# Impact of weight loss on cancer-related proteins in serum: results from a cluster randomised controlled trial of individuals with type 2 diabetes

Caroline J. Bull,[a,b,c,x] Emma Hazelwood,[a,b,x] Danny N. Legge,[c] Laura J. Corbin,[a,b] Tom G. Richardson,[a,b] Matthew Lee,[d] James Yarmolinsky,[a,b] Karl Smith-Byrne,[e] David A. Hughes,[a,b] Mattias Johansson,[d] Ulrike Peters,[f] Sonja I. Berndt,[g] Hermann Brenner,[h,i,j] Andrea Burnett-Hartman,[k] Iona Cheng,[l] Sun-Seog Kweon,[m,n] Loic Le Marchand,[o] Li Li,[p] Polly A. Newcomb,[f,q] Rachel Pearlman,[r] Alex McConnachie,[s] Paul Welsh,[t] Roy Taylor,[u] Mike E. J. Lean,[v] Naveed Sattar,[t] Neil Murphy,[d] Marc J. Gunter,[d,w] Nicholas J. Timpson,[a,b] and Emma E. Vincent[a,b,c,*]

[a]MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK
[b]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
[c]School of Translational Health Sciences, Dorothy Hodgkin Building, University of Bristol, Bristol, UK
[d]Section of Nutrition and Metabolism, International Agency for Research on Cancer, WHO, Lyon, France
[e]Cancer Epidemiology Unit, Oxford Population Health, University of Oxford, UK
[f]Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA
[g]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA
[h]Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany
[i]Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany
[j]German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany
[k]Institute for Health Research, Kaiser Permanente Colorado, Denver, CO, USA
[l]Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, USA
[m]Department of Preventive Medicine, Chonnam National University Medical School, Gwangju, Korea
[n]Jeonnam Regional Cancer Center, Chonnam National University Hwasun Hospital, Hwasun, Korea
[o]University of Hawaii Cancer Center, Honolulu, HI, USA
[p]Department of Family Medicine, University of Virginia, Charlottesville, VA, USA
[q]School of Public Health, University of Washington, Seattle, WA, USA
[r]Division of Human Genetics, Department of Internal Medicine, The Ohio State University Comprehensive Cancer Center, Columbus, OH, USA
[s]Robertson Centre for Biostatistics, Institute of Health and Wellbeing, University of Glasgow, Glasgow, G12 8QQ, UK
[t]School of Cardiovascular and Metabolic Health, University of Glasgow, Glasgow, UK
[u]Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, UK
[v]Human Nutrition, School of Medicine, Dentistry and Nursing, College of Medical, Veterinary & Life Sciences, University of Glasgow, Glasgow, UK
[w]Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, UK

## Summary

**Background** Type 2 diabetes is associated with higher risk of several cancer types. However, the biological intermediates driving this relationship are not fully understood. As novel interventions for treating and managing type 2 diabetes become increasingly available, whether they also disrupt the pathways leading to increased cancer risk is currently unknown. We investigated the effect of a type 2 diabetes intervention, in the form of intentional weight loss, on circulating proteins associated with cancer risk to gain insight into potential mechanisms linking type 2 diabetes and adiposity with cancer development.

**Methods** Fasting serum samples from participants with diabetes enrolled in the Diabetes Remission Clinical Trial (DiRECT) receiving the Counterweight-Plus weight-loss programme (intervention, N = 117, mean weight-loss 10 kg, 46% diabetes remission) or best-practice care by guidelines (control, N = 143, mean weight-loss 1 kg, 4% diabetes remission) were subject to proteomic analysis using the Olink Oncology-II platform (48% of participants were female; 52% male). To identify proteins which may be altered by the weight-loss intervention, the difference in protein levels between groups at baseline and 1 year was examined using linear regression. Mendelian randomization (MR) was performed to extend these results to evaluate cancer risk and elucidate possible

*Corresponding author. MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK.
E-mail address: emma.vincent@bristol.ac.uk (E.E. Vincent).
[x]These authors contributed equally and are joint first authors.

biological mechanisms linking type 2 diabetes and cancer development. MR analyses were conducted using independent datasets, including large cancer meta-analyses, UK Biobank, and FinnGen, to estimate potential causal relationships between proteins modified during intentional weight loss and the risk of colorectal, breast, endometrial, gallbladder, liver, and pancreatic cancers.

**Findings** Nine proteins were modified by the intervention: glycoprotein Nmb; furin; Wnt inhibitory factor 1; toll-like receptor 3; pancreatic prohormone; erb-b2 receptor tyrosine kinase 2; hepatocyte growth factor; endothelial cell specific molecule 1 and Ret proto-oncogene (Holm corrected $P$-value <0.05). Mendelian randomization analyses indicated a causal relationship between predicted circulating furin and glycoprotein Nmb on breast cancer risk (odds ratio (OR) = 0.81, 95% confidence interval (CI) = 0.67–0.99, $P$-value = 0.03; and OR = 0.88, 95% CI = 0.78–0.99, $P$-value = 0.04 respectively), though these results were not supported in sensitivity analyses examining violations of MR assumptions.

**Interpretation** Intentional weight loss among individuals with recently diagnosed diabetes may modify levels of cancer-related proteins in serum. Further evaluation of the proteins identified in this analysis could reveal molecular pathways that mediate the effect of adiposity and type 2 diabetes on cancer risk.

---

**Research in context**

**Evidence before this study**

It is known that individuals with type 2 diabetes are at an increased risk of developing cancer. It is thought that increased adiposity may explain at least some of this risk, and previous studies have found that having an increased body weight alters the levels of certain proteins measured in blood (known as circulating proteins). Previous studies have also explored whether the altered levels of circulating proteins explain the increased cancer risk. However, few studies have evaluated the effect of weight-loss by individuals with type 2 diabetes on circulating proteins, or how the circulating proteins altered by weight loss relate to cancer risk.

**Added value of this study**

In this study, we investigated the impact of a weight-loss intervention on circulating proteins with a known link to cancer. We identified nine proteins which were altered following the intervention. We then evaluated these proteins for evidence of a causal role in increasing the risk of developing six obesity-driven cancers, using a genetic epidemiological method called Mendelian randomization.

**Implications of all the available evidence**

Based on all available evidence, it seems that weight management in individuals recently diagnosed with type 2 diabetes delivered in a primary care setting may influence biological pathways relevant to cancer, which may have implications for both policy implementation and future type 2 diabetes management research.

---

## Introduction

Type 2 diabetes is a chronic disease characterised by hyperinsulinemia and insulin resistance.[1] Diabetes is on the rise globally; with prevalence having quadrupled between 1980 and 2017, it is estimated that over 600 million people could be afflicted worldwide by 2045.[2] This mirrors the increase in obesity, which is the main modifiable risk factor for type 2 diabetes.[3] First line treatment for type 2 diabetes usually comprises of oral hypoglycaemic agents, including metformin.[4] However, recent research suggests that intentional weight-loss, even alongside withdrawal of antidiabetic and antihypertensive medication, may be at least as effective at achieving diabetes remission and restoration of normal insulin metabolism.[5–10]

Type 2 diabetes is known to increase the risks of several diseases including certain cancers, in particular colorectal, breast, endometrial, gallbladder, liver and pancreatic cancer.[11] However, the biological mechanisms thought to mediate the relationship between type 2 diabetes and cancer have not been fully elucidated. Additionally, the efficacy of weight-loss in a cohort of people with type 2 diabetes in reducing risk of cancer is unknown. We and others have shown that increased

body mass index (BMI), which is closely linked to type 2 diabetes incidence, is associated with widespread metabolic and proteomic alterations.[12–19] Perturbations in circulating biomarkers induced by obesity and diabetes may result in long-term systemic exposure of cells and tissues to an abnormal and dysregulated metabolic environment to which tumours must adapt to satisfy the bioenergetic and biosynthetic demands of chronic cell proliferation.[20,21] Circulating biomarkers are therefore plausibly linked to the likelihood of cancer developing and to the characteristics of the resulting tumour.

