communications medicine

A Nature Portfolio journal

Article

https://doi.org/10.1038/s43856-025-01347-y

Genetics identifies obesity as a shared risk factor for co-occurring multiple long-term conditions

Ninon Mounier[1], Bethany Voller[1], Jane A. H. Masoli[1,2], João Delgado[1], Frank Dudbridge[3], Luke C. Pilling[1], Timothy M. Frayling[1,4] & Jack Bowden[1,5] ✉ On behalf of the GEMINI Consortium*

Abstract

Background Multimorbidity, the co-occurrence of multiple long-term conditions (LTCs), is an increasingly important clinical problem, but little is known about the underlying causes. We investigate the role of a critical multimorbidity risk factor, obesity, as measured by body mass index (BMI), in explaining shared genetics amongst 71 common LTCs.

Methods In a population of northern Europeans, we estimated genetic correlation, between LTCs and partial genetic correlations after adjustment for the genetics of BMI. We used multiple causal inference methods to confirm that BMI causally affects individual LTCs, and their co-occurrence. Finally, we quantified the population-level impact of intervening and lowering BMI on the prevalence of 15 key common multimorbid LTC pairs.

Results BMI partially explains some of the shared genetics for 740 LTC pairs (30% of all pairs considered). For a further 161 LTC pairs, the genetic similarity between the LTCs was entirely accounted for by BMI genetics. This list included diabetes and osteoarthritis and gout and osteoarthritis: Causal inference methods confirmed that higher BMI acts as a common risk factor for a subset of these pairs, and therefore BMI-lowering interventions would likely reduce their prevalence. For example, we estimated that a 1 standard deviation or 4.5 unit decrease in BMI would result in 17 fewer people with both chronic kidney disease and osteoarthritis per 1000 who currently have both LTCs.

Conclusions Our genetics-centred approach quantifies the contribution of obesity to multimorbidity. Our method for calculating full and partial genetic correlations is published as an R package *{partialLDSC}*.

Plain language summary

More than half of people over 65 have several long-term health conditions at the same time. This is becoming a bigger issue in the UK, but we don't fully understand why some people develop many conditions.We looked at how body weight, measured by body mass index (BMI), affects the shared genetic risks for 71 common health problems such as diabetes, heart disease, arthritis and depression. Using data from people with northern European ancestry, we studied how much the same genes are linked to different conditions — both before and after taking the genetics of BMI into account.We found that BMI explains some of the shared genetic risks between many health conditions, and all of the shared risk for some, such as diabetes and osteoarthritis.Our results suggest that helping people lower their BMI could reduce the number of long-term health problems they experience, allowing more people to live longer and healthier lives.

Multimorbidity, defined as the coexistence of two or more long-term conditions (LTCs), is an important public health challenge. The prevalence of multimorbidity differs across geographic regions, age groups and between genders. It is also higher in more deprived individuals and is often associated with lower quality of life and increased healthcare costs[1]. Many observational studies have focused on defining and measuring multimorbidity[2,3]. However, disparate definitions have led to variations in its characterisation[4]. For example, a thorough definition of chronicity and a list of chronic LTCs have been proposed by ref. 5 and approaches to cluster LTCs and identify patterns of multimorbidity have recently been developed[6,7]. These approaches have limitations, as they often involve single datasets and modalities, and cluster identification might differ depending on the algorithm[8].

Investigating the role of common epidemiological risk factors for multimorbidity is vitally important, as it can provide a better understanding of the mechanisms underlying the co-occurrence of LTCs and help to develop efficient prevention strategies. For example, using observational data to study the relationship between socio-economic status and multimorbidity, it has been shown that lower education level and higher deprivation were associated with increasing risk of multimorbidity[9]. A large

[1]Department of Clinical and Biomedical Sciences, Faculty of Health and Life Sciences, University of Exeter, Exeter, UK. [2]Royal Devon University Healthcare NHS Foundation Trust, Barrack Road, Exeter, UK. [3]Department of Population Health Sciences, University of Leicester, Leicester, UK. [4]Department of Genetic Medicine and Development, Faculty of Medicine, University of Geneva, Geneva, Switzerland. [5]Novo Nordisk Research Centre Oxford, Roosevelt Drive, Headington, Oxford, UK. *A list of authors and their affiliations appears at the end of the paper. ✉e-mail: J.Bowden2@exeter.ac.uk

multi-cohort prospective study provided evidence of associations between obesity and a wide range of LTCs, as well as with the number of LTCs developed, highlighting the potentially important role of obesity in multimorbidity[10]. However, it is well known that confounding between a risk factor and an LTC could explain all the apparent association, and the process of designating a risk factor as an intervenable cause of an LTC requires further lines of evidence.

We have previously assessed the genetic similarity between 2546 pairs of 72 LTCs[11], allowing us to re-examine associations from observational epidemiology. Using genetic predictors of a trait, rather than observational measures, reduces the impact of inherent issues such as confounding, measurement error and reverse causation[12]. These issues are particularly problematic for studies of multimorbidity. Over the last decade, genetic approaches have also greatly benefited from the increase in sample sizes, and methods that allow data from more than one cohort to be incorporated into the analysis, enabling researchers to work with data linking tens of thousands of cases for common LTCs from a large range of health exposures. This genetics-centred approach has been used to estimate the costs of obesity on healthcare systems[13]. Our previous work showed evidence of widespread genetic correlation across LTCs, with obesity being highly genetically correlated with a broad range of LTCs[11]. Depending on context, obesity can be considered a risk factor for multiple LTCs or an LTC in its own right. In this work, we focus on using statistical genetics methods to quantify the role of obesity as a common risk factor for multimorbidity.

Genetic correlation is said to be driven by pleiotropy when a genetic locus affects several traits, and can reflect different scenarios: a direct relationship between the two traits (vertical pleiotropy – when the two traits are part of a causal cascade), a common biological process or the effect of a common risk factor on both traits (horizontal pleiotropy – when the two traits have no direct effect on each other)[14]. Often, genetic correlations are driven by a combination of both vertical and horizontal pleiotropy[15].

In this present work, we propose an approach to help differentiate these mechanisms in multimorbidity, focusing on obesity as a well-known risk factor for several long-term conditions[16]. We chose the most common, general clinical measure of obesity for our primary investigation: body mass index (BMI). To provide additional insight on the role of central obesity, the Waist Hip Ratio (WHR) was also used as a secondary measure. Using data from the GEMINI collaborative for 71 long-term common LTCs comprising 2485 distinct pairs, we propose an approach to compare pairwise unadjusted genetic correlations to partial genetic correlations that accounts for BMI genetics, formally testing whether BMI explains a significant portion of the genetic correlation for a given LTC pair. Because our analysis leverages standard genetic summary statistics from 13 disease-specific GWAS consortia that are not adjusted for BMI, our approach is simple and powerful. We also view it as less susceptible to collider bias that could arise from such explicit adjustment[17]. Using a smaller bespoke set of genetic variants, we also applied causal inference methods to further elucidate the causal (biological) mechanisms through which BMI affects the genetic correlation between LTCs, quantifying its pivotal role as a common risk factor in multimorbidity.

Methods

Data resources

We used GWAS summary statistics, derived from individuals of European descent, for 71 common and heritable LTCs, encompassing 13 distinct disease domains, grouped according to the International Classification of Diseases (ICD): such as cardiovascular or respiratory domains[18]. These GWAS data are described in detail in ref. 19; relevant diagnostic and analytical code is available on the project GitHub pages [https://github.com/GEMINI-multimorbidity]. LTCs were defined by adapting existing diagnostic code lists with input from clinical experts. LTCs were selected for genetic analyses if reaching a prevalence greater than 0.5% in people over 65 in each of two large population-based cohorts in the UK and Spain[20,21]. *Heritability* measures the proportion of phenotypic variance explained by genetics. We estimated this in the UK Biobank[22] to identify a subset of LTCs with a genetic basis. Finally, for each condition, we used the largest sample size available by combining GWAS data from up to three sources: UK Biobank, FinnGen, and condition-specific consortium data. To identify these data, we used the GWAS Catalog (https://www.ebi.ac.uk/gwas) and 13 disease-specific public repositories linked to GWAS studies. We contacted authors in each case to identify to further ensure that disease definitions aligned and that the participants were of European ancestry to enable synthesis and comparison with UKB and FinnGen. (details available in Supplementary Data 1). For many of these LTCs, the meta-analysed data used represents the largest sample size used to date. We refer to these meta-analysed GWAS summary statistics as "GEMINI summary statistics" for the rest of the paper. GWAS summary statistics can be downloaded from Zenodo (https://doi.org/10.5281/zenodo.14284046).

We also used two different sets of GWAS summary statistics for BMI. First, we used the largest data to date ($N \approx 700,000$), combining results from the GIANT Consortium and UK Biobank[23], to estimate full and partial genetic correlations (the definition of which will be clarified below). We also used earlier results from the GIANT Consortium ($N \approx 340,000$ – minimal overlap with the GEMINI summary statistics)[24], to perform causal inference analyses, after ensuring that the partial genetic correlations results were consistent with the ones obtained using the larger dataset. GWAS summary statistics were downloaded from the "GIANT" website (https://giant-consortium.web.broadinstitute.org/GIANT_consortium_data_files).

The GWAS summary statistics for the 71 LTCs, as well as BMI, were pre-processed using the munging function from the LD score regression (LDSC) package[25], and for all LDSC analyses, we used LD-scores estimated from the 1000G EUR reference panel.

Covariance and correlation

Covariance measures the extent to which the observed value of one quantity predicts the value of another quantity. For example, in the case of two such quantities – osteoarthritis (OA) and type 2 diabetes (T2D) diagnoses in a study population - a positive covariance would indicate that having T2D increases the probability of having OA. *Correlation* is simply a scaled version of covariance that lies between −1 and 1. If all individuals in a population with T2D also have OA, the two would have a correlation of 1. If all individuals in a population with T2D do not have OA, the two would have a correlation of −1.

