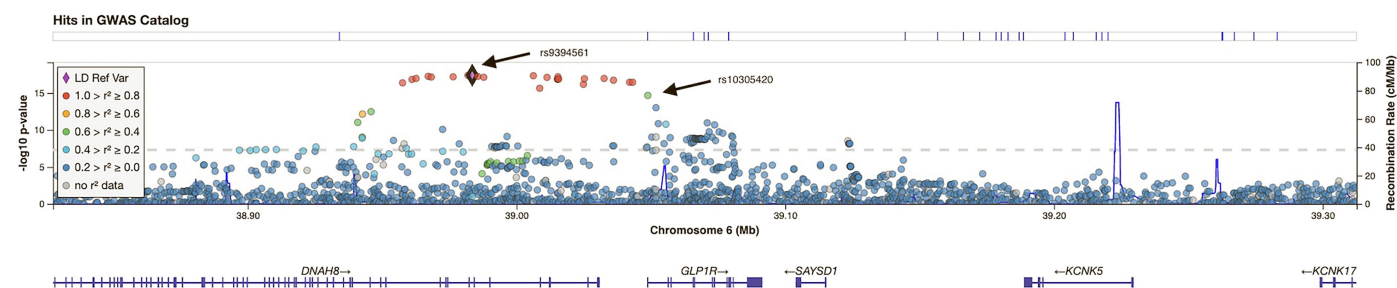
Extended Data Fig. 5 | Colocalization analysis between three traits in the *GLP1R* locus. Colocalization was performed in a 1 Mb window centered on rs10305420. a) $\Delta\text{BMI}_{\%}$ vs nausea. Note that negative BMI change numbers

represent BMI loss, and therefore a negative correlation implies increased drug efficacy with increased nausea. b) $\Delta\text{BMI}_{\%}$ vs vomiting. c) Nausea vs Vomiting.



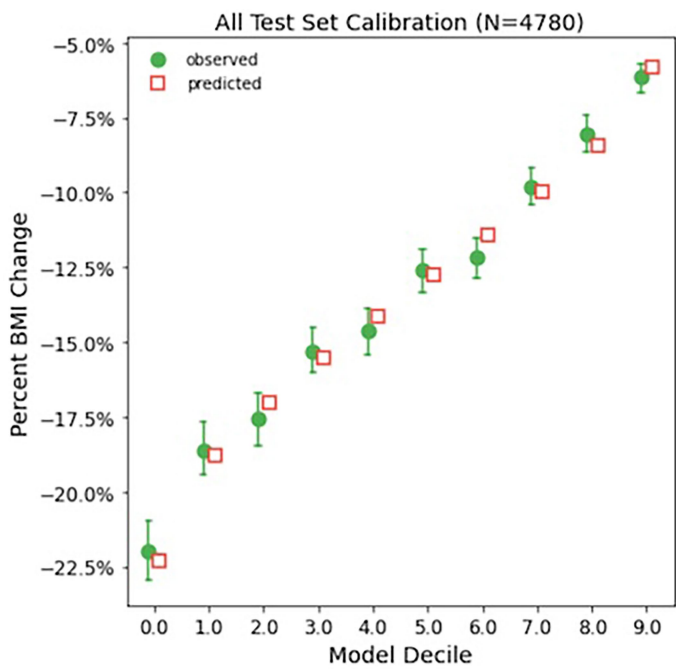
Extended Data Fig. 6 | Estimated effect sizes of rs1800437 across study phenotypes. Point estimates are represented as circles, and horizontal bars indicate 95% confidence intervals. Estimates shown for both tirzepatide (cyan)

and semaglutide (red) treated populations. For the side effect phenotypes, the effect size represents log-odds ratio, whereas for the efficacy phenotype, the effect size represents percentage BMI change.

