Integrated Analyses of Growth Differentiation Factor-15 Concentration and Cardiometabolic Diseases in Humans

Susanna Lemmelä1*, Eleanor M. Wigmore2*, Christian Benner1, Aki Havulinna1,3, Rachel M. Y. Ong4, Tibor Kempf5, Kai C. Wollert5, Stefan Blankenberg6, Tanja Zeller6, James E. Peters7,8, Veikko Salomaa3, Maria Fritsch9, Ruth March10, Aarno Palotie1,11,12, Mark Daly1,11,12, Adam S. Butterworth4,8,13,14, Mervi Kinnunen1, Dirk S. Paul2,4,13, Athena Matakidou2

Affiliations:
1Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Finland.
2Centre for Genomics Research, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Cambridge, UK.
3Finnish Institute for Health and Welfare, Helsinki, Finland.
4British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.
5Department of Cardiology and Angiology, Hannover Medical School, Germany.
6Clinic for General and Interventional Cardiology, University Heart Centre, Hamburg, Germany.
7Department of Immunology and Inflammation, Faculty of Medicine, Imperial College London, London, UK.
8Health Data Research UK Cambridge, Wellcome Genome Campus and University of Cambridge, Cambridge, UK.
9Bioscience Renal, Research and Early Development Cardiovascular, Renal and Metabolism (CVRM), BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden.
10Precision Medicine, Oncology R&D, AstraZeneca, Cambridge, UK.
11Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital, Boston, MA, USA.
12Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA.
13British Heart Foundation Centre of Research Excellence, University of Cambridge, Cambridge, UK.
14National Institute for Health Research Blood and Transplant Research Unit in Donor Health and Genomics, University of Cambridge, Cambridge, UK.

*Susanna Lemmelä
FIMM,
University of Helsinki
Helsinki
Finland
susanna.lemmela@helsinki.fi

*Eleanor M. Wigmore
Centre for Genomics Research, Discovery Sciences
AstraZeneca
Cambridge
UK
eleanor.wigmore@astrazeneca.com

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Growth differentiation factor 15 (GDF15) is a stress response cytokine that is elevated in several cardiometabolic diseases and has attracted interest as a potential therapeutic target. To further explore the association of GDF15 with human disease, we conducted a broad study into the phenotypic and genetic correlates of GDF15 concentration in up to 14,099 individuals. Assessment of 772 traits across 6,610 participants in FINRISK identified associations of GDF15 concentration with a range of phenotypes including all-cause mortality, cardiometabolic disease, respiratory diseases and psychiatric disorders as well as inflammatory markers. A meta-analysis of genome-wide association studies (GWAS) of GDF15 concentration across 3 different assay platforms (n=14,099) confirmed significant heterogeneity due to a common missense variant rs1058587 in GDF15, potentially due to epitope-binding artefacts. After conditioning on rs1058587, statistical fine-mapping identified 4 independent putative causal signals at the locus. Mendelian randomisation (MR) analysis did not find evidence of a causal relationship between GDF15 concentration and cardiometabolic traits. Using reverse MR, we identified a potential causal association of body mass index on GDF15 (IVW $p_{FDR}$=0.0072). Taken together, our data do not support a role for elevated GDF15 concentrations as a causal factor in human cardiometabolic disease but support its role as a biomarker of metabolic stress.
[Main Text:]

Introduction

Obesity accounts for an estimated 2.8 million deaths worldwide and 2.3% of global disability-adjusted life years (Organisation). Obesity has been causally linked to a variety of cardiometabolic risk factors and diseases, including fasting insulin, systolic blood pressure and type 2 diabetes (Holmes et al., 2014). Therefore, interventions to reduce obesity are recommended in the management of these diseases (Yumuk et al., 2016). Currently, the most effective obesity intervention is bariatric surgery (Jammah, 2015); however, its clinical utility is limited as it is highly invasive and accompanied by a high risk of long-term complications such as gastroesophageal reflux disease and nutritional deficiencies (Mesureur & Arvanitakis, 2017). New therapies such as GLP-1 receptor agonists aim to target obesity via appetite suppression but are associated with common side effects including nausea and diarrhoea (Wilding et al., 2021). The development of additional therapeutic strategies targeting obesity as an underlying cause and pathology of cardiometabolic disease represents an area of unmet clinical need.

GDF15, a distant member of the transforming growth factor-β family (Bootcov et al., 1997; Emmerson, Duffin, Chintharlapalli, & Wu, 2018), is upregulated in response to cellular stress and several diseases, including cancer. It has been suggested that GDF15 functions as a stress response agent implicated in organ injury (Kempf et al., 2006; V. W. W. Tsai, Husaini, Sainsbury, Brown, & Breit, 2018; Zimmers et al., 2005). GDF15 was originally identified to play a role in tumour-induced anorexia/cachexia in mice through appetite regulation (Johnen et al., 2007). The body weight reduction is considered to be mainly driven by food intake inhibition mediated by its receptor GDNF-family receptor alpha like (GFRAL) in distinct areas of the brainstem (Emmerson et al., 2017; Hsu et al., 2017; Mullican et al., 2017; Yang et al., 2017). GDF15 plasma levels are elevated in animal models of obesity (Lockhart, Saudek, & O’Rahilly, 2020; Xiong et al., 2017), and administration of recombinant protein robustly lowers body weight in obese and diabetic animals, including non-human primates (Mullican et al., 2017; Xiong et al., 2017). The beneficial effects on glycemic control is considered to be mediated mainly by the body weight loss. Genetic overexpression of Gdf15 results in decreased body weight and increased resistance to obesity.
associated with high-fat diets, whereas Gdf15 and Gfral deficient mice are more susceptible to diet-induced obesity (Mullican et al., 2017; O'Rahilly, 2017; Tran, Yang, Gardner, & Xiong, 2018). Recent preclinical work has identified GDF15’s role as a sentinel protein, upregulated in response to various ingested toxins triggering nausea-related behaviour in rodents assessed by pica and CTA (Conditioned Taste Aversion) (Borner, Wald, et al., 2020; Patel et al., 2019).

There is a lack of data supporting the treatment effect of GDF15 in humans. However, observational studies have revealed strong positive correlations of GDF15 plasma levels with body mass index (BMI), insulin resistance, age and mean arterial blood pressure (in obese individuals) (V. W. Tsai et al., 2015; Vila et al., 2011). Higher GDF15 levels have also been associated with all-cause mortality as well as mortality associated with heart failure and acute myocardial infarction, cancer, advanced heart failure and end-stage chronic kidney disease (Adela & Banerjee, 2015; Khan et al., 2009; Nair et al., 2017; Wiklund et al., 2010). Metformin therapy, a type 2 diabetes treatment, has been shown to depend on GDF15 to lower body weight in mice and plasma GDF15 levels increased up to 40% upon metformin therapy in humans (Coll et al., 2020; Day et al., 2019). Recent data suggest that GDF15 is involved in hyperemesis gravidarum (Fejzo, Arzy, Tian, MacGibbon, & Mullin, 2018), supporting the hypothesis that GDF15 triggers anorexia and subsequent weight loss, at least partly, through the induction of malaise (Borner, Shaulson, et al., 2020). Mendelian randomisation (MR) has previously been applied to explore the causal relationship of GDF15 levels and cardiometabolic diseases, with causal associations reported for high-density lipoprotein (HDL) cholesterol and bone mineral density (BMD) (Cheung, Tan, Au, Li, & Cheung, 2019; Folkersen et al., 2020). However, these analyses were based on small genetic studies for GDF15 and have not been replicated.

