



UNIL | Université de Lausanne

Unicentre

CH-1015 Lausanne

<http://serval.unil.ch>

Year : 2021

Characterizing obesity and its consequences on health

Sulc Jonathan

Sulc Jonathan, 2021, Characterizing obesity and its consequences on health

Originally published at : Thesis, University of Lausanne

Posted at the University of Lausanne Open Archive <http://serval.unil.ch>

Document URN : urn:nbn:ch:serval-BIB_0FA20CE347758

Droits d'auteur

L'Université de Lausanne attire expressément l'attention des utilisateurs sur le fait que tous les documents publiés dans l'Archive SERVAL sont protégés par le droit d'auteur, conformément à la loi fédérale sur le droit d'auteur et les droits voisins (LDA). A ce titre, il est indispensable d'obtenir le consentement préalable de l'auteur et/ou de l'éditeur avant toute utilisation d'une oeuvre ou d'une partie d'une oeuvre ne relevant pas d'une utilisation à des fins personnelles au sens de la LDA (art. 19, al. 1 lettre a). A défaut, tout contrevenant s'expose aux sanctions prévues par cette loi. Nous déclinons toute responsabilité en la matière.

Copyright

The University of Lausanne expressly draws the attention of users to the fact that all documents published in the SERVAL Archive are protected by copyright in accordance with federal law on copyright and similar rights (LDA). Accordingly it is indispensable to obtain prior consent from the author and/or publisher before any use of a work or part of a work for purposes other than personal use within the meaning of LDA (art. 19, para. 1 letter a). Failure to do so will expose offenders to the sanctions laid down by this law. We accept no liability in this respect.



UNIL | Université de Lausanne

Faculté de biologie
et de médecine

Unisanté

Département formation, recherche et innovation

Characterizing obesity and its consequences on health

Thèse de doctorat ès sciences de la vie (PhD)

présentée à la

Faculté de biologie et de médecine
de l'Université de Lausanne

par

Jonathan SULC

Master de l'Université de Lausanne

Jury

Prof. Matthias Stuber, Président

Prof. Zoltán Kutalik, Directeur de thèse

Prof. Bogdan Draganski, Co-directeur de thèse

Prof. Sven Bergmann, Expert

Prof. Bart Deplancke, Expert

Lausanne

(2021)



UNIL | Université de Lausanne

Faculté de biologie
et de médecine

Unisanté

Département formation, recherche et innovation

Characterizing obesity and its consequences on health

Thèse de doctorat ès sciences de la vie (PhD)

présentée à la

Faculté de biologie et de médecine
de l'Université de Lausanne

par

Jonathan SULC

Master de l'Université de Lausanne

Jury

Prof. Matthias Stuber, Président
Prof. Zoltán Kutalik, Directeur de thèse
Prof. Bogdan Draganski, Co-directeur de thèse
Prof. Sven Bergmann, Expert
Prof. Bart Deplancke, Expert

Lausanne
(2021)



UNIL | Université de Lausanne

Faculté de biologie
et de médecine

Ecole Doctorale

Doctorat ès sciences de la vie

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président·e	Monsieur	Prof.	Matthias	Stuber
Directeur·trice de thèse	Monsieur	Prof.	Zoltán	Kutalik
Co-directeur·trice	Monsieur	Prof.	Bogdan	Draganski
Expert·e·s	Monsieur	Prof.	Sven	Bergmann
	Monsieur	Prof.	Bart	Deplancke

le Conseil de Faculté autorise l'impression de la thèse de

Monsieur Jonathan Sulc

Maîtrise universitaire ès Sciences en sciences moléculaires du vivant, Université de Lausanne

intitulée

**Characterizing obesity
and its consequences on health**

Lausanne, le 12 novembre 2021

pour le Doyen
de la Faculté de biologie et de médecine

Prof. Matthias STUBER
Directeur du CIBM-CHUV
RAD-CHUV-BH07
Rue du Bugnon 46
CH-1011 Lausanne
Tél. +41 21 314 75 34
matthias.stuber@chuv.ch

Prof. Matthias Stuber

JONATHAN SULC

CHARACTERIZING
OBESITY

&

ITS CONSEQUENCES
ON HEALTH

I wish to express my deepest gratitude to my supervisor, Professor Zoltán Kutalik, for his support, his guidance, and his patience throughout this entire undertaking. His scientific expertise and wisdom have fueled many discussions that taught me so much while making me realize how much I have left learn.

I would like to thank my co-supervisor, Professor Bogdan Draganski, for his valuable input in the early stages of my thesis. Although circumstances eventually led to a less direct involvement, our meetings were always a source of renewed inspiration for me.

I wish to thank the colleagues and collaborators from the Statistical Genetics Group who helped with the projects and papers presented herein, foremost among which is Anthony Sonrel, whose masters work helped to shape the early phase of the first project. It has been a pleasure to work with all of them.

I am immensely thankful to family and friends, who kept me going through the trials and tribulations of these past years. I am particularly grateful for the unconditional support from and (nearly) endless patience of my partner, Aline, as well as the encouragement of my parents, Mary and Milan.

Abstract

The definition of obesity as a mere ‘excess of fat mass’ belies its complexity. Its quantification alone is far from trivial and it is now clear that that’s not enough. The heterogeneity present in the ‘obese’ phenotype is not simply a cosmetic variation but an accumulation of profound biological differences caused by the interplay of genetic and environmental factors with substantial implications for a person’s health and well-being. Recognizing this complexity has led to the description of many aspects contributing to obesity: establishing their nature, discovering the underlying causes, and understanding their consequences.

Although the study of a complex system through the examination of its parts in isolation is a typical scientific approach, this fails to account for their interdependency and makes it difficult to distinguish the contributions that are specific to each rather than shared among them. By applying principal component analysis, we identified four orthogonal axes of genetically-defined variation in body shape. These independent components arise from distinct genetic bases expressed in different tissues and impact health in separate ways. In particular, this revealed a body mass-neutral component affecting body fat distribution which increased the risk of lipotoxicity-related diseases.

The deleterious consequences of obesity are qualitatively well established but their quantification is less trivial. Some have been found to be specific to overweight individuals while others seem to affect everyone. Stratification can provide some indications as to such differences between groups, but this relies on the definition of an arbitrary threshold and the approximate linearity of the effect within each stratum. Instead, we’ve developed PolyMR, a Mendelian randomization-based method for the inference of non-linear effects through polynomial approximation. We then showed that most effects of obesity-related anthropometric traits are strongly non-linear.

Finally, I showed that failing to account for interaction effects with, e.g., sex attenuates and biases the estimated effects. I propose an extension of PolyMR which enables the simultaneous modeling of interaction effects and demonstrate its effectiveness.

Résumé

La simplicité de la définition de l'obésité comme «excès de matière grasse» ne fait pas justice à la complexité qu'elle cache. Sa mesure n'est déjà pas aisée et il est désormais clair que ça n'est pas suffisant. L'hétérogénéité présente dans le phenotype communément appelé «obésité» n'est pas une simple particularité cosmétique, mais l'accumulation d'importantes différences biologiques, résultats des interactions de nombreux facteurs génétiques et environnementaux avec des conséquences considérables pour la santé et le bien-être de la personne. L'acceptation de cette complexité a entraîné l'étude de nombreux aspects de l'obésité, de la description de leur nature à la découverte de leurs causes et conséquences. Bien que l'étude des systèmes complexes par l'analyse de ses composantes individuels est une approche scientifique typique, elle ne permet pas de tenir compte de leur interdépendance et la distinction entre leurs contributions spécifiques ou partagées en est rendue difficile. En appliquant une analyse à composantes principales, nous avons identifié quatre axes orthogonaux de variations à base génétique dans la morphologie du corps. Ces composantes indépendantes sont le résultat de bases génétiques différentes exprimées dans des organes distincts et avec des conséquences spécifiques sur la santé. Nous avons en particulier identifié une composante neutre en terme de masse corporelle, mais affectant la répartition de matière grasse qui augmente les risques de maladies liées à la lipotoxicité.

Les conséquences néfastes de l'obésité sont qualitativement bien démontrées, mais leur quantification est moins aisée. Certaines sont limitées aux personnes en surpoids, alors que d'autres semblent affecter tout le monde proportionnellement. L'analyse stratifiée peut fournir des indications quant aux différences entre les groupes, mais cela nécessite la fixation d'un seuil arbitraire et l'approximation de l'effet dans les deux groupe reste linéaire. Nous avons développé PolyMR, une méthode basée sur la randomisation Mendélienne pour l'inférence d'effets non-linéaires par approximation polynomiale. Nous avons ensuite montré que la majorité des effets de mesures corporelles liées à l'obésité sont fortement non-linéaires.

Finalement, j'ai montré que de manquer de considérer les interactions des effets avec, par exemple, le sexe peut introduire un biais dans les estimations et une atténuation du signal. Je propose également une extension à PolyMR qui permette d'inclure les interactions dans la modélisation et en démontre l'efficacité.

Contents

1	<i>Introduction</i>	9
1.1	<i>Obesity: Moving beyond body mass index</i>	10
1.2	<i>Genetics of obesity and body shape</i>	16
1.3	<i>Integrative approaches to characterizing body shape</i>	18
1.4	<i>Causal inference</i>	22
2	<i>Obesity and its consequences</i>	29
2.1	<i>Heterogeneity, genetics, and consequences: A review</i>	29
2.2	<i>Through the lens of integrative analysis</i>	48
2.2.1	<i>Choice of methods</i>	62
2.2.2	<i>Choice of variance-covariance matrix</i>	65
2.2.3	<i>Obesity and the brain</i>	66
2.2.4	<i>Contributions</i>	67
3	<i>Non-linear causal inference</i>	69
3.1	<i>Polynomial MR, simulations and application</i>	69
3.1.1	<i>Methods</i>	70
3.1.2	<i>Results</i>	75
3.1.3	<i>Discussion</i>	80
3.1.4	<i>Supplementary Figures</i>	86
3.2	<i>Further applications and extensions</i>	86
3.2.1	<i>Non-linear effects of composite traits</i>	86
3.2.2	<i>Non-linear increase in disease risk</i>	88
3.2.3	<i>Extending PolyMR to model interactions</i>	89
4	<i>Minor contributions</i>	99
4.1	<i>Quantifying gene by environment interactions</i>	99

5	<i>Discussion</i>	101	
5.0.1	<i>Non-linear effects of composite traits</i>	101	
5.0.2	<i>Obesity and disease risk</i>	102	
5.1	<i>Non-linear effects and interactions</i>	103	
5.2	<i>Limitations of polynomial regression</i>	108	
5.3	<i>Towards a more comprehensive view of obesity</i>	110	

List of Figures

- 1.1 **The main types of adiposity**, based on location, are subcutaneous adipose tissue (SAT), which is stored beneath the skin over the entire body, and visceral adipose tissue (VAT), which is deposited in and around organs in the abdomen. 15
- 1.2 **Factor Analysis (FA) and principal component analysis (PCA)** offer slightly different strategies for the decomposition of variance. (A) FA models the factors as shared components giving rise to the observed variables (X1 and X2) and allowing for individual variance, whereas (B) PCA decomposes all variance into principal components (PCs). This difference is illustrated in the equations, where (C) FA typically defines the observed variables as a function of factors, while (D) PCA defines PCs as a function of the observed variables. Illustrated graphically, (E) the residual variance of FA is attributed to the individual error terms, while (F) in PCA this yields a second PC. 20
- 1.3 **Mendelian randomization relies on assumptions** enabling the accurate estimation of effects which may have been difficult to test in practice, such as the effect of HDL cholesterol on cardiovascular disease. Solid lines represent true underlying causal effects, dashed lines show associations. SNP: single nucleotide polymorphism; LDL-c: low-density lipoprotein cholesterol; HDL-c: high-density lipoprotein cholesterol; CVD: cardiovascular disease. 24
- 2.1 **The top 4 body principal components (PCs)** explain more than 99% of the total variance. The silhouettes illustrate the expected phenotypes at the lower (left) and upper (right) extremes of the PC scale. The values in gray indicate the variance explained by the PC. 48
- 3.1 **PolyMR is able to recover the shape of the causal function.** The true causal function is shown in green (solid line). The observed association model is shown in orange (short-dashed) while that obtained using PolyMR is shown in purple (long-dashed). The hulls around the model curves show the 95% coverage hull across 1000 simulations. Shown here are (a) high polygenicity (10'000 causal SNPs accounting for the heritability of 0.3); and (b) a sigmoid causal effect ($f_{\alpha}(X) = 0.1 \cdot \frac{1}{1+e^{-2 \cdot X}}$). The Y-axis shows the expected association with/effect of the exposure on the outcome, relative to the outcome level at the mean population exposure. 76

- 3.2 **Bias as a function of exposure across settings.** In settings with a polynomial causal function $f_{\alpha}(\cdot)$, slight bias from non-genetic confounding was induced under certain combinations of high polygenicity, high heritability, or strong confounding. The bias found in non-polynomial settings was expected due to the polynomial approximation approach. 76
- 3.3 **PolyMR provided greater accuracy in the estimation of causal functions.** The root mean square errors (RMSEs) are shown for both PolyMR and LACE. Each point is the mean RMSE for a given setting with the error bars showing the 95% confidence interval (CI) of the true mean across simulations. Settings were split into polynomial and non-polynomial causal functions. Arrows in the polynomial plot indicate RMSEs which exceed the bounds of the plot. 77
- 3.4 **Most tested causal effects have strong non-linear components in the UK Biobank.** The red points show the mean outcome plotted against the median exposure for each of 100 bins, split by covariate-adjusted exposure level. The red curve (solid) is the multivariable regression model whereas the teal one (dashed) corresponds to the estimated causal function obtained using PolyMR. The hulls around both curves correspond to the 95% confidence interval. 79
- 3.5 **Non-linear causal effects may be observed in several situations.** (A) It may occur biologically in what could be considered *true non-linearity*, (B) as an interaction between the exposure and a correlated variable, or (C) due to the scale on which the exposure is measured. The asterisks (*) and/or wavy arrows mark the source of non-linearity in each case. 83
- 3.6 **The confidence intervals (CIs) of PolyMR are correctly calibrated for the returned function order.** The size of the theoretical 95% CI was calculated for each simulation at each percentile of the exposure, yielding a distribution of CI sizes. The empirical CIs come from the distribution of the estimated causal function. Both panels show the CIs for the weak quadratic effect setting, with either (A) all results or (B) only those results where the correct order of the causal function was determined (920 out of 1000 simulations). 86
- 3.7 **PCs 1–3 mostly increased SBP, though with qualitatively different effects (teal, dashed).** The values on the y-axis are given relative to the expected outcome for an exposure level equal to that of the population mean. The points represent the mean difference in observed outcome for each centile of the population ranked by exposure level, with the red line (solid) showing the observed association. The values in parentheses indicate the variance (in body shape) explained by each PC. 87
- 3.8 **PCs 1–3 had distinct effects on total cholesterol levels in blood (blue, dashed), despite showing similar inverted-U shaped observational associations (red, solid).** The values on the y-axis are given relative to the expected outcome for an exposure level equal to that of the population mean. The points represent the mean difference in observed outcome for each centile of the population ranked by exposure level. The values in parentheses indicate the variance (in body shape) explained by each PC. 87

- 3.9 **Example interaction scenario** in (linear) summary statistics–Mendelian randomization, where the effect of exposure X is altered by the presence of covariate C. This covariate could also be a confounder, i.e. affecting both exposure and outcome, in addition to interacting with the exposure. Solid lines represent causal effects, dashes show associations. 90
- 3.10 **Adjusting for sex can have different consequences** based on the exposure–outcome relationship. The sex-specific and –combined effects of body fat percentage (BFP, left) and waist-to-hip ratio (WHR, right) on LDL cholesterol illustrate the two extremes. The horizontal shift between the sex-specific curves (green and blue) for BFP indicates an interaction with sex, and, in this case, adjusting produces a relatively similar sex-combined curve (red). The absence of such a *horizontal* shift for WHR suggests no interaction with sex, in which case adjusting for sex produces a biased (flattened) sex-combined causal function. Note that the vertical (mis)alignment is simply due to the y-axis being defined relative to the expected effect at the (respective) population mean exposure values. 91
- 3.11 **Modeling the interaction directly** recovers the respective sex-specific curves (solid lines), which are highly similar to those obtained from stratification (dashed). The hulls shown are those of the interaction modeling. 94
- 3.12 **The effects of body fat percentage (BFP) on (A) glucose levels in blood and (B) risk of depression show very different sex-specificity.** The increase in blood glucose from higher BFP in men (teal, dashed) is much greater than that inferred in women (red, solid), whereas the effects on depression were identical across the respective sex-specific exposure distributions (i.e. no significant sex-interaction coefficient). The estimated effect (y-axis) is here shown relative to that estimated for the sex-specific population mean exposure. 95
- 3.13 **Some inferred causal functions showed considerable irregularity**, with many inflection points, as shown here for the effects of (A) BMI on the risk of depression and (B) weight on glucose. The inferred (sex-interacting) effect is plotted (y-axis) as a function of the absolute exposure value for each sex, compared to the expected effect at the sex-specific population mean exposure. 95
- 3.14 **PCs 1–3 show similar effects on glucose levels in blood**, with no effect in normal-weight individuals but monotonic increasing effects for larger values. 98
- 5.1 **Graph of the constituents of BMI and their effects on depression** (A) as intended when using BMI as exposure for MR and (B) the more likely underlying model of causality. Solid arrows represent causal effects, while dashed lines illustrate associations. 107

1

Introduction

The following document details the work I accomplished during my PhD under the supervision of Profs. Zoltán Kutalik and Bogdan Draganski, summarizes the output thereof, and discusses its implications within the wider context of obesity, genetics, and causal inference. The first part of this thesis focused directly on the variation in body shape which leads to such heterogeneity within the ‘obese’ phenotype: from understanding the structure of this variability, to identifying the underlying genetic determinants, biological mechanisms, and tissues involved, to pinpointing the specific consequences of its various aspects through the use of *Mendelian randomization* (MR). The second part focused more on the limitations of MR, specifically in the presence of non-linear causal effects, developing an extension to enable the modeling of such effects, and revealing the high prevalence of non-linearity in the consequences of obesity.

This introduction encapsulates our current understanding of obesity¹ and the heterogeneity therein², methods to explore complex traits, and methods to infer causality between phenotypes or events. The second chapter summarizes my contributions to the aforementioned projects, as well as some aspects of these which were not published. The final chapter discusses the relevance of these results taken together, existing challenges in these topics and possible avenues of future research.

¹ Throughout the manuscript I use the term *obesity* in the broad sense to mean excess adiposity.