Partial genetic covariance and correlation

In this paper, we extensively study the genetic covariance and correlation between two LTCs. That is, the degree to which genetic variants that predict one LTC also predict the other LTC. Furthermore, we focus on understanding how the genetic covariance between two LTCs changes when we remove the genetic effects of a common risk factor, which in our case is obesity, measured primarily by BMI and by WHR as an additional sensitivity analysis.

Using the Schur complement, the partial (or 'adjusted') genetic covariance between LTCs k and l ($\hat{\rho}_{g\{k,l|x\}}$), which corresponds to their genetic covariance while holding the genetic effects of the trait x (in our case, BMI) constant, can be defined as follows:

$$\hat{\rho}_{g\{k,l|x\}} = \hat{\rho}_{g\{k,l\}} - \frac{\hat{\rho}_{g\{k,x\}} * \hat{\rho}_{g\{x,l\}}}{\hat{h}^2_{\{x\}}} \quad (1)$$

where $\hat{\rho}_{g\{k,l\}}$ is the genetic covariance between condition l and k, $\hat{\rho}_{g\{k,x\}}$ is the genetic covariance between condition k and trait x, $\hat{\rho}_{g\{x,l\}}$ is the genetic covariance between trait x and condition l, and $\hat{h}^2_{\{x\}}$ is the heritability estimate for trait x.

Genetic covariance and heritability estimates were obtained from GWAS summary statistics using cross-trait LD Score regression (LDSC), as previously described in ref. 26.

Similarly, the partial heritability for LTCs k and l ($\hat{h}^2_{\{k|x\}}$ and $\hat{h}^2_{\{l|x\}}$, respectively) can be estimated:

$$\hat{h}^2_{\{k|x\}} = \hat{h}^2_{\{k\}} - \frac{\hat{\rho}_{g\{k,x\}} * \hat{\rho}_{g\{x,k\}}}{\hat{h}^2_{\{x\}}}$$
$$\hat{h}^2_{\{l|x\}} = \hat{h}^2_{\{l\}} - \frac{\hat{\rho}_{g\{l,x\}} * \hat{\rho}_{g\{x,l\}}}{\hat{h}^2_{\{x\}}} \quad (2)$$

where $\hat{h}^2_{\{k\}}$ is the heritability estimate for condition k and $\hat{h}^2_{\{l\}}$ is the heritability estimate for condition l.

Finally, the (unadjusted) genetic correlation ($\hat{r}_{g\{k,l\}}$) and the partial genetic correlation ($\hat{r}_{g\{k,l|x\}}$) between LTCs k and l can be derived:

$$\hat{r}_{g\{k,l\}} = \frac{\hat{\rho}_{g\{k,l\}}}{\sqrt{\hat{h}^2_{\{k\}}\hat{h}^2_{\{l\}}}}$$
$$\hat{r}_{g\{kl|x\}} = \frac{\hat{\rho}_{g\{k,l|x\}}}{\sqrt{\hat{h}^2_{\{k|x\}}\hat{h}^2_{\{l|x\}}}} \quad (3)$$

Our implementation of LDSC to calculate the unadjusted and partial correlations followed the LDSC approach taken in Genomic SEM[27].

To test if the partial (or adjusted) genetic correlation estimate is different from the unadjusted genetic correlation estimate, the following test statistic was used:

$$t_{\{k,l\}} = \frac{\hat{r}_{g\{k,l\}} - \hat{r}_{g\{k,l|x\}}}{\sqrt{Var(\hat{r}_{g\{k,l\}}) + Var(\hat{r}_{g\{k,l|x\}}) - 2Cov(\hat{r}_{g\{k,l\}}, \hat{r}_{g\{k,l|x\}})}} \quad (4)$$

where the denominator, which represents the standard error of the numerator, contains terms for the variance of the unadjusted genetic correlation ($Var(\hat{r}_{g\{k,l\}})$), the variance of the partial genetic correlation ($Var(\hat{r}_{g\{k,l|x\}})$) and the covariance between the two ($Cov(\hat{r}_{g\{k,l\}}, \hat{r}_{g\{k,l|x\}})$). Our specific contribution is to obtain this using a bespoke block-jackknife approach. When using SNP-trait associations, the standard jackknife technique cannot readily be used because of the underlying LD structure and the correlation between adjacent SNPs. Therefore, we implemented the following procedure, defining J jackknife blocks, to estimate the variance of each element of the genetic covariance matrix, ρ_g. Consider $\hat{\rho}_g$ the estimate obtained from the entire sample. For each block j, we created the j-th jackknife subsample by removing the j-th block from the GWAS summary statistics. We used this subsample to estimate the leave-one-out value ($\hat{\rho}_g^{\,j}$), and derive the pseudo value ($\tilde{\rho}_g^{\,j}$), when omitting block j: $\tilde{\rho}_g^{\,j} = J\hat{\rho}_g - (J-1)\hat{\rho}_g^{\,j}$. The variance of the pseudo-values ($\tilde{s}^2$) was then used to obtain an estimate of the variance of the genetic covariance matrix estimate: $Var(\hat{\mathrm{P}}) = \frac{\tilde{s}}{J}$. Further details of an empirical investigation which informed our choice of 200 for the block size are provided in Supplementary Note 1.

Similarly, we defined the partial genetic covariance matrix ($\hat{\rho}_{g|x}$), the genetic correlation matrix ($\hat{r}_g$) and the partial genetic correlation matrix ($\hat{r}_{g|x}$) and then calculated leave-one-out and pseudo-values to estimate their variance. This enabled us to rigorously test for a difference between the two.

We estimated partial genetic correlations between all pairs of LTCs, using two different sets of GWAS summary statistics for BMI[23,24]. False discovery rate (FDR) correction ($Q - value < 0.05$) was applied for both unadjusted and partial genetic correlations, and the differences between the two, to account for multiple testing. LTC pairs were then first classified according to the difference between the two (FDR-corrected statistically significant difference or not). Significant LTC pairs were then further classified into four categories:

- Both unadjusted and partial genetic correlations were statistically non-significant (FDR of 5%);
- Both unadjusted and partial genetic correlations were statistically significant;
- Only unadjusted genetic correlations were statistically significant;
- Only partial genetic correlations were statistically significant.

Here, we focus on the results with evidence of genetic similarity before and/or after adjustment for BMI. LTC pairs were defined as *within-domain* if both LTCs belong to the same disease domain, or *cross-domain* if the two LTCs belong to different disease domains.

Our analysis pipeline has been implemented in an R package - partialLDSC - and is available on GitHub [https://github.com/GEMINI-multimorbidity/partialLDSC]. All analyses have been performed using version 0.1.1.

Causal inference using Mendelian randomisation

To better understand the causal relationship between BMI and each individual LTC, we performed two-sample Mendelian Randomisation (MR) analyses. MR is a causal inference method that typically uses only the most strongly associated genetic variants with an exposure as instrumental variables to estimate the causal effect of the exposure on an outcome. We used this technique to estimate the causal effect of intervening to lower BMI on each individual LTC. Our primary analysis in each case was the IVW method, which is a summary data analogue of the general two-stage least squares approach applied to individual-level data – it assumes that all included variants are either valid instruments or, collectively, any pleiotropic effects approximately cancel out. This was supplemented with well-known sensitivity analysis approaches (weighted median, mode and MR-Egger) that allow the assumption of balanced pleiotropy to be relaxed in different ways. For further information on this established methodology, see Supplementary Figs. 1 and 2 and Supplementary Note 2.

Removing the BMI causal effect from genetic correlations

For LTCs found to be strongly causally affected by BMI (23 LTCs, Supplementary Note 2), we used a recently proposed Bayesian GWAS approach to derive direct effect estimates[28,29]. The direct effects were estimated by taking out the causal effect of BMI from the observed association effect between each genetic variant and the condition. They correspond to the direct association between genetic variants and the condition, not via the BMI pathway, and can be seen as BMI-corrected GWAS summary statistics. We ran the analysis for each condition using the bGWAS R package (https://github.com/n-mounier/bGWAS - version 1.0.3)[28], selecting instruments for BMI using a p-value threshold of $5*10^{-8}$ with default values used for other parameters. We then used the summary statistics for the direct effects to re-estimate genetic correlations (denoted as bGWAS genetic correlations, $\hat{r}'_{g\{k,l\}}$) for 253 pairs. We only reported the results for the 246 pairs for which we detected a statistically significant difference between the unadjusted and the partial genetic correlation estimates. We then compared partial genetic correlations to bGWAS genetic correlations to understand if the difference between unadjusted and partial genetic correlations is likely to be driven by the causal effect of BMI on LTCs, or solely by the correlation between BMI and the LTCs, which could be due to other mechanisms.