In this study, we used data from several large biobanks to conduct a systematic and extensive phenotypic and genotypic analysis of GDF15 with cardiometabolic traits and diseases, and to ascertain the causal relationship between GDF15 levels and cardiometabolic traits using MR and protein-truncating variant (PTV) analysis.
Results

Association of GDF15 plasma levels with 676 disease outcomes

To systematically assess the relationship between GDF15 plasma levels and clinical phenotypes, we utilised a large Finnish biobank, FINRISK. The FINRISK cohort comprises a cross-sectional population survey carried out over a 40-year period in Finland. GDF15 plasma levels were available for 6,610 participants from the 1997 recruitment cohort linked to 676 disease outcomes and were included in this analysis.

The median GDF15 concentration (measured using an immunoluminometric assay) was 796ng/L (interquartile range, IQR= 664-986) (Supplementary Figure S1). First, we examined the association of baseline characteristics with GDF15 plasma levels (Supplementary Table S1). Age explained 28% of the observed GDF15 variance, with current smoking and BMI accounting for 1.9% and 0.08% of the variance, respectively (Supplementary Table S2). Gender did not show a significant association. All subsequent analyses have been corrected for these covariates (age, gender, smoking and BMI). Analyses corrected for age and gender alone are also presented for comparison (Supplementary Table S3 and S4), although results were similar.

We then investigated potential associations of GDF15 plasma levels with a range of disease phenotypes (both prevalent and incident), focusing on the phenotypes defined by the FinnGen consortium (T. Lahtela E, Havulinna AS, & Consortium., 2020), as these endpoints have undergone additional clinical validation (see Methods). A total of 676 disease endpoints were examined as dependent variables in association analyses with GDF15 plasma levels (the independent variable). After multiple-testing correction using false discovery rate (FDR, P_{FDR} < 0.05), GDF15 was significantly associated with 80 disease endpoints (Table 1 and Supplementary Table S5).

The most significant clinical associations observed were with all-cause mortality (logistic regression odds ratio, OR=1.8, CI=1.7 – 1.9, p-value=9.8x10^{-27}) and mortality due to cardiac diseases (OR=1.8, CI=1.6 – 1.9, p-value=7.7x10^{-14}). We also observed significant associations between GDF15 levels and type 2 diabetes (OR=1.5, CI=1.3 – 1.6, p-value=2.2x10^{-9}). Further,
GDF15 plasma levels were significantly associated with cardiovascular diseases (OR=1.2, CI=1.1–1.2, p-value=5.6x10^{-6}), as well as subtype endpoints such as atherosclerosis, hypertension, peripheral artery disease and stroke, and chronic kidney disease (OR=2.46, CI=2.1–2.8, p-value=4.4x10^{-6}) (Table 1). Additional positive associations were observed with respiratory diseases, such as chronic obstructive pulmonary disease (COPD) and pneumonia, and psychiatric disorders, including schizophrenia and mental and behavioural disorders due to psychoactive substance use. Cancer phenotypes were also associated with GDF15 levels, including malignant neoplasm of the respiratory system and intrathoracic organs (OR=1.6, CI=1.3–1.8, p-value=0.0015) and malignant neoplasm of the bronchus and lung (OR=1.6, CI=1.3–1.9, p-value=0.0029).

**GDF15 plasma concentration associates with prevalent and incident type 2 diabetes and independently predicts all-cause mortality and cardiometabolic diseases**

To determine the prognostic potential of GDF15 in predicting incident diabetes, we investigated its plasma levels in relation to prevalent and incident type 2 diabetes in the FINRISK cohort. Type 2 diabetics (prevalent cases) had a 1.3-fold increase in median GDF15 levels compared to non-diabetic controls (median difference=222ng/L). Furthermore, GDF15 levels were significantly higher in prevalent diabetics (diagnosed at plasma sampling, n=37; median=1,204ng/L) compared to both incident cases (diagnosed after plasma sampling, n=500; median=992ng/L; Wilcoxon p-value=0.008) and non-diabetic controls (n=6003; 784ng/L; Wilcoxon p-value=5.2x10^{-11}) (Supplementary Figure S2). GDF15 plasma levels were also significantly higher in incident cases than in non-diabetic controls (Wilcoxon p-value=6.2x10^{-51}). Together, these data indicate that increased GDF15 plasma levels could represent an early biomarker in pre-diabetic individuals.

Next, we carried out a Cox regression survival analysis on all-cause mortality, type 2 diabetes and cardiovascular disease. During a 10-year follow-up period, 393 (6%) study subjects had died, 97 (2%) developed diabetes and 438 (7%) developed cardiovascular disease. Models accounted for blood pressure medication, smoking, total cholesterol, HDL cholesterol, BMI, prevalent cardiovascular disease and mean systolic blood pressure. Our results revealed that GDF15 plasma concentration is an independent predictor of all-cause mortality (hazard ratio, HR=1.7, p-value=9.2x10^{-12}), type 2 diabetes (HR=1.40, p-value=0.02) and cardiovascular disease (HR=1.4,
Individuals with GDF15 plasma levels above 967ng/L (the upper quartile) were more than twice as likely to develop type 2 diabetes compared to individuals with lower levels (HR=2.2, CI=1.4–3.5, \( p\)-value=7.1x10^{-4}).

**Association of GDF15 plasma levels with 96 quantitative biomarkers**

We then performed an association analysis of GDF15 plasma levels with 96 quantitative biomarkers in FINRISK. After multiple-testing correction (\( P_{FDR}<0.05 \)), significant associations were observed with 45 biomarkers, most notably, with blood markers known to be associated with inflammation (Supplementary Table S7).

The most significantly associated biomarkers were mid-regional pro-adrenomedullin (MR-proADM, \( \beta=0.24, p\)-value=1.2x10^{-91} ), C-reactive protein (CRP, \( \beta=0.22, p\)-value=4.6x10^{-64} ) and hepatocyte growth factor (HGF, \( \beta=0.17, p\)-value=4.2x10^{-43} ). All three markers have been previously reported to be non-specifically elevated in a number of human diseases and have known prognostic value (Jackson et al., 2016; Madonna, Cevik, Nasser, & De Caterina, 2012; Matsumoto, Umitsu, De Silva, Roy, & Bottaro, 2017; Miller, Redfield, & McConnell, 2007; Peacock, 2014). Significant associations were also observed between GDF15 and components of metabolic syndrome, specifically, levels of triglycerides (\( \beta=0.13, p\)-value=2.4x10^{-22} ), fasting insulin (\( \beta=0.10, p\)-value=1.2x10^{-13} ) and waist-to-hip ratio (\( \beta=0.049, p\)-value=7.0x10^{-10} ).

**Genome-wide association studies of plasma levels of GDF15 quantified using 3 different assays in 2 independent cohorts**

We performed genome-wide association studies (GWAS) to identify the genetic determinants of GDF15 plasma levels using two independent cohorts (FINRISK and INTERVAL) and three GDF15 quantification methodologies (immunoluminometric assay in FINRISK and Olink and SomaScan proteomic assays in INTERVAL; see Supplementary Materials). Results from individual GWAS analyses are reported in Supplementary Tables S8-10. We then performed a meta-analysis of the genome-wide significant variants identified in the GWAS performed in the FINRISK, INTERVAL-SomaScan and INTERVAL-Olink cohorts to generate a combined statistic and heterogeneity value. The analysis identified four genetic variants (rs1059369, rs1054221, rs1054222, and rs1054223).
rs1227734, rs189593084) at the GDF15 locus to be associated with GDF15 plasma levels (p-value<5x10^{-8}) (Supplementary Table S11).