² This is partly based on the review co-first authored with Dr. Thomas W. Winkler in 2020 and published in *Current Diabetes Reports* [1].

1.1 Obesity: Moving beyond body mass index

OBESITY IS DEFINED BY THE EXCESSIVE ACCUMULATION OF FAT, with the potential to cause numerous deleterious effects on health, increasing the risk of many noncommunicable diseases such as type 2 diabetes (T2D) [e.g. 2] and cardiovascular disease (CVD) [reviewed in 3], as well as increasing the mortality and morbidity of infectious diseases such as COVID-19 [4–6]. Combined with the high prevalence worldwide³, obesity is thought to cause approximately 5 million deaths (8.9% PAF⁴) and account for 160 million disability-adjusted life years⁵ (6.3% PAF) worldwide (estimates for 2019) [9]. This imposes a considerable burden on public health and incurs a massive economic cost as well. For example, the estimates for the health expenditures attributable to obesity in the US in 2005 range from \$86 to \$210 billion, representing 9.1–20.6% of the total [10, 11].

Although obesity is *defined* by the accumulation of fat mass, it is usually quantified using body mass index (BMI)⁶, which measures excess weight as a surrogate for adiposity. According to WHO guidelines [7], BMI above 25 qualifies an individual as overweight while above 30 would qualify one as obese, though additional categories [12] and population-specific cutoffs have been proposed [13]. BMI provides a reasonably accurate predictor of many of the adverse health effects of obesity at the population level [14, 15], however its use in the diagnosis of obesity and the prediction of disease risk implicitly relies on certain assumptions about (1) *the relationship between height and normal body weight*, (2) *body composition*, and (3) *the homogeneity of the effects of adiposity*⁷. These simplifications are not necessarily inappropriate, but they do merit careful consideration rather than implicit acceptance.

HEIGHT AND WEIGHT are, of course, not inde-

³ Prevalence of obesity vastly increased over the last decades, reaching an estimated 39% of adults overweight and 13% obese in 2016 [7]

⁴ The population attributable fraction, or PAF, is the proportion of the total (e.g. deaths) which would be avoided in the absence of the exposure (e.g. obesity) [8].

⁵ Disability-adjusted life years (DALYs) account for the overall burden of diseases as the combination of years lost due to ill health, disability, or early death.

⁶ Calculated as weight (kg) / height² (m²).

⁷ I.e. assuming that all adipose tissue has the same effect, regardless of location or type.

pendent and one's weight is expected to increase with height. BMI assumes that body weight should (all other things being equal) be proportional to the height squared, which serves as the very formula that defines BMI. Mathematically, however, the volume (and therefore the weight) of an object increases relative to the cube of the linear scaling factor, i.e. multiplying all dimensions by 2 would increase the weight by $2^3 = 8$ -fold. Of course, people do not simply scale equally in all dimensions, however the exponent 2 is generally considered to under-correct for height⁸, leading to underestimation of excess weight in shorter people and overestimation in taller people [18, 19]. Alternative formulas have been proposed, such as using 2.5 as exponent for height instead of 2 (i.e. $weight(kg)/height(m)^{2.5}$) which seems to better approximate excess weight across a wider range of individuals [20], and could address this limitation to some extent (although these have never been widely adopted). Even with such improvements in the modeling of *excess weight* as a function of height, the use of BMI to quantify obesity relies on the assumption that this reflects a disproportionate increase in *adiposity* rather than increases in other tissues, such as muscle mass. The validity of this assumption depends on many factors which may contribute to large variations in *body composition* across individuals.

BODY COMPOSITION broadly refers to the proportions of different tissue types which compose the human body and varies from person to person. This can be considered at any level of detail but for our purposes we will consider two categories of tissue: *fat mass* and everything else, termed *lean mass*.⁹ Linking these two concepts together is *body fat percentage* (BFP) calculated, exactly as its name suggests, by dividing the amount of fat mass by the total. Given that fat mass is ultimately most directly responsible for adverse effects of obesity on health,

⁸ Interestingly, this inaccuracy of BMI as a predictor of excess weight is not a design flaw, rather it was never intended to serve this purpose. Adolphe Quetelet set forth the formula for BMI, then known as Quetelet's index, in his *description* of the relationship between height and weight in the average man (of the mid-nineteenth century) but never postulated that this reflected body composition or adiposity in any way [16–18].

⁹ Other tissue types typically considered include bone, water, and muscle mass.

its quantification within the body is a critical step towards predicting the risks imposed by an individual's physical condition, however no improvements in the formula for BMI can address the fact that BFP cannot be measured from height and weight alone¹⁰. BMI-based diagnosis of obesity therefore relies on an assumption of 'average' body composition, from which fat mass can be estimated, and any deviation from the assumed body composition will lead to bias and loss of accuracy. The question is: how reasonable is this assumption?

There are of course extremes where large deviations from the population average may provide incorrect or misleading indications as to a person's health risks, particularly in athletes and bodybuilders, who have proportionally greater muscle mass. For example, in his prime Arnold Schwarzenegger's BMI was above 30 [23], which would have qualified him as 'obese' despite his health risks being quite obviously different from those of most obese people. This has led to adapted recommendations, such as those from the Centers for Disease Control and Prevention (CDC) that athletes consult with a "trained healthcare provider" rather than follow standard BMI-based guidelines [24]. While this is a widely acknowledged limitation of the effectiveness of BMI for *specific* individuals, several studies have shown that BMI is insufficiently sensitive to assess adiposity at the individual level even in the general population [25–27]. This should hardly be surprising given that many factors affect body composition independent of height and weight¹¹. This has led to the development of alternate BMI cutoffs for the diagnosis of obesity, such as the population-/ethnicity-specific threshold applied in China (28) [35] or Japan (25) [36]. These alternate cutoffs remain a stopgap measure which improves the performance of BMI in a specific population but lacks flexibility and ultimately only focuses on adjusting for a single factor (here ethnicity). Ultimately, reliable assessment of fat mass (or BFP) cannot depend on BMI.

¹⁰ Body composition-sensitive metrics have been proposed, such as the fat/fat-free mass index (FMI/FFMI) which use the formula for BMI but substitute fat or fat-free mass estimates for weight [21, 22], though these also require body composition estimates and are not widely used.

¹¹ Factors affecting body composition include sex [28], age [28], ethnicity [28], diet [29, 30], lifestyle [31], genetics [32], and gut microbiota [33, 34]

Many other methods have been developed and used to measure BFP specifically or body composition in general, each with their own challenges and limitations, but the current leading methods in terms of accuracy are dual-energy X-ray absorptiometry (DXA), magnetic resonance imaging (MRI), and computerized tomography (CT). These techniques enable imaging of the entire body, from which tissue types can be inferred and body composition estimated. Although these methods are largely considered gold standards in body composition analysis, the values returned will differ slightly depending on the method of choice (e.g. DXA tends to measure ~5kg less fat mass than the other two) [37]. This can be a slight limitation for the *exact* measurement of these quantities, however in the context of obesity this is not necessarily required: it is enough that the fat/lean mass estimates with a given method enable a meaningful comparison across individuals, i.e. that a difference in reported values reflects a true difference in mass in the appropriate direction¹² and this has been shown to be the case [37, 38]. This enables the quantification of adiposity/obesity, allowing the comparison of individuals and the prediction of associated risk. Despite these clear advantages, these methods are rarely used in clinical practice due to onerous requirements for time, expensive machinery, and specialized personnel. Body composition is therefore usually estimated using another method: bio-electrical impedance analysis (also termed bioimpedance or BIA).

BIA exploits the fundamental differences in tissue properties between fat and lean mass, specifically its *electrical resistivity*. The impedance (resistance) of an individual's body when exposed to a weak electric current (1–10 μ A) depends on several factors, among which is body composition: lean mass contains more water, leading to greater electrical conductivity, i.e. reduced impedance/resistance. Combined with other factors such as height, weight,

¹² Mathematically, this implies a monotonic function, which has the property of preserving the rank of (here) individuals.

age, and sex, the measured impedance can be used to estimate lean mass, whereby fat mass and BFP can be deduced. Its efficiency, low cost, and ease of use have made BIA one of the most widespread methods for the estimation of body composition and is now included in many standard scales destined for private use. One of the main drawbacks of BIA is its sensitivity to many factors, such as dehydration, prior moderate intensity physical activity, or extreme BMI values (below 18 kg/m² or above 40 kg/m²) [39–41], which may bias results. Although some studies have reported high accuracy for BIA-derived estimates, provided certain guidelines are strictly adhered to [42, 43], the accuracy in general is modest but useful [41, 44].

That the risk of adverse health effects increases with the amount of adipose tissue is well established and greater accuracy in its quantification can only aid in the diagnosis of obesity and the prediction of disease risk. However, there remains considerable variation in the metabolic response across individuals that cannot be explained by body composition alone. There are many factors which are known or suspected to contribute to these health discrepancies, such as differences in lifestyle and diet, although these mostly act directly upon health or through increased obesity and are not our focus here.¹³ Of greater relevance to us is the fact that body fat is more than simply a number, that individuals with identical BFP may have very different *body fat distribution*, and that the type of adiposity may change how it affects health.

THE INFLUENCE OF BODY FAT DISTRIBUTION ON metabolic health is not yet fully understood, but it is now widely accepted to contribute significantly to health disparities between individuals with similar body composition [15, 45–47]. One of the foremost hypotheses in the field to explain the causal link between body fat distribution and metabolic health is that of *adipose tissue expandability* [reviewed in 48]. This postulates that excess

¹³ It is likely that certain factors interact with obesity to determine health outcomes, a topic discussed in [Sec. 5.1](#).

calories are preferentially stored in *subcutaneous adipose tissue* (SAT, Fig. 1.1) with minimal impact on metabolism or health and that this tissue is capable of expanding to accommodate additional fat. However, this tissue's capacity for expansion is finite and when exceeded the surplus calories are liable to be deposited ectopically in and around organs as *visceral adipose tissue* (VAT) or in skeletal muscle. These ectopic fat deposits are thought to contribute much more directly to the averse health effects of lipotoxicity [45–47] through mechanisms such as low-grade inflammation [49, 50]. Certain research results have further hinted at possible protective effects of SAT [51, 52], proposed to act through pathways such as the segregation of harmful compounds or the release of protective adipokines, though these have not yet been fully demonstrated.

The predominantly abdominal location of unhealthy fat deposits (VAT) has shifted the attention to *abdominal obesity* and metrics to quantify it. Waist circumference (WC) is an obvious choice for this and has been (and, to some extent, still is) used for this purpose. It is a useful measure *at the individual level*, since a *change* in WC is likely to reflect an increase or decrease in abdominal fat, making it useful to gage the evolution of that person's risk of disease. On the other hand, *at the population level* it depends on many other factors such as bone morphology or height which make it less reliable as an indicator for inter-individual comparison of abdominal obesity. Waist-to-hip ratio (WHR)¹⁴ is another metric for abdominal adiposity more widely used at the population level, enabling useful comparisons between individuals and the definition of practical guidelines as to desirable values. Now, as might be expected from the conceptual proximity of *abdominal* and *overall* obesity, WHR tends to increase with BMI¹⁵ but, crucially, it still provides additional information about disease risk independent of BMI [15]. WHR-centered analyses of abdominal obesity have already provided meaningful

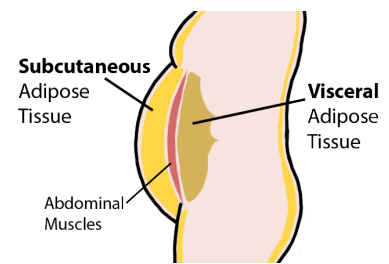


Figure 1.1: The main types of adiposity, based on location, are subcutaneous adipose tissue (SAT), which is stored beneath the skin over the entire body, and visceral adipose tissue (VAT), which is deposited in and around organs in the abdomen.

¹⁴ Calculated as the unit-free ratio of the circumference of the waist to that of the hips.

¹⁵ The correlation coefficient between WHR and BMI is around $r \approx 0.5$.

insight into some of the biological determinants of the location of fat deposits and the mechanisms through which they lead to disease [52], however such metrics are not sufficient to provide a full picture of obesity-related health consequences.

The previously mentioned methods for the high-definition measurement of body composition, namely DXA, MRI, and CT, intrinsically provide much-needed information about the location of adipose tissue in addition to the quantity, and even BIA-derived measurements can be used to approximate body composition in specific parts of the body (e.g. torso or left leg) rather than overall. The gain in accuracy from these methods comes at the cost of sample size and the availability of these measurements has so far been quite limited.¹⁶ These methods have nevertheless been instrumental in the formulation of the adipose tissue expandability hypothesis and such metrics are beginning to provide the means to explore the determinants, particularly genetic, which give rise to such inter-individual variability in body fat distribution.

¹⁶ Although large cohorts such as the UK Biobank are beginning to accumulate dozens of thousands of samples, increasing power beyond what has been possible up until now.

1.2 Genetics of obesity and body shape

In the field of genetics, traits are frequently discussed in terms of *complexity*, with *complex traits* being affected by numerous weak genetic (and possibly environmental) factors, as opposed to *Mendelian traits*, which depend on few genetic variants. There are examples of common Mendelian traits, such as blood type or lactose intolerance, but those studied are generally of public interest for their pathogenicity and are typically caused by rare, *high penetrance* variants.¹⁷ Their study usually involves pedigree analysis in affected families. On the other hand, identifying genetic variants associated with complex traits generally requires the regression¹⁸ of the phenotype of interest onto each individual single nucleotide polymorphism (SNP) across the genome, a process known as *genome-wide association*

¹⁷ Penetrance refers to the likelihood of a phenotype manifesting itself due to the presence of a genetic variant. High penetrance indicates that the presence of the allele is usually sufficient for the effect to occur.

¹⁸ Depending on the phenotype studied, this will usually be *logistic regression* for binary traits such as diseases or *linear regression* for continuous traits.

analysis (GWAS¹⁹) [reviewed in 53].

GWASs come with many challenges, the foremost of which arises from the typically weak genetic associations which characterize complex traits, combined with the large number of tests involved in genome-wide screening, quickly requiring very large sample sizes to detect ever-smaller effect sizes.²⁰ Recent years have seen the creation of multiple large-scale biobanks with the genetic and phenotypic information of hundreds of thousands of individuals [e.g. the UK Biobank, 54], which has enabled the discovery of thousands of genomic loci involved in many traits and diseases [55]. Sample size remains a limitation for the discovery of SNPs with low explained variance²¹ and some traits which are difficult to study using population-based biobanks, such as those with low heritability (e.g. lifespan [56]), where measurements are costly (e.g. body fat distribution) or invasive (e.g. gene expression in brain tissue), or are otherwise difficult to assess/estimate (e.g. social anxiety). Samples cannot always be combined at the individual-level (i.e. simply pooling the data from multiple cohorts) due to privacy concerns, but summary statistics (namely SNP-trait association strength and standard error) across multiple studies can be combined through meta-analysis.

Within the context of obesity, such studies have uncovered hundreds of SNPs associated with BMI and WHR, as well as WHR adjusted for BMI (WHRadjBMI) [57, 58], most of which individually explained little variance due to small effects or the rarity of the effect allele, jointly explaining 3.9–6.0% of the total trait variance. The genomic loci, and indeed the biological pathways and tissues involved, are largely distinct for BMI and WHRadjBMI, suggesting separate mechanisms for the accumulation of body mass and the distribution of adiposity (although larger sample sizes have found overlap in SNPs with smaller effect sizes [57, 58]). The modulation of BMI seems to be largely de-

¹⁹ GWAS is the common abbreviation (instead of GWAA) and the term I will use throughout the manuscript. The ‘S’ stands for *study*, though the term is commonly used to indicate the analysis itself.

²⁰ Note: Although I mention ‘effect’ sizes throughout this document, this does not imply a causal effect of, e.g., a GWAS hit, rather to the size of the effect which is tagged by it.

²¹ *Explained variance* is the phenotypic variance in the population which can be explained by a given SNP. Low explained variance may be due to small effect sizes or low minor allele frequency.

pendent on the central nervous system through mechanisms involving appetite regulation and energy balance [59, 60], while that of WHRadjBMI appears to involve the digestive system and adipose tissue directly through adipogenesis and insulin signaling [52, 61]. Whereas the genetic basis for BMI appears to be largely shared between sexes, many SNPs show sexually dimorphic effects on WHRadjBMI with a majority of them being stronger in women [58]. Several SNPs have been found to be suggestive of non-abdominal adiposity in women, specifically increasing hip circumference, whereas in men they increased both hip and waist circumference [1]. Their effects in women are suggestive of an increased SAT-to-VAT ratio (as has been shown for similar SNPs [52]), though it remains unclear whether this shift is sex-specific or simply detectable in women due to a greater propensity for SAT deposition on the hips.

With the multiplicity of possible metrics available to describe human physiology, from the standards of BMI and WHR to high precision DXA measurements, it follows logically that many of them show high correlation. Which makes sense: for example, any shift in body fat distribution will be characterized by an increase in adiposity in one or more locations, mirrored by decreased adiposity elsewhere, affecting body composition and mass in all locations. This makes it difficult to study the causes or consequences of any of these traits in isolation as would be required for the typical scientific approach. However, the very interdependency which makes the study of single traits difficult can be leveraged through *multi-trait analysis*.

1.3 Integrative approaches to characterizing body shape

The combined analysis of multiple phenotypes can leverage the shared variance to maximize statistical power and/or exploit trait specificity to disentangle the etiology of, e.g., obesity consequences. Here,

I will mainly describe two types of approaches which are commonly used in exploratory analyses of related variables: *clustering* into groups based on similarity and *the analysis of covariance structure* for the creation of composite traits.²²

CLUSTERING-BASED APPROACHES broadly aim to split a set of objects²³ into meaningfully different groups, termed *clusters*, based on their similarity. The resulting clusters are comprised of objects which are more similar to one another than to objects of other clusters. Varying definitions of similarity and constraints on the formation of clusters have given rise to numerous methods for achieving them, such as *K-means*²⁴ or *hierarchical clustering*.²⁵ These methods have, for example, been used in hypothesis-driven approaches to uncover genetic variants associated with favorable metabolic profiles despite increasing adiposity [51, 62, 63].

Typical clustering-based methods assign objects to mutually exclusive clusters, which may be desirable in the extent where it provides more easily interpretable cluster membership but may be limiting as well. For example, genetic variants may reasonably be expected to contribute to multiple biological pathways with potential relevance to the variables used in the classification. Other methods have been proposed to allow for probabilistic cluster membership or potentially overlapping clusters [64, 65], which can overcome these limitations to some extent. An alternative to the classification of these objects is the analysis of their *covariance*.