A systematic review of weight-loss trials reported a significant reduction in the risk of cancer mortality, highlighting the success of these interventions for cancer prevention.[22] However, whether weight-loss can mitigate the increased cancer risk seen in individuals with type 2 diabetes has not been extensively studied. The collection of blood samples in diabetes intervention trials permits the assessment of intermediate endpoints (e.g., molecular traits such as proteins) which may be associated with disease status or progression toward a secondary endpoint (e.g., cancer). Further, using Mendelian randomization (an epidemiological technique whereby genetic variants are used to proxy exposures of interest) it is possible (given specific assumptions) to estimate a potential causal relationship between circulating proteins, identified in the examination of trial data, and disease using *cis* protein quantitative trait loci (pQTLs).[23,24] In this study, we set out to leverage the methodological properties of a weight-loss trial and to focus on circulating proteins, which can inform tissue crosstalk and physiological status, and which have the potential to increase understanding as to possible pathways connecting variation in adiposity, type 2 diabetes, and disease. Proteins may be particularly interesting factors to consider in this regard as they represent over 90% of drug targets.[25] Proteins for which a causal role in cancer aetiology can be established may represent intervention targets to reduce disease risk.

The primary aim of this study was to identify proteins that were altered following intentional weight loss in a primary care randomised controlled trial framework (for trial protocol see).[26] The causal relationship between intervention-associated proteins and six obesity-driven cancers was subsequently evaluated using genetic epidemiological methods.[27]

## Methods

### Study design and participants

Samples and clinical data were obtained from 313 participants of the Diabetes Remission Clinical Trial (DiRECT, ISRCTN registry number 03267836).[8] Details of the study have been published elsewhere.[5,8,26] In brief, the trial recruited individuals aged 20–65 years who had been newly diagnosed with type 2 diabetes, had a body-mass index of 27–45 kg/m$^2$, and were not receiving insulin through GP practices. GP practices were invited to participate in multiple Health Board areas in Scotland and in the Newcastle-upon-Tyne NHS Foundation Trust area in the North East of England by the Primary Care Research Network (PCRN), followed by recruitment by the primary research team (for trail protocol see[26]). Recruitment took place between 25th July 2014 and 5th August 2016. Participants were recruited across 49 primary care practices in Scotland and the Tyneside region of England. The intervention included withdrawal of antidiabetic and antihypertensive drugs. This was followed by a total diet replacement (825–853 kcal/day formula diet (Counterweight-Plus) for 3–5 months), and stepped food reintroduction phase (two to eight weeks), and finally structured support for long-term weight-loss maintenance.[26] Participants were assigned either to a weight management programme (intervention) or best-practice care by guidelines (control) (see *Randomisation and masking* for more information). Phenotypic data were collected at baseline, 12 months, and 2 years. At 12 months, participants in the intervention arm experienced a mean weight-loss of 10 kg, and 46% of participants were in diabetes remission (defined as glycated haemoglobin of less than 6.5% after at least 2 months off all antidiabetic medications).[8] Participants in the control arm experienced a mean weight-loss of 1 kg, and 4% of participants were in diabetes remission.[8] Sex was collected as self-reported by study participants, and the impact of sex on the intervention effects was not investigated due to limited power.

### Randomisation and masking

Practices were randomly assigned (1:1) to provide either a weight management programme (intervention) or best-practice care by guidelines (control), using a minimisation method to maintain balance according to study site (Tyneside or Scotland) and practice list size (>5700 or ≤5700). Cluster randomisation at the GP practice level was performed to avoid contamination between groups and allow consistent advice from practice nurses/dietitians. Randomisation was performed by the Robertson Centre for Bio-Statistics, University of Glasgow, independently of the primary research team. See previous publications for further detail.[5,8,26]

### Proteomic data

In total, a subset of 574 available serum samples from 313 participants (260 of whom had paired samples at baseline and 12 months) collected during the trial were analysed using the Olink Proteomics Oncology II assay (Olink Target 96 Oncology II (v.7004); http://www.olink.com). The proteins included in the panel were specifically selected for their relevance to cancer based on public-access bioinformatic databases.[28] Prior to blood draw, participants were asked to fast overnight. Blood samples were drawn into a serum separator tube and

stored at 4 °C before processing. Samples were centrifuged at 2000*g* for 15 min at 4 °C within 5 h of collection. Serum was separated and stored in 0.5 mL aliquots at –80 °C. Olink Target 96 platforms require 40 μL serum and use paired oligonucleotide labelled antibodies to bind and measure the relative abundance of 92 proteins simultaneously using a proximity extension assay.[29] Samples were sent for analysis by Olink (Uppsala, Sweden). All analysts were blind to intervention/control arm status.

Normalised protein expression (NPX) data were returned by Olink. One NPX unit reflects a doubled protein concentration. Samples deviating >0.3 NPX units from the median were excluded. We applied further quality control (QC) steps using the metaboprep R package: samples and features were excluded based on extreme missingness (80%) before missingness was recalculated and exclusions made using a 20% threshold, additional exclusions were based on sample total sum abundance (>5 SD from the mean) and PCA of independent features (>5 SD from the mean).[30] Protein measurements below the limit of detection were retained with their recorded value with the caveat that data below the limit of detection have a higher risk to be in the non-linear phase of the S-curve of the NPX unit which may bias estimates. All protein measures were standardised and normalised prior to analyses using rank-based inverse normal transformation (RNT).

### Statistical analysis
#### Evaluation of the effect of the DiRECT intervention on serum protein levels
To evaluate the effect of the intervention on serum protein levels, outcomes (i.e., protein levels at 12 months follow-up) were compared between groups with linear models applied to the RNT data. Where protein measurements at baseline or follow-up were missing (unquantified) for an individual, that individual was excluded from the analysis for that specific protein and therefore, the analysed sample size varied slightly across proteins (N = 258–260). The maximum sample size was 260 paired samples. Models were adjusted for study centre and clinical practice list size along with the baseline measurement of the outcome (protein), age and sex, all fitted as fixed effects. Our primary exposure, analysed for all available data was group (trial arm) allocation; therefore, our estimates reflect the intervention effect in a clinical setting accounting for non-compliance. Effect estimates represent the difference in serum protein level seen in the intervention group relative to the control group while adjusting for protein levels at baseline (on the RNT standard deviation (SD) scale). The Holm method[31] was used to adjust *P*-values for multiple testing with an adjusted *P*-value of <0.05 used as a heuristic for strong evidence of an effect.

We also examined the observational association between oncology-related proteins and BMI at baseline

(N = 260, mean BMI 34.5 kg/m$^2$, SD 4.4 kg/m$^2$) to interrogate whether intervention effects were likely reflective of changes in body composition or other aspects of the intervention (e.g., dietary exposures). Associations between protein levels and BMI at baseline were estimated using linear regression adjusted for age, sex, study centre and practice list size. Model effect estimates represent the estimated difference in protein level (SD units) per unit change in the respective independent variable.