Two variants had a strengthened signal in the combined meta-analysis, rs1054221 and rs1227734 (Supplementary Table S11). However, five genetic associations, which were genome-wide significant for two of the three GWAS, displayed substantial heterogeneity, e.g. rs16982345 (heterogeneity $I^2=99.8\%$, heterogeneity $p$-value=7.6x10^{-180}; Supplementary Table S11). By exploring the LD structure between the heterogeneous variants (Supplementary Table S12), we found that these variants were all located within the same LD block that included the missense variant rs1058587 (p.H202D). Heterogeneity in GWAS results of GDF15 levels has been observed in other studies (Jiang et al., 2018; Pietzner et al., 2020) and significant variations in GDF15 antibody binding affinity dependent on rs1058587 have been previously reported (Fairlie et al., 2001).

Mitigation of epitope binding effects due to the rs1058587 missense variant

To remove any potential epitope effects, we applied conditional analysis on rs1058587 when linking plasma GDF15 concentrations with disease phenotypes and quantitative traits. A total of 59 significant disease associations and 43 biomarker associations were identified. These associations largely did not differ from the unconditioned results, although 21 fewer disease associations reached significance (Supplementary Table S13-14).

To minimize confounding of the GWAS results from rs1058587 epitope effects, we conditioned the association signal on chromosome 19 on rs1058587 in both FINRISK (Supplementary Table S15) and INTERVAL-SomaScan (Supplementary Table S16). INTERVAL-Olink did not have a significant association at this LD block. We then meta-analysed these conditioned GWAS. Manhattan and Quantile-Quantile (QQ) plots of this GWAS meta-analysis are shown in Table 2, Supplementary Figure S4 and Supplementary Table S15. We found 146 significant associations in this region, with rs1227734 identified as the strongest association (beta=0.50, $p$-value=2.5x10^{-187}).

Fine-mapping this conditioned meta-analysis revealed four association signals with a causality probability of 0.86. The top configuration consisted of variants rs1054221 (beta=-0.50, $p$-
value=2.4x10^{-184}), rs3787023 (beta=0.04, p-value=0.0017), rs138515339 (beta=-0.27, p-
value=3.2x10^{-11}) and rs141542836 (beta=0.61, p-value=1.9x10^{-40}) and had a regional heritability
of 7%. Functional annotation and eQTL information can be found in Supplementary Table S1
and the Supplementary materials, respectively.

Mendelian randomisation analysis reveals no causal relationship between GDF15 plasma
levels and cardiometabolic traits

Observed associations of GDF15 plasma levels with obesity-related diseases have been presented
here and found in previous reports (V. W. Tsai et al., 2015; Vila et al., 2011) and variants in the
GDF15 gene region have been associated with cardiovascular traits, cholesterol, waist-hip-ratio
(WHR) and BMI (Wang et al., 2021). We therefore applied MR to assess the relationship between
genetically-determined GDF15 plasma levels and BMI, WHR, glucose and type 2 diabetes. We
also included HDL cholesterol and BMD as additional outcomes in our MR analysis due to the
positive MR results in previous studies (Cheung et al., 2019; Folkersen et al., 2020).

We applied two-sample MR (Bowden, Davey Smith, & Burgess, 2015) using LD clumping (R^2 <
0.1, within 1.5Mb of either side of the lead SNP) and identified eight independent genetic
instruments from the conditioned meta-analysis on chromosome 19: rs1227734, rs62122429,
rs1059022, rs181004295, rs75119307, rs11673678, rs138185133 and rs73923175 (regional
heritability=0.071). Association statistics between the conditioned instrumental variables (IVs)
and exposure (GDF15) were taken from the above meta-analysis, and the association statistics
between the IVs and outcome were extracted from publicly available GWAS summary statistics
(see Methods). These values were applied to a fixed effect inverse variance weighted (IVW) MR,
as well as MR-Egger and MR-PRESSO methods (Verbanck, Chen, Neale, & Do, 2018) , in order
to detect horizontal pleiotropy. MR using the conditioned variants did not find significant
associations of genetically-determined GDF15 plasma levels with any of the examined traits. A
nominal association was found with WHR but this did not pass multiple-testing correction (Table
3). Horizontal pleiotropy was detected in the association of genetically-determined GDF15 plasma
levels with BMI (Global pleiotropy test p-value=0.012), HDL cholesterol (Global pleiotropy test
p-value=0.021) and BMD (Global pleiotropy test p-value=0.010) using MR-PRESSO but not MR-
We did not replicate findings of causality of genetically-determined GDF15 plasma levels with BMD and HDL cholesterol, despite two independent studies suggesting otherwise (Cheung et al., 2019; Folkersen et al., 2020).

**Reverse Mendelian randomisation analysis identifies BMI as a causal factor for GDF15 plasma levels**

We applied reverse two-sample MR (as described above) using GDF15 plasma levels as the outcome variable to assess associations with BMI, WHR, glucose, diabetes, HDL cholesterol and BMD. GDF15 GWAS summary statistics were taken from the conditioned meta-analysis of GDF15 levels in FINRISK and INTERVAL. LD clumping was used to identify the genetic instruments (see Methods). We found a significant association between higher genetically-predicted BMI and higher GDF15 plasma levels (IVW estimate=0.089, pFDR=0.0072) but not any other tested trait (Table 4). Sensitivity analyses in MR-PRESSO but not MR-Egger analysis confirmed the association with BMI (p-value>0.05). MR-Egger identified horizontal pleiotropy with WHR, however, MR-PRESSO identified horizontal pleiotropy with every trait other than BMD. The finding of horizontal pleiotropy in MR-PRESSO but not MR-Egger suggests that the causal of an apparent causal effect of BMI on GDF15 plasma levels warrants replication in larger cohorts.

**Effect of GDF15 protein-truncating variants on cardiometabolic traits in 302,388 participants of UK Biobank**

Carriers of *GDF15* protein-truncating variants (PTVs) present an opportunity to explore the phenotypic consequences of predicted loss-of-function (LOF) of GDF15 on human disease. We analysed whole-exome sequencing data from 302,388 participants from the UK Biobank, of whom 109 carried *GDF15* PTVs in the heterozygous state (Supplementary Table S19). We assessed differences in BMI, WHR, glucose, BMD, HDL cholesterol and type 2 diabetes between carriers and non-carriers using the Mann-Whitney U test. The analysis was restricted to unrelated individuals of European ancestry, leaving 91 carriers of *GDF15* PTVs (male N=42, female N=49, mean age=56.3) and 40,000 randomly selected European-ancestry non-carriers (male N=20,000,
female N=20,000, mean age=56.9). In line with the MR analyses we observed no differences in BMI (mean difference=0.1, \( p \)-value=0.90), WHR (male: mean difference=0.01, \( p \)-value=0.64, female: mean difference=0.01, \( p \)-value=0.55), glucose (mean difference=0.14, \( p \)-value=0.32), BMD (mean difference=0.02, \( p \)-value=0.82), HDL cholesterol (mean difference=0.08, \( p \)-value=0.08) or self-reported diabetes (N of diabetes in cases=2, \( p \)-value=0.12), suggesting that mono-allelic \textit{GDF15} LOF does not have a strong effect on these traits.
Discussion

In this study, we present a systematic phenotypic and genotypic analysis of GDF15 with a wide range of health outcomes and biomarkers. In line with previous findings, our analysis confirmed strong correlations of GDF15 plasma levels with a range of clinical parameters (e.g. age, smoking and BMI) and human diseases (e.g. diabetes, cardiovascular and respiratory disease), as well as several inflammatory biomarkers. As preclinical data in mice strongly implicated Gdf15 in the aetiology of obesity and glucose tolerance, we specifically investigated whether human genetic evidence supports these findings. Using data from large biobanks, we found that neither MR nor GDF15 PTV analyses supported a causal role for GDF15 plasma levels within the normal range in influencing BMI, WHR, glucose, diabetes, HDL cholesterol or BMD in humans. Instead, we found that higher BMI may cause increases in GDF15 plasma levels highlighting the role of GDF15 as a likely marker of metabolic stress in humans.