THE COVARIANCE STRUCTURE OF DATA contains information on the shared associations between phenotypes, and patterns therein can provide insight into causal relationships or common causes affecting them. Two commonly used methods to disentangle these are *factor analysis* (FA) and *principal component analysis* (PCA), with slightly different assumptions but similar processes (Fig. 1.2). Both seek to define a set of composite traits, linear

²² A 'composite' trait is one which is defined as a function of several others. In this sense, BMI and WHR could be considered composite traits, however for clarity and simplicity the term 'composite trait' in this manuscript refers to those traits derived from the methods described here rather than manually defined metrics.

²³ The 'objects' to cluster can be anything for which a meaningful similarity (or conversely dissimilarity/distance) metric can be defined. In the context of obesity, this could be individuals, anthropometric traits, genetic variants, etc..

²⁴ K-means clustering stochastically creates a set number (K) of clusters by iteratively assigning each object to the (randomly initiated) nearest center until convergence.

²⁵ Hierarchical clustering iteratively groups the two nearest objects or clusters into a new cluster, providing a tree-shaped structure.

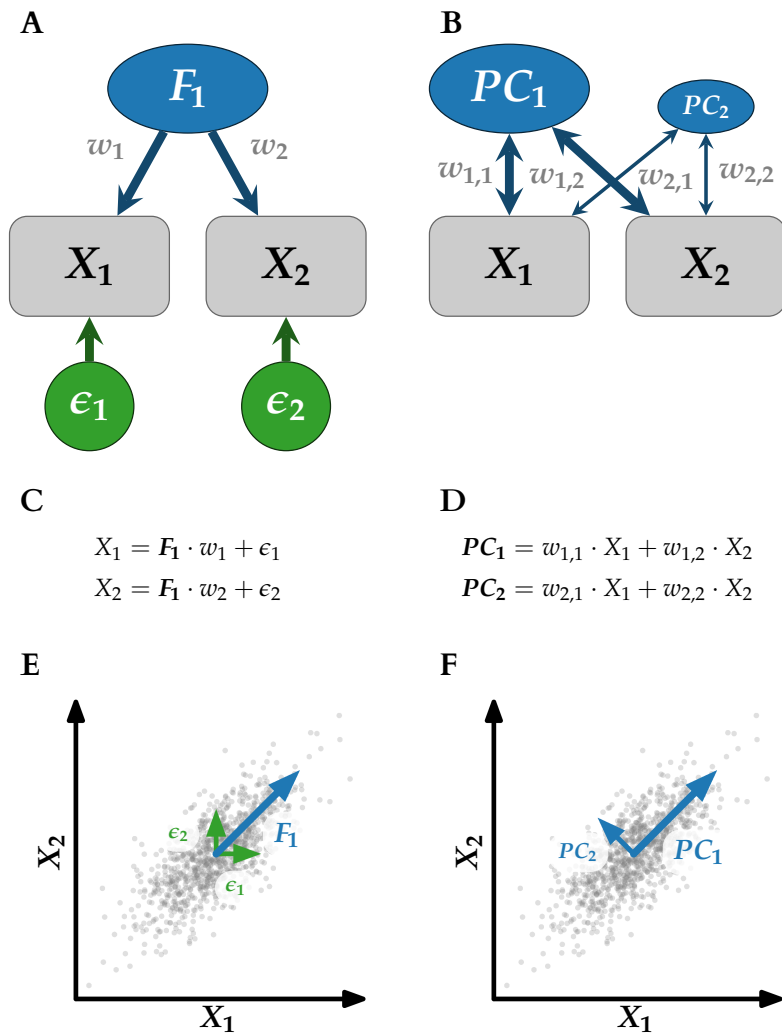


Figure 1.2: **Factor Analysis (FA) and principal component analysis (PCA)** offer slightly different strategies for the decomposition of variance. (A) FA models the factors as shared components giving rise to the observed variables (X_1 and X_2) and allowing for individual variance, whereas (B) PCA decomposes all variance into principal components (PCs). This difference is illustrated in the equations, where (C) FA typically defines the observed variables as a function of factors, while (D) PCA defines PCs as a function of the observed variables. Illustrated graphically, (E) the residual variance of FA is attributed to the individual error terms, while (F) in PCA this yields a second PC.

combinations of the original phenotypes,²⁶ which best explain most of the covariance structure of the original data. FA explicitly assumes the existence of the shared *factors* which give the method its name: latent (unobserved) variables which underlie the observed phenotypes and is the source of the correlation among them. The number of factors is usually defined beforehand based on a priori knowledge or expectations, though in exploratory FA it may be varied to compare results. These factors may affect any number of observed phenotypes and may or may not be independent from one another.²⁷

In PCA, these composite phenotypes, termed *principal components* (PCs), are sequentially chosen

²⁶ Mathematically, these linear combinations are defined by a *rotation matrix*.

²⁷ Independent factors are defined using an *orthogonal rotation matrix*, while non-independent ones use an *oblique rotation matrix*.

to maximize the variance they explain while remaining orthogonal to all other PCs. Depending on the data and the FA algorithm chosen, these methods may provide similar results. The key difference which separates them is that FA allows for part of the variance in an observed variable to remain unexplained by the factors (Fig. 1.2 A, C, E), whereas PCA will partition all observed variance into the necessary number of PCs (Fig. 1.2 B, D, F). It is, however, typical to discard the lower-variance PCs and instead use only the top X PCs to summarize the variance over all included variables, a process known as *dimensionality reduction*. The number of PCs to select is usually determined to explain a certain percentage of the total variance (e.g. select enough PCs to explain 99% of the total variance) or based on a *scree plot*, which displays the eigenvalues of the sorted PCs (which are proportional to the explained variance).

This combination of correlated traits can be leveraged to increase the statistical power to detect loci with concerted effects on multiple phenotypes, as has been shown for anthropometric traits [66], which provides some insight into their co-regulation and biological pathways or tissues involved. These associations can improve our fundamental understanding of the variation in human body shape, however they still lack a critical element to maximize their usefulness in the context of public health: *causality*. Indeed, any association between a shift in adiposity and a corresponding change in metabolic health could indicate an effect of adiposity on metabolism, or it could be the reverse, or they may simply share a common cause. Any health guidelines or drugs designed based on such associations would risk being useless or even harmful, hence the importance of *causal inference*.

1.4 Causal inference

While association analyses serve to detect events which tend to co-occur (or variables which are correlated), causal inference aims to establish whether one of these events occurs as a consequence of the other, or whether a one variable affects another. The study of causation, termed *etiology*, is fascinating in its own right, however it is of vital importance within the context of medicine and public health. Understanding the causes that lead to a given result provides possible means through which to affect the outcome, which is at the core of disease prevention and treatment. The ideal experiment to establish causality would be to observe the outcome of a treatment while knowing precisely what would have occurred in the absence of treatment.²⁸ This would enable the unambiguous deduction that any difference in outcome is due to the treatment and is, of course, impossible. Failing this, the gold standard remains *randomized control trials*.

RANDOMIZED CONTROL TRIALS (RCTs) are a type of intervention study²⁹ aiming to infer causality while minimizing the possibility of bias. Test subjects are *randomly* allocated to two or more groups, one of which is the *control* group, which receives no treatment, while the other(s) receive the treatment(s) of interest, following which the outcome is compared between groups to determine the effect(s) of the treatment. The key assumption underlying the inference of causality is that the groups are *exchangeable*,³⁰ in which case any difference in the outcome can be attributed to the effects of the treatment. RCTs are therefore designed around this assumption, from the eponymous randomization, which serves to avoid selection or allocation bias, to double blinding, which limits the impact of accessory manipulations not directly relevant to the treatment. Although RCTs remain the most reliable method of causal inference, they

²⁸ This alternate scenario is termed the *counterfactual*.

²⁹ *Intervention studies*, as the name suggests, are experiments where the researchers affect the subjects in some way, as opposed to *observational studies*, where the subjects are merely monitored.

³⁰ Exchangeability requires that all variables susceptible of confounding the association between the treatment and outcome are balanced between groups, i.e. that groups could be exchanged (prior to treatment) with no impact on the results.

are often costly and time-consuming. Worse, especially where human subjects are concerned, they may be unethical, impractical, or downright impossible. A more practical alternative has emerged in recent years, enabling causal inference based on observational data alone: *Mendelian randomization*, a method which uses genetic variants as a proxy for RCT.

MENDELIAN RANDOMIZATION (MR)³¹ exploits the naturally occurring, random segregation of genetic alleles at birth as a kind of RCT [67, 68]. The basic idea is that if a phenotype of interest (hereafter termed *exposure*) causally affects an outcome, then any genetic variant associated with a change in the exposure should exhibit a corresponding association with the outcome. Working backwards, MR infers causality by testing whether an exposure-associated variant (known as an *instrumental variable* or IV) is also significantly associated with the outcome. The validity of this inference relies on three core assumptions (Fig. 1.3):

1. **Relevance.** The IV must be associated with a change in the exposure;
2. **No pleiotropy.** The outcome must be independent of the IV, conditional on the exposure;
3. **Exclusion restriction.** The IV must be independent of any confounding of the exposure-outcome relationship.

The first is intrinsic to the definition of an IV: a variant is only useful if it can predict a change in the exposure. The second assumption excludes SNPs tagging any effect acting directly on the outcome without affecting the exposure. This mainly serves to exclude *pleiotropic*³² SNPs or those tagging an effect occurring primarily on the outcome and affecting the exposure through reverse causality. The final assumption is analogous to the exchangeability requirement for RCTs: the probability of having one or more copies of the effect allele must not be associated with any variables which might

³¹ Mendelian randomization is derived from a method originally developed in the field of econometrics called *instrumental variable analysis*.

³² Pleiotropy (or more precisely *horizontal pleiotropy*) occurs when a genetic variant has effects on multiple traits independently.

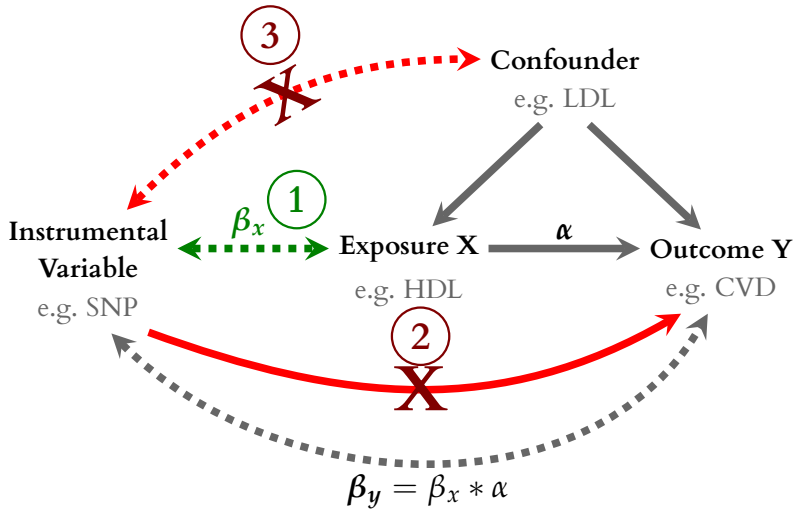


Figure 1.3: **Mendelian randomization relies on assumptions** enabling the accurate estimation of effects which may have been difficult to test in practice, such as the effect of HDL cholesterol on cardiovascular disease. Solid lines represent true underlying causal effects, dashed lines show associations. SNP: single nucleotide polymorphism; LDL-c: low-density lipoprotein cholesterol; HDL-c: high-density lipoprotein cholesterol; CVD: cardiovascular disease.

affect both exposure and outcome. Although this applies to *any* confounder, this is usually considered with respect to population stratification.³³ If these assumptions are satisfied, then the IVs can be used to estimate the causal effect of the exposure on the outcome.

The simplest approach of MR uses two-stage least squares regression on individual-level data. The value the exposure x takes on in an individual can be described as the sum of IV-tagged genetic effects and other contributions (untagged genetic and/or environmental):

$$x = G \cdot \beta_x + \epsilon_x, \quad (1.1)$$

where G is the matrix of IV genotypes, β_x is the vector of genetic effect estimates for the IVs, and ϵ_x is the error term which includes the effects of other factors. The outcome y can be expressed as the causal effect of x plus other IV- and exposure-independent effects:

$$y = x \cdot \alpha + \epsilon_y, \quad (1.2)$$

where α is the causal effect of the exposure on the outcome and ϵ_y is the error term (including other effects). β_x can then be estimated from [Equation 1.1](#) using ordinary least squares regression, with which we can predict the expected exposure level in each individual as $\hat{x} = G \cdot \widehat{\beta}_x$. Regressing

³³ Population stratification occurs due to the assortative (i.e. non-random) mating which occurs in society, causing genetic variants to become more or less frequent within certain strata of the population (due to selection or genetic drift). These variants then become non-causally associated with any variables which differ across the concerned strata, such as geographical location, socio-economic status, or chopsticks usage [69].

the outcome y onto this *predicted* exposure \hat{x} provides an accurate estimate of the causal effect which is free of the observational confounding which would bias results if the outcome were directly regressed on the observed exposure (provided the assumptions are verified, of course), i.e.

$$y = \hat{x} \cdot \alpha + \epsilon'_y \quad (1.3)$$

provides a valid estimate of the causal effect of interest α .

Most widely-used methods today are based on a slightly different approach known as two-sample MR or summary statistics MR [70]. Simply put, if the MR assumptions are respected, the association of an IV with the outcome (β_y) can be described as the combination of its association with the exposure (β_x) and the causal effects of the exposure on the outcome (α), i.e. the product of the two (Fig. 1.3). Therefore, the ratio of the effect estimate of the IV on the outcome to that on the exposure provides an estimate of the causal effect of interest. This can be done for any number of IVs, the estimates of which can then be combined through inverse-variance weighted (IVW) meta-analysis. Equivalently, the IV-outcome effects can be regressed onto the IV-exposure ones (excluding an intercept), where the resulting slope represents the strength of the causal effect. This has the distinct advantage of using summary statistics, which can easily be shared or obtained from publicly available sources, and opens up the possibility of using MR across samples, enabling the testing of effects where the exposure and outcome are not measured in the same individuals.

Many extensions have been proposed for MR (especially summary statistics MR) to address specific limitations or reduce the bias that might be induced by violated assumptions. Below are some of the more notable developments on the topic with a brief summary of each:

MR-Egger [71] is designed to account for IVs with

imbalanced pleiotropy, i.e. where they have independent effects on the exposure and outcome. This is achieved by the inclusion of an intercept in the regression of outcome effects on exposure ones.

Median-based MR [72] reduces bias introduced by a small number of invalid IVs by using a median-based combination of causal effect estimates (instead of the standard IVW meta-analysis).

Mode-based MR [73] uses a similar, majority-based approach as median-based MR, but instead assumes that valid IVs represent the largest group of similar causal effect estimates (expected to be identical in an infinite sample).

MR-PRESSO [74] addresses horizontal pleiotropy through the sequential detection and exclusion of outliers to eliminate or reduce the bias they would introduce.

MR-Clust [75] considers clusters of IVs acting through different mechanisms, which may lead to multiple causal effect estimates representing different biological pathways.

Factorial MR [76] uses genetic IVs and polygenic scores to investigate possible interactions between two IV-tagged exposures.

CAUSE [77] uses structural equation modeling to allow for (some) correlated and uncorrelated pleiotropy.

LHC-MR [78] builds upon a similar framework as that of *CAUSE* [77] (and the LCV model [79]), simultaneously estimating forward and reverse causality between exposure and outcome, as well as confounding.

Multivariate MR [80] accounts for known pleiotropy by including multiple exposures in the effect estimation.

For a comparison of the performance and robustness of some of these methods (and others), see [81]

for univariable approaches or [82] for multivariable ones.

These developments and improvements extend the possibilities of MR, however they have one limitation in common which is rarely mentioned or even considered: they assume the putative causal effect to be *linear*, i.e. that the expected change in the outcome caused by a change in the exposure does not depend on the *value* of the exposure, only the size of the change. For example, this would imply that a 1 kg/m² change in BMI would lead to the same change in blood pressure whether the initial BMI was 20 or 35. Even though this isn't biologically plausible in many cases, this assumption is implicit to any method aiming to estimate a linear causal effect.

This limitation/assumption is not without some MR-based developments aiming to address it, namely *LACE* (Localized average causal effect) [83] and *SpotIV* (Semi-parametric outcome models with possibly invalid IVs) [84]. Both of these approaches make use of the *control function* [85]: the residual variation in the exposure after conditioning on the IV(s), i.e.

$$\epsilon_x = x - G \cdot \beta_x, \quad (1.4)$$

where x is the observed exposure, G is the matrix of IV genotypes, and β_x is a vector of the IV effect estimates. The SpotIV approach is designed to account for invalid IVs and aims to estimate the specific difference in risk between two levels of the exposure, allowing for non-linearity in the causal effect. LACE broadly operates by stratifying the population based on the control function (i.e. the IV-free exposure) and estimating a linear causal effect in each bin. These estimates can then be combined using either a fractional polynomial approach or a piecewise linear one.

Both of these approaches enable the estimation of non-linear effects, but present their own limitations. The SpotIV approach is designed to account for invalid IVs, making it potentially more

robust to (minor) violations of the traditional MR assumptions, but only enables the point estimation of an expected change in risk associated with a specific change in exposure level. This may allow the interrogation of very specific hypotheses but is unsuitable to exploratory investigations of the causal relationship between exposure and outcome. It would theoretically be possible to obtain this point estimation for multiple changes across the exposure distribution, but it is unclear how reliable this would be. Furthermore, we found the current implementation [86] to scale poorly with increasing numbers of IVs or sample size,³⁴ which would likely make it unusable on UK Biobank-sized cohorts (~370'000 unrelated white European individuals) with the number of IVs typically available for complex traits (e.g. 536 independent SNPs associated with BMI [57]); repeating it multiple times to observe the shape of the causal relationship was simply not an option. LACE performs quite well across a wide range of settings and although the selection of the number of bins is somewhat arbitrary, we found the results reasonably robust in UK Biobank-size cohorts. However, its semi-parametric approach does not fully exploit the potential statistical power available. In the second part of this thesis, I focused on the development of a fully parametric approach to increase the statistical power and better approximate the shape of the underlying causal relationship.

³⁴ Initial tests suggest that the runtime increases quadratically with sample size, to an estimated ~150 minutes for 100K individuals for only 7 IVs. The implementation was also not particularly robust to any changes in parameters such as the number of IVs.