Estimating the causal effect of BMI on the co-occurrence of LTCs and estimating the effect of intervening on BMI

To further understand the role of BMI in the co-occurrence of LTCs, we carried out additional analyses for the 15 pairs with the strongest pre-/post-adjustment difference, and no evidence of genetic correlation, after adjusting for BMI. First, we performed an additional GWAS in the UK Biobank, defining cases for our analyses as individuals having been diagnosed with both LTCs (Supplementary Note 3). These GWAS summary statistics were then used to estimate the causal effect of BMI on the co-occurrence of each pair, using the same MR approach as for individual LTCs (Supplementary Note 2). The genetic associations used for MR are in standard deviation units for BMI (https://gwas.mrcieu.ac.uk/datasets/ieu-a-835/, 1 SD = 4.77), and are log odds ratios for the outcomes. Therefore, for each pair p, the causal effect estimate $\hat{\beta}_p$ represents log odds ratios per 1 standard deviation (σ) increase in BMI on the pair, compared to the actual observed BMI in the

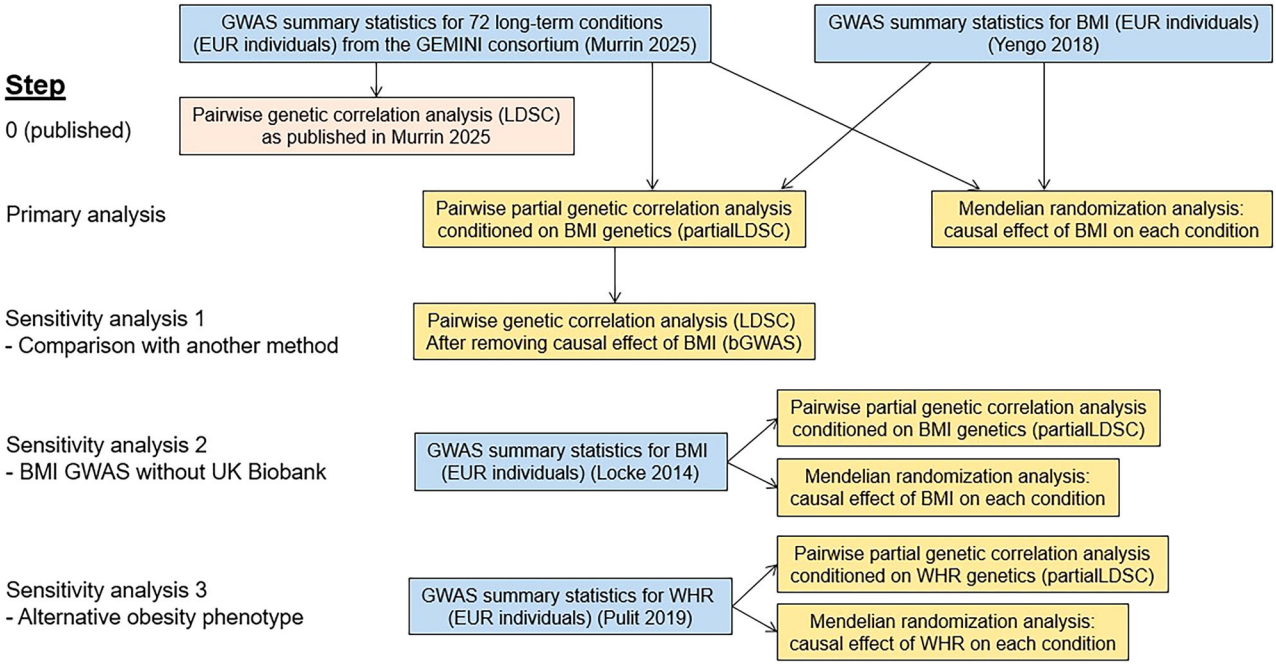

Fig. 1 | The flowchart of the analysis approach taken in our study, including the data and methods used. The flowchart illustrates the sources of input data and methods applied at each step to achieve the project aims.

Table 1 | Summary of the numbers of condition pairs significantly attenuated by BMI genetics

	$\hat{r}_g$ ($Q-value \geq 0.05$)	$\hat{r}_g$ ($Q-value<0.05$)
$\hat{r}_{g\|BMI}$ ($Q-value \geq 0.05$)	308	161
$\hat{r}_{g\|BMI}$ ($Q-value<0.05$)	33	860

Description of the 1362 pairs with statistically significant differences ($Q-value<0.05$ for the test statistics described in Eq. 4), according to the statistical significance of unadjusted ($\hat{r}_g$) and partial ($\hat{r}_{g|BMI}$) genetic correlation estimates.

sample $(\bar{\mu})$:

$$\hat{\beta}_p = \log\left(\frac{oddsP = 1|BMI = \bar{\mu} + \sigma}{oddsP = 1|BMI = \bar{\mu}}\right) \quad (5)$$

This causal effect estimate can then be used to evaluate how intervening to lower BMI would affect the number of people having both LTCs[30]. In practice, for each pair, we estimated the reduction in 1000 individuals ($\hat{R}_p$) of reducing BMI by 1 SD:

$$\hat{R}_p = 1000 \cdot \left(\hat{\pi}_p - \frac{\hat{\beta}_p \hat{\pi}_p}{\left(1 - \hat{\pi}_p + \hat{\beta}_p \hat{\pi}_p\right)}\right) \quad (6)$$

where $\hat{\pi}_p$ is the observed prevalence for pair p.

We used it to derive the number of cases that could be prevented from each pair following a hypothetical intervention in the sample used for our genetic analyses and estimated the impact on the prevalence of such an intervention.

Figure 1 provides a flowchart to illustrate the various steps of our analysis pipeline, including the data and methods used.

Ethics statement

We used published GWAS summary statistics, where the individual studies had obtained the relevant consent to make results available for health research. We used individual-level data from UK Biobank participants in this study for analysis of genetic associations with BMI. The Northwest Multi-Center Research Ethics Committee approved the collection and use of UK Biobank data (Research Ethics Committee reference 11/NW/0382). Participants gave informed consent for the use of their data, health records, and biological materials for health-related research purposes. Access to the UK Biobank was granted under Application Number 14631.

Results

BMI plays an important role in explaining genetic correlations between common long-term conditions

From the 2485 pairs defined from the 71 LTCs, we detected a statistically significant difference when adjusting for BMI genetics for 1362 pairs (Table 1, Supplementary Data 2). Figure 2 shows the difference between the unadjusted genetic correlations and the partial genetic correlation estimates after adjustment for BMI genetics. Supplementary Figs. 3 and 4 show the unadjusted and the partial genetic correlations. The pairs with a statistically significant difference encompassed 64 out of 71 LTCs. For most of these pairs (1078/1362), the partial genetic correlation estimates were weaker than the unadjusted ones, reflecting an attenuation of the genetic correlation when adjusting for BMI genetics.

In most of the pairs (860/1362) for which a statistically significant difference was observed, we found evidence of shared genetics both before and after adjusting for BMI, indicating that while BMI influences the genetic correlation for these pairs, it does not account for all of it. These include 131 *within-domain* pairs and 769 *cross-domain* pairs, many of them encompassing LTCs from the "diseases of the circulatory system", but also from the "diseases of the skin and subcutaneous tissue" and the "diseases of the digestive system" domains. For most of these pairs, the partial genetic correlation estimates were smaller (740/860). The majority of the 120 pairs where the partial genetic correlation estimates were larger were related to anxiety disorder (30), osteoporosis (27), schizophrenia and delusional disorders (12), as well as tinnitus (11). The unadjusted and the partial genetic correlation estimates for the 20 pairs of these 860 with the strongest difference are presented in Fig. 3. Amongst these 20 top pairs, 12 distinct LTCs are represented, with some conditions common between pairs with cholelithiasis in 6 pairs and carpal tunnel syndrome, gout and chronic kidney disease in 7 pairs. This illustrates that BMI genetics tends to explain a larger