Access to the FINRISK cohort provided a valuable opportunity to explore GDF15 plasma level associations with multiple phenotypes within a single large population. We found the strongest associations with all-cause mortality and cardiometabolic diseases (cardiovascular disease and diabetes), as well as cardiometabolic risk factors (e.g. hypertension, triglycerides, BMI), as previously reported (Bonaca et al., 2011; Brown et al., 2002; Daniels, Clopton, Laughlin, Maisel, & Barrett-Connor, 2011; Ho et al., 2012; Kempf et al., 2007; Khan et al., 2009; Lind et al., 2009; Rohatgi et al., 2012; Wiklund et al., 2010; Wollert et al., 2007). Our analyses also found that GDF15 plasma levels are a strong predictor of incident diabetes and an independent predictor of all-cause mortality, cardiovascular disease and diabetes morbidity, suggesting its potential use as a pre-diabetic prognostic biomarker. We found associations with neoplasms of the lung and digestive system but not with other types of cancer. Our data further support the association of GDF15 plasma levels with inflammatory phenotypes and biomarkers (CRP, mid-regional pro-adrenomedullin) (Brown et al., 2007) and uncovered less well-described associations with respiratory disease and psychiatric disorders. Despite previous observations suggesting a role for GDF15 in anorexia (Borner, Shaulson, et al., 2020), we did not identify any association of GDF15 plasma levels with this trait. However, we note that statistical power may have been limited due to the low number of cases (Supplementary Table S5). To identify phenotype associations in an
unbiased way, we performed analyses encompassing all phenotypes available applying covariates uniformly across all analyses. We note that covariates specific to a particular disease might have been omitted introducing bias or leading to inflation or deflation of statistics. We detected a strong association between GDF15 levels and smoking, which was recorded and adjusted in our analyses as a binary phenotype. Smoking is an important contributing factor in multiple diseases and it is therefore likely that adjustment as a continuous variable would have been able to better explain the contribution of smoking to the phenotypic association signals, especially in respiratory disease, lung cancer and some psychiatric disorders. These findings demonstrate that GDF15 plasma levels are a general marker for risk in multiple diseases and are associated with non-specific biomarkers such as CRP.

To define the genetic architecture of plasma GDF15 levels, we performed a GWAS meta-analysis in two large cohorts, FINRISK and INTERVAL, across three different assay platforms. A striking finding of this analysis was the substantial heterogeneity between variants across these studies, consistent with previous reports (Jiang et al., 2018). Exploring the LD between the variants displaying heterogeneity (Supplementary Figure S6) identified that a large proportion resided within an LD block encompassing a common missense variant, rs1058587 (p.H202D). This variant has been previously associated with hyperemesis gravidarum (Fejzo, Sazonova, et al., 2018) but the presence of this missense variant within this locus raises the possibility of epitope artefacts, indeed a previous study has identified epitope effects in this region (Fairlie et al., 2001). Therefore, it is highly likely that the inconsistency in GWAS results across this locus is driven by differences in the properties of the GDF15 assay used in each study (immunoluminometric assay in FINRISK, SomaScan and Olink assays in INTERVAL). Therefore, it may be beneficial to adopt a more consistent method for measuring GDF15 levels within the research environment. Approaches to assay protein levels that don’t involve binding to epitopes, such as mass spectrometry, will be informative for avoiding potential artefacts due to protein-altering variants in the future.

Our analyses did not reveal a causal role for GDF15 plasma levels with any of these cardiometabolic phenotypes. These results were further supported by the lack of association between these phenotypes and GDF15 PTV carriers in the UK Biobank dataset. This is consistent with a recent large-scale, pan-ancestry exome-sequencing study of over 640,000 individuals exploring the associations of rare coding variants with BMI that did not report an association with
GDF15 (Akbari et al., 2021). However, a meta-analysis completed by Cheung et al. (Ho et al., 2012) did identify significant associations of genetically-predicted GDF15 levels with HDL cholesterol (beta=-0.048, p-value=0.001) and estimated BMD (beta=0.026, p-value<0.001) which we did not replicate. The inclusion of differing covariates could have influenced the results of these analyses; our genetic analysis included only age, gender and principal components as covariates whereas the meta-analysis by Cheung et al. included age, gender, systolic blood pressure, antihypertensive medication use, diabetes mellitus and smoking status. Nevertheless, the larger sample size of the study presented here (n=14,099) offered considerably improved statistical power over this previous study (n=3,889).

This lack of causal association with GDF15 plasma levels suggests that GDF15 may behave differently in humans compared to rodents and non-human primates, where there are robust data demonstrating that GDF15 potently reduce body weight. GDF15 levels have also been demonstrated to provoke nausea-related pica and CTA responses in rat and mice, respectively (Borner, Shaulson, et al., 2020; Patel et al., 2019), as well as vomiting in shrews, *suncus murinus*, but studies in non-human primates report no signs of nausea (Xiong et al., 2017). Several clinical trials testing the safety and efficacy of GDF15 therapy have been planned or initiated but no data have yet been disclosed. Nevertheless, body weight loss in pregnant women suffering from hyperemesis gravidarium has been linked with higher plasma GDF15 levels (Fejzo, Arzy, et al., 2018). In FINRISK, information on nausea was not available but diseases with nausea as the main symptom were available (such as gastroesophageal reflux disease) and no significant association with GDF15 plasma levels was found (OR=1.1, p=0.63). In pregnancy GDF15 levels are found at much higher serum levels than normal (A. G. Moore et al., 2000) and it is therefore possible that at higher levels GDF15 may induce nausea. Further studies into the effect of higher-concentration GDF15 in humans and the outcomes of clinical trials will elucidate its role further.

As our analyses did not support a causal role for GDF15 plasma levels in the phenotypes examined, we sought to investigate if genetics supported the role for GDF15 as a consequence of these conditions. We conducted reverse MR using GDF15 plasma levels as the outcome and the previously assessed cardiometabolic traits and diseases as the exposures. We identified a significant causal association of BMI on GDF15 plasma levels, suggesting that higher GDF15
levels could be a consequence of higher BMI, supporting the recently reported role of GDF15 in response to stress (Patel et al., 2019). However, the contradictory findings of horizontal pleiotropy between MR-Egger and MR-PRESSO lead to residual uncertainty in the interpretation of the MR analysis. It is possible that the heterogeneity found in our study caused by the use of different assays is driving this pleiotropic finding. A recent report in a MR study of GDF15 levels performed by the SCALLOP consortium also found a significant causal effect of BMI (IVW effect=0.20, p-value=1x10^{-7}) (Folkersen et al., 2020). Our replication of this finding strengthens the evidence that GDF15 levels are likely a consequence of differences in BMI rather than a cause. Further research will be required to determine conclusively the role of horizontal pleiotropy in this relationship.