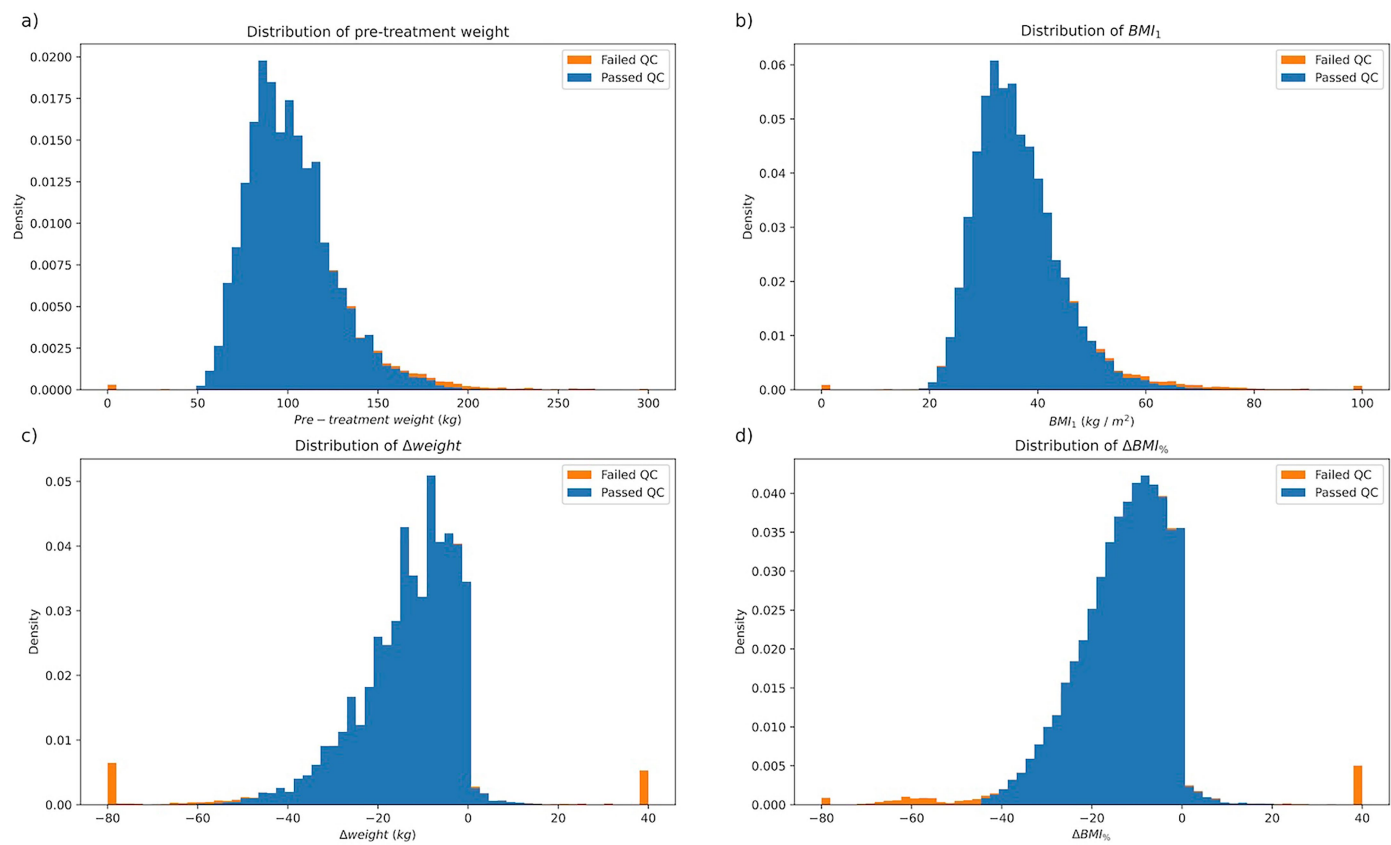


Extended Data Fig. 7 | Million Veterans Project BMI association in the vicinity of the GLP1R locus. This figure shows a regional plot for the BMI, maximum, inv-normal transformed phenotype. The index SNP (rs9394561)

and the missense SNP identified in this study (rs10305420) are highlighted. Colors indicate strength of LD.



Extended Data Fig. 8 | Calibration of efficacy model. This figure shows predicted $\Delta\text{BMI}_{\%}$ compared to observed values, in deciles, as measured in the held-out test set. Predicted values from the model are shown as red squares. Sample means for each decile are shown as green circles (for observed data) or red squares (for sample predictions), with error bars representing 95% confidence intervals.



Extended Data Fig.10 | Distributions of pre- and post-treatment measurements before and after Quality Control (QC) filtering.
The histograms show data that was removed following QC (orange) and data

retained for further analysis (blue). a) Distribution of pre-treatment weight. b) Distribution of pre-treatment BMI. c) Distribution of change in weight in kilograms. d) Distribution of the $\Delta BMI_{\%}$ phenotype.