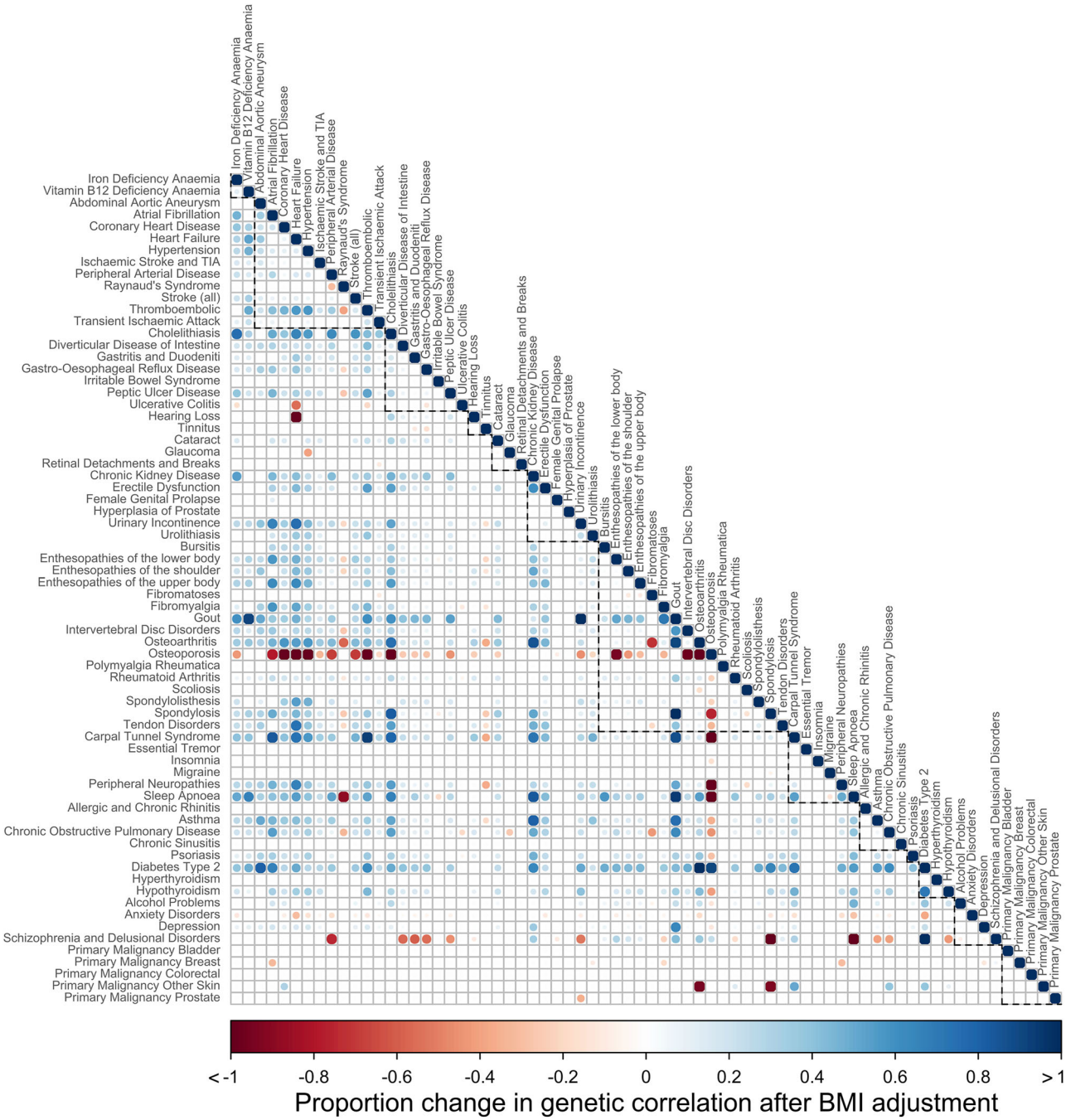

Fig. 2 | Heatmap of the difference between full and partial genetic correlations for all 2485 disease pairs after adjustment for the genetics of BMI. Pairwise partial genetic correlation analysis between 71 conditions (i.e., 2485 disease pairs). Condition pairs where there was no significant (FDR correction) genetic correlation before or after adjustment, or the adjustment for BMI genetics was non-significant (FDR corrected), are coloured white. Data shown is the proportion change in genetic correlation between condition pairs after adjustment for BMI genetics (where the change is statistically significant after FDR correction). See Supplementary Data 2 for tabular data.

proportion of genetic correlation in pairs that contain specific LTCs. Furthermore, these LTCs are also amongst the ones with the strongest genetic correlations with BMI (Supplementary Data 1), and most strongly causally affected by BMI in our MR results (Supplementary Note 2). For 471 of these 860 pairs, our MR analysis showed that BMI is likely causal for both LTCs (Supplementary Data 6), providing further evidence that BMI is the likely source of the genetic correlation.

For more than 10% of the pairs with a significant difference (161/1362), we showed that BMI is likely to account for the vast majority of the genetic correlation. For these pairs, BMI adjustment resulted in a greatly reduced partial genetic correlation consistent with no remaining genetic similarity between the two LTCs. Many of these pairs (76/161) included at least one condition from the "*diseases of the circulatory system*" domain, and almost all the pairs were *cross-domain* pairs (154/161) for example 24 pairs had one condition from the "*diseases of the circulatory system*" domain, and one condition from the "*diseases of the musculoskeletal system and connective tissue*" domain. For instance, we observed complete attenuation of the genetic correlation between: type 2 diabetes and osteoarthritis ($\hat{r}_g = 0.2216(SE = 0.0208)$; $\hat{r}_{g|BMI} = -0.0189\,(SE = 0.0296)$); gout and osteoarthritis ($\hat{r}_g = 0.1920(SE = 0.0229)$; $\hat{r}_{g|BMI} = 0.0191(SE = 0.0292)$); and gout and sleep apnoea ($\hat{r}_g = 0.2313(SE = 0.0266)$; $\hat{r}_{g|BMI} = 0.0118(SE = 0.0355)$) (Supplementary Data 2, Fig. 4). For 93 of these pairs,

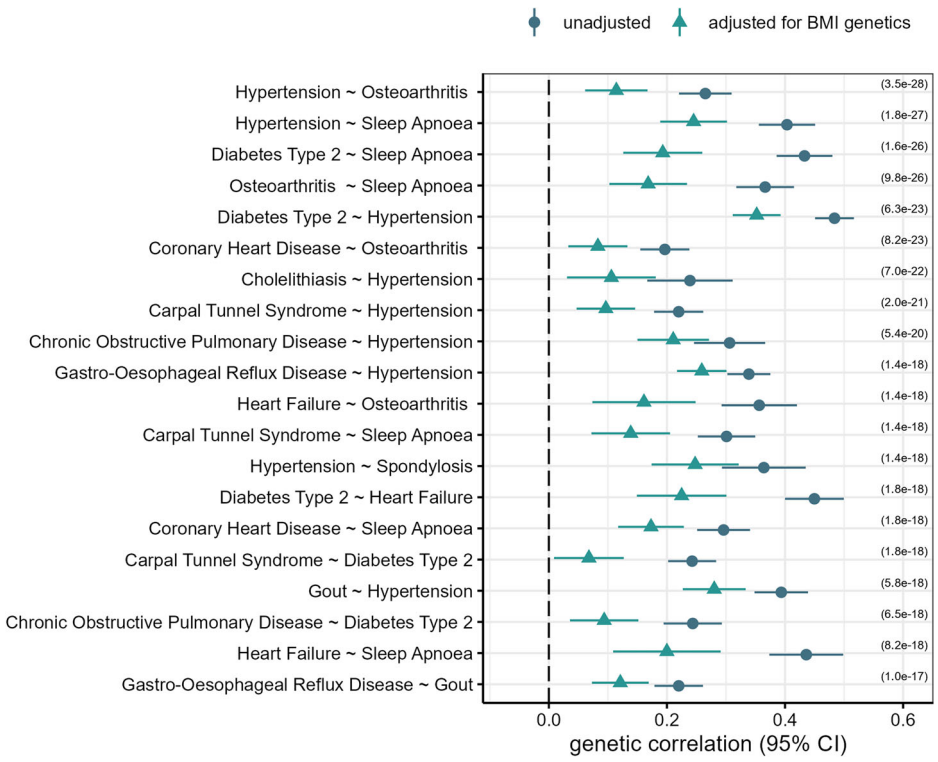

Fig. 3 | Condition pairs where BMI genetics strongly attenuated the genetic correlation, but evidence of shared genetics remained. Unadjusted (blue circle) and partial (blue-green triangle) genetic correlation estimates and 95% confidence intervals for the 20 pairs with the strongest difference ($Q-value$ indicated for each pair on the right) and evidence of shared genetics both before and after adjusting for BMI. See Supplementary Data 2 for tabular data.

including all in Fig. 4, our MR analysis detected a statistically significant causal effect of BMI on both LTCs.

For 33 LTC pairs, there was no evidence (at FDR < 0.05) of genetic correlation, but on adjustment for the genetics of BMI, a residual genetic effect was observed (Supplementary Data 2, Fig. 5). These results are consistent with BMI masking shared genetic mechanisms between the two LTCs. Most of these pairs were cross-domain pairs (28/33), and notably, 14 of these pairs were related to osteoporosis. This can be explained by the fact that lower BMI has been shown to be associated with higher osteoporosis risk[31] whereas for most other LTCs, higher BMI is suspected to be risk-increasing. This is further supported by our MR analysis, as we found a statistically significant causal effect of lower BMI on osteoporosis (Supplementary Note 2). For 19 of these pairs, our MR results showed that BMI is likely causal for both LTCs, but protective for one of them, consistent with the partial genetic correlation being stronger than the unadjusted genetic correlation.

For 1123 LTC pairs, adjusting for BMI genetics had no statistically significant effect on the genetic correlation between them, suggesting that other mechanisms are giving rise to the genetic similarity between these LTCs. For example, the strong genetic correlation between allergic rhinitis and bursitis ($\hat{r}_g = 0.3640(SE = 0.0777)$; $\hat{r}_{g|BMI} = 0.3639(SE = 0.0795)$; difference$Q-value = 0.98$), or between anxiety and schizophrenia ($\hat{r}_g = 0.4417(SE = 0.0234)$; $\hat{r}_{g|BMI} = 0.4414(SE = 0.0232)$; difference $Q-value = 0.91$), remain similar after adjusting for BMI genetics.

These results were consistent with those obtained using less recent BMI data, excluding UK Biobank, with partial correlation estimates that were highly similar (correlation= 0.9974, Supplementary Fig. 5A), with evidence of the analysis using the more recent data being more powered (Supplementary Table 1, Supplementary Data 3, Supplementary Fig. 5B).

BMI acts as a common risk factor in multimorbidity

To compare our partial genetic correlations to another method (bGWAS), we estimated pairwise genetic correlations accounting for the estimated causal effect of BMI on *both* LTCs, for 246 pairs. This comprised the 23 LTCs that were strongly causally affected by BMI and a subset of LTC pairs for which a statistically significant difference between unadjusted and partial genetic correlation estimates was observed. We used bGWAS to take out the causal effect of BMI on each individual LTC and re-estimated pairwise genetic correlations. These differ from the partial correlation estimates since they only remove the BMI genetics estimated to be due to the causal role of BMI on each condition. We observed strong agreement between the partial genetic correlation estimates and the bGWAS correlation estimates, with most of the bGWAS genetic correlation estimates being stronger (less attenuated) than the partial ones (Fig. 6, Supplementary Data 4), suggesting that partial genetic correlation estimates might capture additional associations between BMI and the LTCs that are not due to these causal relationships. See Supplementary Note 4 for details.

For the 15 pairs with the strongest difference and no evidence of genetic correlation after adjusting for BMI genetics, we performed genetic analysis for each pair and further investigated the role of BMI using MR results. We showed that higher BMI had a strong risk-increasing causal effect on every condition within these pairs. In addition, the unadjusted genetic correlation estimates were all positive, and higher BMI was shown to be risk-increasing for all LTCs (Supplementary Data 5). This suggests that the attenuation of the genetic correlation is likely due to BMI acting as a common risk factor. This is further confirmed by the MR analysis on the pairs, with all IVW causal effect estimates being positive and statistically significant ($p-value < 0.05/64$, Bonferroni correction) (Supplementary Data 5).