Using valid genetic instruments in MR is of key importance as an invalid instrument would violate the assumptions of the model and lead to unreliable results. To minimise bias from differential assay binding, we conditioned on rs1058587 in our analyses, which could be overcautious. The association of rs1058587 with adiposity traits and hyperemesis gravidarum raises the possibility that GDF15 exerts a causal effect that is not mediated by altered plasma levels, such as via altered proteoforms and/or altered GFRAL binding. With the use of a wide range of GDF15 assays and the possibility that this missense variant is impacting the assay functionality in different assays to varying degrees, the generation of functional data quantifying assay performance is essential for understanding the contribution of this association to GDF15 plasma levels and improving the power of MR analyses. Similarly, as our GDF15 PTV data is based solely on heterozygous carriers, it is possible that the other allele is compensating for the loss and further assessment of the impact of heterozygote GDF15 PTVs on the presence of the protein will be required to validate our findings.

Taken together, our study provides a systematic and broad investigation of GDF15 phenotypic and genotypic associations, identifying possible epitope artefacts that could affect the data validity of GDF15 assays and introduce systematic bias in genetic analyses. Taking into account these biases, our genetic analyses did not support a causal association between GDF15 plasma levels and obesity and diabetes in humans. Conversely, we found that increased GDF15 plasma levels may be a consequence of higher BMI. We suggest the strong and wide-ranging associations between GDF15 plasma levels and health outcomes observed predominantly represent secondary effects of
the underlying disease process, implicating GDF15 as a stress-induced biomarker. The notion that GDF15 may be merely a predictive marker rather than a causal factor for metabolic diseases draws parallels to CRP in cardiovascular disease. Here, epidemiological studies revealed that elevated plasma CRP levels are consistently associated with increased risks of atherosclerosis and ischemic vascular disease, but Mendelian randomization studies do not show a causal relationship (Davey Smith et al., 2005). Our results do not support GDF15 plasma levels as a causal factor at normal human plasma levels for obesity or its related cardiometabolic diseases.
Materials and Methods

Study population and phenotypes

FINRISK

This study was carried out in accordance with the recommendations of the Declaration of Helsinki. All participants of studies have given written informed consent. FINRISK study was approved by the Ethics Committee of Helsinki and Uusimaa Hospital District.

The FINRISK cohort comprises a cross-sectional population survey carried out over a 40-year period from the year 1972 across multiple regions in Finland. The study aimed to assess the risk factors of chronic diseases and health behaviours in a working age population (Borodulin et al., 2018; Vartiainen et al., 2000) and consisted of 6,000-8,800 individuals per survey. Measurement of GDF15 plasma levels using an immunoluminometric assay was included for participants from the 1997 recruitment cohort. Participants were additionally matched to their electronic health records giving access to longitudinal prescription records, death records and diagnosis history. The cohort characteristics are summarized in Supplementary Table S1. In total, 6,610 Finnish individuals from the FINRISK 1997 cohort with available GDF15 plasma concentrations and up to 676 disease outcomes were included in this analysis.

Nurse assessment/interview, self-report data and blood samples were all collected at various time points during the study (see Supplementary Materials). A group of clinicians working in the FinnGen consortium constructed 676 disease endpoints combining information from multiple health registries (T et al., 2020). Both ICD8, ICD9 and ICD10 codes were utilized in the disease endpoint definitions. Genotyping was carried out for in 6,538 individuals who were recruited in 1997, for more information see Supplementary Materials.

INTERVAL

The INTERVAL study is a prospective cohort study comprising approximately 50,000 participants nested within a pragmatic randomized trial of blood donors [33]. Between 2012 and 2014, blood
donors aged 18 years and older were recruited at 25 NHSBT (National Health Service Blood and Transplant) donor centres across England (see Supplementary Materials). Informed consent was obtained from all participants and the INTERVAL study was approved by the National Research Ethics Service (11/EE/0538).

For the SOMAscan assays, two non-overlapping subcohorts of 2,731 and 831 participants were randomly selected, of which 3,301 participants (2,481 and 820 in the two subcohorts) remained after genetic quality control.

For the Olink assay, protein measurements were conducted in an additional subcohort of 4,998 INTERVAL participants aged over 50 at baseline. 4,987 samples passed quality control for this panel and were included in the analyses.

For information on genotyping in the INTERVAL cohort see Supplementary Materials

UK Biobank

UK Biobank is a population-based cohort consisting of ~500,000 individuals with participants recruited between 2006-2010 and aged between 40-69. Recruitment and cohort information has been previously described (Sudlow et al., 2015). Electronic health records, self-report questionnaires, diet and lifestyle information and biomarker data is available on this cohort. Here, we utilised body mass index (BMI) and waist-hip-ratio (WHR), which were measured during a nurse assessment, as well as diabetes information. Touchscreen questionnaire diabetic patients were applied in this study rather than ICD10 diagnosed in order to increase patient numbers, given the small subgroup of individuals being examined. Details on the exome sequencing are available in the Supplementary Materials.

Laboratory methods for GDF15 measurement

FINRISK
Blood samples were collected after an advisory 4-hour fast, immediately centrifuged and then stored at -70°C until GDF15 measurement which was undertaken using an immunoluminometric assay (ILMA) with a limit of detection of 24 ng/L and a linear range from 200 to 50,000 ng/L (Kempf et al., 2007). The ILMA is technically identical to immunoradiometric assay (IRMA) (Horn, Geldszus, Potter, von zur Muhlen, & Brabant, 1996) except that the GDF15 detection antibody was labelled with acridinium ester and assay results were quantified in a luminometer (Berthold). ILMA and IRMA use an antibody that binds over a sequence of Ala197-Ile308. GDF15 concentrations measured with the ILMA and IRMA are very similar [(ILMA concentration/ILMA concentration) × 100% = 97.8 ± 1.3%] and closely correlated (r² = 0.99, p < 0.001. n = 31 samples) (H). Head-to-head comparison of IRMA with the clinical ‘gold standard’ Roche assay have been previously conducted and reported GDF15 levels were comparable (Wollert et al., 2017).

**INTERVAL**

Blood sample collection procedures have been described previously (C. Moore et al., 2014). Blood samples were collected in 6-ml EDTA tubes and transferred at ambient temperature to UK Biocentre (Stockport, UK) for processing. Plasma was extracted by centrifugation into two 0.8-ml plasma aliquots and stored at -80°C before use.

The procedures for obtaining protein measurements using the SOMAscan assay have also been described previously (Sun et al., 2018). Briefly, SOMAscan is a multiplexed, aptamer-based approach that was used to measure the relative concentrations of 3,622 plasma proteins or protein complexes using modified aptamers (‘SOMAmer reagents’, hereafter referred to as SOMAmers). Aliquots of 150 µl of plasma from INTERVAL baseline samples were sent on dry ice to SomaLogic Inc. (Boulder, Colorado, US) for protein measurement. Modified single-stranded DNA SOMAmers were used to bind to specific protein targets and these are then quantified using a DNA microarray. Protein concentrations are quantified as relative fluorescent units. Quality control (QC) was completed at the sample and SOMAmer levels.
The Olink assay uses pairs of monoclonal antibodies (or a single polyclonal antibody) to bind each protein target, and then uses proximity extension followed by qPCR to quantify protein abundance. Aliquots of plasma from samples taken from INTERVAL participants at the 2-year follow-up survey were shipped on dry ice to Olink Proteomics (Uppsala, Sweden) for assay. Protein concentrations were recorded as normalised relative protein abundances (‘NPX’).