2

Obesity and its consequences

2.1 Heterogeneity, genetics, and consequences: A review

Heterogeneity in obesity: genetic basis and metabolic consequences [1]¹ summarizes our current understanding of obesity and the heterogeneity reflected in the various metrics used to quantify it, with a focus on the underlying genetics, favorable adiposity, and sexual dimorphism. New results provided by Thomas Winkler illustrated the apparent sexual dimorphism involved in the genetics of favorable adiposity, raising further questions concerning the sex-specificity of body fat distribution and its consequences on health. We examined the available evidence for the roles of VAT and SAT in the protective effects of favorable adiposity and the biological mechanisms which may be involved. Finally, we examined two avenues for future study of body fat distribution and its relevance for health, namely high-accuracy techniques (e.g. DXA and MRI) and integrative approaches (e.g. PCA), reviewing their advantages and limitations, past uses, and potential insight they might yield.

This review article was co-first authored by Thomas W. Winkler and myself. I wrote most of the first draft (based on an initial structure provided by Zoltán Kutalik) while Thomas wrote the sections on BMI/WHR genetics and sexual dimorphism. I restructured and finalized the article based on feedback and suggestions from all the other authors.

¹ This review is published in *Current Diabetes Reports* and the preprint version is included below.

Heterogeneity in obesity: genetic basis and metabolic consequences

Jonathan Sulc^{1,2,*}, Thomas Winkler^{3,*}, Iris Heid³, Zoltan Kutalik^{1,2,+}

¹Institute for Primary Care and Public Health, University of Lausanne, Switzerland

²Swiss Institute of Bioinformatics, Lausanne, Switzerland

³Department of Genetic Epidemiology, University of Regensburg, Regensburg, Germany

*co-first authors

+Corresponding author: zoltan.kutalik@unil.ch

Purpose of review. Our review provides a brief summary of the most recent advances towards the identification of the genetic basis of specific aspects of obesity and the quantification of their consequences on health. We also highlight the most promising avenues to be explored in the future.

Recent findings. While obesity has been demonstrated to lead to adverse cardio-metabolic consequences, the underlying mechanisms remain largely unknown. The elucidation of the molecular underpinnings of this relationship is hampered by the extremely heterogeneous nature of obesity as a human trait. Recent technological advances have facilitated a more in-depth characterisation of body composition at large-scale.

Summary. At the pace of current data acquisition and resolution, it is realistic to improve current obesity diagnosis and to advise individuals based on detailed body composition combined with tissue-specific molecular signatures. Individualized predictions of health implications would enable more personalised and effective public health interventions.

Introduction

Obesity has been repeatedly shown to increase the risk of many non-communicable diseases such as type 2 diabetes (T2D, reviewed in [Boles et al. 2017](#)) and cardiovascular disease (CVD, reviewed in [Ortega et al. 2016](#)). It is usually defined using body mass index (BMI, weight [kg] / height² [m²]), where BMI ≥ 30 kg/m² qualifies an individual as obese and BMI ≥ 25 kg/m² as overweight (according to WHO guidelines, <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>). This metric uses excess weight as a surrogate for adiposity, under the assumption that body composition is sufficiently similar between individuals, and provides a fairly reliable predictor of adiposity-related metabolic complications and disease risk in most cases ([Mokdad et al. 2003](#); [Pischon et al. 2008](#)). Although BMI can approximate overall excess body fat, it fails to capture other metabolically relevant aspects of adiposity, leading to considerable disparity in health outcomes between individuals with similar BMI. In particular, differences in body fat distribution, both the location and subtype of adipose tissue used to store excess calories, have been shown to affect health-related outcomes ([Pischon et al. 2008](#); [Kaess et al. 2012](#); [Rosenquist et al. 2013](#); [Abraham et al. 2015](#)).

The link between body fat distribution and metabolic consequences is still somewhat unclear, but one of the most widely supported hypotheses is that excess calories are primarily stored in subcutaneous adipose tissue (SAT) with minimal metabolic impact ([Virtue and Vidal-Puig 2010](#)). This accumulation causes the expansion of peripheral fat stores, however their capacity for expansion can be exceeded, causing a mild form of lipodystrophy where residual calories accumulate in ectopic fat deposits in and around organs (visceral adipose tissue, VAT) or in skeletal muscle. This ectopic adiposity is thought to play a more direct role in the metabolic consequences of obesity. Its predominantly abdominal location has drawn the focus to central adiposity, which can be assessed by waist circumference or waist-to-hip ratio (WHR). WHR in particular has been shown to be a strong predictor for disease risk and mortality independent of

BMI (Pischon et al. 2008). This distinction between overall and central adiposity has enabled a better characterization of obesity, and a first step towards disentangling the healthy adiposity from unhealthy.

Genetics of BMI and WHR

BMI and WHR are among the most commonly used measures of obesity and multiple genetic studies have helped to understand the biological mechanisms underlying these traits. GWAS combining ~700,000 individuals from the UK Biobank (Sudlow et al. 2015) and the Genetics of Anthropometric Traits (GIANT) consortium have identified 536 genomic loci associated with BMI and 346 with WHR adjusted for BMI (WHRadjBMI, **Table 1**, Yengo et al. 2018; Pulit et al. 2019). These loci explained a total of 6.0% and 3.9% of the phenotypic variation in the population, respectively. Approximate conditional analyses highlighted multiple independent signals at many of the identified loci (**Table 1**). For example, four independent association signals were detected at the *TBX15-WARS2* locus for WHRadjBMI, some of which were specific to men or women, suggesting a complex genetic architecture at many of the identified loci (Shungin et al. 2015).

The effect of most genetic variants on obesity traits identified by GWAS are rather moderate, resulting in an average increase of ~0.3 kg per allele (for a 1.70m tall person) for BMI-associated common variants (Locke et al. 2015). Still, the cumulative effect of common variants can improve the prediction of the likelihood of obesity (Shungin et al. 2015; Locke et al. 2015). In meta-analyses of coding variants based on ExomeChip GWAS, 14 independent low-frequency and rare associations were detected for BMI and 9 for WHRadjBMI (Turcot et al. 2018; Justice et al. 2019). Although these rare variants only explain a small additional proportion of the phenotypic variation due to their rarity, they exhibit effect sizes up to ~10 times larger than those of common variants. For example the largest effect for BMI was observed for a nonsense mutation in the *MC4R* gene that is present in 1 in 5,000 individuals and resulted in a ~7 kg increase in body weight per allele

(for a 1.70 m tall person, Turcot et al. 2018). Importantly, genes at identified loci overlapped with genes that are known for severe monogenic forms of obesity, such as *MC4R*, *BDNF*, *BBS4* and *POMC* (Locke et al. 2015). Together this indicates that many genes may affect obesity both through small effects of common variants and large effects of rare ones.

The strongest genetic factors for BMI and WHRadjBMI appear to be largely distinct (Shungin et al. 2015; Locke et al. 2015), however larger sample sizes have uncovered additional loci with smaller effect sizes, 105 of which are associated with both phenotypes (Yengo et al. 2018; Pulit et al. 2019). Concordantly, the tissues and mechanisms involved appear to be different for BMI and WHRadjBMI. Pathway and tissue specificity analyses suggest that genetic variants affecting BMI act primarily through the central nervous system, involving appetite regulation and energy balance (Locke et al. 2015; Turcot et al. 2018). In contrast, genetic effects on WHRadjBMI seem to act primarily through adipose tissue, and involve adipogenesis and insulin signaling. The genomic loci which were common to both phenotypes were enriched for variants associated with opposite effects on BMI and WHR (unadjusted for BMI), contrary to the positive phenotypic correlation between the two (Pulit et al. 2019). This is due to the fact that GWAS on WHRadjBMI is best powered to identify loci affecting hip (but not waist) circumference, thereby increasing BMI while decreasing WHR (Winkler et al. 2018). Although many of these loci affect both sexes, the dependence of adiposity distribution on sex is apparent in many sexually dimorphic genetic effects.

Sexual dimorphism in BMI and WHR

Examining the sex specificity of the genetics of body fat distribution can further our understanding of the mechanisms involved, whether specific to either or common to both. Sex-stratified GWAS in up to 330,000 individuals of European ancestry were unable to detect sexual dimorphism at any locus for BMI but found 44 loci with significantly different effects on

WHRadjBMI in men and women (Randall et al. 2013; Winkler et al. 2015). Although some loci showed opposite effects in both sexes, most were significantly stronger in women.

Similar results were found for the 346 loci associated with WHRadjBMI in larger sample sizes, where a majority of the 53 sexually dimorphic loci had stronger effects in women than in men (Pulit et al. 2019). Among these dimorphic loci, 28 showed directionality of effects which was consistent with favorable, non-central adiposity (increased BMI and decreased WHR), but only in women (**Figure 1**). Using sex-specific summary results for waist and hip circumference for European ancestry individuals from the UK Biobank (www.nealelab.is/uk-biobank, N > 350,000), we observed that while the BMI-increasing alleles of the 28 loci had stronger effects on hips in women, they were enriched for effects increasing both waist and hip circumference in men (**Figure 1**). This suggests that genetically-driven WHR is a stronger indicator of favorable adiposity in women compared to men.

Since sex differences in fat distribution are largely influenced by sex hormones, the sexually dimorphic genetic factors may interact with hormonal levels to regulate gene expression and gene activity (Wells 2007; Kirchengast 2010; Brown and Clegg 2010; Muraleedharan and Jones 2010; Ma et al. 2015). Indeed, recent work in mice has shown that sex-specific differences in mitochondrial activity mediated the sexual dimorphic effects of the genetic variants around *LYPLAL1* on fat accumulation and insulin resistance (Norheim et al. 2019). Although some autosomal genes have been shown to have tissue-specific sex differences in expression in humans (Kassam et al. 2019), the extent of their contribution to the sexual dimorphism of body fat distribution remains unclear. Identifying the genes and tissues mediating differences in adiposity distribution will be crucial to understanding not only the differences between sexes, but the drivers of heterogeneity in obesity as a whole.

Deeper characterisation of favourable adiposity

While BMI and WHR are useful as measures of obesity through their ubiquitous availability in large studies, they remain insufficient for the in-depth investigation of the quantity and distribution of metabolically-relevant adiposity deposits. Imaging methods on the other hand, namely dual-energy X-ray absorptiometry (DXA) and magnetic resonance imaging (MRI), do allow the precise measurement of localized body fat mass, including SAT and VAT and other ectopic fat deposits. While the cost of these methods is prohibitive in large cohorts, their availability in smaller cohorts has been particularly useful in the further characterization of obesity-related genetic factors and uncovering the tissues and mechanisms involved.

Using these, we were able to investigate the tissues involved in the genetic effects of 24 loci increasing non-central obesity, i.e. increased BMI but decreased WHR, and found them to increase SAT in particular, with non-significant effects on either waist circumference or VAT (Winkler et al. 2018). They were, however, associated with a significant decrease in pericardial adipose tissue (PAT). By applying Mendelian randomization (MR), a method which uses genetic variants as instrumental variables to approximate randomized controlled trials (Burgess, Butterworth, and Thompson 2013), we were further able to show that these changes in adipose tissue distribution were protective of CVD and T2D. Follow-up analyses performed in additional data from the UK Biobank and the Ectopic fat (Chu et al. 2017), DIAGRAM (Morris et al. 2012), MAGIC (Dupuis et al. 2010) and Global Lipids Genetics consortia (Dupuis et al. 2010; Willer et al. 2013) confirmed this link with SAT, but found no association with either waist circumference or VAT. It remains unclear whether the protective effects of these genetic variants was mediated entirely through a reduction in PAT or whether other unmeasured fat deposits were also decreased. It has also been suggested that SAT may have protective effects, such as through the secretion of beneficial adipokines (Manolopoulos, Karpe, and Frayn 2010; Rydén et al. 2014;

Lotta et al. 2018). The possible positive contribution of SAT to metabolic health through mechanisms other than the sequestration of lipids requires additional investigation.

Combining BMI with metabolic traits through hierarchical clustering revealed a similar set of genetic variants with profiles consistent with a shift from favorable to unfavorable adiposity (Yaghoobkar et al. 2014a). Despite being associated with lower BMI, these loci were associated with poor metabolic health, including higher triglyceride and lower HDL levels, consistent with a mild form of lipodystrophy. Consistent with the adipose expandability hypothesis, the detrimental effects were associated with increased visceral-to-subcutaneous adipose tissue ratio and liver fat. Further investigation revealed these genetic variants to be associated with increased risk of T2D, CVD and hypertension (Yaghoobkar et al. 2016a). Interestingly, while this was combined with a shift in body fat from the trunk to the legs in women, decreasing central obesity and WHR, the shift was in the opposite direction for men, confirming that WHR is a poor indicator of favorable adiposity in men.

A similar approach has led to the discovery of 53 loci associated with increased insulin resistance phenotypes despite decreasing BMI (Lotta et al. 2017). The decrease in adiposity was found to be driven by reduced SAT, while ectopic fat, namely VAT and liver fat, was increased (Ji et al. 2019a). This shift in body fat distribution was also associated with a metabolic profile similar to a mild form of lipodystrophy (e.g. increased triglyceride levels and blood pressure). Furthermore, a polygenic risk score composed of these genetic variants was found to contribute to familial partial lipodystrophy type 1, highlighting the similarities of the genetic basis of obesity-induced and severe forms of insulin resistance. Other loci have been shown to be associated with ectopic fat traits (Chu et al. 2017) and functional analysis of two genes therein, namely *ATXN1* and *UBE2E2*, revealed a role in adipocyte differentiation. These results further support the hypothesis of peripheral adipose tissue dysregulation and ectopic fat deposits as primary drivers of the adverse metabolic consequences of obesity.

Disentangling the consequences specific to adiposity subtype and distribution is hampered by the limited availability of imaging-derived phenotypes, relegating their study to the secondary characterization of predefined aspects of adiposity. However, the correlation of many aspects of body morphology and their dependence on factors such as age and gender makes it possible to accurately estimate fat/lean mass composition of body parts through the use of bioimpedance.

Estimates of body composition

Exploiting the fact that adipose tissue has higher electric resistivity than lean mass, measures of bio-electric impedance across different parts of the body can be combined with easily available data such as age, gender, and anthropometric traits to provide remarkably accurate estimates of fat/lean mass in various parts of the body. Bioimpedance-based estimates are quite highly correlated with DXA or MRI measures ($r^2 \approx 0.91$), however they tend to be slightly biased in individuals with extreme values of BMI (below ~ 18 or above ~ 40 , (Sun et al. 2005; Achamrah et al. 2018).

The accurate estimation of fat and lean mass in large studies such as the UK Biobank has led to the identification of 98 genetic variants associated with body fat distribution across arms, legs and trunk, 29 of which were novel (Rask-Andersen et al. 2019). Genetic regulation for adiposity in trunk and legs was found to be largely shared, with less overlap in variants with the arms. Approximately one third of all variants had a stronger effect in women than in men.

Although this is typically done for body parts, combinations of the same anthropometric traits and bioimpedance could also provide reasonable estimates for other DXA/MRI traits such as VAT or SAT, which are suspected to be etiologically implicated in metabolic consequences of obesity. Estimating these in the UK Biobank through multivariate linear regression, we found that the accuracy is somewhat lower than estimates for body parts where bioimpedance can be measured (VAT $r^2 \approx 0.67$, abdominal SAT $r^2 \approx 0.74$). Averaging multiple instances of bioimpedance

measurements can reduce the variability slightly and improve accuracy (VAT $r^2 \approx 0.71$, abdominal SAT $r^2 \approx 0.86$). Additional analyses will be required to determine whether these estimates capture etiologically relevant variance beyond what is provided by other measurements. Extending this further, however, the combination of multiple phenotypes into composite traits need not be restricted to the estimation of specific components of body composition.

Exploratory analysis of variance in body morphology

Combining multiple relevant phenotypes into composite traits can be highly effective in increasing statistical power, describing specific aspects of obesity and allowing their detailed characterisation in terms of genetic basis and impact on metabolic health. Methods aimed at reducing data dimensionality, such as principal component analysis (PCA), are particularly suited to the task of combining multiple traits and understanding the underlying architecture. Applying PCA to six anthropometric phenotypes, namely BMI, WHR, and the constituents of both, (Ried et al. 2016) were able to summarize over 99% of the variance into four principal components (PCs), which provide the main axes of variation in body shape (as defined by the six traits). These axes could be summarized as overall obesity, increased height and WHR, increased height and hip circumference (decreasing WHR), and increased BMI and weight. Subsequent genome-wide analysis showed that the genetic factors underlying each PC are largely distinct from each other with little overlap in genome-wide significant hits, which is consistent with different mechanisms affecting these phenotypes. Although there was substantial overlap with hits for the six original traits, their combination nevertheless allowed the discovery of 6 new loci. Unfortunately, the mechanisms and pathways underlying these associations have not yet been investigated.

Similarly, canonical correlation analysis (CCA), a method to infer relations between datasets, has also been shown to be suitable to detect variants associated with combinations of phenotypes (Cichonska et al. 2016a). By applying this to metabolic traits, (Ji et al. 2019b) identified variants associated with changes in metabolic profiles rather than individual traits. They combined these results with summary statistics for body fat percentage using hierarchical clustering to obtain a set of variants which were associated with healthy metabolic profiles despite higher body fat. In addition to the traits used in discovery, the variants were found to be associated with higher BMI and subcutaneous fat, but reduced liver fat and lower visceral-to-subcutaneous adipose tissue ratio. Consistent with previous results, this was associated with reduced risk of T2D, heart disease, and hypertension.

Integrative approaches to characterization of adiposity

As the number of genetic loci associated with individual anthropometric traits has increased, the overlap between them has revealed the complexity of their shared etiology. This has shifted the focus from single traits to combinations thereof, providing a more integrative view of genetic mechanisms involved in adiposity and a promising avenue for future research. Depending on the traits selected in the exploratory phase, the focus of integrative approaches ranges from understanding the shared genetics underlying body shape (as in Ried et al. 2016) to better characterizing specific aspects of adiposity and their consequences (as in Yaghootkar et al. 2014).

Widely-available anthropometric traits, especially BMI and WHR, are almost invariably included, either in the construction of composite traits or in downstream analyses. The inclusion of additional body measures is highly dependent on both the focus of the endeavor and data

availability. DXA- or MRI-based measures of specific fat deposits known or suspected to be biologically relevant (e.g. liver fat) or even in other parts of the body can add a significant amount of information, but their limited availability may be restrictive. Bioimpedance-based estimates of body composition can also provide useful insight into the subtle differences in adiposity subtypes.