For these pairs with directionally consistent individual causal effects, we also estimated how intervening on BMI would affect the number of people having both LTCs. To do so, we calculated the reduction in the number of cases for 1000 individuals expected by a one SD BMI reduction. In the UK Biobank sample used in the analysis, the mean BMI is approximately 27, and its standard deviation is 4.77. A 1 SD intervention applied to the population can therefore be viewed as taking the population mean BMI from 27 (considered moderately overweight) to 22.23 (considered to be 'normal'). Results were highly correlated with the observed prevalence of the pairs in our sample, and lowering BMI had a stronger effect on pairs with higher prevalence (Supplementary Data 5). Results for the 5 pairs with a prevalence above 1% are presented in Table 2. For example, 16 out of 1000

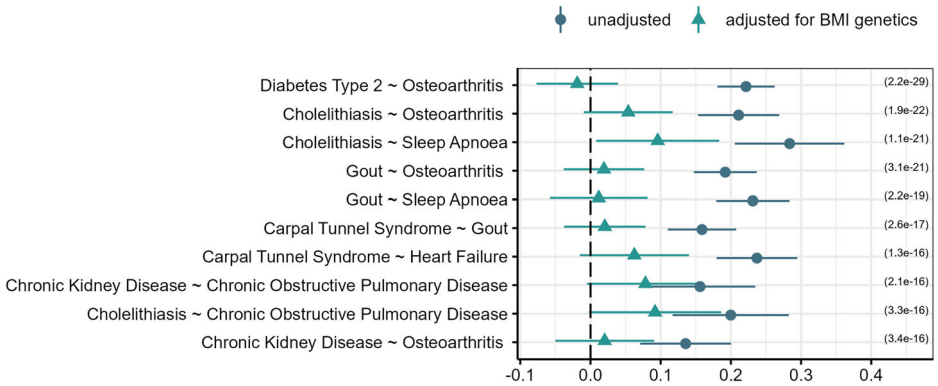

Fig. 4 | Top 10 condition pairs where there was no longer a genetic correlation between the conditions after accounting for BMI genetics. Unadjusted (blue circle) and partial (blue-green triangle) genetic correlation estimates and 95% confidence intervals for the 10 pairs with the strongest difference ($Q-value$ indicated for each pair on the right) and evidence of shared genetics only before adjusting for BMI. See Supplementary Data 2 for tabular data.

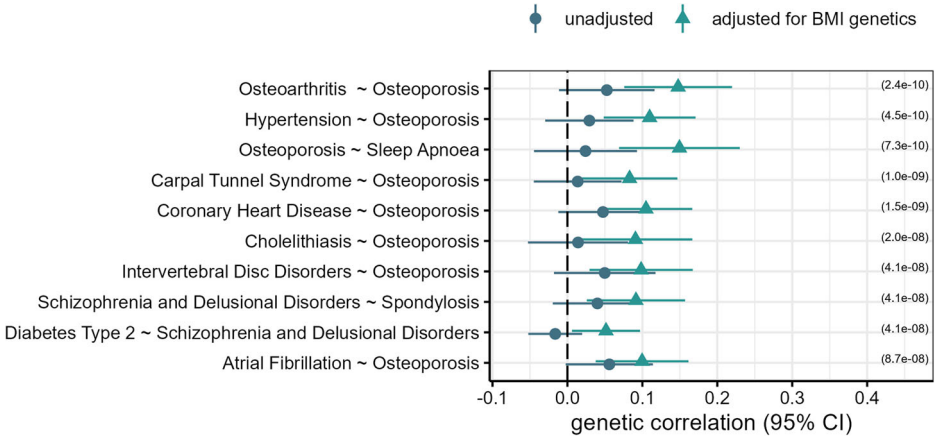

Fig. 5 | Top 10 condition pairs where there was no genetic correlation between the conditions originally, but there was after accounting for BMI genetics. Unadjusted (blue circle) and partial (blue-green triangle) genetic correlation estimates and 95% confidence intervals for the 10 pairs with the strongest difference ($Q-value$ indicated for each pair on the right) and evidence of shared genetics only after adjusting for BMI. See Supplementary Data 2 for tabular data.

people having both chronic kidney disease and osteoarthritis, and 9 out of 1000 people having both type 2 diabetes and osteoarthritis, would not have both LTCs after a one SD BMI intervention (1 SD = 4.77). In the UK Biobank sample used for our analysis, this would translate into a 4.6% reduction in prevalence for the co-occurrence of chronic kidney disease and osteoarthritis, and a 3.3% reduction in prevalence for the co-occurrence of type 2 diabetes and osteoarthritis. We postulate that the impact of such an intervention in the general population is likely to be stronger, since the prevalence of these pairs of LTC in the UK Biobank is lower than in the general population (Supplementary Data 5), due to healthy volunteer bias in the UK Biobank[32].

Highly concordant analysis using WHR

To provide additional robustness and insight, we compared unadjusted genetic correlations to their correlation after adjustment for WHR for all 2485, using genetic summary statistics for WHR in Pulit et al.[33]. Whereas previously we found that 1362 pairs had a statistically significant difference in genetic correlation after adjusting for BMI genetics, we found that 1370 had a statistically significant difference in genetic correlation after adjusting for WHR genetics. 1072 of these (78%) were the same (this includes 12 pairs where the genetic correlation is fully attenuated after adjustment for WHR genetics, but had some residual genetic correlation when using BMI). Our analysis did identify 298 pairs where WHR genetics significantly attenuates the genetic correlation, but adjustment for BMI genetics did not. The top 20/298 pairs where this occurred are shown in Fig. 7. Full partial LDSC results for genetic correlation before/after WHR adjustment are shown in Supplementary Data 7, Supplementary Data 8 contains updated MR results.

Discussion

Clinical trials and Mendelian randomisation studies have shown that obesity is very likely to cause many individual LTCs[34]. However, the role of higher BMI in the co-occurrence of two LTCs and the extent to which it accounts for some or all of the co-occurrence is less certain. In this study, we combined genetics data from very large studies with a powerful statistical analysis technique to estimate the extent to which higher BMI influences the genetic similarity between pairs of LTCs. Our analyses indicated that for a large proportion of pairs of LTCs, higher BMI is likely to explain part of the genetic correlation, but that other factors are likely to also contribute. For 55% of LTC pairs tested, we observed a change in the genetic correlations when accounting for BMI genetics. Using causal inference methods, we confirmed that BMI acts as a common risk factor for a subset of these pairs, and that intervening on BMI could help reduce the prevalence of pairs of multimorbid LTCs.

Our analyses indicated that for 36% (740/2485) of LTC pairs we studied, higher BMI could explain a significant proportion of their genetic correlation. Of these, 161 pairs were identified for which the estimates were consistent with a null partial genetic correlation, suggesting that the genetic similarity between these pairs of LTCs is *entirely* explained by the BMI. These included many pairs that spanned traditional disease domains, especially those involving osteoarthritis and metabolic diseases, and COPD and metabolic diseases. For some of these pairs, previous studies have observed a similar strong attenuating effect of BMI, such as for type 2 diabetes and osteoarthritis and gout and sleep apnoea[35,36].

Thirty-three pairs (about 1%) of LTCs showed no evidence of unadjusted genetic similarity, but after accounting for the genetic effect of BMI, a partial genetic correlation emerged. These results suggest that shared genetic causes may exist for some pairs, but they could be masked by BMI having effects in opposite directions. Most of these pairs were related to osteoporosis, which may be explained by the fact that the direction of the effect of BMI on osteoporosis is opposite to the one on most other LTCs[37]. We observed, for example, a positive genetic correlation between osteoporosis and stroke, only after accounting for BMI genetics. Because genetic risk

