



OPEN

Using trials of caloric restriction and bariatric surgery to explore the effects of body mass index on the circulating proteome

Lucy J. Goudswaard^{1,2,3✉}, Madeleine L. Smith^{1,2}, David A. Hughes^{1,2}, Roy Taylor⁴, Michael Lean⁵, Naveed Sattar⁶, Paul Welsh⁶, Alex McConnachie⁷, Jane M. Blazeby¹, Chris A. Rogers⁸, Karsten Suhre⁹, Shaza B. Zaghlool⁹, Ingeborg Hers³, Nicholas J. Timpson^{1,2,10} & Laura J. Corbin^{1,2,10}

Thousands of proteins circulate in the bloodstream; identifying those which associate with weight and intervention-induced weight loss may help explain mechanisms of diseases associated with adiposity. We aimed to identify consistent protein signatures of weight loss across independent studies capturing changes in body mass index (BMI). We analysed proteomic data from studies implementing caloric restriction (Diabetes Remission Clinical trial) and bariatric surgery (By-Band-Sleeve), using SomaLogic and Olink Explore1536 technologies, respectively. Linear mixed models were used to estimate the effect of the interventions on circulating proteins. Twenty-three proteins were altered in a consistent direction after both bariatric surgery and caloric restriction, suggesting that these proteins are modulated by weight change, independent of intervention type. We also integrated Mendelian randomisation (MR) estimates of the effect of BMI on proteins measured by SomaLogic from a UK blood donor cohort as a third line of causal evidence. These MR estimates provided further corroborative evidence for a role of BMI in regulating the levels of six proteins including alcohol dehydrogenase-4, nogo receptor and interleukin-1 receptor antagonist protein. These results indicate the importance of triangulation in interrogating causal relationships; further study into the role of proteins modulated by weight in disease is now warranted.

The circulating proteome includes thousands of proteins naturally secreted from cells or present because of cell damage or cell death¹. These proteins include cytokines, growth factors and hormones and have been shown to be modified by environmental factors and risk pathways such as obesity and conditions such as cancer^{2,3}. Protein levels can also predict the risk of certain outcomes⁴, and taken together have an exciting potential to provide biomarkers with potential clinical utility. Exploring protein changes associated with excess adiposity by using causal analyses could help with identifying targets to prevent or reduce adverse health outcomes such as type 2 diabetes (T2D), coronary artery disease (CAD), musculoskeletal diseases and many types of cancer⁵⁻⁷. With the widespread availability of proteomic datasets in increasingly large population samples, such investigations have become practically feasible.

Integrating evidence from independent sources which each have their own specific limitations and potential biases (known as triangulation)⁸, can be an important tool in overcoming specific limitations to any one analytical technique/study design. In this context, there are both surgical and non-surgical approaches to weight loss which

¹Population Health Sciences, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK. ²MRC Integrative Epidemiology Unit, Bristol, UK. ³Physiology, Pharmacology & Neuroscience, University of Bristol, Biomedical Sciences Building, University Walk, Bristol BS8 1TD, UK. ⁴Newcastle Magnetic Resonance Centre, Translational and Clinical Research Institute, Newcastle University, Newcastle Upon Tyne NE4 5PL, UK. ⁵Human Nutrition, School of Medicine, Dentistry and Nursing, College of Medical, Veterinary & Life Sciences, University of Glasgow, Glasgow G31 2ER, UK. ⁶School of Cardiovascular and Medical Science, University of Glasgow, Glasgow G12 8TA, UK. ⁷Robertson Centre for Biostatistics, School of Health and Wellbeing, University of Glasgow, Glasgow G12 8QQ, UK. ⁸Bristol Medical School, Bristol Trials Centre, University of Bristol, Bristol BS8 1NU, UK. ⁹Department of Biophysics and Physiology, Weill Cornell Medicine - Qatar, Doha, Qatar. ¹⁰These authors contributed equally: Nicholas J. Timpson and Laura J. Corbin. ✉email: lucy.goudswaard@bristol.ac.uk

present opportunities for the examination of the circulating proteome in the context of variation in body mass index (BMI). Clinical trials have implemented interventions in an attempt to induce weight loss; these interventions include bariatric surgery (also referred to as surgical weight loss) and caloric restriction. Bariatric surgery and caloric restriction have been shown to achieve weight loss in people with obesity, with bariatric surgery on average achieving greater weight loss⁹. Previous studies have characterized the effects of individual weight loss interventions on circulating proteins at a relatively small scale^{10–19}. Despite these scenarios offering independent records of weight loss, there has not been a comparison of broad proteomic effects of caloric restriction and bariatric surgery, nor either of these in comparison to Mendelian randomisation (MR) results investigating the likely impact of BMI. Consistent evidence across independent trials may help in pointing towards common BMI-driven signals.

In order to study the causal effects of BMI, previous MR studies are useful as these studies have examined the effect of BMI on the plasma proteome using population-based samples^{20,21}. MR is a technique that uses a genetic variant or variants to act as proxies for modifiable risk factors of interest, such as BMI, in order to estimate effects of that risk factor on an outcome²². Although MR helps to overcome confounding and reverse causation, MR also has limitations. For example, the assumptions of MR²³ may be violated, especially in instances where traits that are instrumented are complex traits such as BMI. In this case, genetic variants may then not only yield apparent differences in circulating protein levels through variation in BMI, but by other biological processes which are statistically indistinguishable from BMI effects. Further, it is well known that large sample sizes are required for MR²⁴, which has only recently become achievable for collections of proteomic data^{25,26}.

The aim of this study was to determine the effect of BMI on the proteome. To achieve this aim there were four main objectives of this study. First, use proteomic data from a caloric restriction randomised controlled trial (RCT) to identify the effect this has on circulating proteins. Second, use proteomic data from a bariatric surgery RCT to identify the effect of surgical weight loss on circulating proteins. Third, compare the effects across study designs by identifying proteins with consistent evidence across the surgical and diet-based interventions and incorporating (previously published) BMI to protein abundance MR estimates from a UK-based blood donor cohort (INTERVAL), as a further line of evidence to determine whether protein effects are likely to be driven by BMI. Finally, for proteins that have evidence across the three study designs for a role of BMI in their regulation, utilise existing information about drug targets to characterise the function of proteins in health and disease.

Results

An overview of the results that will be presented is shown in Fig. 1.

Diabetes Remission Clinical Trial (DiRECT) participant characteristics

Participant characteristics for those included in the caloric restriction trial and in the current proteomic analysis are displayed in Table 1, where $N = 292$. There were more males than females in the study (55% males in the intervention group and 62% males in the control group) and this is possibly a reflection of the patient population more generally²⁷. Participants were older and had higher high-density lipoprotein (HDL)-cholesterol in the control group than those in the intervention group. However, other characteristics were similar across treatment and control groups. Baseline BMI in the control group and intervention group was 35.1 kg/m^2 (SD 4.6 kg/m^2) and 34.3 kg/m^2 (SD 4.3 kg/m^2) respectively. All participants had a T2D diagnosis (mean 3 years, SD 1.8 years) and the majority were on medication for control of their diabetes. Mean HbA1c levels in the control group were 60 mmol/mol (SD 14 mmol/mol) and 58 mmol/mol (SD 12 mmol/mol) in the intervention group, while the target HbA1c level for a patient with T2D is $< 48 \text{ mmol/mol}$ ²⁸. The participants in the control group had a mean reduction in BMI of 0.34 kg/m^2 (SD 1.3 kg/m^2) and the intervention group had a mean reduction in BMI of 3.50 kg/m^2

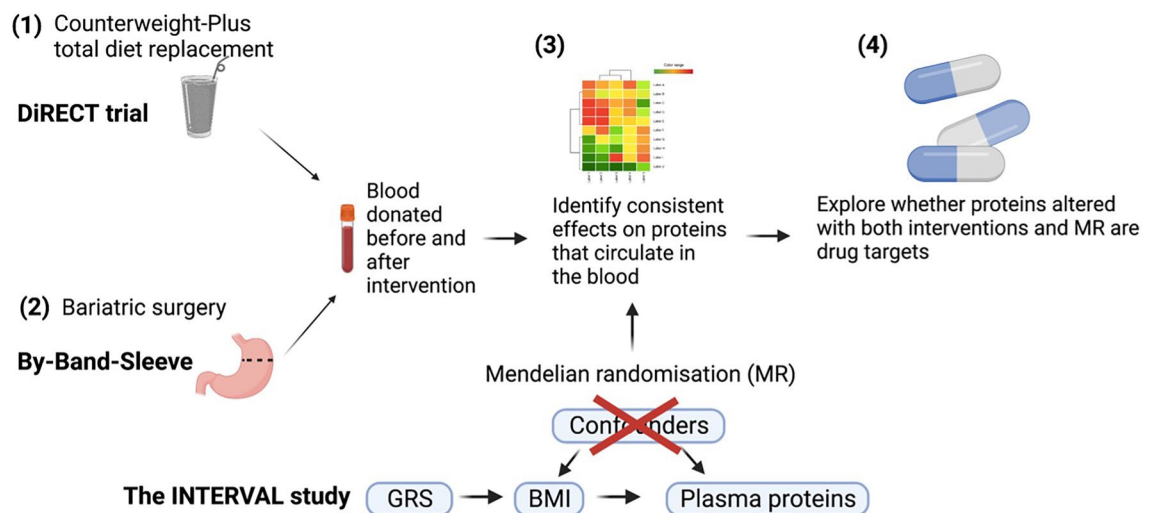


Figure 1. Study overview. GRS = genetic risk score (used as the instrumental variable). BMI = body mass index.