**Genome-wide association analysis**

**FINRISK**

For the FINRISK cohort genome-wide association analyses were performed for 5817 individuals with plasma GDF15 concentrations available. A normal distribution of GDF15 plasma concentrations was achieved through applying an inverse normal transformation. Multidimensional scaling was done for genetic data of both studies using PLINK version 1.07 (Purcell et al., 2007). Only good quality autosomal markers passing the following criteria: imputation informativeness >0.7, and MAF>0.001, were included in further evaluations. Frequentist test with method ‘expected’, assuming an additive genetic model, was performed using SNPTEST v2 (Marchini, Howie, Myers, McVean, & Donnelly, 2007). Results were adjusted for the first five principal components of the genetic data to account for the population stratification and for gender, age and genotyping set. The Quantile-Quantile and Manhattan plots and regional plots were created using R-2.11 to visualize genome-wide association results. The genomic positions indicated throughout this study are based on NCBI human genome build 37.

**INTERVAL**

The GWAS for INTERVAL-SOMAscan has previously been described in detail (Sun et al., 2018). Relative protein abundances were first natural log-transformed within each subcohort. Log-transformed protein levels were then adjusted in a linear regression for age, sex, duration between blood draw and processing (binary, ≤1 day/>1day) and the first three principal components of ancestry from multi-dimensional scaling. A normal distribution of the protein residuals from this linear regression was achieved through rank-inverse normal transformation and these were used as phenotypes for association testing. Simple linear regression assuming an additive genetic model
was used to test for genetic associations. Genotype uncertainty was accounted for by carrying out association tests on allelic dosages (‘method expected’ option) using SNPTEST v2.5.2 [51].

For the INTERVAL-Olink GWAS, normalized protein levels (‘NPX’) were first regressed on age, sex, plate, time from blood draw to processing (in days), and season (categorical: ‘Spring’, ‘Summer’, ‘Autumn’, ‘Winter’). The residuals from this linear regression were rank-inverse normalized. The rank-inverse normalized residuals were then fitted in a linear regression model adjusted for ancestry by including the first three components of multi-dimensional scaling as covariates. GWAS was conducted using SNPTEST v2.5.2.

**Fine mapping**

Fine mapping was performed using FINEMAP program (Benner et al., 2016). FINEMAP can potentially identify sets of variants with more evidence of being causal than those highlighted by a stepwise conditional analysis. The FINEMAP provides (1) a list of potential causal configurations together with their posterior probabilities and Bayes factors and, (2) for each variant, the posterior probability and Bayes factor of being causal. FINEMAP was applied with its default settings. LD between SNPs for the conditional analysis was combined using the formula: (N1 * LD1 + N2 * LD2 + N3 * LD3) / (N1 + N2 + N3), where N represents the number of individuals in the GWAS in each cohort, LD represents the $R^2$ value and 1, 2 and 3 represent the three different cohorts: FINRISK, INTERVAL-SomaScan and INTERVAL-Olink.

**Meta-analysis**

Meta-analysis was completed on the chromosome 19 locus around the GDF15 gene in METAL (Willer, Li, & Abecasis, 2010) using the Inverse-variance Weighted method. A total of 3,799 SNPs were found in common across FINRISK, INTERVAL-SomaScan and INTERVAL-Olink.

**Statistical analysis**

Association study between GDF15 levels and disease endpoints and quantitative biomarkers
To examine the association of GDF15 with disease endpoints and quantitative biomarkers, logistic and linear regression models were used, respectively. Inverse-variance transformed GDF15 levels were used in the analyses due to right-skewed distribution. Association of GDF15 with disease endpoints and quantitative biomarkers were examined in 1) age- and sex 2) age, sex and BMI 3) age, sex and smoking -adjusted models. Analyses using linear- and logistic regression models were performed using R2.11 (see URLs). False discovery rate (FDR) was applied to test for multiple test correction.

Survival analysis

All analyses were performed in R version 3.4.0. Cox proportional hazard regression model was conducted to identify predictors of outcomes during 10 years. The survival curves for a Cox proportional hazards model were used to illustrate the timing of the death, type 2 diabetes and CVD events during 10 years in relation to GDF15 quartiles and statistical assessment between upper quartile and other quartiles was performed using cox proportional hazard regression model. To inspect the validity of the cox-model, the test of the proportional hazards assumption for a Cox regression model was used. Hosmer-Lemeshow Goodness of Fit (GOF) test was used to determine whether models were calibrated with similar expected and observed event rates in both low- and high-risk individuals.

Wilcoxon test for prevalent and incident diabetes

Wilcoxon rank sum test (Mann–Whitney U test) was performed, in R using function “wilcox.test” from stats package, to test whether the distributions of incident type 2 diabetes case group, prevalent type 2 diabetes case group and control group were systematically different from one another.

Mendelian Randomisation

We applied MR (Bowden et al., 2015) to explore causality of GDF15 with BMI, WHR, glucose, type 2 diabetes, HDL cholesterol and BMD. We additionally explored these relationship in reverse;
with GDF15 as an outcome and other traits as the exposure. GWAS summary statistics for GDF15 were taken from the conditional meta-analysis of FINRISK and INTERVAL for the forward MR and from FINRISK only for the reverse MR. GWAS in INTERVAL being completed on chromosome 19 only whereas GWAS in FINRISK were completed on all chromosomes. Summary statistics for the other traits were obtained from large public data resources: BMI (n=590,827) (Yengo L et al., 2018), WHR (n=485,486) (Shungin et al., 2015), glucose (n=314,916) (Neale, 2018), type 2 diabetes (n cases=74,124, n controls=824,006) (Mahajan et al., 2018), HDL cholesterol (n=315,133) (Neale, 2018) and BMD (n=426,824) (Morris et al., 2019). For BMI, MR with GDF15 as the exposure was applied using summary statistics from UK Biobank only (n=361,194) as the LD clumped variants were not available in the larger meta-analysis, that was utilised for reverse MR. We utilised LD clumping to identify independent genetic variants in summary statistics and applied these as genetic instruments in MR analyses. MR analysis was run in R (version 4.0.2) using the package ‘MendelianRandomization’. The inverse variance weighted method (IVW) from this package was applied to test for causality and a sensitivity analysis using MR-Egger to testing for horizontal pleiotropy. A finding of significant causality is indicated with the IVW p-value and a significant MR-Egger intercept p-value indicates a finding of horizontal pleiotropy. MR-PRESSO (Mendelian randomization pleiotropy residual sum and outlier) (Verbanck et al., 2018) was additionally applied as a further sensitivity analysis to test for horizontal pleiotropy and was run in R using the ‘MRPRESSO’ package. The Global p-value indicates a significant finding of horizontal pleiotropy which instigates removal of outliers from the analysis. MR-PRESSO then re-runs IVW testing and gives a significant distortion p-value to indicate if the removal of outliers reduced the horizontal pleiotropy found in the analysis.

GDF15 PTV analysis

Exome data

Individuals with GDF15 PTVs were identified by extracting variants annotated as protein truncating variants (including exon loss variants, frameshift variants, start lost, stop gained, stop lost, splice acceptor variants, splice donor variant, rare amino acid variant, transcript ablation, gene fusion and bidirectional gene fusion). Supplementary Table S19 demonstrates the frequency of
LOF variants in each ancestry. From the individuals that did not carry GDF15 PTV variants 20,000 males and 20,000 females were randomly extracted as controls.