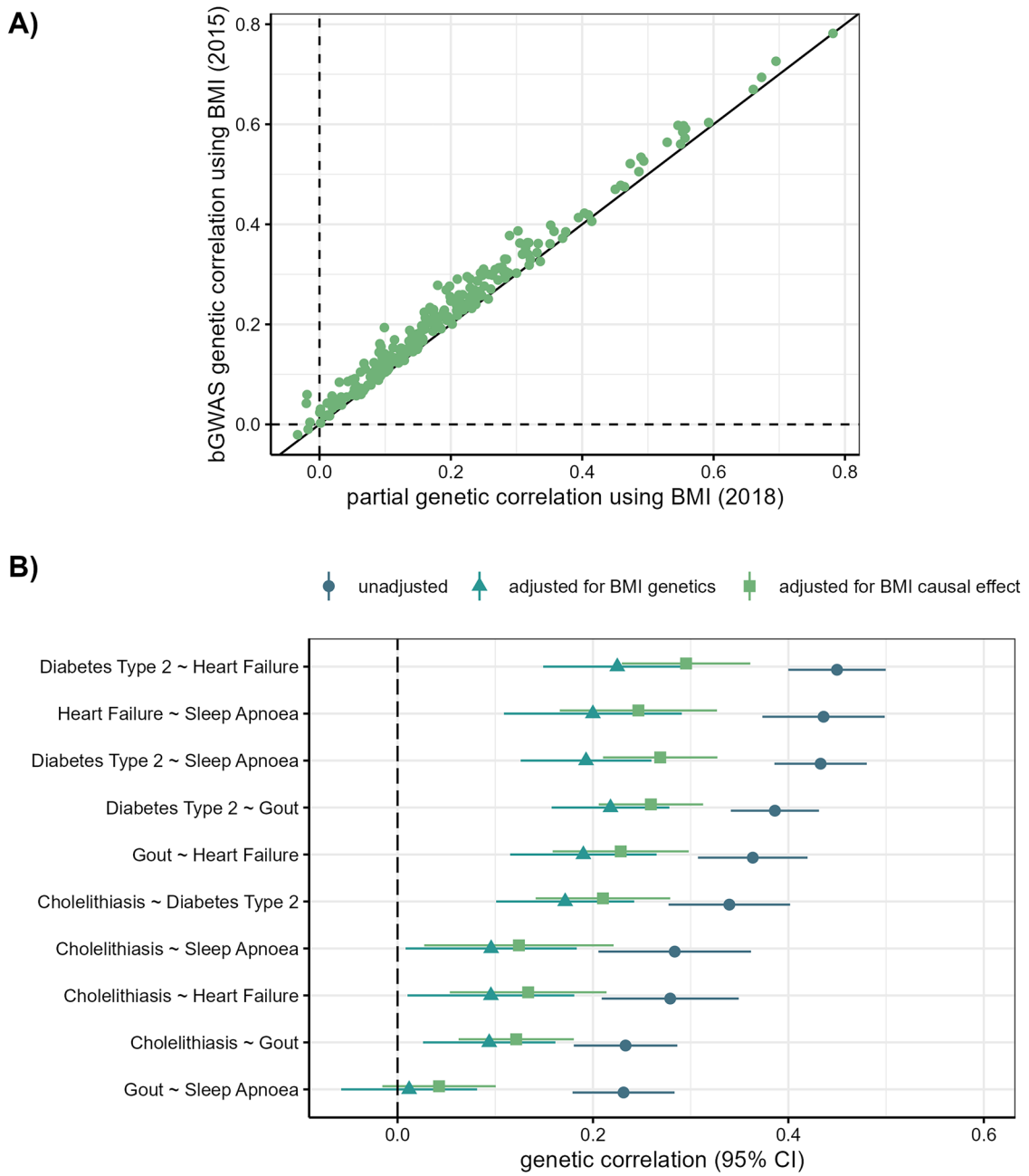

Fig. 6 | Comparison of two methods to adjust the genetic correlations between each condition pair for the genetics of BMI. A Partial genetic correlation estimates (using BMI data from 2018, x-axis) and bGWAS-adjusted (which removes the causal effect of BMI prior to estimating genetic correlations) genetic correlation estimates (estimated from direct effects obtained using bGWAS and BMI data from 2015, y-axis). **B** Unadjusted (blue circle), partial (blue-green triangle) and bGWAS (green square) genetic correlation estimates and 95% confidence intervals for the 10 pairs encompassing the 5 conditions that are the most strongly affected by BMI. See Supplementary Data 2 and 3 for tabular data.

usually reflects life-long exposure, this correlation suggests the presence of shared risk factors for osteoporosis and metabolic traits, although we cannot rule out a genetic correlation resulting from reduced mobility and bone load reduction in stroke patients[38].

Our results corroborate the role of BMI as a common risk factor for various pairs of LTCs. We have used causal inference approaches to test the causal relationship between BMI and the different LTCs, and we used the bGWAS approach to re-estimate genetic correlation estimates, adjusting only for the part of the genetic correlation that is driven by the causal effect of BMI on the LTCs. Overall, we observed a strong agreement between partial and bGWAS genetic correlation estimates, providing further evidence that BMI does act as a common risk factor. However, the bGWAS genetic correlation estimates were in general stronger, suggesting that the correlation between BMI and the LTCs is not entirely explained by the causal effect and that more complex relationships may exist. In addition, the effects derived from bGWAS would only capture linear causal effects, so any non-linear effects would not be accounted for, possibly explaining why the bGWAS genetic correlation estimates were stronger. We also showed that BMI has a strong, risk-increasing, causal effect on some pairs of LTCs, and that lowering BMI could help reduce the prevalence of these pairs.

Many of the LTC pairs include those from across traditional clinical domains, and for most of these pairs, the attenuation after accounting for BMI was as strong as that for within-domain pairs. For example, based on our genetic approach, the role of BMI in explaining the co-occurrence of COPD and several metabolic LTCs and osteoarthritis and several metabolic LTCs, is just as strong as the role of BMI in explaining the co-occurrence of

Table 2 | Estimation of the effect of intervening on BMI

Condition 1	Condition 2	$\hat{r}_g$ (SE)	$\hat{r}_{g\|BMI}$ (SE)	$\hat{\beta}_{IVW}$ (SE; $p-value$)	$\hat{R}$	πUK Biobank	$\hat{\pi}_R$ UK Biobank
Chronic kidney disease	Osteoarthritis	0.1356 (0.0330)	0.0202 (0.0359)	0.38 (0.0835; 4.2e-06)	16.76	2.752	2.706
Diabetes type 2	Osteoarthritis	0.2216 (0.0208)	−0.0189 (0.0296)	0.71 (0.1820; 9.4e-05)	9.69	3.434	3.401
Cholelithiasis	Osteoarthritis	0.2110 (0.0295)	0.0538 (0.0323)	0.63 (0.0775; 6.0e-16)	8.83	2.406	2.385
Asthma	Chronic kidney disease	0.0829 (0.0353)	0.0158 (0.0361)	0.35 (0.1132; 2.1e-03)	7.10	1.092	1.085
Gout	Osteoarthritis	0.1920 (0.0229)	0.0191 (0.0292)	0.61 (0.0963; 3.2e-10)	6.47	1.657	1.646

Mendelian randomisation and BMI intervention results for 5 pairs with a prevalence above 1% in our sample. For each pair of LTC (condition 1 and condition 2), the unadjusted genetic correlation estimate and standard error ($\hat{r}_g$), the partial genetic correlation estimate and standard error ($\hat{r}_{g|BMI}$), as well as the causal effect estimate ($\hat{\beta}_{IVW}$), its standard error and the corresponding p-value, of BMI on the pair, and the reduction in the number of cases with both LTCs when reducing BMI by one SD for 1000 individuals ($\hat{R}$), and the prevalence of cases with both LTCs in percentage in UK Biobank before (π UK Biobank) and after an hypothetical intervention ($\hat{\pi}_R$ UK Biobank), are reported.

two metabolic LTCs. The reasons for these findings require further investigation, but our results suggest that interventions aimed at weight loss should monitor these additional LTCs more closely, or conversely, patients with a combination of obesity and a metabolic condition should be monitored for less obvious non-metabolic LTCs.

The definition of obesity, and other obesity-related measures and phenotypes that can be used as proxies in public health research, is critical[39], and which one is the best predictor for disease risk might depend on the condition[40–46]. For that reason, we believe that results using BMI, as a continuous phenotype that can be easily measured in the general population, are less likely to be biased and are easier to interpret. However, to add further robustness to our findings, we also observed a broadly similar picture when using WHR as an alternative measure of obesity.

Our work not only has important implications for research that aims at identifying shared biological pathways between pairs of LTCs but also has direct implications for potential intervention. We showed that intervening on BMI would directly impact the prevalence of pairs of LTCs, such as type 2 diabetes and chronic kidney disease, or type 2 diabetes and osteoarthritis. This is particularly important with the advent of effective weight loss pharmaceutical treatments that could be used to reduce the co-occurrence of LTCs through a better weight management strategy.

When assessing if two long-term conditions' genetic correlation is driven by BMI genetics, and looking as we did across 2485 distinct pairs, we took the view that it was sensible and practical to do this at the 'global' level, by averaging over all mechanisms underlying BMI. However, obesity is a multidimensional quantity and two individuals with the same BMI could have very different body types, health status and subsequent disease risk. We recommend that interesting LTC pairs highlighted from our broader analysis strategy are perfect candidates for further in-depth study to uncover the possible mechanisms at play. To provide a tool for just this purpose, in closely related work recently published by the GEMINI group, Liang et al.[47] develop an extended summary data MR method to uncover heterogeneity in the joint causal effects of an exposure on two LTCs that is driven by subsets of genetic variants. These subsets are conceptualised as encoding latent subcomponents of the exposure. Our exemplar study estimated the causal effect of body fat percentage (BFP) on Type 2 diabetes and osteoarthritis using 487 SNPs previously identified by Martin et al.[42]. Liang et al. identified four distinct genetic clusters: the most pertinent cluster (positive causal effect of increased BFP on OA but negative causal effect on OA risk) and Cluster 1 (positive causal effects of increased BFP on T2D *and* OA risk). These results add weight to the growing hypothesis that increased body fat exerts heterogeneous effects on T2D, with metabolically favourable (i.e., ectopic) fat lowering T2D risk and metabolically unfavourable (i.e., visceral) fat increasing T2D risk, but increased body fat is uniformly associated with increased OA risk due to its load-bearing effect.

An obvious limitation of our general approach for exploring genetically correlated LTCs is the consideration of a single risk factor. In future work, we will extend the approach to enable adjustment for the genetics of multiple obesity risk factors simultaneously, most obviously BMI, BFP, WC and WHR together. This will provide further insights into the most important measure driving the correlation within each pair.

Another limitation of our current approach is its focus on disease pairs in isolation. A more complete model would attempt to uncover the complex network of causal relationships between all 71 LTCs. Extending our current genome-wide correlation workflow from pairs to look even at triplets or quadruplets of LTCs would necessitate around 60,000 and 1 M analyses, respectively, which clearly isn't scalable. However, promising recently proposed algorithms such as MRSL[48] can learn causal network structures from the relatively modest starting point of all bi-directional causal estimates between each LTC pair (i.e., 2*2485). We will look to incorporate this approach as future work.