Extended Data Table 1 | Demographics and treatment response characteristics of study population

	All study participants	Ozempic	Wegovy	Compounded semaglutide	Mounjaro	Zepbound	Compounded tirzepatide
Sample size, n	27,885	11,971	3,821	3,777	4,650	2,321	1,345
Percent of total	100%	42.90%	13.70%	13.50%	16.70%	8.30%	4.80%
Median age (s.d.)	52 (13.4)	53 (13.9)	49 (12.8)	49 (12.8)	54 (13.0)	50 (12.8)	50 (12.4)
Female, n (%)	22,977 (82.4%)	9,664 (80.7%)	3,221 (84.3%)	3,302 (87.4%)	3,710 (79.8%)	1,903 (82.0%)	1,177 (87.5%)
Ancestry							
European, n (%)	21,822 (78.3%)	9,102 (76.0%)	2,980 (78.0%)	3,038 (80.4%)	3,738 (80.4%)	1,869 (80.5%)	1,095 (81.4%)
Latino, n (%)	3,604 (12.9%)	1,741 (14.5%)	457 (12.0%)	464 (12.3%)	537 (11.5%)	240 (10.3%)	165 (12.3%)
African American, n (%)	1,174 (4.2%)	526 (4.4%)	208 (5.4%)	119 (3.2%)	188 (4.0%)	102 (4.4%)	31 (2.3%)
Other, n (%)	1,285 (4.6%)	602 (5.0%)	176 (4.6%)	156 (4.1%)	187 (4.0%)	110 (4.7%)	54 (4.0%)
Efficacy							
Median pre-treatment BMI, kg/m ² (s.d.)	35.1 (7.6)	35.3 (7.8)	36.0 (7.0)	33.1 (7.0)	36.2 (7.9)	35.6 (7.1)	32.4 (7.2)
Median pre-treatment weight, kg (s.d.)	98.4 (23.7)	99.8 (24.2)	99.8 (22.6)	91.6 (21.4)	101.6 (24.7)	99.8 (22.3)	89.8 (22.3)
Median months on treatment (s.d.)	8.3 (10.2)	10.0 (12.7)	7.6 (8.2)	6.3 (6.1)	10.6 (7.9)	5.6 (4.3)	7.3 (5.9)
Mean dose, mg (s.d.)	N/A (N/A)	1.1 (0.7)	1.5 (0.8)	0.9 (0.6)	8.4 (4.1)	7.4 (3.8)	7.5 (4.0)
Median post-treatment BMI, kg/m ² (s.d.)	30.2 (7.4)	30.8 (7.6)	31.0 (7.0)	29.1 (6.8)	29.8 (7.6)	30.3 (6.9)	27.6 (6.7)
Median post-treatment weight, kg (s.d.)	84.4 (22.8)	86.2 (23.5)	86.2 (22.2)	81.2 (20.6)	84.4 (23.5)	84.4 (21.6)	77.1 (20.6)
Median BMI change, kg/m ² (s.d.)	-4.1 (3.7)	-3.7 (3.6)	-4.3 (3.8)	-3.5 (3.1)	-5.3 (4.3)	-4.6 (3.6)	-4.1 (3.3)
Median % change BMI (s.d.)	-11.7% (9.5)	-10.7% (9.2)	-11.9% (9.6)	-10.5% (8.4)	-15.1% (10.5)	-13.1% (9.1)	-12.8% (8.9)
Reason for treatment							
To improve blood sugar, n (%)	10,689 (23.1%)	5,315 (28.9%)	684 (12.1%)	667 (12.6%)	2,497 (32.2%)	409 (11.4%)	273 (13.4%)
To reduce cardiovascular risk, n (%)	7,180 (15.5%)	2,628 (14.3%)	1,041 (18.4%)	789 (14.9%)	1,189 (15.3%)	692 (19.3%)	367 (18.0%)
For weight loss, n (%)	26,799 (57.8%)	9,892 (53.8%)	3,712 (65.5%)	3,661 (69.0%)	3,794 (49.0%)	2,281 (63.5%)	1,310 (64.1%)
Other, n (%)	1,671 (3.6%)	544 (3.0%)	230 (4.1%)	190 (3.6%)	269 (3.5%)	210 (5.8%)	94 (4.6%)
Moderate / severe side effects							
Constipation, n (%)	7,612 (27.3%)	3,189 (26.6%)	1,150 (30.1%)	1,069 (28.3%)	1,167 (25.1%)	553 (23.8%)	300 (22.3%)
Diarhea, n (%)	3,461 (12.4%)	1,636 (13.7%)	458 (12.0%)	423 (11.2%)	515 (11.1%)	208 (9.0%)	102 (7.6%)
Nausea, n (%)	7,006 (25.1%)	3,165 (26.4%)	1,110 (29.0%)	991 (26.2%)	897 (19.3%)	456 (19.6%)	193 (14.3%)
Vomiting, n (%)	2,353 (8.4%)	1,187 (9.9%)	405 (10.6%)	284 (7.5%)	253 (5.4%)	109 (4.7%)	34 (2.5%)
Loss of appetite, n (%)	16,417 (58.9%)	5,918 (49.4%)	2,352 (61.6%)	2,344 (62.1%)	2,980 (64.1%)	1,627 (70.1%)	912 (67.8%)
Unpleasant taste sensation, n (%)	2,132 (7.6%)	981 (8.2%)	302 (7.9%)	253 (6.7%)	316 (6.8%)	143 (6.2%)	75 (5.6%)
Stomach pain, n (%)	3,492 (12.5%)	1,702 (14.2%)	454 (11.9%)	445 (11.8%)	484 (10.4%)	189 (8.1%)	105 (7.8%)
Loss of pleasure from eating, n (%)	7,180 (25.7%)	2,723 (22.7%)	1,032 (27.0%)	957 (25.3%)	1,303 (28.0%)	641 (27.6%)	375 (27.9%)
Muscle Weakness, n (%)	2,277 (8.2%)	1,027 (8.6%)	294 (7.7%)	296 (7.8%)	348 (7.5%)	159 (6.9%)	94 (7.0%)
Facial skin sagging ('Ozempic face'), n (%)	1,204 (4.3%)	554 (4.6%)	144 (3.8%)	109 (2.9%)	240 (5.2%)	70 (3.0%)	58 (4.3%)
Sagging skin around buttocks ('Ozempic butt'), n (%)	2,202 (7.9%)	852 (7.1%)	291 (7.6%)	233 (6.2%)	479 (10.3%)	181 (7.8%)	132 (9.8%)