Baseline variable		Intervention group (mean \pm SD or n (%))	Intervention group N	Control group (mean \pm SD or n (%))	Control group N	P-value for difference
Age		53.0 (7.5)	146	56.2 (7.1)	146	2.3×10^{-4}
Sex	Male	55%	146	62%	146	0.28
	Female	45%		38%		
BMI (kg/m ²)		35.1 (4.6)	146	34.3 (4.3)	146	0.10
Systolic blood pressure (mmHg)		133 (17)	146	137 (16)	146	0.03
HbA1c (mmol/mol)		60 (14)	146	58 (12)	146	0.13
Glucose (mmol/l)		9.2 (3.3)	146	8.8 (2.6)	144	0.23
Insulin (mIU/ml)		25 (15)	146	23 (14)	144	0.26
Cholesterol (mmol/l)		4.3 (1.2)	146	4.3 (1.2)	143	0.91
HDL (mmol/l)		1.1 (0.3)	146	1.2 (0.3)	143	9.5×10^{-3}
Triglycerides (mmol/l)		2.1 (1.4)	146	1.9 (0.9)	143	0.34
Diabetes duration (y)		3.0 (1.7)	146	3.0 (1.8)	146	0.81
Number of anti-diabetic medications		1.14 (0.94)	146	1.09 (0.83)	146	0.60
Centre	Scotland	78%	146	57%	146	1.8×10^{-4}
	Tyneside	22%		43%		
List size	> 5700	62%	146	50%	146	0.06
	\leq 5700	38%		50%		

Table 1. Baseline characteristics of the DiRECT trial participants who underwent the intervention (total diet replacement). Sample size (N) is indicated and is up to 146 where there is no missing data. A two-tailed unpaired student's t-test was used to compare continuous variables and a Chi-squared test was used to compare differences in categorical variables.

(SD 2.8 kg/m²). Characteristics were mostly similar across males and females (Supplementary Table 1), except females had higher levels of total cholesterol and HDL-cholesterol and a higher BMI (35.3 kg/m², SD 4.6 kg/m² in females vs 34.3 kg/m² (SD 4.2 kg/m²) in males, $p=0.04$). The cohort is representative of diabetes patterns in the wider community as women tend to have a higher BMI than men to develop T2D²⁹.

DiRECT caloric restriction effect on circulating protein levels

A linear mixed model was used to estimate the effect of the caloric restriction intervention on proteins, providing an estimate of the change in protein levels in normalised SD units. There was evidence that caloric restriction, which was associated with a reduction in weight/BMI, affected circulating levels of 216 out of 4601 proteins (5%) at a conservative Bonferroni adjusted $p < 1.2 \times 10^{-5}$ (Supplementary Table 2). Of these, levels of 120 proteins increased after the intervention and levels of 96 proteins decreased (Fig. 2). Proteins that had the strongest evidence for an increase following the intervention include chondroadherin (CHAD, beta = 0.97 normalised SDs, 95% CI 0.80 to 1.13, $p = 3.2 \times 10^{-26}$), osteomodulin (OMD, beta = 0.87 SDs, 95% CI 0.71 to 1.02, $p = 2.05 \times 10^{-23}$) and Immunoglobulin superfamily DCC subclass member 4 (IGDC4, beta = 0.80 SDs, 95% CI 0.65 to 0.94, $p = 2.4 \times 10^{-23}$). Proteins most strongly decreased include aminoacylase-1 (beta = -0.97 SDs, 95% CI -1.15 to -0.80, $p = 2.6 \times 10^{-23}$), fatty acid-binding protein (heart) (FABP3, beta = -0.89 SDs, 95% CI -1.06 to -0.73, $p = 2.8 \times 10^{-22}$) and fumarylacetoacetase (FAAA, beta = -0.95 SDs, 95% CI -1.14 to -0.76, $p = 2.6 \times 10^{-19}$). Full results are available in Supplementary Table 2. In some cases, proteins that have previously been reported as BMI-associated had betas that were directionally consistent with reported effects but did not pass our multiple testing threshold for association, for example CRP (beta = -0.42 SDs, 95% CI -0.64 to -0.22, $p = 7.9 \times 10^{-5}$).

By-Band-Sleeve characteristics at baseline

A subset of participants (N = 125) from a bariatric surgery trial (By-Band-Sleeve) were included in the proteomics study. Participants had study visits at baseline (randomisation) and ~ 36 months post-randomisation (median = 33 months post-surgery, range = 21 to 44 months). Most participants in the By-Band-Sleeve proteomics study were female (69%) (Table 2). They had a mean age of 50 years (SD 10 years) and a mean BMI at baseline of 45 kg/m² (SD 8 kg/m²). Age and BMI were similar in males and females. Forty-three participants (36%) had been diagnosed with T2D with a mean duration of 6 years and 4 months (SD 5 years 5 months). There was evidence for some sex differences across participant characteristics: males had higher rates of T2D (61% vs 26% females, $p < 2.2 \times 10^{-4}$). There was also a difference between cardiometabolic traits, where females had higher levels of low density lipoprotein (LDL)-cholesterol, HDL-cholesterol and total cholesterol, whereas males had higher levels of triglycerides. As the trial is ongoing at the time of writing, the exact reduction in BMI cannot be reported. Based on existing literature, the expected reduction in BMI across bariatric surgery subtypes is around 10 kg/m²³⁰. Published results by the By-Band-Sleeve Trial Management Group and Investigators provides further baseline information on all participants in the By-Band-Sleeve trial (not just the subset whose samples were sent for proteomic analysis)³¹.

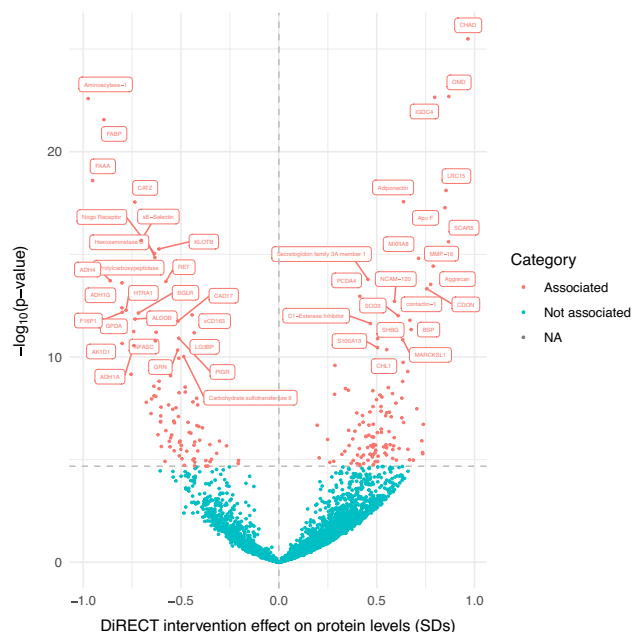


Figure 2. Volcano plot displaying the linear mixed model results in DiRECT. Proteins were categorised as associated if $p < 1.2 \times 10^{-5}$. “Associated” proteins are labelled where labels are non-overlapping. Full protein names can be found in Supplementary Table 2.

Characteristic	All participants (mean ± SD or n (%))	N	Male, N = 36	Female, N = 82	p-value
Age (years)	50 (10)	118	50 (9)	50 (10)	0.76
Body mass index (BMI) (kg/m ²)	45 (8)	118	46 (7)	45 (8)	0.27
Baseline systolic blood pressure (mmHg)	136 (15)	118	138 (15)	136 (15)	0.70
Baseline diastolic blood pressure (mmHg)	84 (7)	118	85 (8)	83 (7)	0.30
Smoking category		118			0.069
Never	49 (42%)		14 (39%)	35 (43%)	
Ex-smoker	64 (54%)		18 (50%)	46 (56%)	
Rarely or occasionally	5 (4.2%)		4 (11.1%)	1 (1.2%)	
Diabetic (diet, oral treated or insulin)	42 (36%)	118	22 (61%)	21 (26%)	2.2×10^{-4}
Time since diabetes diagnosis (months)	76 (65)	118	72 (49)	80 (79)	0.63
Ethnicity		118			0.70
English/Welsh/Northern Irish/Scottish/British	110 (93%)		33 (92%)	77 (94%)	
Other (Any other white background, Irish, Gypsy/Traveller, White and Black Caribbean)	8 (7%)		3 (8%)	5 (6.1%)	
Employment status		118			8.3×10^{-3}
Full-time	38 (32%)		16 (44%)	22 (27%)	
Part-time	27 (23%)		5 (14%)	22 (27%)	
Self-employed	9 (7.6%)		6 (17%)	3 (3.7%)	
Other (Homemaker/Student/Retired/Unable to work/Unemployed)	44 (37.2%)		9 (25%)	35 (43%)	
Income		114			0.91
< £10,000–20,000	53 (46%)		18 (51%)	35 (44%)	
£20,001–£40,000	38 (33%)		11 (31%)	27 (34%)	
> £40,001	23 (20%)		6 (17%)	17 (22%)	
Baseline triglycerides (mmol/L)	1.59 (0.76)	118	1.83 (0.74)	1.48 (0.75)	3.6×10^{-3}
Baseline LDL-cholesterol (mmol/L)	2.71 (0.96)	117	2.39 (0.89)	2.86 (0.96)	0.014
Baseline HDL-cholesterol (mmol/L)	1.21 (0.27)	118	1.03 (0.17)	1.29 (0.27)	8.7×10^{-7}
Baseline total cholesterol (mmol/L)	4.65 (1.10)	118	4.25 (0.94)	4.83 (1.12)	9.3×10^{-3}

Table 2. Characteristics of participants included in the analysis for the By-Band-Sleeve trial. Sample size (N) is up to 118 where there was no missing data. Categorical variables were compared using a Chi-squared test and continuous variables were compared using a Wilcoxon signed rank test.