#### *Mendelian randomization and colocalisation analyses of intervention-associated proteins with cancer risk*
In addition to the primary trial-based results characterising protein differences associated with randomised intervention, Mendelian randomization and colocalisation analyses were conducted using independent datasets to estimate the causal relationship between intervention-associated (Holm corrected *P*-value <0.05) proteins and colorectal, breast, endometrial, gallbladder, liver, and pancreatic cancer risk.[27] To obtain genetic instruments for intervention-associated proteins, association estimates for the top *cis* protein quantitative trait loci (pQTL, SNP with the smallest *P* value within 1 Mb from the gene encoding the protein as defined by Sun et al.[32]) passing a multiple-testing corrected threshold of $P < 3.4 \times 10^{-11}$ were obtained from a genome-wide association study (GWAS) of 1462 circulating proteins detected by the Olink Explore 3072 assay, conducted in up to 54,306 UK Biobank participants. These data are given in NPX values and therefore represent an arbitrary unit on the log2 scale. As such, values are comparable across individuals for the same protein, but not across proteins or across analyses.[32] Participants who had been diagnosed with blood cancer, or those with leukocyte count $>200 \times 10^9$/L or $>100 \times 10^9$/L with 5% immature reticulocytes, haemoglobin concentration >20 g/dL, haematocrit >60%, and platelet count $>1000 \times 10^9$/L were excluded. Therefore, it is possible that some of the participants included in this GWAS had recently been diagnosed with other cancer types and were undergoing cancer treatment, which could impact their serum protein levels and is therefore a limitation of using this GWAS. However, given the sample size of 54,306 and the exclusion criteria, it is unlikely that these individuals, if present, represent a substantial proportion of the GWAS. pQTLs were functionally evaluated using the Ensembl Variant Effect Predictor (VEP, based on RefSeq and Ensembl).[33] Suitable genetic instruments were available for eight out of the nine intervention-associated proteins (all except pancreatic prohormone (PPY)).

Summary genetic association data were obtained from six cancer-specific GWAS. Association estimates for colorectal cancer in 52,775 cases and 45,940 controls were obtained from the ColoRectal Transdisciplinary Study (CORECT), the Colon Cancer Family Registry (CCFR), and the Genetics and Epidemiology of

Colorectal Cancer (GECCO) consortium. Association estimates provided for the analysis did not include UK Biobank study participants to avoid potential overlap between protein and colorectal cancer summary statistics.[34] Association estimates for liver cancer were obtained from a GWAS of liver cancer (specifically hepatocyte carcinoma) conducted in up to 648 cases and 259,583 controls in FinnGen.[35] Summary genetic data for pancreas cancer were obtained from a GWAS of 1249 cases and 259,583 controls in FinnGen.[35] Summary genetic association data were obtained for breast cancer from a GWAS of 133,384 cases and 113,789 controls of European ancestry.[36] Summary genetic association data were obtained for endometrial cancer from a GWAS of 12,906 cases and 108,979 controls of European ancestry.[37] Summary genetic data for gallbladder cancer were obtained from the UK Biobank, which included 195 cases and 456,134 controls of European ancestry.[38] In the Mendelian randomization analysis of protein levels and gallbladder cancer, both GWAS were obtained using UK Biobank participants. Therefore, there is potential for sample overlap between the studies (i.e., the same individuals being included in both exposure and outcome datasets). Sample overlap in two sample Mendelian randomization analyses can bias estimates towards the confounded observational estimate in the presence of weak instrument bias.[39] Any analyses with at least weak evidence for an effect of protein levels on gallbladder cancer ($P < 0.05$) were repeated using an alternative GWAS of gallbladder performed in FinnGen participants (84 cases, 259,584 controls), avoiding any sample overlap.[40]

Wald ratio estimates[41] were calculated to estimate the causal effect of intervention-associated proteins on cancer risk. Summary statistics were harmonised using the harmonise_data function within the TwoSampleMR R package (version 0.5.7).[42] Three core assumptions must be satisfied for Mendelian randomization analyses to be valid: (i) that the genetic instrument is associated with the exposure under investigation ("relevance"); (ii) that the genetic instrument does not share a common cause with the outcome under investigation ("exchangeability"); (iii) that the genetic instrument has no direct effect on the outcome ("exclusion restriction").[23] F-statistics were computed to appraise the "relevance" assumption. We examined F-statistics for evidence of weak instrument bias, defined as an F-statistic less than 10.[43] We also examined if *cis* variants used to instrument intervention-associated proteins were nominally associated ($P$-value $< 0.05$) and directionally concordant in a GWAS of proteins detected using a different technique (SomaScan v4 assay).[44] The online calculator available at https://sb452.shinyapps.io/power/ was used for Mendelian randomization power calculations.[44]

We also performed genetic colocalisation to examine whether proteins and cancer outcomes share the same causal variant at a given locus, which is necessary for (though not alone sufficient for evidence of) a causal relationship between the traits. Colocalisation was performed using the *coloc* R package (version 5.1.0.1).[45,46] *coloc* uses approximate Bayes factor computation to estimate posterior probabilities that associations between two traits represent: (i) no genetic association in the region for either trait ($H_0$), (ii) a genetic association in the region for the first trait ($H_1$), (iii) a genetic association in the region for the second trait ($H_2$), (iv) associations for both traits, but different causal variants ($H_3$) and (v) associations for both traits with a shared single causal variant ($H_4$).[45] Priors were selected based on a window size of 10 Mb and a p1 of 0.000001, p2 of 0.000001, and p12 of 0.000001 using the online calculator available at: https://chr1swallace.shinyapps.io/coloc-priors/. SomaScan association estimates from Pietzner et al.[44] were used for colocalisation analyses as full summary statistics were not available from Sun et al.[32] at the time of publication. Variants within 10 Mb of pQTLs used to instrument serum proteins for Mendelian randomization estimates were included in colocalisation analyses. A posterior probability of >0.80 for $H_4$ (i.e., a shared causal variant between the traits) was used to indicate evidence of colocalisation.[47]

### Ethics

DiRECT is registered with the ISRCTN registry, number 03267836. Ethics approval for the DiRECT trial was granted by West 3 Ethics Committee in January 2014 (reference number 13/WS/0314), with approvals by the National Health Service (NHS) health board areas in Scotland and clinical commissioning groups in Tyneside, in accordance with the Declaration of Helsinki. All participants provided informed consent.

### Role of funders

The funders had no role in study design, data collection, data analyses, interpretation, or writing of the manuscript.

## Results

### Effect of the weight-loss intervention on serum proteins

Following data QC, 260 participants (117 intervention, 143 control, Supplementary Table S1) had trial and proteomic data available measured in samples at baseline and 12 months. The minimum sample size for the analysis of an individual protein was 258. A large proportion of samples (57%) were below the level of detection for one protein (FAS-associated death domain protein, FADD). Baseline characteristics were similar between the control and intervention groups (Table 1). Overall, 9 proteins were modified by the intervention in the linear model analysis (Holm corrected $P$-value <0.05) —3 of which were increased, and 6 reduced by the

| | Intervention group (n = 117) | | Control group (n = 143) | |
|---|---|---|---|---|
| | Mean/N | SD/% | Mean/N | SD/% |
| Sex | | | | |
| Female | 49 | 42% | 55 | 38% |
| Male | 68 | 58% | 88 | 62% |
| Age (years) | 53.7 | 7.1 | 56.3 | 6.8 |
| Female | 54.0 | 7.0 | 56.5 | 7.0 |
| Male | 53.5 | 7.0 | 56.3 | 7.0 |
| Body mass index (kg/m$^2$) | 34.8 | 4.5 | 34.3 | 4.3 |
| Female | 35.7 | 4.4 | 34.8 | 4.4 |
| Male | 34.2 | 4.4 | 34.0 | 4.4 |
| Weight (kg) | 100.3 | 16.9 | 98.9 | 16.3 |
| Female | 91.4 | 16.3 | 90.4 | 16.3 |
| Male | 107.0 | 16.3 | 104.0 | 16.3 |
| Fasting glucose (mmol/L) | 9.3 | 3.2 | 8.8 | 2.6 |
| Female | 9.6 | 2.9 | 8.7 | 2.9 |
| Male | 9.1 | 2.9 | 8.9 | 2.9 |
| Total cholesterol (mmol/L) | 4.3 | 1.1 | 4.3 | 1.1 |
| Female | 4.7 | 1.1 | 4.4 | 1.1 |
| Male | 4.0 | 1.1 | 4.2 | 1.1 |
| HDL cholesterol (mmol/L) | 1.1 | 0.3 | 1.2 | 0.3 |
| Female | 1.2 | 0.3 | 1.3 | 0.3 |
| Male | 1.1 | 0.3 | 1.1 | 0.3 |
| Triglycerides (mmol/L) | 2.0 | 1.5 | 1.9 | 0.9 |
| Female | 1.8 | 1.2 | 1.8 | 1.2 |
| Male | 2.2 | 1.2 | 2.1 | 1.2 |

Summary statistics based on the analysis sample (N = 260). SD = standard deviation.