The incorporation of molecular traits or disease outcomes related to obesity, in addition to or instead of anthropometric traits, provides a more targeted perspective of co-occurring symptoms. By focusing on the adiposity-related phenotypes which are most clinically relevant, this approach is particularly suited to a supervised approach aimed at further characterizing known or suspected aspects of obesity with certain expected consequences, such as favorable adiposity or dysregulation of SAT similar to lipodystrophy. However, the inclusion of secondary phenotypes not causally linked to aspects of body morphology may bias results away from the relevant mechanisms towards environmentally confounded correlations. Furthermore, inferring causality between any of the selected phenotypes, primary or secondary, is likely to be complicated by their inclusion in the exploratory phase, leading to circular arguments. Care should also be given to the selection of traits to include in the exploratory phase, as oversampling traits related to any aspect of obesity will inherently bias the derived components toward the selected aspect unless appropriately weighted.

In addition to the selection of phenotypes themselves, the nature of the data itself merits consideration as both phenotypic and genetic data may provide useful insight, but the requirements, interpretation and downstream analyses will vary accordingly. Phenotypic data may provide insight into overall correlations between traits, including the influence of genes and environment, as well as any interaction they may have. An advantage of phenotype-based composite traits is that they can be used in any type of downstream analysis much like any other phenotype, including genetic ones such as GWAS (Ried et al. 2016). The main limitations of this approach are likely to be the (potentially undesirable) environmental effects included, as well as the requirement for individual-level data. On the other hand, using genetic associations as a

starting point may remove some of the influence of environmental factors but requires reliable effect estimates and, implicitly, a strong genetic basis and/or large sample sizes. Public availability of many GWAS results makes this an attractive option for many anthropometric traits, but genetics-based composite traits may require additional adjustments, such as accounting for phenotypic correlation in the linear combination of genetic effects to produce interpretable effect sizes. Either type of data is suitable to many methods for the analysis of data structure and covariance.

Clustering methods broadly provide a categorization of elements based on their properties and hierarchical clustering has already proven successful in furthering our understanding of adiposity subtypes, their genetic basis and their effects on health (Yaghoobkar et al. 2014, 2016; Ji et al. 2019). This type of clustering does, however, have limitations, such as the resulting clusters being mutually exclusive and non-probabilistic. In light of the high prevalence of pleiotropy, it is reasonable to expect that some of the variants may contribute to more than one mechanism relevant to adiposity and/or metabolic health. Forcing these into non-overlapping clusters could potentially obscure secondary mechanisms in the etiology of adiposity-related health consequences. Alternative clustering methods allowing potentially overlapping and/or non-deterministic clusters, such as the Iterative Signature Algorithm (Bergmann, Ihmels, and Barkai 2003), may provide complementary insight into these effects.

The covariance structure of data is particularly indicative of shared mechanisms and its analysis can shed light on both etiology and consequences of shifts in adiposity distribution. PCA, as was used by (Ried et al. 2016c), is one of the most widely used techniques for this but factor analysis (FA) could prove to be a valid alternative. While PCA decomposes the data into uncorrelated components while preserving most of the covariance structure, FA identifies a predefined number of latent variables which are assumed to give rise to the observed data. The principle behind FA may in fact be closer to the biological hypothesis that changes in the subtypes of adiposity are affecting both body morphology and metabolic health and may yield more

biologically meaningful results. CCA has similarly already proven useful in the creation of composite traits (Cichonska et al. 2016b; Ji et al. 2019d), but this method can be used more generally to co-analyze datasets using cross-covariance matrices. This provides optimal linear combinations of two sets of variables to obtain the maximum correlation and could conceivably be employed to find composite adiposity traits co-occurring with a set of metabolic abnormalities and/or obesity-related diseases.

Conclusion

While the ever-increasing sample sizes of GWAS have led to the discovery of dozens or even hundreds of loci associated with many anthropometric traits, this has only highlighted the complexity of the biological mechanisms underlying body morphology. Integrative and comparative analyses of multiple aspects of obesity have provided critical insight into the shared and distinct genetic mechanisms underpinning the emergence of heterogeneity in obesity and its consequences on health. Many genetic factors have already been found to be relevant to the shift between healthy and unhealthy adiposity, highlighting the role of subcutaneous adipose tissue as a relatively benign or even beneficial fat storage. Additional investigations will be required to further elucidate the mechanisms involved, which may be instrumental in preventing obesity-related health deterioration. Furthermore, the discovery of these and additional genetic markers facilitate the classification of individuals into more homogeneous subgroups of obesity and will subsequently provide a better characterization of individual-level health risks associated with excessive adiposity and allow more targeted interventions in patients at greatest risk of complications.

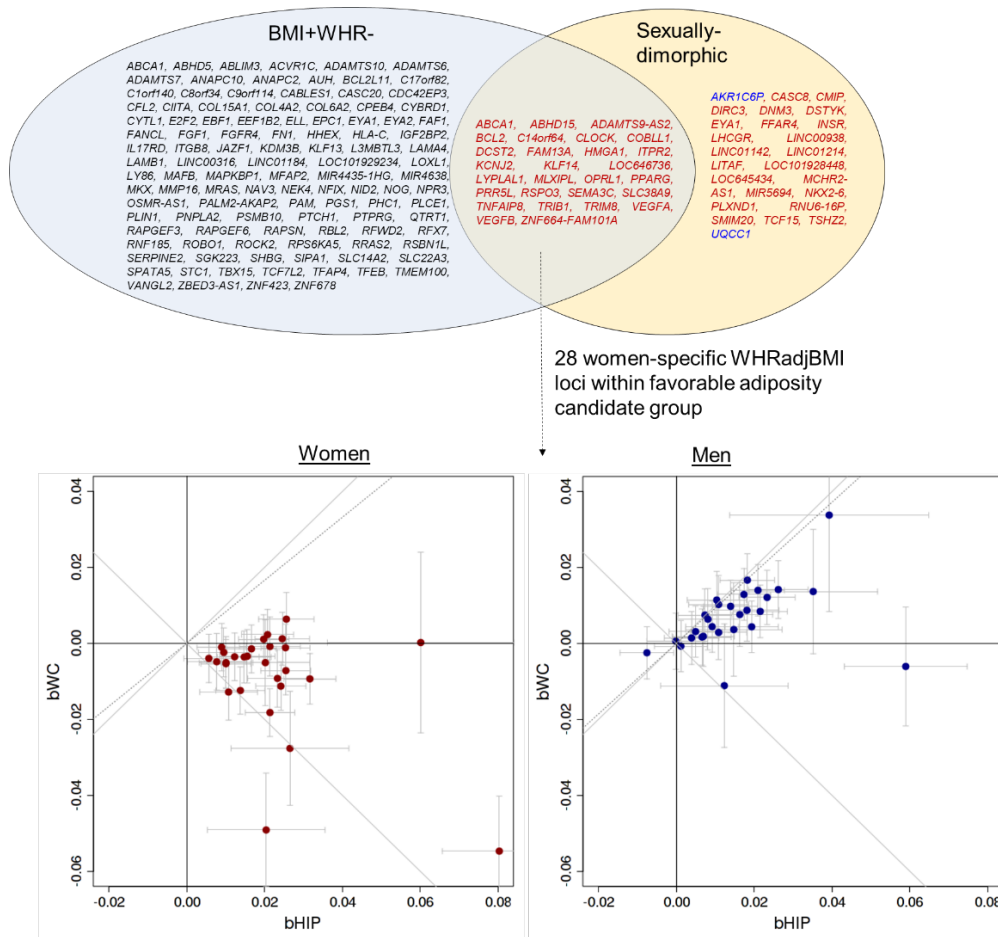
Tables & Figures

Table 1. Overview of GIANT consortium genome-wide association studies for BMI and WHRadjBMI from 2010 to 2019.

Trait	Reference	Year	Description	Discovery Sample size	# GWS loci	# Secondary signals	Explained var	# Sex diff loci
BMI	Speliotes et al	2010	46 studies (HapMap)	123,865	32	-	1.45%	0
	Locke et al	2015	125 studies (HapMap / MetaboChip)	339,224	97	6	2.70%	2
	Turcot et al	2018	123 studies (ExomeChip)	526,508	14 ^a	-	<0.1%	1
	Yengo et al	2018	Locke et al + UK Biobank	795,612	536	405	6.00%	-
WHRadjBMI	Heid et al	2010	32 studies (HapMap)	77,167	14	-	1.03%	7
	Shungin et al	2015	101 studies (HapMap / MetaboChip)	224,459	49	16	1.40%	20
	Justice et al	2019	74 studies (ExomeChip)	344,369	9 ^a	-	0.10%	3
	Pulit et al	2019	Shungin et al + UK Biobank	694,649	346	117	3.90%	53

^a Number of rare or low frequency variants (MAF < 5%)

Figure 1. Favorable adiposity candidate versus sexually dimorphic loci. Based on the 346 WHRadjBMI loci from Pulit et al 2019, the Venn diagram compares 137 favorable adiposity candidate loci (positive BMI and negative WHR effect based on nominal significant association, $P < 0.05$, "BMI+WHR-") with the 53 loci with significant sex-difference in their effect on WHRadjBMI. For the 28 overlapping loci, the scatter plots compare genetic effects on waist and hip circumference by sex using summary results from UK Biobank (<http://www.nealelab.is/uk-biobank>).



References

- Abraham, Tobin M., Alison Pedley, Joseph M. Massaro, Udo Hoffmann, and Caroline S. Fox. 2015. "Association between Visceral and Subcutaneous Adipose Depots and Incident Cardiovascular Disease Risk Factors." *Circulation* 132 (17): 1639–47.
- Achamrah, Najate, Guillaume Colange, Julie Delay, Agnès Rimbert, Vanessa Folope, André Petit, Sébastien Grigioni, Pierre Déchelotte, and Moïse Coëffier. 2018. "Comparison of Body Composition Assessment by DXA and BIA according to the Body Mass Index: A Retrospective Study on 3655 Measures." *PloS One* 13 (7): e0200465.
- Bergmann, Sven, Jan Ihmels, and Naama Barkai. 2003. "Iterative Signature Algorithm for the Analysis of Large-Scale Gene Expression Data." *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics* 67 (3 Pt 1): 031902.
- Brown, L. M., and D. J. Clegg. 2010. "Central Effects of Estradiol in the Regulation of Food Intake, Body Weight, and Adiposity." *The Journal of Steroid Biochemistry and Molecular Biology* 122 (1-3): 65–73.
- Burgess, Stephen, Adam Butterworth, and Simon G. Thompson. 2013. "Mendelian Randomization Analysis with Multiple Genetic Variants Using Summarized Data." *Genetic Epidemiology* 37 (7): 658–65.
- Chu, Audrey Y., Xuan Deng, Virginia A. Fisher, Alexander Drong, Yang Zhang, Mary F. Feitosa, Ching-Ti Liu, et al. 2017. "Multiethnic Genome-Wide Meta-Analysis of Ectopic Fat Depots Identifies Loci Associated with Adipocyte Development and Differentiation." *Nature Genetics* 49 (1): 125–30.
- Cichonska, Anna, Juho Rousu, Pekka Marttinen, Antti J. Kangas, Pasi Soininen, Terho Lehtimäki, Olli T. Raitakari, et al. 2016a. "metaCCA: Summary Statistics-Based Multivariate Meta-Analysis of Genome-Wide Association Studies Using Canonical Correlation Analysis." *Bioinformatics* 32 (13): 1981–89.
- Dupuis, Josée, Claudia Langenberg, Inga Prokopenko, Richa Saxena, Nicole Soranzo, Anne U. Jackson, Eleanor Wheeler, et al. 2010. "New Genetic Loci Implicated in Fasting Glucose Homeostasis and Their Impact on Type 2 Diabetes Risk." *Nature Genetics* 42 (2): 105–16.
- Ji, Yingjie, Andrianos M. Yiorkas, Francesca Frau, Dennis Mook-Kanamori, Harald Staiger, E. Louise Thomas, Naeimeh Atabaki-Pasdar, et al. 2019a. "Genome-Wide and Abdominal MRI Data Provide Evidence That a Genetically Determined Favorable Adiposity Phenotype Is Characterized by Lower Ectopic Liver Fat and Lower Risk of Type 2 Diabetes, Heart Disease, and Hypertension." *Diabetes* 68 (1): 207–19.
- Justice, Anne E., Tugce Karaderi, Heather M. Highland, Kristin L. Young, Mariaelisa Graff, Yingchang Lu, Valérie Turcot, et al. 2019. "Protein-Coding Variants Implicate Novel Genes Related to Lipid Homeostasis Contributing to Body-Fat Distribution." *Nature Genetics* 51 (3): 452–69.
- Kaess, B. M., A. Pedley, J. M. Massaro, J. Murabito, U. Hoffmann, and C. S. Fox. 2012. "The Ratio of Visceral to Subcutaneous Fat, a Metric of Body Fat Distribution, Is a Unique Correlate of Cardiometabolic Risk." *Diabetologia* 55 (10): 2622–30.
- Kassam, Irfahan, Yang Wu, Jian Yang, Peter M. Visscher, and Allan F. McRae. 2019. "Tissue-Specific Sex Differences in Human Gene Expression." *Human Molecular Genetics* 28 (17): 2976–86.
- Kirchengast, Sylvia. 2010. "Gender Differences in Body Composition from Childhood to Old Age: An Evolutionary Point of View." *Journal of Life Sciences*.
<https://doi.org/10.1080/09751270.2010.11885146>.
- Locke, Adam E., Bratati Kahali, Sonja I. Berndt, Anne E. Justice, Tune H. Pers, Felix R. Day, Corey Powell, et al. 2015. "Genetic Studies of Body Mass Index Yield New Insights for Obesity Biology." *Nature* 518 (7538): 197–206.

- Lotta, Luca A., Pawan Gulati, Felix R. Day, Felicity Payne, Halit Ongen, Martijn van de Bunt, Kyle J. Gaulton, et al. 2017. "Integrative Genomic Analysis Implicates Limited Peripheral Adipose Storage Capacity in the Pathogenesis of Human Insulin Resistance." *Nature Genetics* 49 (1): 17–26.
- Lotta, Luca A., Laura B. L. Wittemans, Verena Zuber, Isobel D. Stewart, Stephen J. Sharp, Jian 'an Luan, Felix R. Day, et al. 2018. "Association of Genetic Variants Related to Gluteofemoral vs Abdominal Fat Distribution With Type 2 Diabetes, Coronary Disease, and Cardiovascular Risk Factors." *JAMA: The Journal of the American Medical Association* 320 (24): 2553–63.
- Manolopoulos, K. N., F. Karpe, and K. N. Frayn. 2010. "Gluteofemoral Body Fat as a Determinant of Metabolic Health." *International Journal of Obesity* 34 (6): 949–59.
- Ma, Xiuquan, Paul Lee, Donald J. Chisholm, and David E. James. 2015. "Control of Adipocyte Differentiation in Different Fat Depots; Implications for Pathophysiology or Therapy." *Frontiers in Endocrinology*. <https://doi.org/10.3389/fendo.2015.00001>.
- Morris, Andrew P., Benjamin F. Voight, Tanya M. Teslovich, Teresa Ferreira, Ayellet V. Segrè, Valgerdur Steinthorsdottir, Rona J. Strawbridge, et al. 2012. "Large-Scale Association Analysis Provides Insights into the Genetic Architecture and Pathophysiology of Type 2 Diabetes." *Nature Genetics* 44 (9): 981–90.
- Muraleedharan, Vakkat, and T. Hugh Jones. 2010. "Testosterone and the Metabolic Syndrome." *Therapeutic Advances in Endocrinology and Metabolism* 1 (5): 207–23.
- Norheim, Frode, Yehudit Hasin-Brumshtein, Laurent Vergnes, Karthickeyan Chella Krishnan, Calvin Pan, Marcus M. Seldin, Simon T. Hui, et al. 2019. "Gene-by-Sex Interactions in Mitochondrial Functions and Cardio-Metabolic Traits." *Cell Metabolism* 29 (4): 932–49.e4.
- Pischon, T., H. Boeing, K. Hoffmann, M. Bergmann, M. B. Schulze, K. Overvad, Y. T. van der Schouw, et al. 2008. "General and Abdominal Adiposity and Risk of Death in Europe." *The New England Journal of Medicine* 359 (20): 2105–20.
- Pulit, Sara L., Charli Stoneman, Andrew P. Morris, Andrew R. Wood, Craig A. Glastonbury, Jessica Tyrrell, Loïc Yengo, et al. 2019. "Meta-Analysis of Genome-Wide Association Studies for Body Fat Distribution in 694 649 Individuals of European Ancestry." *Human Molecular Genetics* 28 (1): 166–74.
- Randall, Joshua C., Thomas W. Winkler, Zoltán Kutalik, Sonja I. Berndt, Anne U. Jackson, Keri L. Monda, Tuomas O. Kilpeläinen, et al. 2013. "Sex-Stratified Genome-Wide Association Studies Including 270,000 Individuals Show Sexual Dimorphism in Genetic Loci for Anthropometric Traits." *PLoS Genetics* 9 (6): e1003500.
- Rask-Andersen, Mathias, Torgny Karlsson, Weronica E. Ek, and Åsa Johansson. 2019. "Genome-Wide Association Study of Body Fat Distribution Identifies Adiposity Loci and Sex-Specific Genetic Effects." *Nature Communications*. <https://doi.org/10.1038/s41467-018-08000-4>.
- Ried, Janina S., Janina Jeff M, Audrey Y. Chu, Jennifer L. Bragg-Gresham, Jenny van Dongen, Jennifer E. Huffman, Tarunveer S. Ahluwalia, et al. 2016a. "A Principal Component Meta-Analysis on Multiple Anthropometric Traits Identifies Novel Loci for Body Shape." *Nature Communications* 7 (November): 13357.
- Rosenquist, Klara J., Alison Pedley, Joseph M. Massaro, Kate E. Therkelsen, Joanne M. Murabito, Udo Hoffmann, and Caroline S. Fox. 2013. "Visceral and Subcutaneous Fat Quality and Cardiometabolic Risk." *JACC. Cardiovascular Imaging* 6 (7): 762–71.
- Rydén, Mikael, Daniel P. Andersson, Ingrid B. Bergström, and Peter Arner. 2014. "Adipose Tissue and Metabolic Alterations: Regional Differences in Fat Cell Size and Number Matter, but Differently: A Cross-Sectional Study." *The Journal of Clinical Endocrinology and Metabolism* 99 (10): E1870–76.
- Shungin, Dmitry, Thomas W. Winkler, Damien C. Croteau-Chonka, Teresa Ferreira, Adam E. Locke, Reedik Mägi, Rona J. Strawbridge, et al. 2015. "New Genetic Loci Link Adipose and