Mann Whitney U test

Analysis was completed in R (version 3.2.4) using “wilcox.test” and models compared BMI, WHR (in males and females separately), diabetic status (self-reported), BMD and HDL in GDF15 PTV carriers and non-carriers.

Acknowledgements: We thank the participants and investigators in the UK Biobank study who made this work possible (Resource Application Number 26041). We thank the UK Biobank Exome Sequencing Consortium (UKB-ESC) members AbbVie, Alnylam Pharmaceuticals, AstraZeneca, Biogen, Bristol-Myers Squibb, Pfizer, Regeneron and Takeda for funding the generation of the data and Regeneron Genetics Center for completing the sequencing and initial quality control of the exome sequencing data. We thank the AstraZeneca Centre for Genomics Research Analytics and Informatics team for processing and analysis of sequencing data.

Participants in the INTERVAL randomised controlled trial were recruited with the active collaboration of NHS Blood and Transplant England (www.nhsbt.nhs.uk), which has supported field work and other elements of the trial. DNA extraction and genotyping was co-funded by the National Institute for Health Research (NIHR), the NIHR BioResource (http://bioresource.nihr.ac.uk) and the NIHR Cambridge Biomedical Research Centre (BRC-1215-20014) [*]. SomaLogic assays were funded by Merck and the NIHR Cambridge BRC (BRC-1215-20014) [*]. The academic coordinating centre for INTERVAL was supported by core funding from: NIHR Blood and Transplant Research Unit in Donor Health and Genomics (NIHR BTRU-2014-10024), UK Medical Research Council (MR/L003120/1), British Heart Foundation (SP/09/002; RG/13/13/30194; RG/18/13/33946) and the NIHR Cambridge BRC (BRC-1215-20014) [*]. A complete list of the investigators and contributors to the INTERVAL trial is provided in reference [**]. The academic coordinating centre would like to thank blood donor centre staff and blood donors for participating in the INTERVAL trial.
This work was supported by Health Data Research UK, which is funded by the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), British Heart Foundation and Wellcome.

This work was conducted as part of an alliance between the University of Cambridge and the AstraZeneca Centre for Genomics Research (AZ Ref: 10033507).

The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from the GTEx Portal on 08/12/2021.

*The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.


**Funding:** Rachel Ong is co-funded by the NIHR Cambridge Biomedical Research Centre (BRC-1215-20014) [*]. VS has been supported by the Finnish Foundation for Cardiovascular Research.

**Author contributions:**

**Competing interests:** EMW, RM, DSP and AM are employees of AstraZeneca. RMYO is currently an employee of GlaxoSmithKline (although was not when this work was carried out). AB reports grants outside of this work from Biogen, BioMarin, Bioverativ, Merck, Novartis, Pfizer and Sanofi and personal fees from Novartis. VS has received honoraria from Novo Nordisk and Sanofi for consulting. He also has ongoing research collaboration with Bayer Ltd (All outside this work).
REFERENCES


H, T. Dissertation.


Organisation., W. H. Global Health Observatory (GBO) data: Obesity.


### Tables

**Table 1. Disease endpoints associated ($p_{FDR}$<1x10⁻⁵) with GDF15 plasma levels in FINRISK.**

<table>
<thead>
<tr>
<th>Disease Endpoint</th>
<th>Cases / Controls</th>
<th>OR (95% CI)</th>
<th>$p_{FDR}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>All-cause mortality</td>
<td>1057/5481</td>
<td>1.79 (1.68-1.90)</td>
<td>7.5x10⁻²⁴</td>
</tr>
<tr>
<td>Death due to cardiac causes</td>
<td>471/6067</td>
<td>1.76 (1.61-1.90)</td>
<td>3.0x10⁻¹¹</td>
</tr>
<tr>
<td>Atherosclerosis, excluding cerebral, coronary and PAD</td>
<td>379/6159</td>
<td>1.67 (1.51-1.82)</td>
<td>3.2x10⁻⁸</td>
</tr>
<tr>
<td>Diabetes mellitus type 2</td>
<td>567/5971</td>
<td>1.48 (1.36-1.60)</td>
<td>3.2x10⁻⁸</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>592/5946</td>
<td>1.44 (1.33-1.56)</td>
<td>9.8x10⁻⁸</td>
</tr>
<tr>
<td>Diseases of arteries, arterioles and capillaries</td>
<td>505/6033</td>
<td>1.48 (1.35-1.61)</td>
<td>3.6x10⁻⁸</td>
</tr>
<tr>
<td>Other chronic obstructive pulmonary disease</td>
<td>221/6317</td>
<td>1.84 (1.63-2.04)</td>
<td>5.7x10⁻⁷</td>
</tr>
<tr>
<td>Chronic obstructive pulmonary disease</td>
<td>235/6303</td>
<td>1.77 (1.57-1.97)</td>
<td>9.6x10⁻⁷</td>
</tr>
<tr>
<td>Pneumonia (excl. viral and due to other infectious organisms not elsewhere classified)</td>
<td>779/5759</td>
<td>1.34 (1.24-1.44)</td>
<td>9.6x10⁻⁷</td>
</tr>
<tr>
<td>Type 2 diabetes without complications</td>
<td>485/6053</td>
<td>1.44 (1.31-1.57)</td>
<td>1.2x10⁻⁶</td>
</tr>
<tr>
<td>All pneumonia</td>
<td>789/5749</td>
<td>1.33 (1.23-1.43)</td>
<td>1.9x10⁻⁶</td>
</tr>
<tr>
<td>Chronic kidney disease</td>
<td>66/6472</td>
<td>2.46 (2.14-2.79)</td>
<td>4.4x10⁻⁶</td>
</tr>
<tr>
<td>Influenza and pneumonia</td>
<td>833/5795</td>
<td>1.30 (1.21-1.40)</td>
<td>4.4x10⁻⁶</td>
</tr>
<tr>
<td>Type 2 diabetes with renal complications</td>
<td>35/6503</td>
<td>2.97 (2.57-3.38)</td>
<td>8.1x10⁻⁶</td>
</tr>
<tr>
<td>Sequelae of cerebrovascular disease</td>
<td>209/6329</td>
<td>1.67 (1.47-1.86)</td>
<td>9.2x10⁻⁶</td>
</tr>
<tr>
<td>Alcoholic liver disease</td>
<td>56/6482</td>
<td>2.26 (1.95-2.57)</td>
<td>9.2x10⁻⁶</td>
</tr>
</tbody>
</table>

Results are adjusted for age, gender, smoking and BMI. Abbreviations: OR, odds ratio; CI, confidence interval; COPD, chronic obstructive pulmonary disease; PAD, peripheral artery disease; SAH, subarachnoid haemorrhage. ICD codes for these disease endpoints have been published (T et al., 2020).
Table 2. Meta-analysis of GDF15 GWAS conditioned on rs1058587 in FINRISK and INTERVAL.