This present work focused on individuals of European descent, mostly because of data availability, and analyses are needed in other ethnic groups, as it is known that both BMI distribution and its effect on LTCs may vary depending on ethnicity[49,50]. It is important to note that the relative paucity of

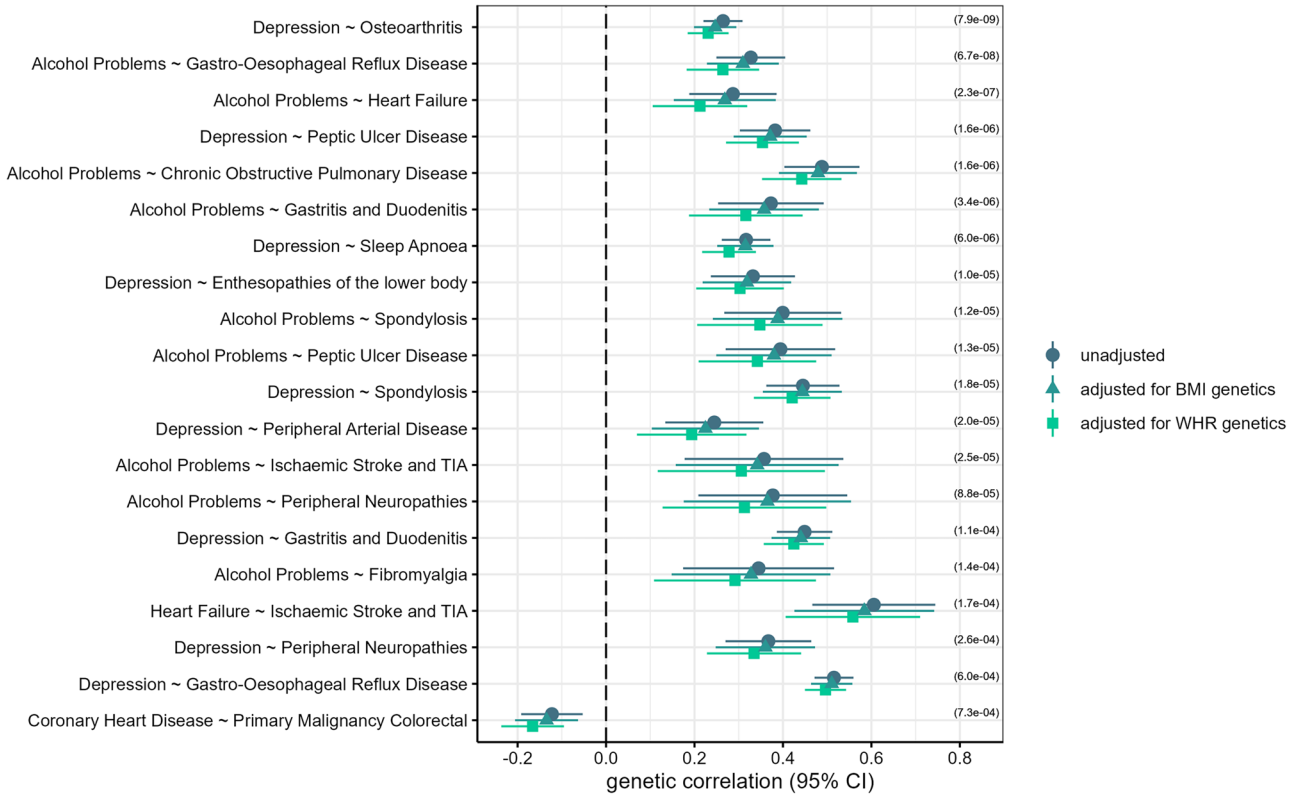

Fig. 7 | Condition pairs where WHR genetics attenuated the genetic correlation, but BMI did not. Genetic correlation estimates and 95% confidence intervals for the top 20 out of 298 LTC pairs ordered by significance of the difference of before/after adjustment for WHR, which did not have a significant difference between before/after adjustment for BMI (after removing the sex-specific conditions, with the BMI effects also). See Supplementary Data 2 and 7 for tabular data.

data from people of non-European ancestry has a disproportionately large effect on the utility of genetics for multimorbidity, because missing data from only one condition will affect multiple pairs.

Conclusion

Our work presents a clear, logical pathway for interrogating LTC pairs that are genetically correlated, finding pairs whose genetic correlation changes dramatically when adjusting for the genetics of a risk factor, and then using Mendelian randomisation to further test for and quantify causal relationships. Mendelian randomisation is not, however, a panacea. It should be part of a broader suite of analyses, as we have done here, because it is fallible. Given the large number of genome-wide significant SNPs available for our chosen risk factors (BMI and WHR), our MR analyses are not affected by weak instrument bias, but genetic pleiotropy[51] (variants affecting the LTC outcome through pathways other than BMI or WHR) is a real possibility. Finally, our MR-estimates rely on a single measurement of the exposure and therefore estimate average effects across a person's adult lifetime. Future work based on longitudinal study data[52], with multiple measures of BMI across the life course, would provide a richer basis for understanding the role of obesity on long-term disease using traditional methods and Mendelian randomisation.

Data availability

GWAS summary statistics from the GEMINI consortium are available on Zenodo (https://doi.org/10.5281/zenodo.14284046). GWAS summary statistics for BMI and WHR were downloaded from the "GIANT" website (https://giant-consortium.web.broadinstitute.org/GIANT_consortium_data_files). Source data for Fig. 2 can be found in Supplementary Data 1. Source data for Figs. 3–5 can be found in Supplementary Data 2. Source data for Fig. 6 can be found in Supplementary Data 2 and Supplementary Data 3. Source data for Fig. 7 can be found in Supplementary Data 2 and Supplementary Data 7.

Code availability

We have released the 'partialLDSC' method as an R package (https://github.com/GEMINI-multimorbidity/partialLDSC, https://doi.org/10.5281/zenodo.12721532). We used v0.1.1 in this manuscript. We have made the main analysis scripts to perform the partialLDSC analysis, perform the bGWAS adjustment, and perform the MR analyses available in the following public GitHub repository: https://github.com/GEMINI-multimorbidity/Mounier2025_BMI_attenuation (this repository is archived on Zenodo https://doi.org/10.5281/zenodo.17725332)[53].

Received: 22 July 2024; Accepted: 18 December 2025;
Published online: 04 February 2026

References

1. Marengoni, A. et al. Aging with multimorbidity: a systematic review of the literature. *Ageing Res. Rev.* **10**, 430–439 (2011).
2. Johnston, M. C., Crilly, M., Black, C., Prescott, G. J. & Mercer, S. W. Defining and measuring multimorbidity: a systematic review of systematic reviews. *Eur. J. Public Health* **29**, 182–189 (2019).
3. Ho, I. S. S. et al. Measuring multimorbidity in research: Delphi consensus study. *BMJ Med.* **1**, e000247 (2022).
4. Chowdhury, S. R., Chandra Das, D., Sunna, T. C., Beyene, J. & Hossain, A. Global and regional prevalence of multimorbidity in the adult population in community settings: a systematic review and meta-analysis. *EClinicalMedicine* **57**, 101860 (2023).
5. Calderón-Larrañaga, A. et al. Assessing and measuring chronic multimorbidity in the older population: a proposal for its operationalization. *J. Gerontol. A Biol. Sci. Med Sci.* **72**, 1417–1423 (2017).
6. Zemedikun, D. T., Gray, L. J., Khunti, K., Davies, M. J. & Dhalwani, N. N. Patterns of multimorbidity in middle-aged and older adults: an analysis of the UK Biobank data. *Mayo Clin. Proc.* **93**, 857–866 (2018).