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
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☐ | ☒ | 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
<i>Give P values as exact values whenever suitable.</i> |
| ☒ | ☐ | 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 Phenotype data was collected via the 23andMe Research platform. No additional software was used.

Data analysis Standard GWAS analyses were performed using the 23andMe Research platform, as described in numerous peer-reviewed publications (<https://research.23andme.com/publications/>). Additional analyses were conducted in R version 4.4.0, python3, and statsmodels 0.14.1.

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](#)

The full set of de-identified summary statistics can be made available to qualified investigators who enter into an agreement with 23andMe that protects participant confidentiality. Interested investigators should visit the 23andMe Dataset Access Program at <https://research.23andme.com/dataset-access/>. Individual-level data

from 23andMe participants used in the analyses are not publicly available due to participant confidentiality and in accordance with the IRB-approved protocol under which the study was conducted.

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	In the context of this study, biological sex was determined on the based on the number of X and Y chromosomes detected in the genetic sample.
Reporting on race, ethnicity, or other socially relevant groupings	Ancestry groups were determined by a genetic ancestry classification algorithm, as detailed in Durand, E. Y. et al. A scalable pipeline for local ancestry inference using tens of thousands of reference haplotypes. bioRxiv 2021.01.19.427308 (2021) doi:10.1101/2021.01.19.427308.
Population characteristics	Details of the population characteristics are described in Table 1.
Recruitment	Participants in this study were recruited from the customer base of 23andMe Research Institute, a nonprofit medical research organization. Participants provided informed consent and volunteered to participate in the research online, under a protocol approved by the external AAHRPP-accredited Salus IRB (https://www.versitclinicaltrials.org/salusirb). Participants were included in the analysis on the basis of consent status verified at the time data analyses were initiated. Given this approach to recruitment, the participants represent a self-selected cohort which may induce the generalizability of the findings.
Ethics oversight	The research protocol is approved by the external AAHRPP-accredited Salus IRB (https://www.versitclinicaltrials.org/salusirb)

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](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

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

Sample size	Sample size was determined on the basis of the number of respondents to online surveys and was judged appropriate to adequately power genome-wide association studies. A freeze of the data collection was taken in August 2025, with all subsequent analyses derived from that data freeze.
Data exclusions	Only 23andMe Research participants that consented to participate in research and completed the online survey were included. Additional exclusion criteria for specific analyses are detailed in the manuscript for QC purposes.
Replication	Replication was performed in the All of Us and UK Biobank cohorts, as detailed in the manuscript. The All of Us replication successfully replicated the efficacy finding in the manuscript. The UK Biobank replication did not replicate the finding, but had limited power and differed substantially in terms of the medications under consideration.
Randomization	This is an observations population study for which randomization is not applicable.
Blinding	This is an observations population study for which blinding is not applicable.

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	Involvement 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	Involvement in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks

n/a

Novel plant genotypes

n/a

Authentication

n/a