<table>
<thead>
<tr>
<th>SNPs</th>
<th>LD block</th>
<th>FINRISK</th>
<th>INTERVAL-SomaScan</th>
<th>INTERVAL-Olink</th>
<th>Meta-analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>beta</td>
<td>p-value</td>
<td>beta</td>
<td>p-value</td>
</tr>
<tr>
<td>rs16982345</td>
<td>1</td>
<td>A/G</td>
<td>-0.28 0.55</td>
<td>0.30 0.18</td>
<td>0.01 0.72</td>
</tr>
<tr>
<td>rs1058587</td>
<td>1</td>
<td>G/C</td>
<td>-0.011 0.72</td>
<td>-0.017 0.56</td>
<td>0.04 0.072</td>
</tr>
<tr>
<td>rs3787023</td>
<td>1</td>
<td>A/G</td>
<td>0.06 0.0011</td>
<td>-0.018 0.54</td>
<td>0.04 0.078</td>
</tr>
<tr>
<td>rs1055150</td>
<td>1</td>
<td>G/C</td>
<td>0.06 0.0012</td>
<td>-0.017 0.56</td>
<td>0.04 0.080</td>
</tr>
<tr>
<td>rs1059369</td>
<td>1</td>
<td>A/T</td>
<td>0.06 0.0010</td>
<td>-0.018 0.54</td>
<td>0.04 0.078</td>
</tr>
<tr>
<td>rs1054221</td>
<td>2</td>
<td>C/T</td>
<td>0.38 3.4x10^-37</td>
<td>0.62 7.8x10^-83</td>
<td>0.52 9.4x10^-74</td>
</tr>
<tr>
<td>rs1227734</td>
<td>2</td>
<td>T/C</td>
<td>0.38 3.9x10^-37</td>
<td>0.62 1.8x10^-83</td>
<td>0.51 7.1x10^-74</td>
</tr>
<tr>
<td>rs189593084</td>
<td>3</td>
<td>A/C</td>
<td>-0.33 8.4x10^-14</td>
<td>-0.61 0.011</td>
<td>-0.47 0.0053</td>
</tr>
</tbody>
</table>

For comparison identical variants to those listed in Supplementary Table S10 are shown here (these variants were identified by fine mapping unconditioned GWAS results from FINRISK and INTERVAL). LD blocks were defined as SNPs that had LD > 0.1 with the lead variant (most significantly associated variant).
**Table 3. Mendelian randomisation results for exposure genetically-determined GDF15 plasma levels with cardiometabolic outcomes.**

<table>
<thead>
<tr>
<th>Methods</th>
<th>IVW (fixed)</th>
<th>MR-Egger</th>
<th>MR-PRESSO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate (SE)</td>
<td>$p_{VDR}$ (raw $p$-value)</td>
<td>Estimate $p$-value</td>
</tr>
<tr>
<td>BMI</td>
<td>-0.006 (0.054)</td>
<td>1.0 (0.92)</td>
<td>0.043 (0.10)</td>
</tr>
<tr>
<td>WHR</td>
<td>0.018 (0.007)</td>
<td>0.078 (0.013)</td>
<td>0.023 (0.014)</td>
</tr>
<tr>
<td>Glucose</td>
<td>0.019 (0.018)</td>
<td>1.0 (0.30)</td>
<td>0.010 (0.034)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>0.004 (0.007)</td>
<td>1.0 (0.52)</td>
<td>-0.003 (0.012)</td>
</tr>
<tr>
<td>HDL</td>
<td>-0.005 (0.004)</td>
<td>1.0 (0.23)</td>
<td>-0.010 (0.007)</td>
</tr>
<tr>
<td>BMD</td>
<td>0.008 (0.009)</td>
<td>1.0 (0.41)</td>
<td>-0.005 (0.017)</td>
</tr>
</tbody>
</table>

A significant finding of pleiotropy is indicated by the intercept $p$-value in MR-Egger and the Global $p$-value in MR-PRESSO. For MR-PRESSO the outlier test is only run if horizontal pleiotropy is observed. The distortion $p$-value represents whether the outlier removal significantly reduces the horizontal pleiotropy.
Table 4. Reverse Mendelian randomization with GDF15 plasma levels as outcome.

<table>
<thead>
<tr>
<th>Methods</th>
<th>SNPs</th>
<th>IVW (fixed) Estimate (SE)</th>
<th>p&lt;sub&gt;DR&lt;/sub&gt; (raw p-value)</th>
<th>MR-Egger Estimate</th>
<th>p-value</th>
<th>Intercept</th>
<th>p-value</th>
<th>MR-PRESSO Raw estimate</th>
<th>Raw p-value</th>
<th>Outlier Estimate</th>
<th>Outlier p-value</th>
<th>Global p-value</th>
<th>Distortion p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI</td>
<td>1140</td>
<td>0.089 (0.027)</td>
<td>0.0072 (0.0012)</td>
<td>0.11</td>
<td>0.19</td>
<td>0.81</td>
<td>(0.081)</td>
<td>0.089</td>
<td>0.0012</td>
<td>0.001</td>
<td>0.027</td>
<td>&lt;0.00013</td>
<td>0.92</td>
</tr>
<tr>
<td>WHR</td>
<td>577</td>
<td>0.021 (0.038)</td>
<td>0.83 (0.58)</td>
<td>-0.23</td>
<td>0.054</td>
<td>0.027</td>
<td>(0.12)</td>
<td>-0.021</td>
<td>0.58</td>
<td>-</td>
<td>-</td>
<td>0.0049</td>
<td>-</td>
</tr>
<tr>
<td>Glucose</td>
<td>67</td>
<td>-0.024 (0.060)</td>
<td>0.83 (0.69)</td>
<td>-0.11</td>
<td>0.44</td>
<td>0.51</td>
<td>(0.14)</td>
<td>0.024</td>
<td>0.69</td>
<td>-</td>
<td>-</td>
<td>0.030</td>
<td>-</td>
</tr>
<tr>
<td>Diabetes</td>
<td>259</td>
<td>-0.001 (0.014)</td>
<td>0.97 (0.97)</td>
<td>-0.015</td>
<td>0.66</td>
<td>0.64</td>
<td>(0.034)</td>
<td>0.001</td>
<td>0.71</td>
<td>-</td>
<td>-</td>
<td>0.033</td>
<td>-</td>
</tr>
<tr>
<td>HDL</td>
<td>431</td>
<td>-0.056 (0.059)</td>
<td>0.68 (0.34)</td>
<td>-0.006</td>
<td>0.95</td>
<td>0.51</td>
<td>(0.097)</td>
<td>-0.056</td>
<td>0.34</td>
<td>0.065</td>
<td>0.058</td>
<td>0.0033</td>
<td>?</td>
</tr>
<tr>
<td>BMD</td>
<td>1131</td>
<td>0.020 (0.017)</td>
<td>0.68 (0.24)</td>
<td>0.046</td>
<td>0.19</td>
<td>0.41</td>
<td>(0.035)</td>
<td>0.002</td>
<td>0.24</td>
<td>-</td>
<td>-</td>
<td>0.089</td>
<td>-</td>
</tr>
</tbody>
</table>

A significant finding of pleiotropy is indicated by the intercept p-value in MR-Egger and the Global p-value in MR-PRESSO. For MR-PRESSO the outlier test is only run if horizontal pleiotropy is observed. The distortion p-value represents whether the outlier removal significantly reduces the horizontal pleiotropy.
Figures

Figure 1. Forest plots of Cox proportional hazard for independent predictors of (a) all-cause mortality, (b) cardiovascular disease and (c) diabetes.
GDF15 is highlighted in red and variables are ordered by highest hazards ratio. Abbreviations: BMI, body mass index; GDF15, growth differentiation factor-15; CHD, coronary heart disease; STR, stroke; HDL, high density lipoprotein.