*Table 1*: Baseline characteristics of participants with type 2 diabetes included in the study (N = 260).

intervention (Fig. 1 and Table 2, full results Supplementary Table S2). The strongest evidence for a positive association was seen for glycoprotein Nmb (GPNMB, Beta: 0.61, SE: 0.11, Holm corrected *P*-value = 3.78e-08) with 8% of variance in glycoprotein Nmb at 12 months explained by group allocation. The strongest evidence for a negative association was seen for furin (FURIN, Beta: –0.54, SE: 0.11, Holm corrected *P*-value = 6.28e-07), with 7% of the variance in furin at 12 months explained by group allocation. Of the 9 intervention-associated proteins, 5 were associated with BMI at baseline in a direction consistent with the effect of weight-loss on the proteins (*P* < 0.05; furin, hepatocyte growth factor, toll-like receptor 3, receptor tyrosine-protein kinase erbB-2 and proto-oncogene tyrosine-protein kinase receptor Ret, Supplementary Table S3).

### Effect of modified proteins on cancer risk
Details of genetic variants used to instrument oncology-related proteins are presented in Supplementary Table S4. A single genome-wide significant *cis* pQTL was available for eight of the nine intervention-associated proteins (no genome-wide significant

associations were detected for Receptor tyrosine-protein kinase erbB-2 (ERBB2) by Sun et al.[32]). For one of these proteins (furin), the *cis* pQTL was borderline nominally associated (*P*-value = 0.05) in a GWAS of proteins detected using a different technique (SomaScan v4 assay), suggesting possible proxying of binding effects.[44] Using the Ensembl Variant Effect Predictor it was estimated that the selected pQTLs for *GPNMB, FURIN, TLR3* and *HGF* had consequential effects on gene transcripts (Supplementary Table S4). The minimum F-statistic for protein genetic instruments was 166, suggesting analyses should avoid weak instrument bias (Supplementary Table S4). Power to detect an odds ratio (OR) of 1.2 was >80% for five out of six cancers for at least one protein (all except gallbladder cancer; mean power = 65%; median power = 95%, per SD change oncology-related protein (α = 0.05)) (Supplementary Table S5).

A Bonferroni correction was applied to correct for multiple testing across genetic analyses, resulting in a *P*-value threshold of < 0.001 (0.05/48 statistical tests; eight proteins against six cancer end points) with this threshold being used to define "strong evidence". Findings where *P* ≥ 0.001 and *P* < 0.05 were interpreted as "weak evidence".

Amongst all protein to cancer associations investigated, there was weak evidence to support an effect of furin and glycoprotein Nmb on cancer risk (Fig. 2 and Supplemental Table S6). We found weak evidence that increased serum furin reduces breast cancer risk (OR 0.81, 95% CI = 0.67–0.99, *P* = 0.03 per SD increase in serum furin), and that increased glycoprotein Nmb reduces breast cancer risk (OR 0.88, 95% CI = 0.78–0.99, *P* = 0.04 per SD increase in serum glycoprotein Nmb). Genetic colocalisation analysis suggested that serum furin and glycoprotein Nmb were unlikely to share a causal variant with risk of breast cancer (posterior probability of a shared causal variant (H$_4$): ≤80%, Supplemental Table S7).

### Discussion
There is strong evidence that individuals with type 2 diabetes are at increased risk of at least six cancer types.[1] Multiple studies have investigated the influence of obesity on circulating intermediate traits, but few studies have examined the impact of weight-loss on circulating proteins, particularly in individuals diagnosed with type 2 diabetes, or how circulating proteins relate to type 2 diabetes-related cancer risk. In this study, we investigated the impact of a type 2 diabetes weight-loss intervention (Counterweight-Plus) on oncology-related proteins in serum. Under a conservative correction for multiple testing, nine proteins were associated with intervention status (weight-loss): glycoprotein Nmb, furin, Wnt inhibitory factor 1, toll-like receptor 3, pancreatic prohormone, erb-b2 receptor
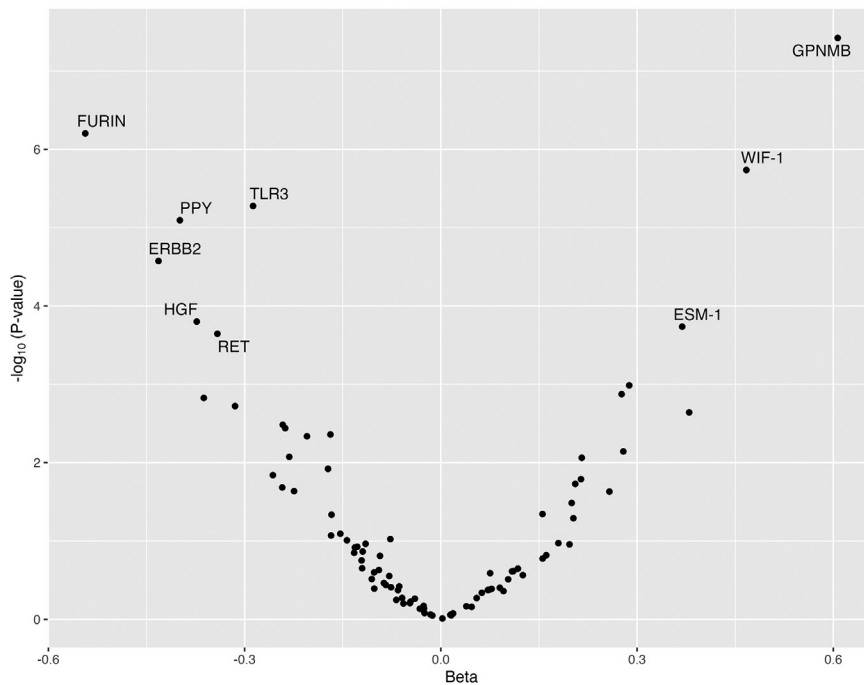
**Fig. 1:** Comparison of serum protein levels at 12 months follow-up between intervention groups (n = 260). Betas reflect the mean difference between allocation groups (protein measures were standardised and normalised prior to analyses using rank-based inverse normal transformation) while adjusting for protein levels at baseline. Labelled proteins are those that pass a Holm corrected P = 0.05. Values greater than 0 indicate an increase in serum protein seen in the intervention arm and values less than 0 reflect a decrease in serum protein levels. GPNMB: Glycoprotein Nmb; FURIN: Furin; WIF1: WNT Inhibitory Factor 1; TLR3: Toll-like receptor 3; PPY: Pancreatic prohormone; ERBB2: Receptor tyrosine-protein kinase erbB-2; HGF: Hepatocyte growth factor; ESM-1: Endothelial cell-specific molecule 1; RET: Proto-oncogene tyrosine-protein kinase receptor Ret.

tyrosine kinase 2, hepatocyte growth factor, endothelial cell specific molecule 1, and Ret proto-oncogene. We then evaluated these proteins for an observational association with BMI at baseline, and for evidence of a causal effect on risk of type 2 diabetes-related cancers in Mendelian randomization analyses.