Bariatric surgery effect on protein levels

A linear mixed model was implemented to estimate the effect of bariatric surgery on protein levels, providing an estimate of the change in protein (in normalised SD units) following the intervention. The serum levels of 191 proteins out of 1472 (13%) were associated with the bariatric surgery intervention ($p < 6.2 \times 10^{-5}$) in the By-Band-Sleeve study, with 118 proteins showing higher levels and 73 proteins showing lower levels post-surgery (Fig. 3, Supplementary Table 3). Of all proteins associated with the intervention, the maximum percentage of samples below the specified limit of detection (LOD) was 28% (mean = 1.0%, range = 0–28%). The protein most strongly increased after bariatric surgery was insulin-like growth factor binding protein 2 (IGFBP2, beta = 1.00 SDs, 95% CI = 0.85 to 1.15, $p = 4.3 \times 10^{-25}$). Other proteins that increased following bariatric surgery include insulin-like growth factor binding protein 1 (IGFBP1, beta = 0.77 SDs, 95% CI 0.63 to 0.92, $p = 1.4 \times 10^{-18}$), guanylin (GUCA2A), pregnancy-specific beta-1 glycoprotein 1 (PSG1) and ficolin-2 (FCN2, beta = 0.69 SDs, 95% CI 0.52 to 0.85, $p = 2.4 \times 10^{-13}$). Bariatric surgery reduced the levels of other proteins, such as the scavenger receptor cysteine rich domain-containing group B protein (SSC4D, beta = -0.74 SDs, 95% CI -0.87 to -0.62, $p = 3.1 \times 10^{-21}$), cadherin-related family member 2 (CDHR2) and B-cell receptor CD22. Proteins with an established relationship with BMI such as leptin²⁰ were reduced with the bariatric surgery intervention (beta = -0.70 SDs, 95% CI -0.86 to -0.53, $p = 5.2 \times 10^{-13}$). Full results are provided in Supplementary Table 3.

Comparison of intervention effects on protein levels

There were 989 unique proteins matched by UniProt ID when combining the 4601 protein measurements from the DiRECT trial and 1472 protein measurements from the By-Band-Sleeve trial. Supplementary Table 4 provides a summary of the effect estimates for shared proteins. Among the 989 proteins in the merged data set, 81 proteins (39 with a positive direction, and 42 with a negative direction) were associated with caloric restriction and 130 proteins (84 with a positive direction and 46 with a negative direction) were associated with bariatric surgery. A total of 25 proteins were detectably associated with both interventions: 23 of these were consistent in direction of change and two had opposite effects. 56 of the 81 proteins associated only with caloric restriction, and 105 of the 130 proteins associated only with bariatric surgery (Fig. 4). The proteins with the strongest effects and that had consistent directions of effect include intervention-increasing effects on IGFBP1/2, osteomodulin (OMD), and intervention-lowering effects on pro-inflammatory proteins such as interleukin-1 receptor antagonist protein (IL-1Ra), scavenger receptor cysteine-rich type 1 protein M130 (sCD163) and E-selectin (sE-selectin). Proteins were measured in the two studies using different proteomic technologies, platforms run by SomaLogic (in DiRECT) and Olink (in By-Band-Sleeve). Among the 23 proteins with consistent effects across interventions, 19 feature in a recent analysis comparing quantification of proteins by these platforms; of these, all 19 were found to be positively correlated across platforms³² (mean $r = 0.71$, range = 0.35 to 0.96).

Two of the proteins associated with both interventions had opposite directions of effect. The CAD17 protein was found to be increased by bariatric surgery but decreased by caloric restriction, and the HPCL1 protein was reduced by bariatric surgery but elevated by caloric restriction (Fig. 4). Correlation data between SomaLogic and Olink platforms for these two proteins with discordant effects was not available from Haslam et al.³² since the Olink panel used in that study had fewer unique proteins ($N = 972$)³² than the ones measured in this study.

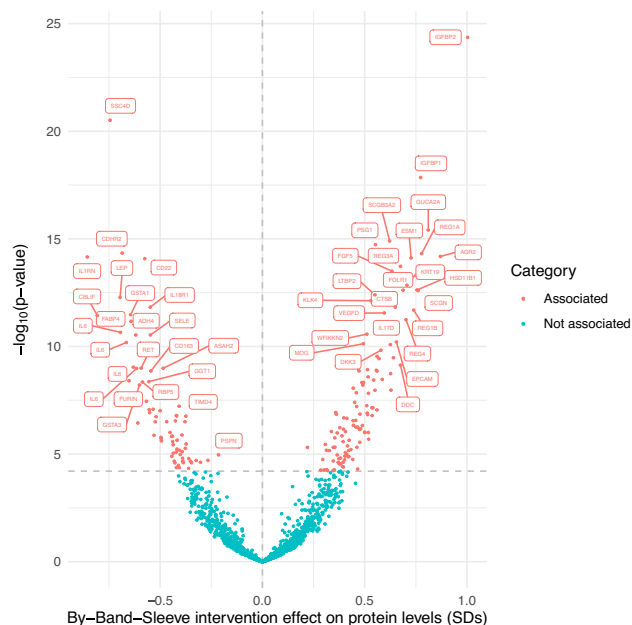


Figure 3. Volcano plot of the change in protein levels after bariatric surgery using a linear mixed model. Proteins were categorised as associated if $p < 6.2 \times 10^{-5}$. Full protein names and effect estimates can be found in Supplementary Table 3.

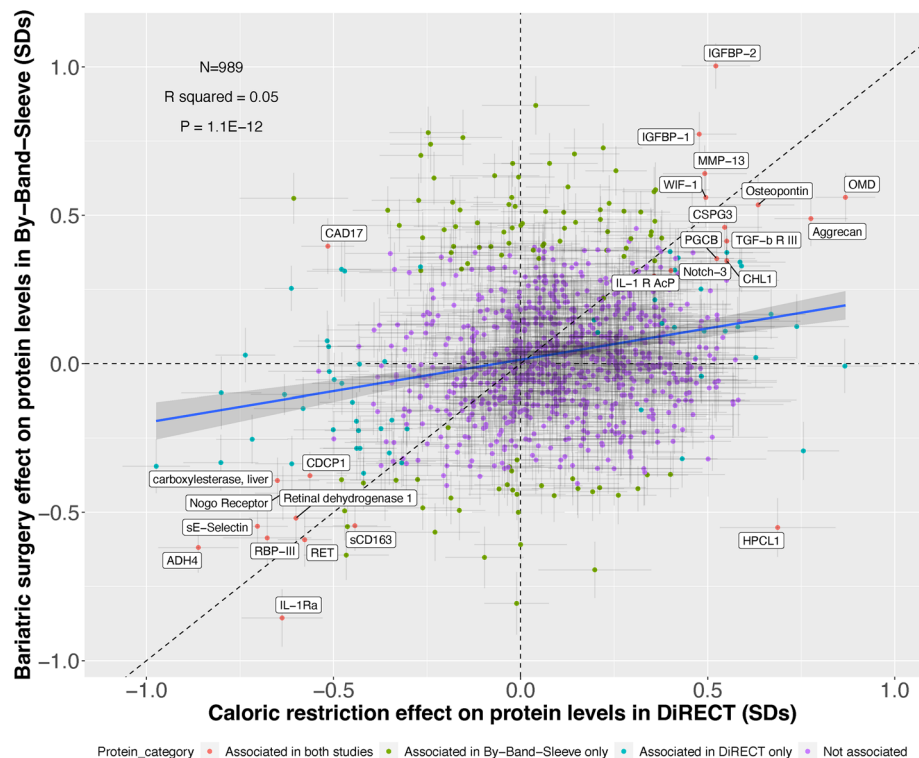


Figure 4. Comparison of estimates for the change in protein levels following either the total diet replacement in DiRECT or bariatric surgery in By-Band-Sleeve. Proteins denoted as "associated" passed the pre-defined adjusted p-value thresholds in one or both analyses. Effect estimates are the change in protein levels in normalised SD units comparing endpoint (post intervention) levels to baseline. The diagonal dashed line indicates $y = x$ and the blue solid line is the regression line with 95% confidence intervals. Proteins that were associated in both studies are labelled.

Comparing intervention effects on protein levels with MR estimates

Previously published one-sample MR results provided estimates of the difference in protein levels per 1 SD ($\sim 4.8 \text{ kg/m}^2$) higher BMI²⁰. These estimates are provided in Supplementary Table 5. The estimates for the effect of each intervention on proteins were compared with one sample BMI to protein MR estimates to identify whether intervention effects are consistent with the BMI-associated causal effect estimates²⁰. To provide an estimate to match the intervention effects, the MR estimates were multiplied by -1 to indicate the mean difference in protein in normalised SDs per normalised SD lower BMI. For example, if the MR estimate suggests that lower BMI results in a lower level of a protein, we would predict that the protein would be reduced with the interventions. Out of the 23 proteins which displayed consistent effects with both interventions, 20 proteins had available MR estimates. As the underlying populations have different BMI distributions, and the trials likely induce a different degree of weight loss, the estimates presented only allow for interpretation of direction effect, rather than direct comparison of the magnitudes of effect. For all but two proteins, the direction of effect of the BMI to protein MR estimate was consistent in direction with the intervention effect estimates (Fig. 5A,B). Six of these MR estimates had 95% confidence intervals that did not overlap the null, providing a short list of consistently associated proteins across all three studies (DiRECT, By-Band-Sleeve, and INTERVAL) (Table 3).

The two proteins, CAD17 and HPCL1, that were associated with both interventions but in opposite directions did not have published MR estimates available. Furthermore, it is unclear whether these seemingly opposing effects represent meaningful heterogeneity across interventions and/or study samples or if the discordance of effect is a result of differences in the way these proteins are measured by the two platforms (no correlation statistics are available from Haslam et al.³² in this case).

Are proteins altered by body composition druggable targets?