7. Ho, I. S.-S. et al. Examining variation in the measurement of multimorbidity in research: a systematic review of 566 studies. *Lancet Public Health* **6**, e587–e597 (2021).
8. Nichols, L. et al. In simulated data and health records, latent class analysis was the optimum multimorbidity clustering algorithm. *J. Clin. Epidemiol.* **152**, 164–175 (2022).
9. Pathirana, T. I. & Jackson, C. A. Socioeconomic status and multimorbidity: a systematic review and meta-analysis. *Aust. N. Z. J. Public Health* **42**, 186–194 (2018).
10. Kivimäki, M. et al. Body-mass index and risk of obesity-related complex multimorbidity: an observational multicohort study. *Lancet Diabetes Endocrinol.* **10**, 253–263 (2022).
11. Murrin, O. et al. A systematic analysis of the contribution of genetics to multimorbidity and comparisons with primary care data. *EBioMedicine* **113**, 105584 (2025).
12. Kraft, P., Chen, H. & Lindström, S. The use of genetic correlation and mendelian randomization studies to increase our understanding of relationships between complex traits. *Curr. Epidemiol. Rep.* **7**, 104–112 (2020).
13. Dixon, P., Hollingworth, W., Harrison, S., Davies, N. M. & Davey Smith, G. Mendelian Randomization analysis of the causal effect of adiposity on hospital costs. *J. Health Econ.* **70**, 102300 (2020).
14. van Rheenen, W., Peyrot, W. J., Schork, A. J., Lee, S. H. & Wray, N. R. Genetic correlations of polygenic disease traits: from theory to practice. *Nat. Rev. Genet.* **20**, 567–581 (2019).
15. Verbanck, M., Chen, C.-Y., Neale, B. & Do, R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat. Genet.* **50**, 693–698 (2018).
16. Larsson, S. C. & Burgess, S. Causal role of high body mass index in multiple chronic diseases: a systematic review and meta-analysis of Mendelian randomization studies. *BMC Med.* **19**, 320 (2021).
17. Munafò, M. R., Tilling, K., Taylor, A. E., Evans, D. M. & Davey Smith, G. Collider scope: when selection bias can substantially influence observed associations. *Int. J. Epidemiol.* **47**, 226–235 (2018).
18. World Health Organisation. *International Statistical Classification of Diseases and Related Health Problems (ICD)*. https://icd.who.int/ (2019).
19. Murrin, O. et al. A systematic analysis of the contribution of genetics to multimorbidity and comparisons with primary care data. *EBioMedicine* **113**, 105584 (2025)
20. Recalde, M. et al. Data resource profile: the information system for research in primary care (SIDIAP). *Int. J. Epidemiol.* **51**, e324–e336 (2022).
21. Herrett, E. et al. Data resource profile: Clinical Practice Research Datalink (CPRD). *Int. J. Epidemiol.* **44**, 827–836 (2015).
22. Sudlow, C. et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
23. 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 (2018).
24. Locke, A. E. et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197–206 (2015).
25. Bulik-Sullivan, B. K. et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
26. Bulik-Sullivan, B. et al. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
27. Grotzinger, A. D. et al. Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat. Hum. Behav.* **3**, 513–525 (2019).
28. Mounier, N. & Kutalik, Z. bGWAS: an R package to perform Bayesian genome wide association studies. *Bioinformatics* **36**, 4374–4376 (2020).
29. Pirastu, N. et al. Using genetic variation to disentangle the complex relationship between food intake and health outcomes. *PLoS Genet.* **18**, e1010162 (2022).
30. Murad, M. H. et al. Methods for deriving risk difference (absolute risk reduction) from a meta-analysis. BMJ 381, e073141 (2023).
31. Espallargues, M. et al. Identifying bone-mass-related risk factors for fracture to guide bone densitometry measurements: a systematic review of the literature. *Osteoporos. Int.* **12**, 811–822 (2001).
32. Fry, A. et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am. J. Epidemiol.* **186**, 1026–1034 (2017).
33. Pulit, S. L. et al. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. *Hum. Mol. Genet.* **28**, 166–174 (2019).
34. Kivimäki, M. et al. Body-mass index and risk of obesity-related complex multimorbidity: an observational multicohort study. *Lancet Diabetes Endocrinol.* **10**, 253 (2022).
35. Duclos, M. Osteoarthritis, obesity and type 2 diabetes: the weight of waist circumference. *Ann. Phys. Rehabil. Med.* **59**, 157–160 (2016).
36. Zeng, Z. et al. Assessing the causal associations of obstructive sleep apnea with serum uric acid levels and gout: a bidirectional two-sample Mendelian randomization study. *Semin. Arthritis Rheum.* **57**, 152095 (2022).
37. Espallargues, M. et al. Identifying bone-mass-related risk factors for fracture to guide bone densitometry measurements: a systematic review of the literature. *Osteoporos. Int.* **12**, 811 (2001).
38. Carda, S., Cisari, C., Invernizzi, M. & Bevilacqua, M. Osteoporosis after stroke: a review of the causes and potential treatments. *Cerebrovasc. Dis.* **28**, 191–200 (2009).
39. Chooi, Y. C., Ding, C. & Magkos, F. The epidemiology of obesity. *Metabolism* **92**, 6–10 (2019).
40. Sulc, J., Winkler, T. W., Heid, I. M. & Kutalik, Z. Heterogeneity in obesity: genetic basis and metabolic consequences. *Curr. Diab Rep.* **20**, 1 (2020).
41. Sulc, J. et al. Composite trait Mendelian randomization reveals distinct metabolic and lifestyle consequences of differences in body shape. *Commun. Biol.* **4**, 1064 (2021).
42. Martin, S. et al. Disease consequences of higher adiposity uncoupled from its adverse metabolic effects using Mendelian randomisation. *eLife* https://doi.org/10.7554/eLife.72452 (2022).
43. Emerenziani, S. et al. Role of overweight and obesity in gastrointestinal disease. *Nutrients* **12**, 111 (2019).
44. Bianchini, F., Kaaks, R. & Vainio, H. Overweight, obesity, and cancer risk. *Lancet Oncol.* **3**, 565–574 (2002).
45. Silva Junior, da G. B., Bentes, A. C. S. N., Daher, E. D. F. & de Matos, S. M. A. Obesity and kidney disease. *J. Bras. Nefrol.* https://doi.org/10.5935/0101-2800.20170011 (2017).
46. Williamson, D. F. et al. Intentional weight loss and mortality among overweight individuals with diabetes. *Diabetes Care.* **23**, 1499–1504 (2000).
47. Liang, X. et al. Using clustering of genetic variants in Mendelian randomization to interrogate the causal pathways underlying multimorbidity. *Genet Epidemiol*. **49**, e22582 (2025)
48. Hou, L. et al. MRSL: a causal network pruning algorithm based on GWAS summary data. *Brief Bioinform.* https://doi.org/10.1093/bib/bbae086 (2024).
49. Colin Bell, A. Ethnic differences in the association between body mass index and hypertension. *Am. J. Epidemiol.* **155**, 346–353 (2002).
50. Caleyachetty, R. et al. Ethnicity-specific BMI cutoffs for obesity based on type 2 diabetes risk in England: a population-based cohort study. *Lancet Diabetes Endocrinol.* **9**, 419–426 (2021).
51. Hemani, G., Bowden, J. & Davey Smith, G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum. Mol. Genet.* **27**, R195–R208 (2018).

52. Pott, J. et al. Mendelian Randomization with longitudinal exposure data: simulation study and real data application. Preprint at *medRxiv* https://doi.org/10.1101/2025.04.25.25326352 (2025).
53. Pilling, L. C. GEMINI-multimorbidity/Mounier2025_BMI_attenuation: Mounier (2025) BMI attenuation v1.0. *Zenodo* https://doi.org/10.5281/zenodo.17725332 (2025).

Acknowledgements

This work was supported by the UK Medical Research Council [grant number MR/W014548/1]. This study was supported by the National Institute for Health and Care Research (NIHR) Exeter Biomedical Research Centre (BRC), the NIHR Leicester BRC, the NIHR Oxford BRC, the NIHR Peninsula Applied Research Collaboration, and the NIHR Health Tech Research Centre. J.M. is funded by an NIHR Advanced Fellowship (NIHR302270). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. We acknowledge all members of the GEMINI Consortium. This includes researchers Olivia Murrin, Carlos Gallego-Moll, Albert Roso-Llorach, Lucía A Carrasco-Ribelles, Chris Fox, Louise M Allan, Xiaoran Liang, Ruby M Woodward, Deniz Türkmen, Kate Boddy, Jose M Valderas, Sara M Khalid, Sally E Lamb, David Melzer, and Concepción Violán, in addition to patient representatives Mary Mancini and Leon Farmer. The whole GEMINI team was central to developing the LTC definitions, observational associations in population representative primary care cohorts, genetic associations between LTC pairs, and developing and refining the research question(s) and results. This research has been conducted using the UK Biobank Resource under Application Number 14631. We want to acknowledge the participants and investigators of the FinnGen study. We also thank the authors and study participants for the published GWAS consortium meta-analyses utilised in this report (see ref. 19 and the project GitHub pages [https://github.com/GEMINI-multimorbidity] for details and citation). The authors would like to acknowledge the use of the University of Exeter High-Performance Computing (HPC) facility in carrying out this work. For the purpose of open access, the author has applied a 'Creative Commons Attribution (CC BY)' licence to any Author Accepted Manuscript version arising from this submission.

Author contributions

Conceptualisation: N.M., L.P., F.D., T.F., J.B. Data curation: N.M., B.V., L.P. Formal analysis: N.M., B.V., L.P. Funding acquisition: J.M., J.D., F.D., L.P., J.B., T.F. Interpretation: N.M., B.V., J.M., J.D., F.D., L.P., T.F., J.B. Investigation: N.M., J.B. Resources: N.M., B.V., L.P. Software: N.M. Supervision: J.B., T.F., L.P., F.D. Writing – Original Draft: N.M., L.P., T.F., J.B. Review & editing: N.M., B.V., J.M., J.D., F.D., L.P., T.F., J.B.

Competing interests

J.B. is a part-time employee of Novo Nordisk Research Centre Oxford, limited. T.F. has consulted for several pharmaceutical companies. All other authors have no disclosures to declare.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43856-025-01347-y.

Correspondence and requests for materials should be addressed to Jack Bowden.

Peer review information *Communications Medicine* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

On behalf of the GEMINI Consortium

Ninon Mounier[1], Bethany Voller[1], Jane AH Masoli[1,2], João Delgado[1], Frank Dudbridge[3], Luke C. Pilling[1], Timothy M. Frayling[1,4], Jack Bowden[1,5] ✉, Olivia Murrin[1], Linus Tata[6], Carlos Gallego-Moll[7,8], Albert Roso-Llorach[7,8], Lucía A. Carrasco-Ribelles[7,8,9], Chris Fox[9], Louise M. Allan[3], Frank Dudbridge[3], Ruby M. Woodward[3], Xiaoran Liang[1], Jose M. Valderas[9,10], Sara M. Khalid[11], Sally E. Lamb[9], Mary Mancini[12], Leon Farmer[12], Kate Boddy[9], David Melzer[1] & Concepción Violán[6,7,13,14,15]

[6]Unitat de Suport a la Recerca Metropolitana Nord, Fundacio Institut Universitari per a la recerca a l'Atencio Primaria de Salut Jordi Gol I Gurina (IDIAPJGol), Barcelona, Spain. [7]Grup de REcerca en Impacte de les Malalties Cròniques i les seves Trajectòries (GRIMTra) (2021 SGR 01537), Institut Universitari d'Investigació en Atenció Primaria Jordi Gol (IDIAPJGol), Barcelona, Spain. [8]Red de Investigación en Cronicidad, Atención Primaria y Prevención y Promoción de la Salut (RICAPPS), Instituto de Salud Carlos III (ISCIII), Madrid, Spain. [9]Department of Health and Community Sciences, Faculty of Health and Life Sciences, University of Exeter, Exeter, UK. [10]Department of Family Medicine, National University Health System, Singapore, Singapore. [11]Centre for Statistics in Medicine, Nuffield of Orthopaedics Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK. [12]Public and Patient involvement representative, Oxford, UK. [13]Unitat de Suport a la Recerca Metropolitana Nord, Institut Universitari d'Investigació en Atenció Primaria Jordi Gol (IDIAP Jordi Gol), Mataró, Spain. [14]Germans Trias i Pujol Research Institute (IGTP), Street, Badalona, Spain. [15]Department of Medicine, Universitat Autònoma de Barcelona, Cerdanyola de Vallès, Spain.