Although the proteins investigated were specifically selected for their biological relevance to cancer,[28] the majority of proteins associated with the weight-loss intervention in this study are typically membrane bound or intracellular and therefore the relevance of serum levels of these proteins to cancer risk is less well

| Protein ID | Protein name | Beta | SE | *P*-value | Adjusted *P*-value |
|---|---|---|---|---|---|
| GPNMB | Glycoprotein Nmb | 0.61 | 0.11 | 3.78e-08 | 3.48e-06 |
| FURIN | Furin | −0.54 | 0.11 | 6.28e-07 | 5.71e-05 |
| WIF1 | WNT Inhibitory Factor 1 | 0.47 | 0.10 | 1.84e-06 | 1.66e-04 |
| TLR3 | Toll-like receptor 3 | −0.29 | 0.06 | 5.30e-06 | 4.72e-04 |
| PPY | Pancreatic prohormone | −0.40 | 0.09 | 8.06e-06 | 7.10e-04 |
| ERBB2 | Receptor tyrosine-protein kinase erbB-2 | −0.43 | 0.10 | 2.66e-05 | 2.32e-03 |
| HGF | Hepatocyte growth factor | −0.37 | 0.10 | 1.58e-04 | 0.01 |
| ESM-1 | Endothelial cell-specific molecule 1 | 0.36 | 0.10 | 1.84e-04 | 0.02 |
| RET | Proto-oncogene tyrosine-protein kinase receptor Ret | −0.34 | 0.09 | 2.26e-04 | 0.02 |

Betas reflect the association of intervention group with in serum protein levels at 12 months follow-up while adjusting for protein levels at baseline (protein measures were standardised and normalised prior to analyses using rank-based inverse normal transformation). Adjusted *P*-value refers to a Holm corrected *P*-value. Listed proteins are those that pass a Holm corrected *P* = 0.05.

*Table 2:* Association of serum protein levels and intervention status at 12 months follow-up.
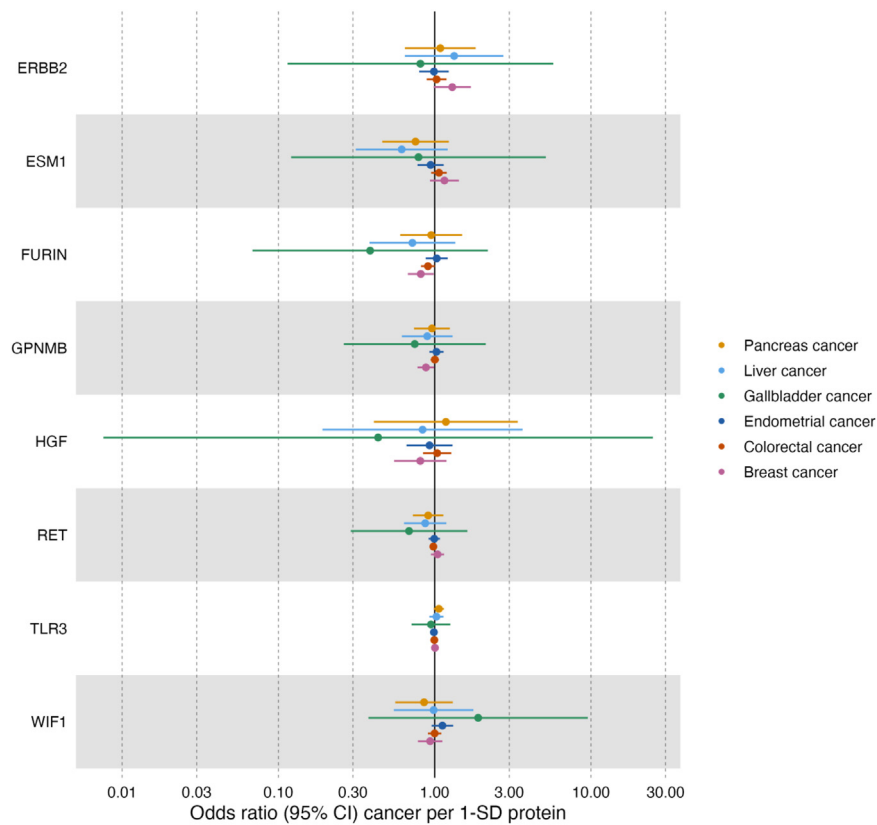
**Fig. 2:** Estimated effect of circulating proteins on cancer risk in Mendelian randomization analyses (odds ratio (OR) and 95% confidence interval; n = 98,715–260,832 depending on cancer). All analyses used the Wald ratio model. Where estimates are missing the genetic instruments for the exposure were not present in the outcome dataset and thus MR could not be performed. GPNMB: Glycoprotein Nmb; FURIN: Furin; WIF1: WNT Inhibitory Factor 1; TLR3: Toll-like receptor 3; PPY: Pancreatic prohormone; ERBB2: Receptor tyrosine-protein kinase erbB-2; HGF: Hepatocyte growth factor; ESM-1: Endothelial cell-specific molecule 1; RET: Proto-oncogene tyrosine-protein kinase receptor Ret; CI = confidence interval.

known. Of the 92 proteins investigated, glycoprotein Nmb, a transmembrane glycoprotein, showed the largest increase in response to weight-loss, but was not associated with BMI at baseline suggesting the observed association may not be driven by BMI specifically and may reflect another aspect of the intervention e.g., dietary exposures. Our Mendelian randomization and colocalisation analyses found little evidence for a causal role for serum glycoprotein Nmb in development of six type 2 diabetes-related cancers (colorectal, breast, endometrial, gallbladder, liver, and pancreatic cancers), though the tissue-specific relevance or predictive utility of serum glycoprotein Nmb in type 2 diabetes-related cancer risk should be further investigated.

A recent pan-cancer analysis reported that increased expression of furin in cancer tissue was associated with poor prognosis.[48] Consistent with a protective effect of weight-loss on cancer risk, we found that serum furin was reduced following the Counterweight Plus intervention and that furin was positively associated with BMI at baseline. However, our Mendelian randomization analyses suggested there might be a protective effect of higher serum furin on breast cancer risk, which would suggest a possible detrimental effect of this reduction in the intervention group. Colocalisation analyses, however, suggested serum furin was unlikely to share a causal variant with breast cancer in this region given low posterior probabilities to support $H_4$ (i.e., shared causal variants). The lack of colocalising signal and estimated protective effect of serum furin on cancer risk in genetic analyses may be attributable to our ability to reliably instrument (proxy) serum furin with a single *cis* SNP (rs6227), a reduced signal in the furin locus in the GWAS of proteins measured using the SomaScan platform,[44] or potential genetic pleiotropy of the variant (this SNP has also been found to influence expression of FES proto-oncogene, tyrosine kinase.)[49]

Circulating levels of toll-like receptor 3, another transmembrane protein, were also found to be reduced with weight-loss. Consistent with the effect of the intervention, we found that toll-like receptor 3 was positively associated with BMI at baseline. Across solid tumours from differing tissues, toll-like receptor 3

expression has been found to be associated with both good and poor prognosis in cancer, possibly via pro-apoptotic pathways and resistance to anti-tumour drugs, respectively.[50] In Mendelian randomization analyses, we observed little evidence for a causal effect of increased serum toll-like receptor 3 on any cancer risk.

In the trial analysis, we found that pancreatic pro-hormone, receptor tyrosine-protein kinase erbB-2, hepatocyte growth factor and proto-oncogene tyrosine-protein kinase receptor Ret were reduced with weight-loss. Of these proteins, hepatocyte growth factor and tyrosine-protein kinase receptor Ret were positively associated with BMI at baseline. Our results for hepatocyte growth factor are consistent with previous cross-sectional studies which have reported a positive correlation between obesity and hepatocyte growth factor in serum.[51,52] Hepatocyte growth factor is known to promote proliferation, migration, invasion and survival of cancer cells.[53] However, the origin of hepatocyte growth factor present in the tumour microenvironment is unknown as it may be synthesised by a variety of tissues throughout the body, including liver or adipose tissue. Mendelian randomization and colocalisation estimates did not support evidence of causality between hepatocyte growth factor and type 2 diabetes-related cancers.