For proteins with evidence of consistent effects across both interventions and in an MR framework, we explored whether the proteins are current drug targets to help understand the relevance of the proteins to comorbidities of BMI. To do this, we searched for the relevant targets using DrugBank³³ and searched protein targets identified by Finan et al.³⁴. IGFBP1 and IGFBP2 have approved drugs to treat disorders such as growth failure³³. ADH4 is a target for nutritional supplementation³³. Other proteins suggested as having the potential to be drug targets include RTN4R³⁴, IL1-Ra³⁴ and brevicin core protein (BCAN/PGCB)³⁴. These do not currently have approved

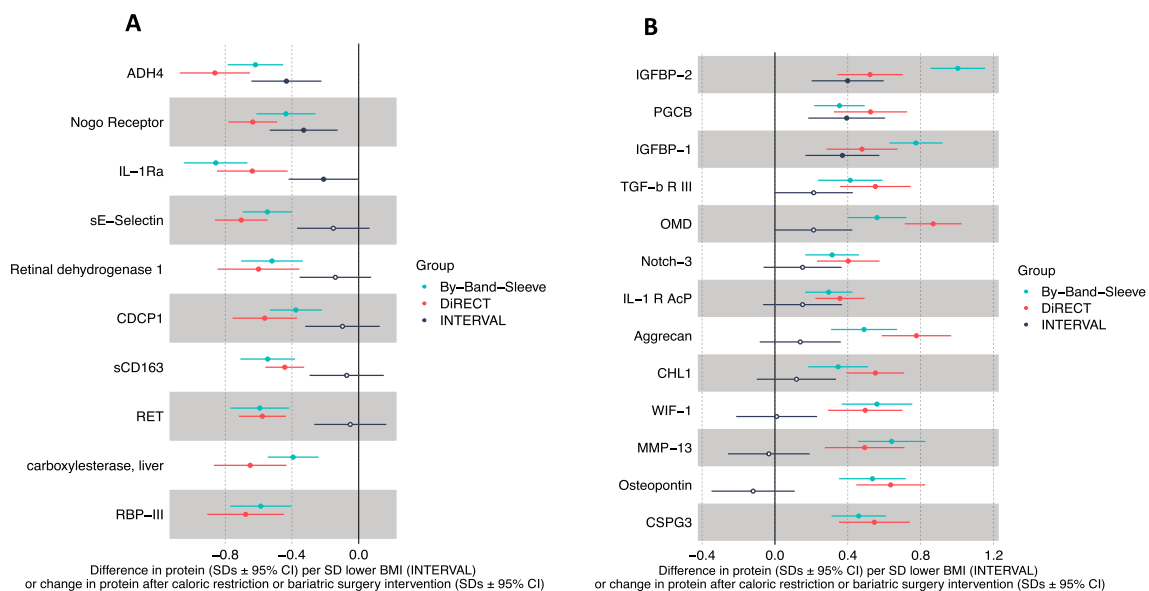


Figure 5. Forest plots comparing the estimates for the effect of interventions on protein levels (based on a linear mixed model) with published MR results from the INTERVAL study (where beta coefficients represent an estimate of the mean protein change per SD lower BMI). The MR estimates from Supplementary Table 5 have been multiplied by -1 so that the units are the difference in protein per SD lower BMI. **(A)** Proteins that were reduced with total diet replacement (TDR) and bariatric surgery. **(B)** Proteins that were increased following TDR and bariatric surgery.

Gene	Target	Protein full name	Direction of relationship with BMI	Current or potential to be small-molecule target ¹	Current or potential to be biologic target ¹	Protein involved in ADME ¹	Tier ¹	Drug which involves protein target and therapeutic use (approved only) ²
<i>IGFBP1</i>	IGFBP1	Insulin-like growth factor binding protein 1	Lower BMI increases protein	Y	Y	N	2	Mecasermin: IGFBP1 is a carrier for this recombinant IGF1. Used in growth failure as a result of insulin-like growth factor 1 deficiency in paediatric patients
<i>ADH4</i>	ADH4	Alcohol dehydrogenase 4	Lower BMI reduces protein	Y	N	Y	1	Ethanol: chronic pain NADH: nutritional supplement
<i>RTN4R</i>	Nogo Receptor	Reticulon-4 receptor	Lower BMI reduces protein	Y	Y	N	3A	NA
<i>IL1RN</i>	IL-1Ra	Interleukin-1 receptor antagonist protein	Lower BMI reduces protein	N	Y	N	3A	NA
<i>IGFBP2</i>	IGFBP2	Insulin-like growth factor binding protein-2	Lower BMI increases protein	Y	Y	N	2	Mecasermin: IGFBP2 is a carrier for this recombinant IGF1. Used in growth failure as a result of insulin-like growth factor 1 deficiency in paediatric patients
<i>BCAN</i>	PGCB	Brevican core protein	Lower BMI increases protein	N	Y	N	3B	NA

Table 3. Exploration of proteins modified by weight loss interventions and in a Mendelian randomisation framework as therapeutic targets. ¹Information from Supplementary Table 1 in Finan et al.³⁴ Y = yes, N = no. Small molecule = chemically derived, biologic = extracted from living organism. ²Information from DrugBank³³. ADME = absorption, digestion, metabolism, excretion³⁴. Tier 1 = target with approved small molecule or biologics, or drug candidate in clinical phase. Tier 2 = target with known bioactive drug-like small molecular binding partners or have high similarity to approved drug targets. Tier 3 = structural similarities to approved drugs (to a lesser degree than Tier 2) or are a protein with known potential to be targeted (such as a G protein-coupled receptor)³⁴. More details can be found by Finan et al.

therapeutic interventions, but based on their structure they may have potential to be a drug target. A summary of these targets and the currently approved drugs is provided in Table 3.

Discussion

This study provides an in-depth characterization of the effect of weight loss on circulating proteins by combining the results from two weight loss interventions that are routinely used by health services: a low-calorie TDR and bariatric surgery. Results demonstrated that the low-calorie diet and bariatric surgery had a broad effect on protein levels. These analyses identified consistent evidence for an effect of weight loss interventions on 23 proteins. Integrating results from a recently published MR study estimating the causal effect of BMI on plasma proteins provided a third line of evidence for a causal role of BMI in the circulating levels of 6 of the 23 proteins, with estimates that were consistent in direction for 18 out of 20 proteins with available MR estimates. Through exploring these proteins as drug targets, we found that some are currently involved in treating growth failure and others have the potential to be therapeutically targetable if they are found to have a role in disease.

Previous studies have explored the effect on the proteome of either caloric restriction^{10–15} or bariatric surgery^{16–19,35,36} as a means of characterising the proteomic features of BMI. However, this study was able to advance this by comparing proteomic effects across two common weight loss treatments³⁷. Our results demonstrated estimates that agreed with effects previously observed by other groups. These changes include a low-calorie diet increasing levels of IGFBP1^{11,14}, IGFBP2^{12,14} and interleukin 1 receptor accessory protein (IL1RAP)^{10,14} as well as reducing levels of RTN4R¹⁴. We also replicated effects observed in bariatric surgery studies, where surgical weight loss led to increased levels of IGFBP1 and IGFBP2 and a reduction in levels of FCN2^{18,19}. IGFBP1 and IGFBP2 are known to be involved in the transport of IGF-1 and have a role in glucose metabolism and insulin sensitivity^{38,39}. Higher levels of IGFBP-2 have been shown to be associated with higher insulin sensitivity, and lower levels of plasma insulin following bariatric surgery, possibly due to increased IGFBP-2 mRNA expression following surgically induced weight loss⁴⁰. The comparison of the effect estimates with existing studies provides us with confidence that we can capture robust proteomic effects resulting from bariatric surgery and caloric restriction and suggest that a comparison of effects across study designs is appropriate.

The ability to identify consistent proteomic effects across independent evidence sources is one of the key strengths of the study. Bariatric surgery and caloric restriction resulted in a reduction of proteins associated with inflammation such as IL-1Ra⁴¹, sCD163 and sE-selectin. sCD163 and sE-selectin are soluble receptors: sCD163 is shed from macrophages/monocytes under inflammatory conditions⁴² and sE-selectin is involved in the adhesion of leukocytes to the endothelial wall and plays a role in atherosclerosis⁴³. Both interventions altered circulating levels of ADH4, which was supported by MR analyses. ADH4 is a protein in the alcohol dehydrogenase family, which catalyses NAD-dependent oxidation and is important in alcohol metabolism. Circulating levels of ADH4 have been shown to be associated with BMI observationally²⁵, however evidence is limited on whether levels of this protein are altered with a weight loss intervention. MR studies have provided evidence that non-alcoholic fatty liver disease (NAFLD) raises levels of ADH4⁴⁴, therefore weight loss reducing ADH4 could be an indicator of an improvement in liver function. Levels of ADH4 were also positively associated with the polygenic risk score for incident T2D⁴⁵, therefore it is possible that ADH4 could be involved in the relationship between adiposity and T2D. This study also found consistent effects across the three study designs that lower BMI raises levels of BCAN. The gene encoding brevican core protein (*BCAN*) is highly expressed in the central nervous system and is thought to have a role in the formation of the brain extracellular matrix⁴⁶. The levels of BCAN were shown to be increased 12 years following bariatric surgery¹⁹; our study supports that the effect can also be seen after a shorter bariatric surgery follow up of three years, and 1 year following a dietary intervention. A recent study suggested that higher levels of BCAN may be linked to a reduction in cardiovascular-specific mortality⁴⁷.

One way to explore the function of proteins altered by BMI is to look at whether drugs which target these proteins have an approved therapeutic use. Proteins such as IGFBP1 and IGFBP2 are currently targeted by drugs to treat growth disorders³³. Other proteins altered by BMI and weight change have been suggested to be potential drug targets, but currently do not have any approved drugs acting at the protein, such as the RTN4R, IL-1Ra and BCAN³⁴. Although no approved drugs, an antagonist for RTN4R was in clinical trials for the treatment of multiple sclerosis (however no development has been reported in such trials). Drugs which mimic IL-1Ra are currently in Phase II clinical trials for the treatment of urinary tract infections (Citeline Pharma Intelligence citeline.informa.com). If these proteins are established as having a role in disease, they could be useful targets for future drug repurposing/development. Previous studies comparing protein levels in individuals with and without T2D⁴⁸ suggests that the changes in levels of proteins, including ADH4, RTN4R, IGFBP2 and BCAN, following the weight loss interventions support a proteomic signature of remission of T2D. Further characterization of the role of these proteins in physiological processes and in disease is required through use of MR and laboratory studies. Despite the possibility that changes in protein composition related to changing BMI may have adverse downstream effects, it is also important to recognise that changes in some (or many) of these proteins may simply be biomarkers for adiposity. The exploration of these proteins as drug targets is just one of many tools that should be employed to understand how a change in circulating proteins relates to health.