- Insulin Biology to Body Fat Distribution." *Nature* 518 (7538): 187–96.
- Sudlow, Cathie, John Gallacher, Naomi Allen, Valerie Beral, Paul Burton, John Danesh, Paul Downey, et al. 2015. "UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age." *PLoS Medicine* 12 (3): e1001779.
- Sun, Guang, Curtis R. French, Glynn R. Martin, Ban Younghusband, Roger C. Green, Ya-Gang Xie, Maria Mathews, et al. 2005. "Comparison of Multifrequency Bioelectrical Impedance Analysis with Dual-Energy X-Ray Absorptiometry for Assessment of Percentage Body Fat in a Large, Healthy Population." *The American Journal of Clinical Nutrition* 81 (1): 74–78.
- Turcot, Valérie, Yingchang Lu, Heather M. Highland, Claudia Schurmann, Anne E. Justice, Rebecca S. Fine, Jonathan P. Bradfield, et al. 2018. "Protein-Altering Variants Associated with Body Mass Index Implicate Pathways That Control Energy Intake and Expenditure in Obesity." *Nature Genetics* 50 (1): 26–41.
- Virtue, Sam, and Antonio Vidal-Puig. 2010. "Adipose Tissue Expandability, Lipotoxicity and the Metabolic Syndrome--an Allostatic Perspective." *Biochimica et Biophysica Acta* 1801 (3): 338–49.
- Wells, Jonathan C. K. 2007. "Sexual Dimorphism of Body Composition." *Best Practice & Research. Clinical Endocrinology & Metabolism* 21 (3): 415–30.
- Willer, Cristen J., Ellen M. Schmidt, Sebanti Sengupta, Gina M. Peloso, Stefan Gustafsson, Stavroula Kanoni, Andrea Ganna, et al. 2013. "Discovery and Refinement of Loci Associated with Lipid Levels." *Nature Genetics* 45 (11): 1274–83.
- Winkler, Thomas W., Felix Günther, Simon Höllerer, Martina Zimmermann, Ruth Jf Loos, Zoltán Kutalik, and Iris M. Heid. 2018. "A Joint View on Genetic Variants for Adiposity Differentiates Subtypes with Distinct Metabolic Implications." *Nature Communications* 9 (1): 1946.
- Winkler, Thomas W., Anne E. Justice, Mariaelisa Graff, Lilda Barata, Mary F. Feitosa, Su Chu, Jacek Czajkowski, et al. 2015. "The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study." *PLoS Genetics* 11 (10): e1005378.
- Yaghootkar, Hanieh, Luca A. Lotta, Jessica Tyrrell, Roelof A. J. Smit, Sam E. Jones, Louise Donnelly, Robin Beaumont, et al. 2016a. "Genetic Evidence for a Link Between Favorable Adiposity and Lower Risk of Type 2 Diabetes, Hypertension, and Heart Disease." *Diabetes* 65 (8): 2448–60.
- Yaghootkar, Hanieh, Robert A. Scott, Charles C. White, Weihua Zhang, Elizabeth Speliotes, Patricia B. Munroe, Georg B. Ehret, et al. 2014a. "Genetic Evidence for a Normal-Weight 'Metabolically Obese' Phenotype Linking Insulin Resistance, Hypertension, Coronary Artery Disease, and Type 2 Diabetes." *Diabetes* 63 (12): 4369–77.
- Yengo, Loic, Julia Sidorenko, Kathryn E. Kemper, Zhili Zheng, Andrew R. Wood, Michael N. Weedon, Timothy M. Frayling, et al. 2018. "Meta-Analysis of Genome-Wide Association Studies for Height and Body Mass Index in ~700000 Individuals of European Ancestry." *Human Molecular Genetics* 27 (20): 3641–49.

2.2 Through the lens of integrative analysis

The core aim of this project was to disentangle the various aspects of body shape, understand their genetic basis, and identify the specific consequences of each on health and lifestyle. The paper has been published in *Communications Biology* [87] and is included below.

Using principal component analysis (PCA), we showed that most (> 99%) of the total genetically-defined variance in 14 traits related to body shape can be summarized using just 4 principal components (PCs, Fig. 2.1), in order of decreasing explained variance: (1) *body size* increased all included traits; (2) *adiposity* decreased stature but increased fat mass; (3) *predisposition to abdominal fat deposition* increased WHR and waist circumference at the expense of hip circumference; and (4) *lean mass* decreased stature but increased lean mass. Enrichment analyses suggest that the biological mechanisms underlying the first two PCs involve genes expressed in the brain and central nervous system, while those of the latter two appear to involve many tissues (among which adipose tissue) and rather concern embryo-/morphogenesis and energy homeostasis. Mendelian randomization (MR) analyses revealed many consequences on health and lifestyle, both shared and distinct, the comparison of which provided useful clarifications to the etiology of certain obesity-related diseases. Finally, we showed that the orthogonality of these PCs could be leveraged to create combined predictors of diseases which outperformed individual traits within the same population and retained much of their predictive accuracy in an ethnically diverse sample.

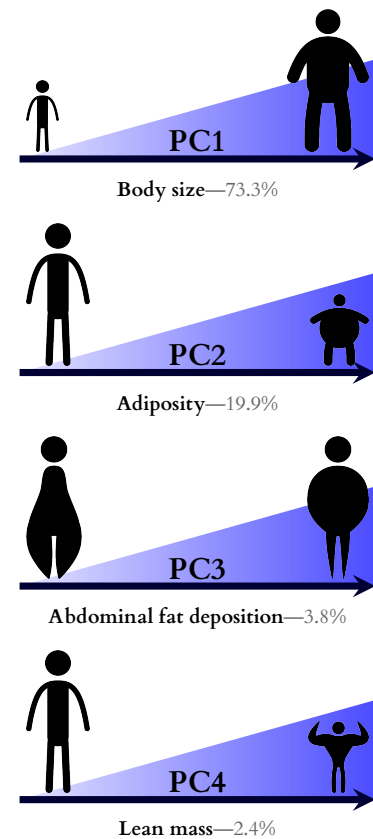










Figure 2.1: The top 4 body principal components (PCs) explain more than 99% of the total variance. The silhouettes illustrate the expected phenotypes at the lower (left) and upper (right) extremes of the PC scale. The values in gray indicate the variance explained by the PC.



Composite trait Mendelian randomization reveals distinct metabolic and lifestyle consequences of differences in body shape

Jonathan Sulc^{1,2}, Anthony Sonrel^{2,3}, Ninon Mounier^{1,2}, Chiara Auwerx^{1,2,4} , Eirini Marouli^{5,6} , Liza Darrous^{1,2} , Bogdan Draganski^{7,8} , Tuomas O. Kilpeläinen⁹ , Peter Joshi¹⁰ , Ruth J. F. Loos^{11,12,13}  & Zoltán Kutalik^{1,2,14} 

Obesity is a major risk factor for a wide range of cardiometabolic diseases, however the impact of specific aspects of body morphology remains poorly understood. We combined the GWAS summary statistics of fourteen anthropometric traits from UK Biobank through principal component analysis to reveal four major independent axes: body size, adiposity, predisposition to abdominal fat deposition, and lean mass. Mendelian randomization analysis showed that although body size and adiposity both contribute to the consequences of BMI, many of their effects are distinct, such as body size increasing the risk of cardiac arrhythmia ($b = 0.06$, $p = 4.2 * 10^{-17}$) while adiposity instead increased that of ischemic heart disease ($b = 0.079$, $p = 8.2 * 10^{-21}$). The body mass-neutral component predisposing to abdominal fat deposition, likely reflecting a shift from subcutaneous to visceral fat, exhibited health effects that were weaker but specifically linked to lipotoxicity, such as ischemic heart disease ($b = 0.067$, $p = 9.4 * 10^{-14}$) and diabetes ($b = 0.082$, $p = 5.9 * 10^{-19}$). Combining their independent predicted effects significantly improved the prediction of obesity-related diseases ($p < 10^{-10}$). The presented decomposition approach sheds light on the biological mechanisms underlying the heterogeneity of body morphology and its consequences on health and lifestyle.

¹University Center for Primary Care and Public Health, University of Lausanne, Lausanne, Switzerland. ²Swiss Institute of Bioinformatics, Lausanne, Switzerland. ³Department of Molecular Life Sciences, University of Zürich, Zürich, Switzerland. ⁴Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland. ⁵William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, London, UK. ⁶Centre for Genomic Health, Life Sciences, London, UK. ⁷Laboratory for Research in Neuroimaging, Department of Clinical Neurosciences, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland. ⁸Neurology Department, Max-Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany. ⁹Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. ¹⁰Centre for Global Health Research, Usher Institute, University of Edinburgh, Edinburgh, UK. ¹¹The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ¹²The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ¹³Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ¹⁴Genetics of Complex Traits, University of Exeter Medical School, University of Exeter, Exeter, UK. ✉email: zoltan.kutalik@unil.ch

Obesity is one of the main risk factors for many non-communicable diseases, such as type 2 diabetes (reviewed in ref. 1) and cardiovascular diseases (reviewed in ref. 2). The associated disease risk and underlying biology of obesity are generally studied through the lens of body mass index (BMI, weight [kg]/height² [m²]), which uses excess body mass as a surrogate for adiposity. This approximation provides a reasonably accurate predictor at the population level^{3,4} but is blind to many aspects of body shape and composition that may be critical to disease etiology. Indeed, disease risk and progression have been shown to be affected by the location and type of the adipose tissue in which excess calories are stored^{3–6}. Abdominal obesity in particular, usually assessed using waist circumference or waist-to-hip ratio (WHR), is associated with increased disease risk and mortality independent of BMI³.

Although the exact mechanisms underlying the consequences of adiposity on health have not been fully elucidated, the predominant hypothesis is that of adipose tissue expandability: that excess calories are preferentially stored in subcutaneous adipose tissue (SAT), which can expand with little to no deleterious impact on health^{7,8}. When its capacity for expansion is exceeded, adipose cells hypertrophy, causing local inflammation, and fat is increasingly stored as ectopic fat in organs and as visceral adipose tissue (VAT) around organs in a process similar to a mild form of lipodystrophy. Ectopic and visceral fat are thought to play a central role in many of the direct consequences of obesity.

Much remains unknown about the genetic and environmental factors affecting body fat distribution and its contribution to health outcomes. Currently available non-invasive imaging techniques such as magnetic resonance imaging (MRI) and dual-energy X-ray absorptiometry (DXA) allow the accurate measurement of adipose mass in different parts of the body and have considerably improved our understanding of the impact of different subtypes of adiposity on health^{9,10}. However, such techniques remain costly and are therefore generally restricted to smaller sample sizes.

An alternative approach is the concurrent analysis of multiple traits, leveraging the co-occurring changes in multiple phenotypes to understand the underlying causes and mechanisms. Analysis of variance methods, such as principal component analysis (PCA), have been used to investigate the complex architecture underlying body morphology, revealing the main axes of phenotypic variation and increasing the statistical power to detect novel loci affecting body morphology¹¹. While this has improved our understanding of the genetic basis underlying common and distinct components of anthropometric traits, their impact on health and quality of life remains unknown. Other approaches such as clustering and canonical correlation analysis have identified single nucleotide polymorphisms (SNPs) associated with healthier metabolic profiles, despite higher BMI and/or body fat percentage^{12–15}. However, these hypothesis-driven approaches (i.e., identifying clusters of functionally similar SNPs based on both obesity measures and health outcomes) are not suited to determine the causality of these correlated differences or the directionality of potential causal effects because the SNP groups have different health consequences by construction.

Here we followed a hypothesis-free approach to isolate independent axes of variations in body shape and size and investigated their health consequences. We performed a PCA on GWAS summary statistics of 14 anthropometric traits from the UK Biobank¹⁶ to extract orthogonal components, each representing different features of body shape (Fig. 1). We show that these measures of body shape can be summarized using four principal components (PCs) affecting body size, adiposity, abdominal fat deposition, and lean mass, respectively. Enrichment analyses highlighted differences in the pathways and tissues involved in

these composite traits, providing insight into the underlying biological mechanisms. We then used robust cross-sex Mendelian randomization (MR) to assess the impact of these independent components on health and lifestyle. While many health and lifestyle consequences were shared with individual traits, these orthogonal PCs allowed us to better disentangle the independent contributions of different aspects of body shape. The results can be explored using the shiny app which can be downloaded by following the instructions at <http://wp.unil.ch/sgg/pca-mr/>. Furthermore, the combination of these PCs improved the prediction accuracy of obesity-related diseases.

Results

PCA of genetic effects. A schematic representation of our composite MR analysis framework is shown in Fig. 1. We selected 14 anthropometric and bioimpedance-derived traits (Supplementary Data 1) as the basis for the PCA, 13 of which were available in the UK Biobank¹⁶ with genome-wide summary statistics made available by the Neale lab (www.nealelab.is/uk-biobank). Summary statistics for WHR were not available in the UK Biobank, therefore we performed a GWAS in the UK Biobank following the same procedure as that used by the Neale lab. SNPs which were then genome-wide significant (GWS, $p < 5 \times 10^{-8}$) for any trait were pruned to be independent. The resulting SNP x traits matrix of effect estimates was subjected to PCA. The loadings of each PC were then rescaled according to the phenotypic correlation between these traits in the UK Biobank so as to obtain standardized effect sizes, i.e. the resulting PC phenotypes would have a variance of 1. SNP-PC associations were then calculated genome-wide and SNPs for each PC were re-pruned individually.

The top four PCs explained more than 99% of the total variance (Fig. 2, Supplementary Data 2). PC1 (73.3% variance) represents an overall increase in body size, with positive weights for all traits indicating a slightly disproportionate increase in body mass compared to height, resulting in higher BMI as well. PC2 (19.9%) shows a decrease in height with an increase in fat mass at the expense of lean mass. PC3 (3.8%) is largely BMI- and body fat mass-neutral, decreasing hip circumference, and increasing waist circumference and hence WHR, reflecting a shift in body fat from hips to the waist. PC4 (2.4%) resulted in decreased height and increased BMI, with an increase in lean mass at the expense of fat mass.

Performing this analysis using male- or female-specific summary statistics produced very similar results, though PC3 explained less variance than PC4 in men (Supplementary Data 3–4, Supplementary Note 1). PCs were also robust to changes in the selection of specific traits. For example, excluding WHR yielded effect estimates which were highly correlated with those obtained using the full set of traits (all $|r| > 0.99$, see Supplementary Note 2), though the altered explained variance also resulted in the reordering of PCs 3 and 4.

PCs 1–3 were highly correlated with weight, BMI, and WHR, respectively, both phenotypically ($r \geq 0.70$, Supplementary Fig. 3) and in terms of their causal effects on the tested outcomes. The (genetically) orthogonal nature of PCs nevertheless resulted in much lower phenotypic correlation with each other than between traits (e.g., the phenotypic correlation between PC1 and PC2 was only 0.36 while that between weight and BMI was 0.90, Supplementary Fig. 4, Supplementary Data 5). Note that the phenotypic realization of PCs are not orthogonal, and their correlation is therefore not zero, due to differences between their environmental and genetic correlation. However, the MR-derived causal effect estimates of PCs are independent and therefore additive (though they may still be correlated), which is not the case for individual traits.

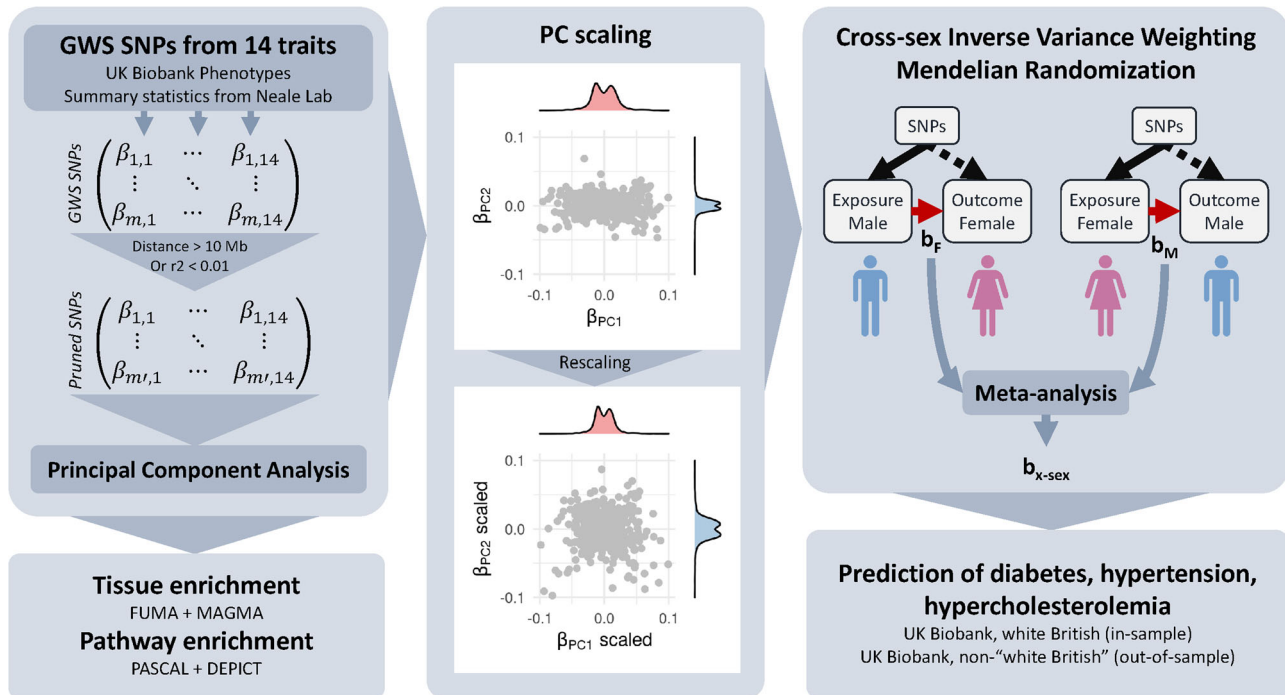


Fig. 1 Overview of the methods. Summary statistics for anthropometric traits from the UK Biobank were pruned for independence before being subjected to principal component analysis (PCA). Principal component-associated SNPs were tested for enrichment in genes expressed in certain tissues or in pathways. The genetic effects on the resulting components were scaled to obtain effect sizes corresponding to a trait with a variance of 1 (standardized). Mendelian randomization was used to determine the impact of these composite traits on lifestyle and health outcomes. Using these effect estimates, the individual risk was predicted in the UK Biobank and accuracy compared to BMI and WHR.