In the trial analysis we found that serum Wnt Inhibitory Factor 1 and endothelial cell-specific molecule 1 (endocan) were increased with weight-loss, though these proteins were not consistently associated with BMI at baseline and follow-up genetic analyses did not identify any effects of serum Wnt Inhibitory Factor 1 or endothelial cell-specific molecule 1 on type 2 diabetes-related cancer risk. Our findings are inconsistent with previous cross-sectional studies that have reported a positive correlation between obesity and serum levels of endothelial cell-specific molecule 1.[54,55] It is possible these earlier estimates may be influenced by residual confounding which should be minimised in our intervention-based approach.

One limitation of our analysis is that we were unable to ascertain whether the intervention-related changes identified were due directly to weight-loss, to the rapid dietary change, to the withdrawal of antidiabetic and antihypertensive medication, or as a result of the food reintroduction stage. However, an association between protein levels and BMI at baseline may be indicative of a driving effect of an association between protein levels and adiposity. Of the nine intervention-associated proteins, there was strong evidence ($P < 0.05$) for five of the proteins for an association with BMI at baseline (furin, hepatocyte growth factor, toll-like receptor 3, receptor tyrosine-protein kinase erbB-2 and proto-oncogene tyrosine-protein kinase receptor RET). This replicates results from a previous analysis using SomaLogic proteomic data from up to 2737 healthy participants other than for toll-like receptor 3, which was not included in the previous analysis, and receptor tyrosine-protein kinase erbB-2, for which they found little evidence of an association ($P > 0.05$).[17] In our analysis, the direction of the association was consistent with the direction of effect of the intervention (i.e., if the intervention decreased protein levels, the association with BMI at baseline was positive and vice versa) for a further protein (Glycoprotein Nmb), despite the $P$-value not reaching the predefined cut-off of 0.05. In the previous analysis, Glycoprotein Nmb showed weak evidence for an association with BMI ($P < 0.05$) in a direction consistent with our analyses, suggesting that we may not have had sufficient sample size to detect this association. Three proteins showed an inconsistent direction of effect between the weight-loss intervention and BMI at baseline (WNT inhibitory factor 1, pancreatic pro-hormone, and endothelial cell-specific molecule 1). For WNT inhibitory factor 1 and endothelial cell-specific molecule 1, the previous analysis found strong evidence ($P < 1.4 \times 10^{-5}$) for an association with BMI in a direction consistent with the effect of the weight-loss intervention in our analysis.

One possible explanation for the lack of evidence for an association with BMI at baseline in our analysis is that the relationship between BMI and protein levels may be non-linear, so the association may not be observable in a population where BMI is ubiquitously high. For the final intervention-associated protein (pancreatic prohormone), the previous analysis found strong evidence ($P < 1.4 \times 10^{-5}$) for an association with BMI in the same direction to that in our analysis, i.e., in a direction inconsistent with the direction of association with the intervention. This suggests that the effect of the intervention may not be driven by change in BMI itself but another aspect (e.g., a non-weight-related response to dietary change). Recent work has highlighted that some aspects of adiposity which are relevant to health outcomes may not be captured by BMI. For instance, body fat distribution has been shown to have differential effects on cardiovascular outcomes depending on anatomical location of adipose tissue.[56,57] Additionally, some differential effects have been seen in the relationship between BMI and cancer risk according to sex.[19,58] Although such measurements were not available in this analysis, future work should evaluate the role of body fat distribution in altering oncology-related protein levels in serum, and whether these effects are likely to differ by sex.

We found little evidence for a shared causal variant between expression of any of the proteins and risk of cancers investigated in colocalisation analyses, including for those proteins which had evidence for an effect in Mendelian randomization analyses. One possible explanation for this is a lack of true causal relationships across expression of the nine proteins and six cancer sites investigated. For instance, in the case of furin, the pQTL used to proxy expression in the Mendelian randomization analyses, rs6227, is more strongly

associated with expression of neighbouring protein FES proto-oncogene, tyrosine kinase than furin.[49] Thus, it is possible that the effect seen in this and other Mendelian randomization analyses could in fact reflect a causal association of neighbouring genes on cancer risk rather than the protein being instrumented. However, there are several other explanations for a lack of strong evidence of colocalisation between expression of the proteins investigated here and cancer risk. Firstly, due to a lack of available full summary statistics in the Sun et al. GWAS,[32] we used an alternative GWAS for the colocalisation analyses which used SomaScan technology to quantify proteins. One limitation of this approach is that several cis pQTLs used to proxy protein levels in the Mendelian randomization analyses are missense variants, which have been well-documented to increase the risk of aptamer binding effects in the SomaScan technology. Secondly, it is likely that the reason for a lack of shared causal variant between exposure and outcome is due to lack of power from sample size in the outcome (i.e., cancer) datasets, which may be limiting the reliability of the colocalisation analyses described here.

A major strength of this study is the use of sequential samples and clinical data from a large and well characterized clinical trial (DiRECT).[5,8,26] We measured 92 oncology-related proteins in paired samples from the same participants at baseline and 12 months. Though it was not possible to measure proteins in all participants, this subset with samples reflects the original trial population, which has been shown to be representative of the general population of people with short duration type 2 diabetes.[59] Due to the importance of potential pathways linking type 2 diabetes and adiposity with cancer risk, we chose to measure proteins with a known link with cancer determined through bioinformatic databases (i.e., those included in Olink's oncology II panel) which limits our ability to identify novel cancer-relevant proteins. Further, as proteins are measured on an arbitrary scale, this limits our ability to assess the impact of the DiRECT intervention on protein concentrations in absolute units. Future studies should explore the use of novel technologies such as Olink Flex[60] which can be used to translate NPX values of certain proteins to concentrations in pg/mL units which are more meaningful. Whilst our study design enabled us to confirm that the Counterweight-Plus intervention influences the levels of oncology-related proteins in serum, it is challenging to attribute changes to specific elements of the intervention (e.g., weight-loss or dietary exposures). We attempted to appraise weight-loss specific effects by performing a cross-sectional analysis of protein levels and BMI at baseline, however, this analysis may be underpowered to detect weak effects given the narrow BMI range of participants recruited by DiRECT, which may also make detection of non-linear associations challenging. Another limitation of our analysis is that in limiting the

analysis to one cohort, the transferability of our findings to other groups is as of yet unknown. To attempt to extend these results beyond the trial population, we conducted Mendelian randomization and colocalisation analyses using independent datasets in order to estimate the causal relationship between intervention-associated proteins and the risk of six cancers where type 2 diabetes is an important cause (pancreas, colorectal, breast, gallbladder, endometrial and liver cancer).[27] As we selected only the top cis variants to instrument proteins, given that trans-acting variants are typically considered to be more prone to introducing horizontal pleiotropy into analyses, it was not possible to conduct sensitivity analyses which require a greater number of genetic variants such as MR-Egger or weighted median methods.[61,62] Finally, for one of the nine intervention-associated proteins, pancreatic prohormone (PPY), no suitable genetic instruments were available, meaning we could not include this protein in our Mendelian randomization analyses.

In summary, we performed an analysis of samples from a weight-loss intervention trial and found that nine cancer-related proteins were associated with intervention status (glycoprotein Nmb; furin; Wnt inhibitory factor 1; toll-like receptor 3; pancreatic prohormone; erb-b2 receptor tyrosine kinase 2; hepatocyte growth factor; endothelial cell specific molecule 1 and Ret proto-oncogene). These changes provide evidence that weight management in individuals recently diagnosed with type 2 diabetes delivered in a primary care setting may influence biological pathways relevant to cancer. Follow-up Mendelian randomization analyses provided weak evidence for a causal role of 2 of these proteins in type 2 diabetes-related cancer, specifically furin and glycoprotein Nmb in breast cancer risk. Further evaluation of proteins associated with weight-loss could help to facilitate the development of cancer prevention strategies.