Although we have data from two comprehensive and well validated technologies, and are making good use of existing data, the two trials included in the current study utilized different proteomic technologies. Where possible, we explored the correlation of proteins across Olink and the SomaScan[®] by incorporating published correlation coefficients and have provided these results in full. Here, a large proportion of proteins which had evidence for consistent effects with both interventions had strong correlations. We therefore believe that the utility of comparing estimates across studies outweighs the limitations that arise by cross-platform comparisons. Second, it is important to note that proteomic profiling was run on plasma samples in the DiRECT study and in serum in By-Band-Sleeve. Serum is the liquid remaining once the blood has been allowed to clot, whereas plasma

is derived from blood that has not clotted as it has been taken into a vacutainer containing an anticoagulant⁴⁹. Proteins may have different absolute abundance in plasma or serum, however the ability to detect changes in protein levels should be similar⁵⁰. The two trials had differing follow-up times, with By-Band-Sleeve at 3-years and DiRECT at 1-year. All participants in DiRECT had T2D, whereas By-Band-Sleeve had a mix of participants with and without T2D. By-Band-Sleeve also had a higher proportion of females. Bariatric surgery and caloric restriction generally induce differing degrees of weight loss, therefore the absolute magnitude of the changes in proteins across interventions are not directly comparable. Whilst these differences in patient populations could explain some of the discordance in protein effects across interventions, the heterogeneity across studies is a strength where observing consistent protein effects. Analyses performed in the current study estimated the average effects across both males and females, however, future work could include exploring sex-specific effects within these studies.

Overall, this study has provided an extensive characterisation of caloric restriction or bariatric surgery interventions at the level of circulating proteins. We have exemplified how triangulation, using two independent trials and an MR study, can be used to identify proteins that are affected by body composition and which may play a role in the considerable pathophysiology associated with excess adiposity. Further MR and laboratory studies are required to investigate the role of these proteins in health and disease.

Methods

Datasets and statistical analysis

DiRECT study overview

Samples analysed were collected from participants enrolled in the DiRECT trial. Participants enrolled were between 20 and 65 years of age, diagnosed with T2D within the previous six years and had a BMI of between 27 and 45 kg/m². Ethics approval was granted by West 3 Ethics Committee in January, 2014, with approvals by the National Health Service (NHS) health board areas in Scotland and clinical commissioning groups in Tyneside⁵¹. A statistical analysis plan was written and access to data was granted by DiRECT trial principal investigators. All methods were performed in accordance with the relevant guidelines and regulations. Participants were excluded if: they were using insulin, had a glycosylated haemoglobin (HbA1c) concentration of $\geq 12\%$ (≥ 108 mmol/mol), had more than 5 kg weight loss in the preceding six months and/or had an estimated glomerular filtration rate of < 30 mL/min per 1.732 m². Other exclusion criteria include malignancy, heart failure, recent myocardial infarction (< 6 months), enrolment in other clinical trials, addiction to illegal drugs, difficulty in understanding the study, current use of drugs to treat obesity, eating disorders, pregnancy or admission to hospital for depression or use of antipsychotic medication⁵¹. General practitioner (GP) practices were assigned to control or intervention, which was dependent on the practice list size (number of patients registered to each practice). This was done to ensure that the intervention/control allocations were balanced across centres and list size (small ≤ 5700 or large > 5700). Therefore, centre and list size are variables used for stratified randomization²⁷. Participants in the control group received best-practice care by guidelines. The intervention group followed the Counterweight-Plus weight management programme⁵². This programme involved a total diet replacement (TDR) phase using a low energy diet (825–853 kcal/day) for 3–5 months. Following the TDR there was a structured food reintroduction phase of 2–8 weeks. Participants then attended monthly weight loss maintenance visits. Those in the intervention group had their antidiabetic and antihypertensive drugs discontinued. In total, there were 306 individuals recruited into the study, with 149 patients included in each intention-to-treat population (in both intervention and control groups) after removal of participants that had been randomised in error or removed consent⁵¹.

Age and sex were self-reported. Height was measured with the Frankfort plane horizontal, with a portable stadiometer (Chasmors Ltd, London). Weight was measured using Class 111 approved calibrated scales (Marsden Group UK). Blood was donated at various timepoints including at baseline (week 0) and at 1 year (~ week 52), when HDL-cholesterol, triglycerides, HbA1c and plasma glucose were measured. Systolic blood pressure was measured with the patient seated, rested and with legs uncrossed for ≥ 5 min. BMI (kg/m²) was calculated by dividing the weight (kg) by the square of the height (m).

DiRECT proteomics and statistical analysis

Blood was taken from participants by venipuncture into 9 mL vacutainers with ethylenediaminetetraacetic acid (EDTA) at baseline and at 1-year post-randomisation. Blood samples were centrifuged to derive plasma samples and plasma was stored at -80 °C. Protein detection was performed by the SomaScan[®] assay by SomaLogic. This was performed on 569 samples from 302 individuals. This technique uses Slow Off-rate Modified Aptamers (SOMAmers) which make direct contact with proteins and quantifies protein levels in relative fluorescence units (RFUs) by using a DNA microarray⁵³. This quantification is a product of both the affinity of the SOMAmer for the target and the concentration of the protein. Measurements returned by SomaLogic had undergone internal processing, where data were hybridised control normalised, intraplate median signal normalised, plate scaled, calibrated and adaptive normalized; further details of these adjustments can be found in the technical note by SomaLogic⁵⁴. There were 5284 proteins included in the array, of which 4601 proteins remained after internal technical quality control (QC), including the removal of non-human proteins. The proteomic data were then subject to a study-level QC using the “metaboprep” R package⁵⁵, with data from both timepoints QC'd together. Although this package was primarily developed for use with metabolomic data, the functions are also applicable to proteomic data. The following input parameters were used for exclusion of protein measurements: extreme missingness ($> 80\%$) for each individual or each protein, user defined missingness of $> 20\%$ for each individual or each protein, protein measurement > 5 SDs from the total peak area (sum of protein level for each individual at proteins with no missingness), and > 5 SDs from the mean of principal components (PCs) PC1 and PC2. These filtering criteria excluded 4 samples based on PC outliers, leaving 565 samples from 300 participants. On

merging with clinical data (as analysed in the primary results paper⁵¹), 292 participants (146 per study group) and 552 samples remained and were therefore included in statistical analyses. In the control group, there were 145 samples at baseline and 143 at endpoint. In the intervention group, there were 143 samples at baseline and 121 samples at endpoint (Fig. 6).

The metaboprep package also calculated the number of independent proteins by using pairwise Spearman's correlation coefficients between proteins (2380 representative proteins based on correlation coefficient of 0.5). The Shapiro–Wilk test was implemented to identify proteins which have a normal distribution ($W \geq 0.95$). Only 644 out of 4601 proteins had W statistics ≥ 0.95 and therefore all data were transformed to meet normality assumptions for analyses. Data were rank-based inverse normal transformed to give a mean of 0 and a standard deviation of 1 and data were adjusted for age and sex. The units of protein measurements are therefore in normalised SD units. Analyses were performed using R version 3.6.1.

The effect of the intervention on plasma proteins was estimated using a linear mixed model (`lmer()` function from the “lme4” R package). Within this model, the timepoint (visit), treatment group, centre and list size were included as fixed effects and the subject was included as a random effect. The centre and list size of each GP practice were stratification variables within the trial, therefore they were included as covariables⁵⁶. The estimate for the effect of the intervention on protein was reported as the interaction coefficient for treatment group (with control group as reference) and timepoint (with baseline as reference), where the direction of effect indicates the change in protein level when comparing endpoint to baseline (i.e. a negative slope indicates the intervention reduces the level of the protein). A p -value was derived by performing an ANOVA of two fitted models, one including and one excluding the interaction term, under the conditions of a maximum likelihood (ML) model. A summary of this analysis is provided in Fig. 6. We used a Bonferroni multiple adjusted p -value of $0.05/2380 = 2.1 \times 10^{-5}$ to guide strength of associations based on the number of representative proteins at a Spearman's correlation coefficient of $\rho = 0.5$. All effect estimates, measures of precision and p -values are presented in full in the supplementary material.

By-Band-Sleeve study overview

By-Band-Sleeve is a multi-centre trial which aims to determine which bariatric surgery type is the most effective for weight loss and quality of life (out of Roux-en-Y gastric bypass, the laparoscopic adjustable gastric band and the sleeve gastrectomy) (NIHR09/127/53, UK) at three years after randomisation. The trial was granted research ethics approval by the Southwest Frenchay Research Ethics Committee (reference 11/SW/0248). Written informed consent was obtained from all participants. The study is funded by the NIHR and aims to report this year. This trial began recruitment (as part of an internal pilot phase) in December 2012 in two centres, where the initial bariatric surgeries included the gastric bypass and gastric band⁵⁷. The sleeve gastrectomy was later added as a third bariatric surgery within the trial and the number of participating centres increased to twelve⁵⁸. The trial completed recruitment in September 2019 with 1341 patients having been randomised. The trial reported the cardiovascular disease history, medication, full blood count measurements and cardiometabolic risk factors of participants. Due to the results of the trial not being published, the exact BMI change that occurred in By-Band-Sleeve cannot yet be reported. A statistical analysis plan was written and access to data was granted by By-Band-Sleeve trial principal investigators. All methods were performed in accordance with the relevant guidelines and regulations.