For brevity and clarity, we mainly describe the PCs in comparison with BMI and occasionally weight or WHR. Other traits had causal effects similar to or weaker than these traits (e.g., the effects of body fat percentage and BMI on the disease had a correlation coefficient $r = 0.95$, $p = 7.6 \times 10^{-43}$) or were less relevant to obesity-related health outcomes (e.g., height).

We found 615 independent GWS SNPs associated with PC1 (body size), 641 with PC2 (adiposity), 354 with PC3 (predisposition to abdominal fat deposition), and 610 with PC4 (lean mass). Among these, respectively 3, 83, 137, and 330 SNPs did not reach genome-wide significance for any individual trait. For comparison, BMI had 532 independent GWS SNPs.

Tissue/pathway enrichment. We tested the PCs for the enrichment of genes expressed in specific tissues using the MAGMA method¹⁷ through the FUMA interface¹⁸ in both GTEx v8¹⁹ and brain development and aging data from BrainSpan²⁰. We also tested for their enrichment of molecular pathway terms from the DEPICT dataset²¹ using PASCAL²². The tissue-enrichment results of the four PCs as well as weight, BMI, and WHR are shown in Fig. 3 and listed in Supplementary Data 6–9.

Loci associated with both PC1 (body size) and PC2 (adiposity) were mainly enriched for genes expressed in the cerebellum ($p \leq 2.2 \times 10^{-5}$) and the pituitary gland ($p \leq 5.7 \times 10^{-4}$). Using data from BrainSpan²⁰, we found PC2 to be further enriched for genes expressed in the brain specifically during the early to late mid-prenatal phases of development ($p \leq 6.1 \times 10^{-4}$).

PC3- (predisposition to abdominal fat deposition) associated SNPs were harbored by genes most enriched for expression in SAT ($p = 3.0 \times 10^{-14}$), followed by VAT ($p = 4.2 \times 10^{-10}$), female reproductive tissues (breast mammary tissue, ecto- and endocervix, and uterus, all $p \leq 2.4 \times 10^{-10}$), nerves (tibial, $p = 2 \times 10^{-15}$), arteries ($p \leq 2.2 \times 10^{-9}$), and digestive system ($p \leq 8.8 \times 10^{-5}$). Note that using sex-specific (e.g., male-specific)

summary statistics and PC loadings with the same gene expression datasets produced a similar enrichment for female-specific tissues. Data from BrainSpan showed PC3 to be enriched for genes expressed in the late prenatal brain ($p = 9.3 \times 10^{-8}$).

PC4 (lean mass) showed similar enrichment to that of PC3, though stronger for the digestive system ($p \leq 1.6 \times 10^{-10}$) and some female reproductive tissues (uterus, ecto- and endocervix, all $p \leq 1.3 \times 10^{-12}$), and weaker for adipose tissue and tibial nerve. PC4 was also enriched for genes expressed in the prostate ($p = 2.9 \times 10^{-4}$). In BrainSpan, PC4 also showed enrichment for genes expressed prenatally in the brain.

Loci associated with BMI showed slightly stronger enrichment than PCs 1 and 2 for genes expressed in the cerebellum, as well as other areas of the adult brain, including the basal ganglia, hippocampus, hypothalamus, amygdala, and frontal cortex. They were also enriched for genes expressed in the mid-prenatal brain, similar to that found for PC2, if slightly weaker.

The analysis of molecular pathways (Supplementary Data 10) showed qualitatively similar enrichment for PCs 1 (341 pathways) and 2 (348), weight (318), and BMI (293). These were mostly terms related to the brain, synapses, behavior, or learning. PC3 had the most numerous enriched terms (628) and the strongest enrichment overall. Most of the enriched terms were not related to brain function, but to embryogenesis and morphology, with many others specifically concerning adiposity, metabolism, and glucose homeostasis. Other terms were related to vascular or heart function, or hormones. PC4 (569) showed some overlap in terms with PC3, mainly in terms related to embryogenesis and morphology. Overall, 1067 pathways were found to be enriched for at least one PC.

Cross-sex MR analysis. We used inverse-variance weighted (IVW) MR to test for causal effects of the 14 anthropometric traits and PCs on disease outcomes, continuous measures of health, lifestyle



Fig. 2 The contributions of each trait to the first four genetic principal components (PCs). The explained variance of each PC is included in parentheses along with the descriptive name used in the main text. The loadings presented here are those typically used in the principal component analysis (PCA), scaled such that the sum of the squared weights is equal to 1 (as opposed to the scaling used to obtain composite traits with a variance of 1). This provides a consistent scale and makes PCs more easily comparable with each other.

factors, and diet, as well as the reverse. To avoid bias in the MR estimates due to sample overlap, we used sex-specific effects from opposite sexes for exposures and outcomes (male exposure–female outcome and female exposure–male outcome, using the same combined-sex PC loadings for exposure PCs). This produces unbiased estimates of the sex-specific causal effects in the sex used in the outcome, provided the strength of the instruments’ association with the exposure was not different between sexes. The two cross-sex causal effect estimates (female-to-male and male-to-female) were then meta-analyzed using inverse-variance weighting. Where the strength of the instruments’ association with the exposure differed between sexes, namely for WHR and PC3 (abdominal fat distribution), the meta-analyzed causal effect estimates may be biased (generally towards the null). Although the effect size estimates may be slightly over- or underestimated, this should not affect the type I error rate (see “Methods”).

Effects on disease risk. PC1 (body size) increased the risk of many diseases (Fig. 4). An increase of one standard deviation

(SD) increased the absolute risk of diabetes by 1.7% (95% CI: 1.3–2.1), that of hypertension by 2.3% (95% CI: 1.4–3.2), as well as many other diseases, such as nerve disorders, diseases of the veins and circulatory system, and prolapsed disc (Supplementary Data 11). Although it also increased the risk of cardiac arrhythmias by 0.93% (95% CI: 0.71–1.1), it did not significantly affect the risk of hypercholesterolemia or heart disease.

PC2 (adiposity) had much stronger effects on many obesity-related diseases (Fig. 4), where a 1 SD increase also increased the absolute risk of diabetes by 2.1% (95% CI: 1.7–2.5), hypertension by 6.8% (95% CI: 5.8–7.7), as well as hypercholesterolemia by 3.4% (95% CI: 2.8–4.0) and ischemic heart disease (IHD) by 1.8% (95% CI: 1.5–2.2). The risk of many other diseases, such as arthrosis and diseases of the nervous system were also increased (Supplementary Data 11).

PC3 (predisposition to abdominal fat deposition), despite being weight- and BMI-neutral, was a risk factor for many of the same obesity-related diseases as PC2 (Fig. 4, Supplementary Data 11). A 1 SD increase in PC3 increased the absolute risk of diabetes by

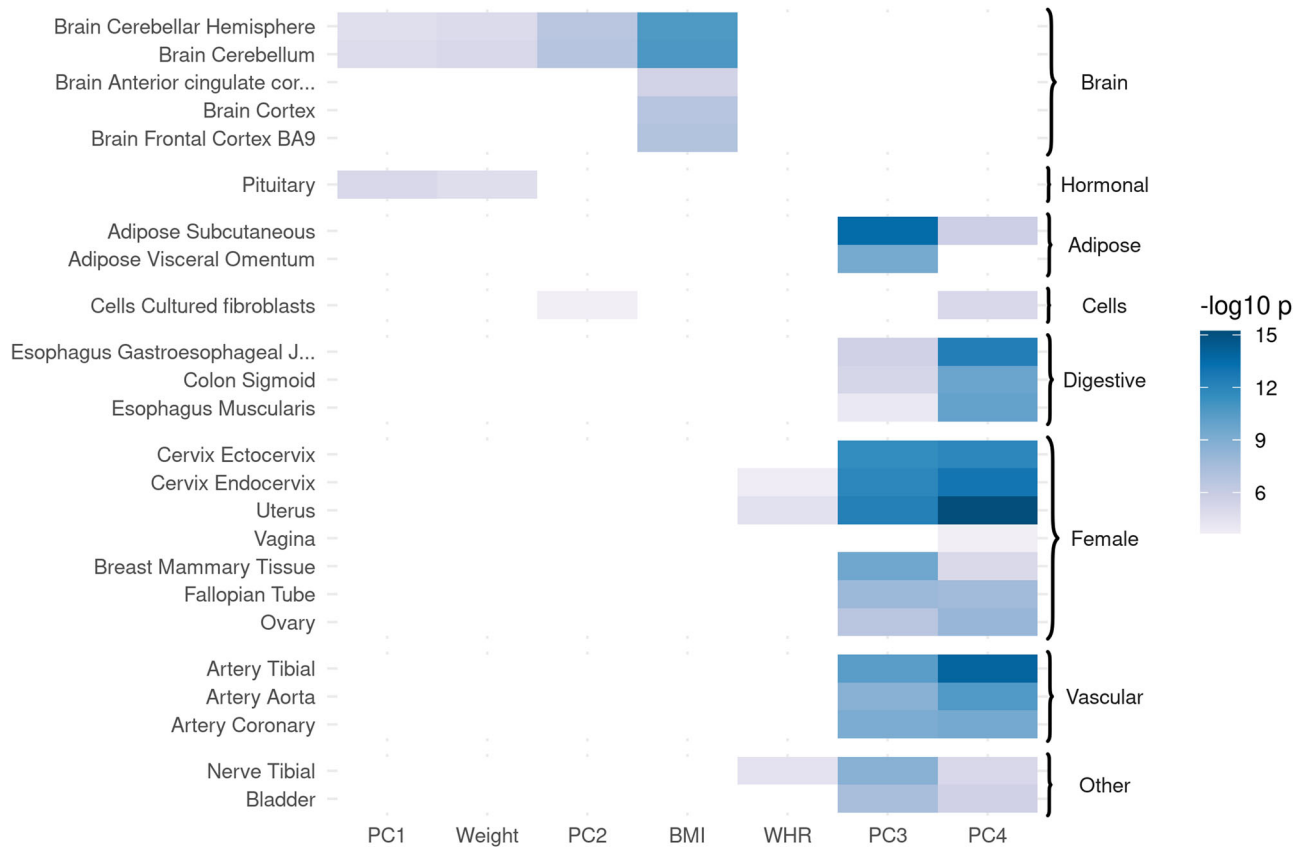


Fig. 3 Body size and accumulation of body fat were mainly enriched for genes expressed in the brain, while the others were enriched for a broader range of tissues. Enrichment of traits and principal components (PCs) for tissue-specific gene expression (negative log 10 *p*-values). Genome-wide SNP effect *p*-values were analyzed using MAGMA on GTEx v8 data (54 tissues). Results not significant after Bonferroni correction are masked in white. Traits with no significant enrichment results are hidden for clarity (full results are available in Supplementary Data 6).



Fig. 4 Single and composite traits increase the risk of multiple diseases. Mendelian randomization causal effects of traits and principal components (PCs) on a selection of diseases on a standardized scale. The 95% confidence interval of the effect is indicated in brackets. Effects that were not significant at the Bonferroni-corrected threshold ($p < 4.3 \times 10^{-5}$) are colored in white. The full list of effects can be found in Supplementary Data 11.

1.6% (95% CI: 1.2–1.9), hypertension by 4.2% (95% CI: 3.3–5.0), hypercholesterolemia by 3.0% (95% CI: 2.4–3.6), and IHD by 1.6% (95% CI: 1.2–2.0).

PC4 (lean mass) had few consequences on health, only significantly increasing the risk of nerve disorders and diseases,

carpal tunnel syndrome, and joint disorders (Fig. 4, Supplementary Data 11).

For comparison, a 1 SD increase in BMI (5.1 kg/m² in women, 4.2 kg/m² in men) increased the risk of diabetes by 3.8% (95% CI: 3.4–4.2), hypertension by 9.9% (95% CI: 8.8–11.0),

hypercholesterolemia by 3.0% (95% CI: 2.3–3.7), and IHD by 1.8% (95% CI: 1.4–2.2). Weight and WHR had similar, if somewhat weaker, effects (see Supplementary Data 11). Although in many cases these effects exceed those of individual PCs, they remain less than the cumulative (summed) effects of the four PCs (Supplementary Figs. 5–25).

Effects on continuous health outcomes. Many continuous health indicators were affected by both PCs and traits, in a manner largely consistent with expectations (Supplementary Data 12).

Consistent with the increased risk of diabetes, PCs 1–3 all increased the levels of glycated hemoglobin in blood, though the effect was strongest for PC2 ($b = 0.13$, 95% CI: 0.11–0.15). Glucose levels were similarly affected, though the effects were weaker. All three PCs also increased triglyceride levels (b between 0.095 and 0.24) and decreased HDL cholesterol (b between -0.22 and -0.16), but only PC1 decreased LDL cholesterol ($b = -0.076$, 95% CI: -0.094 to -0.059) and total cholesterol. They all increased blood pressure (systolic and/or diastolic), with the strongest effects from PC2 ($b > 0.15$). Plasma concentrations of several liver enzymes, such as γ -glutamyltransferase (GGT) and alanine aminotransferase (ALT), were found to be increased as well, possibly indicating liver damage.

PCs 1 and 2 also increased levels of cystatin C and decreased albumin, possibly indicative of impaired kidney function²³, whereas the effects of PC3 were in the opposite direction (decreasing cystatin C and increasing albumin). C-reactive protein (CRP) was also increased by PCs 1 and 2 ($b_1 = 0.17$, 95% CI 0.15–0.20; $b_2 = 0.25$, 95% CI 0.22–0.27) but unaffected by PC3. PCs 1–3 all strongly decreased levels of sex hormone-binding globulin (SHBG) with the strongest effects from PC2 ($b = -0.19$, 95% CI: -0.21 to -0.16), as well as testosterone where the effects of PCs 1 and 3 were stronger ($b_1 = -0.065$, 95% CI -0.086 to -0.043 ; $b_3 = -0.058$, 95% CI -0.082 to -0.034).

PC4 mainly increased creatinine levels while decreasing CRP ($b = -0.090$, 95% CI: -0.11 to -0.067) and the maximum heart rate during fitness test. The levels of triglycerides and the liver function markers GGT and ALT were slightly decreased but only alkaline phosphatase reached Bonferroni-corrected significance ($b = -0.068$, 95% CI: -0.09 to -0.046).

The effects of BMI were comparable to a combination of PCs 1 and 2, increasing levels of glycated hemoglobin ($b = 0.2$, 95% CI 0.17–0.22) and glucose ($b = 0.12$, 95% CI 0.10–0.15), as well as increasing the levels of triglycerides while decreasing both HDL ($b = -0.29$, 95% CI -0.32 to -0.26) and LDL ($b = -0.059$, 95% CI -0.082 to -0.037) cholesterol. Systolic and diastolic blood pressure was also increased, as were levels of CRP ($b = 0.33$, 95% CI 0.3–0.36) and liver function markers such as ALT and GGT. SHBG ($b = -0.24$, 95% CI -0.26 to -0.21) and testosterone ($b = -0.11$, 95% CI -0.13 to -0.084) were also decreased.

Effects on lifestyle factors. PC1 slightly reduced socio-economic status (SES), as shown by increased Townsend deprivation index ($b = 0.033$, 95% CI: 0.017–0.049). This was accompanied by a longer working week, as well as an increase in smoking and alcohol consumption, particularly spirits (Fig. 5, Supplementary Data 13). The duration and frequency of physical activity, as well as walking pace, were all decreased in favor of increased time spent using the computer. PC1 also increased daytime dozing and napping but decreased snoring.

The effects of PC2 on SES were similar but much more pronounced (Fig. 5, Supplementary Data 13), not only associated with increased Townsend deprivation index ($b = 0.057$, 95% CI: 0.037–0.076) and the likelihood of having a job involving heavy physical work ($b = 0.11$, 95% CI: 0.086–0.14), but strongly linked

to decreased income, fluid intelligence score, and education (all $b < -0.14$). Accompanying these were lifestyle changes similar to those of PC1, increasing smoking and the frequency of alcohol consumption, with a decrease in wine in favor of spirits and alcohol being taken more often outside of meals. Although PC2 increased the duration of walks ($b = 0.05$, 95% CI: 0.031–0.069) and vigorous activity, the duration of walking for pleasure was decreased ($b = -0.044$, 95% CI: -0.062 to -0.025), as were several other measures of physical activity, namely the frequency of stair climbing and the usual walking pace. Unlike PC1, PC2 decreased the time spent using the computer in favor of time spent watching TV.

PC3 increased alcohol intake frequency ($b = 0.046$, 95% CI: 0.028–0.065) and napping during the day ($b = 0.035$, 95% CI: 0.018–0.053) (Fig. 5, Supplementary Data 13).

PC4 only associated with an increased length of the working week ($b = 0.043$, 95% CI: 0.024–0.062) (Supplementary Data 13).

The effects of BMI on lifestyle were most similar to those of PC2, linked to decreased SES (Townsend deprivation index $b = 0.078$, 95% CI 0.058–0.099) and physical activity while increasing smoking and alcohol consumption. Additional comparisons of overall effects on lifestyle can be seen in Supplementary Figs. 26–46.

Effects on diet. PC1 was associated with greater reported variation in diet ($b = 0.052$, 95% CI: 0.034–0.069) and increased consumption of healthy foods such as fresh fruit, vegetables, and water (Supplementary Data 14). Consumption of coffee was also increased ($b = 0.087$, 95% CI: 0.068–0.11).

PC2 was similarly associated with reportedly increased variation in diet ($b = 0.084$, 95% CI: 0.065–0.10), as well as salt added to food ($b = 0.045$, 95% CI: 0.026–0.064). Other changes in diet reflect increased consumption of cheaper meats, namely pork and poultry, while decreasing intake of grain products (bread and cereal), cheese, and dried fruits (Supplementary Data 14).

PC3 was associated with increased bread consumption ($b = 0.035$, 95% CI: 0.017–0.053).

PC4 increased fruit intake, both fresh and dried ($b > 0.041$), while decreasing processed meat intake (Supplementary Data 14).

Effects of BMI on diet were similar to those of PC2, such as increased variation in diet ($b = 0.12$, 95% CI: 0.097–0.14) and decreased intake of grain products and cheese (Supplementary Data 14). In addition, some effects were similar to those of PC1, including increased vegetable and fruit consumption, as well as coffee intake ($b = 0.11$, 95% CI: 0.080–0.13).

Sex-specific effects. The genetic effects of the selected IVs on PCs 1, 2, and 4, as well as BMI and weight, were not significantly different between men and women, which made it possible to obtain unbiased sex-specific causal effects (see “Methods”). Briefly, if the IV effects on the exposure do not differ between sexes, summary statistics for the opposite sex can be used for the exposure while using those of the sex of interest as outcome. Those of PC3 (and WHR) were stronger in women ($p_{PC3} = 6.9 \times 10^{-46}$, $p_{WHR} = 1.4 \times 10^{-159}$) and were unsuitable for this purpose. Full results are available in Supplementary Data 15–22 and shown in Supplementary Note 5.