## Appendix A. Supplementary data
Supplementary data related to this article can be found at https://doi.org/10.1016/j.ebiom.2024.104977.

## References
1. Shlomai G, Neel B, LeRoith D, Gallagher EJ. Type 2 diabetes mellitus and cancer: the role of pharmacotherapy. *J Clin Oncol*. 2016;34:4261. https://doi.org/10.1200/JCO.2016.67.4044.
2. Standl E, Khunti K, Hansen TB, Schnell O. The global epidemics of diabetes in the 21st century: current situation and perspectives. *Eur J Prev Cardiol*. 2019;26:7–14. https://doi.org/10.1177/2047487319881021.
3. Valabhji J, Kar P. Rise in type 2 diabetes shows that prevention is more important than ever. *BMJ*. 2023;381. https://doi.org/10.1136/BMJ.P910.
4. Padhi S, Nayak AK, Behera A. Type II diabetes mellitus: a review on recent drug based therapeutics. *Biomed Pharmacother*. 2020;131:110708. https://doi.org/10.1016/J.BIOPHA.2020.110708.
5. Lean MEJ, Leslie WS, Barnes AC, et al. Durability of a primary care-led weight-management intervention for remission of type 2 diabetes: 2-year results of the DiRECT open-label, cluster-randomised trial. *Lancet Diabetes Endocrinol*. 2019;7:344–355. https://doi.org/10.1016/S2213-8587(19)30068-3.
6. Taylor R, Barnes AC, Hollingsworth KG, et al. Aetiology of Type 2 diabetes in people with a 'normal' body mass index: testing the personal fat threshold hypothesis. *Clin Sci*. 2023;137:1333–1346. https://doi.org/10.1042/CS20230586.
7. Al-Mrabeh A, Hollingsworth KG, Shaw JAM, et al. 2-year remission of type 2 diabetes and pancreas morphology: a post-hoc analysis of the DiRECT open-label, cluster-randomised trial. *Lancet Diabetes Endocrinol*. 2020;8:939–948. https://doi.org/10.1016/S2213-8587(20)30303-X.
8. Lean ME, Leslie WS, Barnes AC, et al. Primary care-led weight management for remission of type 2 diabetes (DiRECT): an open-label, cluster-randomised trial. *Lancet*. 2018;391:541–551. https://doi.org/10.1016/S0140-6736(17)33102-1.
9. Steven S, Hollingsworth KG, Al-Mrabeh A, et al. Very low-calorie diet and 6 Months of weight stability in type 2 diabetes: pathophysiological changes in responders and nonresponders. *Diabetes Care*. 2016;39:808–815. https://doi.org/10.2337/DC15-1942.
10. Lim EL, Hollingsworth KG, Aribisala BS, Chen MJ, Mathers JC, Taylor R. Reversal of type 2 diabetes: normalisation of beta cell function in association with decreased pancreas and liver triacylglycerol. *Diabetologia*. 2011;54:2506–2514. https://doi.org/10.1007/S00125-011-2204-7.
11. Pearson-Stuttard J, Papadimitriou N, Markozannes G, et al. Type 2 diabetes and cancer: an umbrella review of observational and mendelian randomization studies. *Cancer Epidemiol Biomark Prev*. 2021;30:1218–1228. https://doi.org/10.1158/1055-9965.EPI-20-1245/671063/AM/TYPE-2-DIABETES-AND-CANCER-AN-UMBRELLA-REVIEW-OF.
12. Holmes MV, Lange LA, Palmer T, et al. Causal effects of body mass index on cardiometabolic traits and events: a mendelian randomization analysis. *Am J Hum Genet*. 2014;94:198–208. https://doi.org/10.1016/j.ajhg.2013.12.014.
13. Wurtz P, Wang Q, Kangas AJ, et al. Metabolic signatures of adiposity in young adults: mendelian randomization analysis and effects of weight change. *PLoS Med*. 2014;11:e1001765. https://doi.org/10.1371/journal.pmed.1001765.
14. Abu Bakar MH, Sarmidi MR, Cheng KK, et al. Metabolomics–the complementary field in systems biology: a review on obesity and type 2 diabetes. *Mol Biosyst*. 2015;11:1742–1774. https://doi.org/10.1039/c5mb00158g.
15. Rauschert S, Uhl O, Koletzko B, Hellmuth C. Metabolomic biomarkers for obesity in humans: a short review. *Ann Nutr Metab*. 2014;64:314–324. https://doi.org/10.1159/000365040.

16  Bell JA, Carslake D, O'Keeffe LM, et al. Associations of body mass and fat indexes with cardiometabolic traits. *J Am Coll Cardiol*. 2018;72:3142–3154. https://doi.org/10.1016/j.jacc.2018.09.066.

17  Goudswaard LJ, Bell JA, Hughes DA, et al. Effects of adiposity on the human plasma proteome: observational and Mendelian randomisation estimates. *Int J Obes*. 2021;45:2221–2229. https://doi.org/10.1038/s41366-021-00896-1.

18  Zaghlool SB, Sharma S, Molnar M, et al. Revealing the role of the human blood plasma proteome in obesity using genetic drivers. *Nat Commun*. 2021;12:1–13. https://doi.org/10.1038/S41467-021-21542-4.

19  Bull CJ, Bell JA, Murphy N, et al. Adiposity, metabolites, and colorectal cancer risk: mendelian randomization study. *BMC Med*. 2020;18:1–16. https://doi.org/10.1186/S12916-020-01855-9/FIGURES/5.

20  DeNicola GM, Cantley LC. Cancer's fuel choice: new flavors for a picky eater. *Mol Cell*. 2015;60:514–523. https://doi.org/10.1016/j.molcel.2015.10.018.

21  Perera RM, Bardeesy N. Pancreatic cancer metabolism: breaking it down to build it back up. *Cancer Discov*. 2015;5:1247–1261. https://doi.org/10.1158/2159-8290.CD-15-0671.

22  Ma C, Avenell A, Bolland M, et al. Effects of weight loss interventions for adults who are obese on mortality, cardiovascular disease, and cancer: systematic review and meta-analysis. *BMJ*. 2017;359:4849. https://doi.org/10.1136/bmj.j4849.

23  Smith GD, Ebrahim S. "Mendelian randomization": can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003;32:1–22.

24  Melzer D, Perry JRB, Hernandez D, Corsi A-M, Stevens K. A genome-wide association study identifies protein quantitative trait loci (pQTLs). *PLoS Genet*. 2008;4. https://doi.org/10.1371/journal.pgen.1000072.

25  Santos R, Ursu O, Gaulton A, et al. A comprehensive map of molecular drug targets. *Nat Rev Drug Discov*. 2016;16:19–34. https://doi.org/10.1038/nrd.2016.230.

26  Leslie WS, Ford I, Sattar N, et al. The diabetes remission clinical trial (DiRECT): protocol for a cluster randomised trial. *BMC Fam Pract*. 2016;17:1–10. https://doi.org/10.1186/S12875-016-0406-2/FIGURES/1.

27  *When could overweight and obesity overtake smoking as the biggest cause of cancer in the UK.* Cancer Research UK; 2018. Available online: https://www.cancerresearchuk.org/sites/default/files/obesity_tobacco_cross_over_report_final.pdf?_gl=1*wawq46*_ga*MTk0Njg5Ni4xNjY2MTc1NjU2*_ga_58736Z2GNN*MTY4MDk3MTcxNi42LjEuMTY4MDk3MTg5Ny41MS4wLjA.&_ga=2.53827752.103249110.1680969463-1946896.1666175656. Accessed April 8, 2023.