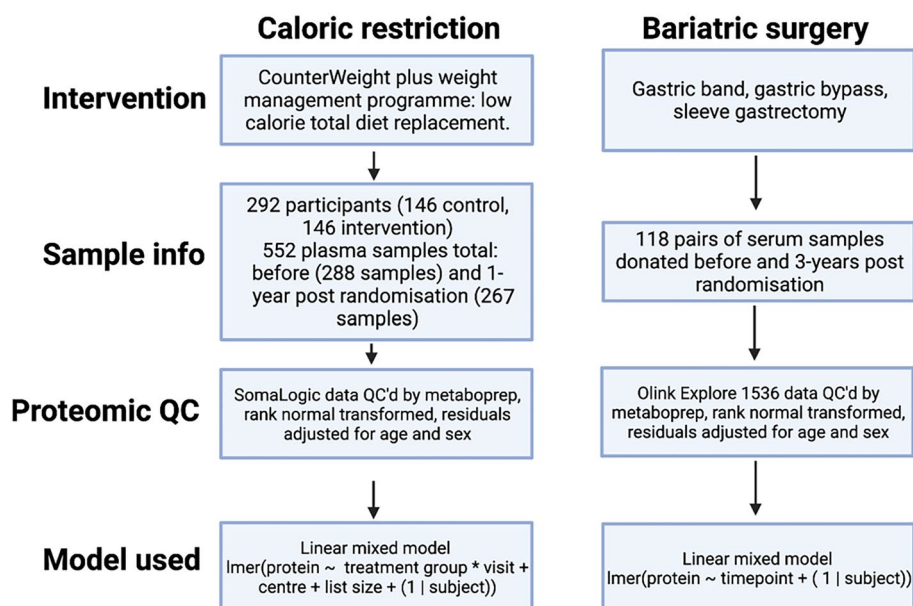


Figure 6. Overview of the two interventions used and a summary of analyses performed.

Ten of the twelve participating centres elected to collect samples for future research with all participants enrolled at these centres being given the option to consent to sample collection at baseline (pre-randomisation) and 36-months post-randomization for this purpose. Samples were collected using 4 ml clot activator gel vacutainers and centrifuged at the specific site then stored at -80°C . Samples were shipped on dry ice. A subset of these samples were used in this study for proteomic profiling. Specifically, only sample pairs (those collected from the same patient before and after surgery) were selected for proteomics analysis. These sample pairs were collected at Musgrove Park Hospital (Taunton, UK) and were available for analysis as of December 2020.

By-Band-Sleeve proteomics and statistical analysis

Samples were thawed and randomly aliquoted across three plates, ensuring pairs of samples (from the same individual) were on the same plate, and sent to Olink in February 2021. This resulted in 250 serum samples from 125 participants being analysed. Within this manuscript, data from the By-Band-Sleeve trial refers to this subset of patients and their data from the trial. Samples were analysed by the Olink Explore 1536 panel (Olink Proteomics, Uppsala, Sweden). Olink uses proximity extension assay (PEA) technology to detect and quantify protein levels⁵⁹. This technology uses pairs of antibodies bound to DNA tags. When the antibodies bind to the protein, the DNA tags hybridise and can be quantified using next generation sequencing Illumina® NovaSeq platform. Proteins are measured in normalized expression (NPX) units which are on a \log_2 scale. This panel provides a maximal readout of 1536 proteins (Supplementary Table 6). Olink detected and returned data for 1472 proteins after excluding proteins that failed internal technical QC procedures. Samples with QC warnings from Olink (internal control deviation of more than ± 0.3 NPX, at least 500 matched counts or deviation of negative controls < 5 SDs of the predefined value) were left in but further study-level QC was implemented using “metaboprep”⁵⁵. Olink flagged proteins which fell below a lower LOD, however as the data were further QC'd by “metaboprep” and subsequently rank-based inverse normal transformed, no proteins were excluded based on the LOD. It has also been reported that including values that are lower than the LOD helps with increasing statistical power and in increasing the normality of the data⁵⁰. Information about the percentage of samples that fell below the lower LOD for each protein are provided alongside the results.

The input parameters for exclusion in “metaboprep” were the same as for the proteomic data in DiRECT. From a total of 250 samples (125 pairs) and 1472 proteins (following Olink QC above), two samples were excluded as PC outliers (leaving 123 complete pairs). As in DiRECT, a Spearman's rho of 0.5 (tree cut height of $1-\rho=0.5$) was selected to determine the number of independent or representative proteins, which totalled 805. The proteomic data were rank-based inverse normal transformed, residuals then adjusted for age and sex, and the residuals used for the main analysis. We removed participants from the analysis if they did not have a surgery date or if the surgery date was after their 36-month post-randomisation appointment, therefore suggesting they had not yet undergone surgery ($N=4$). Participants were also removed from analyses if, after study-level sample QC, they had missing proteomic data at one of the timepoints ($N=1$). This left 118 individuals for the main analyses.

Similar to DiRECT, a linear mixed model was used to assess the effect of timepoint (at baseline and 3 years after randomisation) on plasma proteins, where the timepoint was added as a fixed effects predictor and subject as a random effects predictor. P-values were derived by comparing models with and without timepoint using an ANOVA, where a multiple testing adjusted p-value of $0.05/805$ (number of independent proteins at a correlation of $r=0.5$) = 6.2×10^{-5} was used to guide strength of associations. R version 4.0.3 was used for the analysis of By-Band-Sleeve data.

INTERVAL study

INTERVAL was a trial that aimed to assess the safety and efficacy of reducing the time between blood donations in a population free from major (self-reported) disease. The study enrolled around 50,000 participants. A sub-sample of the cohort ($N=3,301$) also had plasma protein measurements on the SomaScan (SomaLogic) platform. We utilized these MR results which we have previously published to identify proteins with causal evidence for BMI-driven effects²⁰. Within the analysis, a genetic risk score (GRS) for BMI was constructed using 654 SNPs weighted by available betas from summary statistics of the genetic variants associated with BMI from a recent GWAS meta-analysis⁶⁰. Two-stage least squares analysis was used to derive one-sample MR estimates for the effect of BMI on 4034 proteins (3622 unique proteins as some proteins were targeted by more than one SOMAmer). MR analyses were conducted in the 2729 participants with genetic data, BMI, and protein data. MR results provide the average difference in protein in rank normalized SD units per 1 normalized SD (~ 4.7 kg/m) higher BMI.

Comparison analyses

The effects of caloric restriction and bariatric surgery-induced weight loss were compared to identify consistencies in signal. Results were combined by restricting both results to unique UniProt IDs, then merging DiRECT and By-Band-Sleeve results based on the UniProt ID of the protein⁶¹. Consistency was determined by effect estimates having the same direction of effect and the corresponding p-values passing pre-specified thresholds in both studies. Opposing effects were defined as estimates displaying opposite directions of effects (for example, the protein levels were raised with one intervention but reduced with the other) and corresponding p-values passing the pre-specified thresholds. Proteins were categorised as only being associated with one intervention where the p-value only passed the pre-specified threshold in one study. Differing effects or null effects in only one study could point towards intervention-specific effects. As proteins were measured using different technologies in each trial, we also explored whether it is appropriate to compare estimates derived using SomaLogic and Olink. For this, we used published correlation information from Haslam et al. to explore how well protein measurements correlate across platforms³². Haslam et al. calculated Spearman's correlation coefficients for every protein that was detected in both platforms, along with the 95% confidence intervals. We integrated these correlation results

with our results from each intervention to aid interpretation. For example, the correlation data can help infer whether discordant results are likely due to biological effects specific to the intervention and/or study sample, or whether the differences may be arising due to technological differences across platforms, such as the technologies picking up different isoforms or variants of the same protein³². References made herein to ‘correlations across platforms’ refer to these published estimates³².

Proteins with consistent or opposite effects across interventions were compared with published MR results in INTERVAL again using the UniProt ID to merge the information. As we used the MR results to look up pre-specified protein results, we deemed MR estimates as putatively causal effects if the 95% confidence intervals did not cross the null ($p < 0.05$). Proteins which had evidence for a consistent direction of effect across all three study designs were explored as drug targets by searching on DrugBank³³ and using published drug target data from Finan et al.³⁴. We determined if the protein is currently (or has the capability to be) targeted by a drug and extracted what currently approved drugs are used for. This was performed as one way of exploring the possible role of the protein in health and disease.

Data and code availability

The R code for main analyses presented in this manuscript has been made publicly available on GitHub at https://github.com/lucygodswaard/Proteome_comparison_paper. The terms of participant consent in DiRECT does not allow making the study data freely available in its raw form. The data used for analysis will be placed on a research data repository (<https://researchdata.gla.ac.uk/>) with access given to researchers subject to appropriate data sharing agreements. For the By-Band-Sleeve trial, following publication of the main trial results, anonymised individual patient data will be made available upon request to the chief investigator for secondary research, conditional on assurance from the secondary researcher that the proposed use of the data is compliant with the Medical Research Council Policy on Data Sharing regarding scientific quality, ethical requirements, and value for money, and is compliant with the National Institute for Health and Care Research policy on data sharing. A minimum requirement with respect to scientific quality will be a publicly available prespecified protocol describing the purpose, methods, and analysis of the secondary research (e.g., a protocol for a Cochrane systematic review), approved by a UK Research Ethics Committee or other similar, approved ethics review body. Participant identifiers will not be passed on to any third party.

Received: 11 July 2023; Accepted: 8 November 2023

Published online: 29 November 2023

References

- Williams, S. A. *et al.* Plasma protein patterns as comprehensive indicators of health. *Nat. Med.* **25**, 1851–1857. <https://doi.org/10.1038/s41591-019-0665-2> (2019).
- Timpson, N. J. *et al.* C-reactive protein levels and body mass index: elucidating direction of causation through reciprocal Mendelian randomization. *Int. J. Obes. (Lond.)* **35**, 300–308. <https://doi.org/10.1038/ijo.2010.137> (2011).
- Füzéry, A. K., Levin, J., Chan, M. M. & Chan, D. W. Translation of proteomic biomarkers into FDA approved cancer diagnostics: Issues and challenges. *Clin. Proteom.* **10**, 13. <https://doi.org/10.1186/1559-0275-10-13> (2013).
- Malik, P. *et al.* Biomarkers and outcomes of COVID-19 hospitalisations: Systematic review and meta-analysis. *BMJ Evid. Based Med.* **26**, 107–108. <https://doi.org/10.1136/bmjebm-2020-111536> (2021).
- Khan, S. S. *et al.* Association of body mass index with lifetime risk of cardiovascular disease and compression of morbidity. *JAMA Cardiol.* **3**, 280–287. <https://doi.org/10.1001/jamacardio.2018.0022> (2018).
- Garg, S. K., Maurer, H., Reed, K. & Selagamsetty, R. Diabetes and cancer: Two diseases with obesity as a common risk factor. *Diabetes Obes. Metab.* **16**, 97–110. <https://doi.org/10.1111/dom.12124> (2014).
- Kortt, M. & Baldry, J. The association between musculoskeletal disorders and obesity. *Aust. Health Rev.* **25**, 207–214 (2002).
- Lawlor, D. A., Tilling, K. & Davey Smith, G. Triangulation in aetiological epidemiology. *Int. J. Epidemiol.* **45**, 1866–1886. <https://doi.org/10.1093/ije/dyw314> (2016).
- Gloy, V. L. *et al.* Bariatric surgery versus non-surgical treatment for obesity: a systematic review and meta-analysis of randomised controlled trials. *BMJ* **347**, f5934. <https://doi.org/10.1136/bmj.f5934> (2013).
- Geyer, P. E. *et al.* Proteomics reveals the effects of sustained weight loss on the human plasma proteome. *Mol. Syst. Biol.* **12**, 901. <https://doi.org/10.1552/msb.20167357> (2016).
- Figarska, S. M. *et al.* Proteomic profiles before and during weight loss: Results from randomized trial of dietary intervention. *Sci. Rep.* **10**, 7913. <https://doi.org/10.1038/s41598-020-64636-7> (2020).
- Moreno, S. *et al.* The differential plasma proteome of obese and overweight individuals undergoing a nutritional weight loss and maintenance intervention. *Proteom. Clin. Appl.* **12**, 1. <https://doi.org/10.1002/prca.201600150> (2018).
- Piening, B. D. *et al.* Integrative personal omics profiles during periods of weight gain and loss. *Cell Syst.* **6**, 157–170.e158. <https://doi.org/10.1016/j.cels.2017.12.013> (2018).
- Carayol, J. *et al.* Protein quantitative trait locus study in obesity during weight-loss identifies a leptin regulator. *Nat. Commun.* **8**, 2084. <https://doi.org/10.1038/s41467-017-02182-z> (2017).
- Bruderer, R. *et al.* Analysis of 1508 plasma samples by capillary-flow data-independent acquisition profiles proteomics of weight loss and maintenance. *Mol. Cell Proteom.* **18**, 1242–1254. <https://doi.org/10.1074/mcp.RA118.001288> (2019).
- Wewer Albrechtsen, N. J. *et al.* Plasma proteome profiling reveals dynamics of inflammatory and lipid homeostasis markers after Roux-En-Y gastric bypass surgery. *Cell Syst.* **7**, 601–612. <https://doi.org/10.1016/j.cels.2018.10.012> (2018).
- Culnan, D. M., Cooney, R. N., Stanley, B. & Lynch, C. J. Apolipoprotein A-IV, a putative satiety/antiatherogenic factor, rises after gastric bypass. *Obesity (Silver Spring)* **17**, 46–52. <https://doi.org/10.1038/oby.2008.428> (2009).
- Shah, R. V. *et al.* Proteins altered by surgical weight loss highlight biomarkers of insulin resistance in the community. *Arterioscler. Thromb. Vasc. Biol.* **39**, 107–115. <https://doi.org/10.1161/ATVBAHA.118.311928> (2019).
- Yousri, N. A. *et al.* Proteome-wide associations with short- and long-term weight loss and regain after Roux-en-Y gastric bypass surgery. *Obesity (Silver Spring)* **30**, 129–141. <https://doi.org/10.1002/oby.23303> (2022).
- Goudswaard, L. J. *et al.* Effects of adiposity on the human plasma proteome: Observational and Mendelian randomisation estimates. *Int. J. Obes. (Lond.)* **45**, 2221–2229. <https://doi.org/10.1038/s41366-021-00896-1> (2021).

21. Zaghlool, S. B. *et al.* Revealing the role of the human blood plasma proteome in obesity using genetic drivers. *Nat. Commun.* **12**, 1279. <https://doi.org/10.1038/s41467-021-21542-4> (2021).
22. Burgess, S., Butterworth, A. & Thompson, S. G. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet. Epidemiol.* **37**, 658–665. <https://doi.org/10.1002/gepi.21758> (2013).
23. Lawlor, D. A., Harbord, R. M., Sterne, J. A., Timpson, N. & Davey Smith, G. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Stat. Med.* **27**, 1133–1163. <https://doi.org/10.1002/sim.3034> (2008).
24. Freeman, G., Cowling, B. J. & Schooling, C. M. Power and sample size calculations for Mendelian randomization studies using one genetic instrument. *Int. J. Epidemiol.* **42**, 1157–1163. <https://doi.org/10.1093/ije/dyt110> (2013).
25. Sun, B. B. *et al.* Genetic regulation of the human plasma proteome in 54,306 UK Biobank participants. *bioRxiv*, 2022.2006.2017.496443 (2022). <https://doi.org/10.1101/2022.06.17.496443>
26. Ferkingstad, E. *et al.* Large-scale integration of the plasma proteome with genetics and disease. *Nat. Genet.* **53**, 1712–1721. <https://doi.org/10.1038/s41588-021-00978-w> (2021).
27. Kernan, W. N., Viscoli, C. M., Makuch, R. W., Brass, L. M. & Horwitz, R. I. Stratified randomization for clinical trials. *J. Clin. Epidemiol.* **52**, 19–26. [https://doi.org/10.1016/s0895-4356\(98\)00138-3](https://doi.org/10.1016/s0895-4356(98)00138-3) (1999).
28. McGuire, H. *et al.* Management of type 2 diabetes in adults: summary of updated NICE guidance. *BMJ* **353**, i1575. <https://doi.org/10.1136/bmj.i1575> (2016).
29. Paul, S., Thomas, G., Majeed, A., Khunti, K. & Klein, K. Women develop type 2 diabetes at a higher body mass index than men. *Diabetologia* **55**, 1556–1557. <https://doi.org/10.1007/s00125-012-2496-2> (2012).
30. Douglas, I. J., Bhaskaran, K., Batterham, R. L. & Smeeth, L. Bariatric surgery in the United Kingdom: A cohort study of weight loss and clinical outcomes in routine clinical care. *PLoS Med* **12**, e1001925. <https://doi.org/10.1371/journal.pmed.1001925> (2015).
31. Group & B. S. C., Roux-en-Y gastric bypass, gastric banding, or sleeve gastrectomy for severe obesity: Baseline data from the By-Band-Sleeve randomized controlled trial. *Obesity (Silver Spring)* **31**, 1290–1299. <https://doi.org/10.1002/oby.23746> (2023).
32. Haslam, D. E. *et al.* Stability and reproducibility of proteomic profiles in epidemiological studies: comparing the Olink and SOMAscan platforms. *Proteomics*, e2100170 (2022). <https://doi.org/10.1002/pmic.202100170>
33. Wishart, D. S. *et al.* DrugBank: A comprehensive resource for in silico drug discovery and exploration. *Nucl. Acids Res.* **34**, D668–672. <https://doi.org/10.1093/nar/gkj067> (2006).
34. Finan, C. *et al.* The druggable genome and support for target identification and validation in drug development. *Sci. Transl. Med.* **9**, 1. <https://doi.org/10.1126/scitranslmed.aag1166> (2017).
35. Fachim, H. A. *et al.* Relationship between the plasma proteome and changes in inflammatory markers after bariatric surgery. *Cells* **10**, 1. <https://doi.org/10.3390/cells10102798> (2021).
36. Jüllig, M. *et al.* Lower fetuin-A, retinol binding protein 4 and several metabolites after gastric bypass compared to sleeve gastrectomy in patients with type 2 diabetes. *PLoS One* **9**, e96489. <https://doi.org/10.1371/journal.pone.0096489> (2014).
37. Alfadda, A. A. *et al.* A proteomic analysis of excreted and circulating proteins from obese patients following two different weight-loss strategies. *Exp. Biol. Med. (Maywood)* **239**, 568–580. <https://doi.org/10.1177/1535370214523894> (2014).
38. Rajpathak, S. N. *et al.* The role of insulin-like growth factor-I and its binding proteins in glucose homeostasis and type 2 diabetes. *Diabetes Metab. Res. Rev.* **25**, 3–12. <https://doi.org/10.1002/dmrr.919> (2009).
39. Haywood, N. J., Slater, T. A., Matthews, C. J. & Wheatcroft, S. B. The insulin like growth factor and binding protein family: Novel therapeutic targets in obesity & diabetes. *Mol. Metab.* **19**, 86–96. <https://doi.org/10.1016/j.molmet.2018.10.008> (2019).
40. Faramia, J. *et al.* IGFBP-2 partly mediates the early metabolic improvements caused by bariatric surgery. *Cell Rep. Med.* **2**, 100248. <https://doi.org/10.1016/j.xcrm.2021.100248> (2021).
41. Arend, W. P., Malyak, M., Guthridge, C. J. & Gabay, C. Interleukin-1 receptor antagonist: Role in biology. *Annu. Rev. Immunol.* **16**, 27–55. <https://doi.org/10.1146/annurev.immunol.16.1.27> (1998).
42. Møller, H. J. Soluble CD163. *Scand. J. Clin. Lab. Invest.* **72**, 1–13. <https://doi.org/10.3109/00365513.2011.626868> (2012).
43. Lee, C. H. *et al.* Serum E-selectin concentration is associated with risk of metabolic syndrome in females. *PLoS One* **14**, e0222815. <https://doi.org/10.1371/journal.pone.0222815> (2019).
44. Ghodasian, N. *et al.* Electronic health record-based genome-wide meta-analysis provides insights on the genetic architecture of non-alcoholic fatty liver disease. *Cell Rep Med* **2**, 100437. <https://doi.org/10.1016/j.xcrm.2021.100437> (2021).
45. Ritchie, S. C. *et al.* Integrative analysis of the plasma proteome and polygenic risk of cardiometabolic diseases. *bioRxiv*, 2019.2012.2014.876474 (2021). <https://doi.org/10.1101/2019.12.14.876474>
46. Fagerberg, L. *et al.* Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell Proteom.* **13**, 397–406. <https://doi.org/10.1074/mcp.M113.035600> (2014).
47. Sethi, A., Raj, A., Wright, K. & Melamud, E. Plasma Proteomic Determinants of Common Causes of Mortality. *PREPRINT (version 1) available at Research Square* (2023). <https://doi.org/10.21203/rs.3.rs-2626017/v1>
48. Gudmundsdottir, V. *et al.* Circulating protein signatures and causal candidates for type 2 diabetes. *Diabetes* **69**, 1843–1853. <https://doi.org/10.2337/db19-1070> (2020).
49. Issaq, H. J., Xiao, Z. & Veenstra, T. D. Serum and plasma proteomics. *Chem. Rev.* **107**, 3601–3620. <https://doi.org/10.1021/cr068287r> (2007).
50. Wik, L. *et al.* Proximity extension assay in combination with next-generation sequencing for high-throughput proteome-wide analysis. *Mol. Cell Proteom.* **20**, 100168. <https://doi.org/10.1016/j.mcpro.2021.100168> (2021).
51. Lean, M. E. *et al.* Primary care-led weight management for remission of type 2 diabetes (DiRECT): An open-label, cluster-randomised trial. *Lancet* **391**, 541–551. [https://doi.org/10.1016/S0140-6736\(17\)33102-1](https://doi.org/10.1016/S0140-6736(17)33102-1) (2018).
52. Lean, M. *et al.* Feasibility and indicative results from a 12-month low-energy liquid diet treatment and maintenance programme for severe obesity. *Br. J. Gen. Pract.* **63**, e115–124. <https://doi.org/10.3399/bjgp13X663073> (2013).
53. Rohloff, J. C. *et al.* Nucleic acid ligands with protein-like side chains: Modified aptamers and their use as diagnostic and therapeutic agents. *Mol. Ther. Nucl. Acids* **3**, e201. <https://doi.org/10.1038/mtna.2014.49> (2014).
54. SomaLogic. SomaScan® v4 Data Standardization and File Specification Technical Note [White paper]. (2022). <https://somalogic.com/tech-notes/>.
55. Hughes, D. A. *et al.* metaboprep: An R package for preanalysis data description and processing. *Bioinformatics* <https://doi.org/10.1093/bioinformatics/btac059> (2022).
56. Scott, N. W., McPherson, G. C., Ramsay, C. R. & Campbell, M. K. The method of minimization for allocation to clinical trials. A review. *Control Clin. Trials* **23**, 662–674. [https://doi.org/10.1016/s0197-2456\(02\)00242-8](https://doi.org/10.1016/s0197-2456(02)00242-8) (2002).
57. Rogers, C. A. *et al.* The By-Band study: Gastric bypass or adjustable gastric band surgery to treat morbid obesity: Study protocol for a multi-centre randomised controlled trial with an internal pilot phase. *Trials* **15**, 53. <https://doi.org/10.1186/1745-6215-15-53> (2014).
58. Paramasivan, S. *et al.* Enabling recruitment success in bariatric surgical trials: pilot phase of the By-Band-Sleeve study. *Int. J. Obes. (Lond.)* **41**, 1654–1661. <https://doi.org/10.1038/ijo.2017.153> (2017).
59. Olink Proteomics. PEA – a high-multiplex immunoassay technology with qPCR or NGS readout [White paper]. <https://www.olink.com/application/pea-a-high-multiplex-immunoassay-technology-with-qpcr-or-ngs-readout-2/> (2020).
60. Yengo, L. *et al.* Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Hum Mol Genet* **27**, 3641–3649. <https://doi.org/10.1093/hmg/ddy271> (2018).