28  *How have the biomarkers in the Olink panels been selected?–Olink*; 2016. https://olink.com/faq/how-have-the-biomarkers-in-the-proseek-multiplex-panels-been-selected/. Accessed October 14, 2023.

29  *The Olink PEA platform–Olink*; 2016. https://olink.com/application/pea/. Accessed October 14, 2023.

30  Hughes DA, Taylor K, McBride N, et al. metaboprep: an R package for preanalysis data description and processing. *Bioinformatics*. 2022;38:1980–1987. https://doi.org/10.1093/BIOINFORMATICS/BTAC059.

31  Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat*. 1979;6:65–70.

32  Sun BB, Chiou J, Traylor M, et al. Genetic regulation of the human plasma proteome in 54,306 UK Biobank participants. *Popul Analytics Janssen Data Sci*. 2022;20. https://doi.org/10.1101/2022.06.17.496443.

33  Mclaren W, Pritchard B, Rios D, et al. Databases and ontologies deriving the consequences of genomic variants with the Ensembl API and SNP effect predictor. *Bioinformatics*. 2010;26:2069–2070. https://doi.org/10.1093/bioinformatics/btq330.

34  Huyghe JR, Bien SA, Harrison TA, et al. Discovery of common and rare genetic risk variants for colorectal cancer. *Nat Genet*. 2019;51:76–87. https://doi.org/10.1038/s41588-018-0286-6.

35  Kurki MI, Karjalainen J, Palta P, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature*. 2023;613:7944. https://doi.org/10.1038/S41586-022-05473-8.

36  Michailidou K, Lindström S, Dennis J, et al. Association analysis identifies 65 new breast cancer risk loci. *Nature*. 2017;551:92–94. https://doi.org/10.1038/NATURE24284.

37  O'Mara TA, Glubb DM, Amant F, et al. Identification of nine new susceptibility loci for endometrial cancer. *Nat Commun*. 2018;9:1–12. https://doi.org/10.1038/S41467-018-05427-7.

38  Jiang L, Zheng Z, Fang H, Yang J. A generalized linear mixed model association tool for biobank-scale data. *Nat Genet*. 2021;53:1616–1621. https://doi.org/10.1038/S41588-021-00954-4.

39  Haycock PC, Burgess S, Wade KH, Bowden J, Relton C, Smith GD. Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. *Am J Clin Nutr*. 2016;103:965–978. https://doi.org/10.3945/AJCN.115.118216.

40  Kurki MI, Karjalainen J, Palta P, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature*. 2023;613:508–518. https://doi.org/10.1038/s41586-022-05473-8.

41  Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res*. 2017;26:2333–2355. https://doi.org/10.1177/0962280215597579.

42  Hemani G, Zheng J, Elsworth B, et al. The MR-base platform supports systematic causal inference across the human phenome. *Elife*. 2018;7. https://doi.org/10.7554/ELIFE.34408.

43  Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med*. 2008;27:1133–1163. https://doi.org/10.1002/sim.3034.

44  Pietzner M, Wheeler E, Carrasco-Zanini J, et al. Mapping the proteo-genomic convergence of human diseases. *Science*. 2021;374.

45  Giambartolomei C, Vukcevic D, Schadt EE, et al. Bayesian test for colocalisation between Pairs of genetic association studies using summary statistics. *PLoS Genet*. 2014;10:e1004383. https://doi.org/10.1371/JOURNAL.PGEN.1004383.

46  Zuber V, Grinberg NF, Gill D, et al. Combining evidence from Mendelian randomization and colocalization: review and comparison of approaches. *Am J Hum Genet*. 2022;109:767. https://doi.org/10.1016/J.AJHG.2022.04.001.

47  Wallace C. A more accurate method for colocalisation analysis allowing for multiple causal variants. *PLoS Genet*. 2021;17:e1009440. https://doi.org/10.1371/JOURNAL.PGEN.1009440.

48  Zhou B, Gao S. Pan-cancer analysis of FURIN as a potential prognostic and immunological biomarker. *Front Mol Biosci*. 2021;8:648402. https://doi.org/10.3389/fmolb.2021.648402.

49  Aguet F, Barbeira AN, Bonazzola R, et al. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science*. 2020;369:1318–1330. https://doi.org/10.1126/SCIENCE.AAZ1776/SUPPL_FILE/AAZ1776_TABLESS10-S16.XLSX.

50  Muresan XM, Bouchal J, Culig Z, Souček K. Toll-like receptor 3 in solid cancer and therapy resistance. *Cancers*. 2020;12:3227. https://doi.org/10.3390/cancers12113227.

51  Bell LN, Ward JL, Degawa-Yamauchi M, et al. Adipose tissue production of hepatocyte growth factor contributes to elevated serum HGF in obesity. *Am J Physiol Endocrinol Metab*. 2006;291:843–848. https://doi.org/10.1152/AJPENDO.00174.2006/ASSET/IMAGES/LARGE/ZH10100646550005.JPEG.

52  Faber DR, Moll FL, Vink A, et al. Adipose tissue quantity and composition contribute to adipokine concentrations in the subclavian vein and the inferior mesenteric vein. *Int J Obes*. 2012;36:1078–1085. https://doi.org/10.1038/ijo.2011.214.

53  McAinch AJ, Naderpoor N, Loomes K, et al. The role of hepatocyte growth factor (HGF) in insulin resistance and diabetes. *Diabetes Front Endocrinol*. 2018;9:503. https://doi.org/10.3389/fendo.2018.00503.

54  Nedeva I, Kamenov Z. Circulating endocan in subjects with obesity and carbohydrate disturbances. *Metabolism*. 2021;116:154489. https://doi.org/10.1016/j.metabol.2020.154489.

55  Klisic A, Kavaric N, Spasojevic-Kalimanovska V, Kotur-Stevuljevic J, Ninic A. Serum endocan levels in relation to traditional and non-traditional anthropometric indices in adult population. *J Med Biochem*. 2021;40:41. https://doi.org/10.5937/JOMB0-25170.

56  Agrawal S, Wang M, Klarqvist MDR, et al. Inherited basis of visceral, abdominal subcutaneous and gluteofemoral fat depots. *Nat Commun*. 2022;13:1–17. https://doi.org/10.1038/s41467-022-30931-2.

57  Liu Y, Basty N, Whitcher B, et al. Genetic architecture of 11 organ traits derived from abdominal MRI using deep learning. *Elife*. 2021;10. https://doi.org/10.7554/ELIFE.65554.

58  Rask-Andersen M, Ivansson E, Höglund J, Ek WE, Karlsson T, Johansson Å. Adiposity and sex-specific cancer risk. *Cancer Cell*. 2023;41:1186–1197.e4. https://doi.org/10.1016/J.CCELL.2023.05.010.

59  Taylor R, Leslie WS, Barnes AC, et al. Clinical and metabolic features of the randomised controlled diabetes remission clinical trial

(DiRECT) cohort. *Diabetologia*. 2017;61:589. https://doi.org/10.1007/s00125-017-4503-0.

60  *Ensuring quality with flexibility- Olink® Flex validation & verification– Olink*; 2023. https://olink.com/application/ensuring-quality-with-flexibility-olink-flex-validation-verification/. Accessed October 14, 2023.

61  Burgess S, Dudbridge F, Thompson SG. Combining information on multiple instrumental variables in Mendelian randomization:

comparison of allele score and summarized data methods. *Stat Med*. 2016;35:1880–1906. https://doi.org/10.1002/sim.6835.

62  Holmes MV, Richardson TG, Ference BA, Davies NM, Davey Smith G. Integrating genomics with biomarkers and therapeutic targets to invigorate cardiovascular drug development. *Nat Rev Cardiol*. 2021;18:435–453. https://doi.org/10.1038/S41569-020-00493-1.

63  R Core Team. *R: a language and environment for statistical computing.* 2020.