61. Katz, D. H. *et al.* Proteomic profiling platforms head to head: Leveraging genetics and clinical traits to compare aptamer- and antibody-based methods. *Sci Adv* **8**, eabm5164. <https://doi.org/10.1126/sciadv.abm5164> (2022).

Acknowledgements

The DiRECT formula diet was donated by Cambridge Weight Plan. Neither organisation had any input into the study design, data analysis or interpretation. We thank the National Health Service (NHS) Primary Care Research Network and Northeast Commissioning Support for their support and valuable input to recruitment. We thank Elaine Butler, Josephine Cooney, Sara-Jane Duffus, and Philip Stewart from the University of Glasgow for providing technical assistance; Helen Pilkington from the Newcastle upon Tyne Hospitals NHS Foundation Trust for providing research nurse support; and Sarah Weeden and Sarah Cadzow from the Robertson Centre for Biostatistics. We thank Wilma Leslie and research dietitians Louise McCombie, George Thom, Naomi Brosnahan and Alison Barnes for their contributions to the collection and transport of samples. We are enormously grateful to the GP practices, health-care professionals, and volunteers for their participation. The By-Band-Sleeve trial is led by JM Blazeby (Chief Investigator) and C Rogers (lead methodologist). We acknowledge the support of all other members of the NIHR By-Band-Sleeve Management Group: Robert C Andrews, John Bessant, James P Byrne, Nicholas Carter, Caroline Clay, Jenny L Donovan, Eleanor Gidman, Graziella Mazza, Mary O’Kane, Barnaby C Reeves, Nicki Salter, Janice L Thompson, Richard Welbourn and Sarah Wordsworth. We also thank all the By-Band-Sleeve contributors including the investigators, research dietitians and nurses, the independent trial steering committee and data monitoring and safety committee. We are grateful to all the patients who participated in this trial. By-Band-Sleeve is in the follow-up phase.

Author contributions

Conceptualization and design of the study—N.J.T., L.J.C., L.J.G., I.H., J.M.B., C.A.R. Preparation of the data—L.J.G., M.L.S., L.J.C. Performing analyses—L.J.G. Interpretation of the data—L.J.G., L.J.C., N.J.T., D.A.H. Writing first draft—L.J.G. Reviewing and/or editing manuscript—L.J.C., N.J.T., M.L.S., D.A.H., S.Z., K.S., I.H., M.L., N.S., P.W., A.M., R.T., J.M.B., C.A.R.

Funding

LJG is supported by a Transition Fellowship as part of the British Heart Foundation Accelerator Award (AA/18/1/34219) and an Academic Career Development Fund. NJT is a Wellcome Trust Investigator (202802/Z/16/Z), is the PI of the Avon Longitudinal Study of Parents and Children (MRC & WT 217065/Z/19/Z), is supported by the University of Bristol NIHR Biomedical Research Centre (BRC-1215-2001), the MRC Integrative Epidemiology Unit (MC_UU_00011/1) and works within the CRUK Integrative Cancer Epidemiology Programme (C18281/A29019). DAH & LJC are supported by NJT’s Wellcome Investigator Award (202802/Z/16/Z). MLS is supported by the Wellcome Trust through a PhD studentship (218495/Z/19/Z). RT is the CI of the study ‘Reversal of Type 2 diabetes Upon Normalisation of Energy intake in non-obese people’ (ReTUNE) funded by Diabetes UK (17/0005645). NS acknowledges funding support from the British Heart Foundation Research Excellence Award (RE/18/6/34217). JMB is an NIHR Senior Investigator. CAR was funded by the British Heart Foundation until 2016. The DiRECT study was funded as a Strategic Research Initiative by Diabetes UK (award number 13/0004691). The formula diet was donated by Cambridge Weight Plan. Neither organisation had any input into the study design, data analysis or interpretation. By-Band-Sleeve was funded by National Institute of Health and Care Research (NIHR) Health Technology Assessment Programme (HTA 09/127/53). We also acknowledge funding from the MRC ConDuCT-II Hub for Trials Methodology Research and the NIHR Biomedical Research Centre (BRC) at the University of Bristol. This trial was designed and delivered in collaboration with the Bristol Trials Centre, a UKCRC registered clinical trials unit (CTU), which is in receipt of NIHR CTU support funding. The Olink Explore 1536 analysis was funded by the Wellcome Trust (202802/Z/16/Z) and by the National Institute for Health and Care Research Bristol Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care. This research was funded in whole, or in part, by the Wellcome Trust [202802/Z/16/Z]. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. RT has received lecture honoraria from Eli Lilly, Nestle Health and Janssen. NS has received consulting and/or speaker honoraria from Abbott Laboratories, Amgen, AstraZeneca, Boehringer Ingelheim, Eli Lilly, Hanmi Pharmaceuticals, Janssen, Merck Sharp & Dohme, Novartis, Novo Nordisk, Pfizer, Roche Diagnostics, and Sanofi; and grant funding support paid to his University from AstraZeneca, Boehringer Ingelheim, Novartis, and Roche Diagnostics outside the submitted work. All other authors declare no competing interests.

Competing interests

RT has received lecture honoraria from Eli Lilly, Nestle Health and Janssen. NS has received consulting and/or speaker honoraria from Abbott Laboratories, Amgen, AstraZeneca, Boehringer Ingelheim, Eli Lilly, Hanmi Pharmaceuticals, Janssen, Merck Sharp & Dohme, Novartis, Novo Nordisk, Pfizer, Roche Diagnostics, and Sanofi; and grant funding support paid to his University from AstraZeneca, Boehringer Ingelheim, Novartis, and Roche Diagnostics outside the submitted work. All other authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-47030-x>.

Correspondence and requests for materials should be addressed to L.J.G.

Reprints and permissions information is available at www.nature.com/reprints.

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 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

© The Author(s) 2023