None of the PCs had significantly different effects on diet and lifestyle in men and women, though PC2 did have a tendency for stronger effects on disease risk in men (TLS slope = 1.27), mainly driven by diabetes ($b_m = 0.15$, 95% CI 0.12–0.17; $b_f = 0.080$, 95% CI 0.052–0.11; $p_{diff} = 1.8 \times 10^{-3}$) and heart diseases such as IHD ($b_m = 0.11$, 95% CI 0.081–0.103; $b_f = 0.059$, 95% CI 0.037–0.08; $p_{diff} = 4.7 \times 10^{-3}$). However, many blood molecular traits and

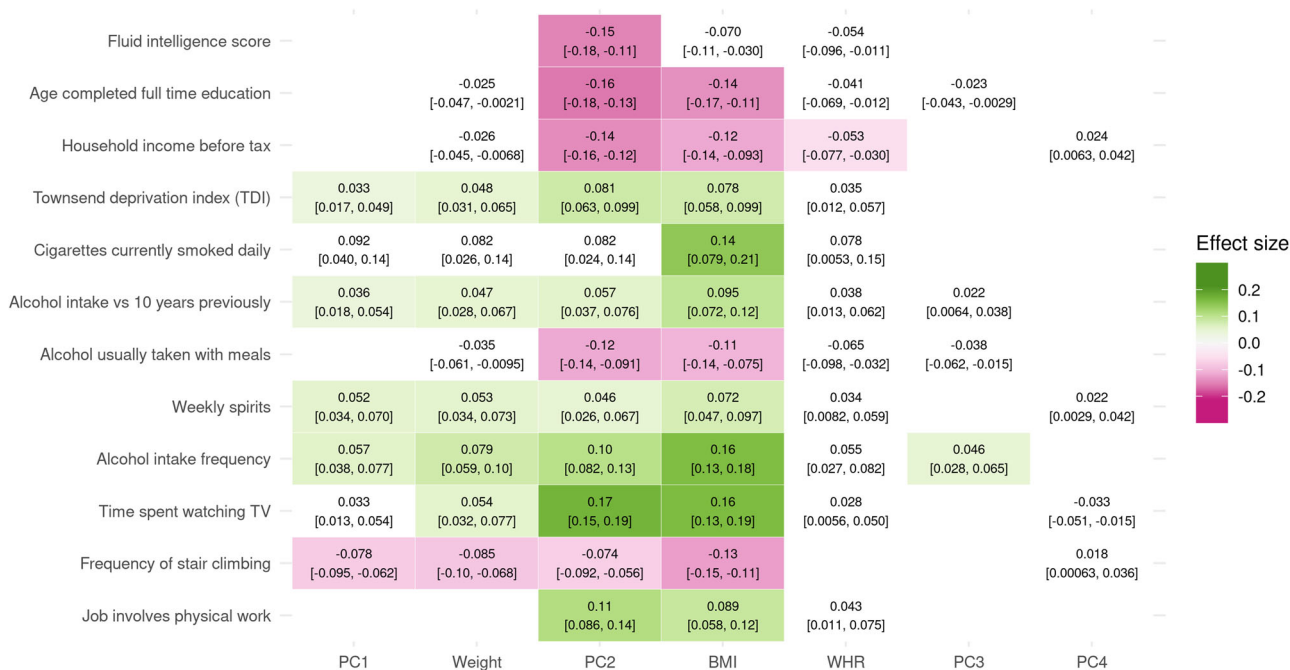


Fig. 5 Single and composite traits affect many aspects of lifestyle. Mendelian randomization causal effects of traits and principal components (PCs) on a selection of lifestyle factors on a standardized scale. The 95% confidence interval of the effect is indicated in brackets. Effects that were not significant at the Bonferroni-corrected threshold ($p < 7.1 \times 10^{-5}$) are colored in white. The full list of effects can be found in Supplementary Data 13.

other continuous health outcomes were differently affected by PCs in men and women.

Testosterone in particular was differently affected between sexes, strongly decreased in men by both PC1 ($b = -0.15$, 95% CI -0.18 to -0.12) and PC2 ($b = -0.18$, 95% CI -0.21 to -0.14), whereas in women it was unaffected by PC1 and increased by PC2 ($b = 0.082$, 95% CI 0.05 – 0.11). SHBG was similarly affected by PC2 in both sexes, but the decrease caused by PC1 was twice as strong in women ($b_m = -0.084$, 95% CI -0.12 to -0.051 ; $b_f = -0.17$, 95% CI -0.21 to -0.14 ; $p_{diff} = 1.5 \times 10^{-4}$).

Other differences include PC1 causing a stronger decrease in albumin in women ($b_m = -0.095$, 95% CI -0.12 to -0.068 ; $b_f = -0.17$, 95% CI -0.20 to -0.15 ; $p_{diff} = 6.0 \times 10^{-5}$) and increasing glycated hemoglobin only in men ($b = 0.12$, 95% CI 0.089 – 0.15). The effects of PC2 in men were stronger for ALT, GGT, and pulse rate, and an increase in total protein was only seen in men.

The effects of BMI were again similar to those of PCs 1 and 2. The increase in glycated hemoglobin was also stronger in men and was accompanied a larger increase in glucose as well ($b_m = 0.16$, 95% CI 0.13 – 0.20 ; $b_f = 0.083$, 95% CI 0.05 – 0.12 ; $p_{diff} = 5.2 \times 10^{-4}$). LDL was decreased only in men ($b = -0.099$, 95% CI -0.13 to -0.067). The opposite effects on testosterone levels were also observed for BMI ($b_m = -0.24$, 95% CI -0.27 to -0.21 ; $b_f = 0.072$, 95% CI 0.035 – 0.11 ; $p_{diff} = 7.9 \times 10^{-36}$).

Bi-directional MR. In addition to determining which environmental factors may affect body shape, we were interested in feedback loops. The stringency of the selection criteria for IVs, namely that they be GWS for the sex used as exposure, resulted in few exposures significantly affecting the body traits or PCs. Several exposures had unrealistically large estimated effect sizes (>0.4), likely due to confounding, and are not considered here. For example, forced vital capacity (FVC), which is dependent on lung—and therefore body—size, showed strong bi-directional associations with most body traits. The main results are summarized here, full results are listed in Supplementary Data 23–26.

While PC1 reduced cholesterol, it was in turn reduced by LDL and total cholesterol ($b \leq -0.068$), as well as apolipoprotein A and diagnosis of hypercholesterolemia. PC1 also showed a positive feedback loop with cystatin C but a negative one with IGF-1, reducing IGF-1 levels in the blood ($b = -0.13$, 95% CI: -0.16 to -0.11) but being increased by it ($b = 0.042$, 95% CI: 0.024 – 0.061).

PC2 showed a positive feedback loop with CRP and mutual negative loop with creatinine. PC2 was also reduced by IGF-1 ($b = -0.071$, 95% CI: -0.089 to -0.052).

PC3 was decreased by hypertension ($b = -0.07$, 95% CI: -0.11 to -0.035), forming a negative feedback loop. It also formed a positive feedback loop with diabetes (as did WHR), although this seems to be due to a negative effect of diabetes on hip circumference (HC, $b = -0.16$, 95% CI: -0.26 to -0.067) rather than an increase in waist circumference (WC). PC3 also had a positive feedback loop with triglycerides (from decreased HC), and mutual negative effects on apolipoprotein A and HDL cholesterol (from decreased WC). We also found the negative effects of PC3 on testosterone and SHBG to be reciprocal ($b \leq -0.058$).

PC4 had a positive feedback loop with creatinine and was also increased by IGF-1 ($b = 0.12$, 95% CI: 0.11 – 0.14). SHBG and testosterone both decreased PC4, though for testosterone this appears to have been driven by an effect in women only ($b = -0.09$, 95% CI: -0.11 to -0.065).

BMI increased CRP in a positive feedback loop, similar to that of PC2, and reciprocal negative effects with LDL and total cholesterol ($b \leq -0.035$), similar to PC1.

DXA-based measures. DXA and MRI technologies provide increased accuracy in measurements of body composition and have been suggested to provide a clearer picture of obesity and its consequences on health²⁴. We were able to impute many such traits based on the subset of ~5000 UK Biobank participants with these phenotypes (Supplementary Data 27). Several of these, including both trunk and android tissue fat percentages (both

predicted with r^2 above 0.75), had effects on disease comparable to BMI and PC2 (correlation $r \geq 0.97$) but with a tendency for slightly larger effects (TLS slopes of 1.17 and 1.16, respectively). However, with fewer GWS SNPs (~28% fewer than BMI), these phenotypes revealed no additional diseases affected by obesity (see Supplementary Data 28).

Prediction of disease risk. To further emphasize the biological relevance of PCs over single traits, we compared the accuracy of prediction of obesity-related disease, namely diabetes, high cholesterol, and hypertension, in the UK Biobank. Given the independence of estimated PC effects, these could be combined in a weighted linear fashion to obtain the 4-PC prediction and were compared with the single trait prediction from BMI and WHR using receiver operating characteristic (ROC) curves, in a subsample of the UK Biobank similar to that used by the Neale lab.

As expected, both BMI and WHR were effective in predicting these obesity-related diseases, as indicated by areas under the ROC curve (AUC) ranging from 0.62 to 0.74 “within-sample.” BMI was more accurate in predicting hypertension (0.66 vs. 0.64, $p = 1.3 \times 10^{-72}$), while WHR was more accurate for high cholesterol (0.65 vs. 0.62, $p = 4.9 \times 10^{-101}$) and slightly more for diabetes (0.74 vs. 0.73, $p = 3.0 \times 10^{-5}$). In all cases, the PC-based predictors significantly outperformed both BMI and WHR with AUCs of 0.68 for hypertension, 0.67 for high cholesterol, and 0.78 for diabetes (all $p < 2.3 \times 10^{-183}$, Supplementary Figs. 67–68, Supplementary Data 29).

To explore the transferability of PCs, we replicated these results out-of-sample in 76,756 non-“white British” UK Biobank participants, who had therefore been excluded from the original summary statistics. Despite the considerable potential for disparities between these samples, the AUCs for high cholesterol and hypertension were almost exactly identical to those obtained above (Fig. 6, Supplementary Fig. 69, Supplementary Data 29), with PC-based predictors significantly outperforming either trait ($p \leq 2.0 \times 10^{-43}$). The accuracy of PCs and BMI in predicting diabetes out-of-sample decreased to 0.75 and 0.68, respectively, and although WHR was no less accurate out-of-sample (AUC = 0.74) it remained less accurate than PCs ($p = 5.1 \times 10^{-11}$).

We also tested whether a weighted linear combination of BMI and WHR (analogous to that of the 4-PC predictor) would improve the prediction, and found that in most (5 out of 6) cases the AUC of the combined predictor performed worse than the better of the two traits alone. In all cases, this predictor did not exceed the accuracy of the PCs ($p < 3.1 \times 10^{-12}$). This is likely due to the fact that these risk factors are not independent and hence their conferred risk is not simply additive.

Discussion

By combining multiple traits through PCA, we found that more than 99% of genetically determined variation in body shape and size (as defined by the 14 selected traits) can be summarized using four PCs affecting (1) body size, (2) adiposity, (3) predisposition to abdominal fat deposition, and (4) lean mass. PCs 1–3 showed some similarity with weight, BMI, and WHR, respectively, and the latter two especially have been widely studied in the context of obesity. However, these traits are highly correlated and partially redundant, making it difficult to understand their individual contributions to health outcomes, whereas PCs are orthogonal by design and their effects additive. The PCs obtained here also share some similarities with the average PCs (avPCs) obtained by Ried et al.¹¹ by meta-analyzing PCs from individual-level phenotypic data. Specifically, avPC1 increased all body measures (though the impact of height was weaker than for PC1); avPC2 showed opposing directions for height and BMI/WHR (though WHR had much more importance than BMI); avPC3 was dominated by WHR (however both height and BMI contributed as well); and loadings for avPC4 showed an increase in weight and BMI despite decreased waist and hip circumference, suggesting increased density of body mass and consistent with decreased body fat percentage. The robustness of the obtained PCs to changes in the trait selection or the use of male- or female-specific data supports the hypothesis that these represent true biological mechanisms underlying the shared variance across these anthropometric traits. Furthermore, the molecular basis of PCs 3 and 4 in particular appear to be more homogenous, as shown by enrichment for genes expressed in many tissues, such as SAT and VAT but also in the digestive, reproductive, and vascular systems, which were

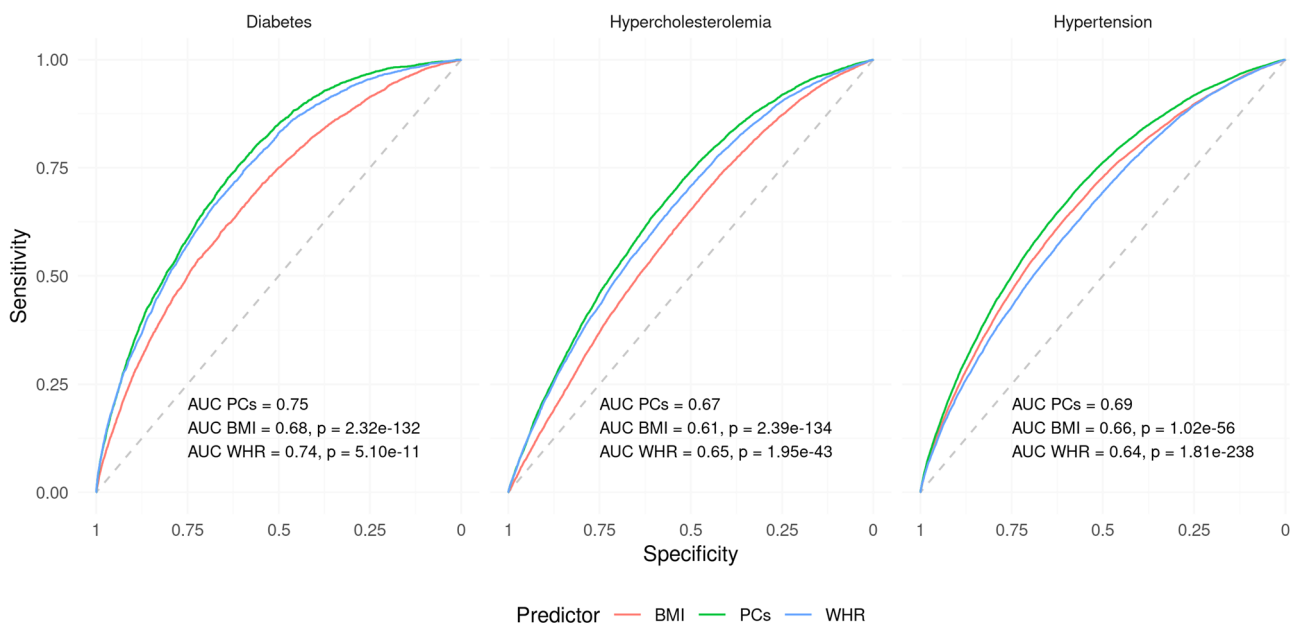


Fig. 6 Principal components (PCs) improve prediction of obesity-related diseases out-of-population. Receiver operating characteristic (ROC) curves for PC-, BMI-, and WHR-based prediction of diabetes, hypercholesterolemia, and hypertension out-of-sample/-population. The indicated p -values for the difference between the PC- and single trait-based curves were obtained using the DeLong method. AUC area under the curve.

not picked up by individual traits. This is also reflected in the much broader spectrum of pathways enriched only for these PCs.

Leveraging the orthogonality of PCs, we can dissect the etiology of obesity-related health and lifestyle consequences by comparing the effects of different PCs. Both PCs 1 and 2 embody the excessive accumulation of body mass and their enrichment for genes expressed in the brain and involved in neuronal signaling is consistent with findings from BMI²⁵ suggesting that behavioral changes are likely one of the major factors underpinning heritable susceptibility to obesity (reviewed in ref. 26). Shared effects between PC1 and PC2 highlight diseases whose etiology involves elements common to both, such as metabolic overexertion in the case of diabetes²⁷ or the physical burden of a larger body in the case of arthrosis²⁸. Differences between their effects can provide insight into disease etiology which single traits cannot. For example, Hyppönen et al.²⁹ performed a phenome-wide MR analysis to assess the effects of BMI on a number of diseases and found that BMI increased the risk of IHD, cardiac ar-/dysrhythmia, and diseases of the veins (phlebitis and thrombophlebitis), among others. However, our results show that these do not all arise through the same mechanisms, IHD being increased by PCs 2 and 3 but not PC1, consistent with the predominant role of adipose tissue-related dyslipidemia in this disease⁸, whereas cardiac arrhythmia and diseases of the veins were only increased by PC1. BMI is also a known risk factor for prolapsed disc³⁰, though it is unknown whether this occurs due to mechanical overstraining, dysregulation of the metabolic and immune systems, or a combination of these³¹. That PC1 alone increased the risk of prolapsed disc strongly supports the role of mechanical stress over any form of lipotoxicity.

The decreased SES associated with PC2 is consistent with what has been reported for BMI^{32,33}, although recent results suggest that this is due to residual population stratification and non-genetic familial effects^{34,35}. Many other differences in lifestyle can be considered as concurrent with this change in SES (e.g., reduced education and income) whereas others are likely secondary to these (e.g., alcohol taken outside of meals³⁶ and increased consumption of cheaper meats). Despite a moderate association with lower SES, PC1 was not associated with any of these secondary/concurrent changes, though the increased time spent using a computer may be indicative of white-collar occupations rather than the physical jobs associated with PC2. The similarity of the effects of both PCs on smoking and alcohol consumption suggests mechanisms independent of SES. For example, increased smoking may reflect the use of smoking as a strategy for weight loss, leveraging the appetite suppressant effects of nicotine³⁷, and is consistent with the observational correlation found between the number of cigarettes smoked and the risk of obesity³⁸. Although we cannot exclude residual population stratification, the increased alcohol consumption found for both PC1 and PC2 suggests a directionality of causal effects which is rarely considered, as obesity is typically viewed as a consequence of alcohol consumption rather than the cause³⁹. The lack of robust genetic instruments to explore the distinct effect of different alcoholic beverages render MR analysis suboptimal to resolve such reverse causations. The increased variation in diet and salt added to food may reflect a greater desire for palatable foods, possibly due to altered or reduced activation of the reward system which has been shown to occur in obese individuals (reviewed in ref. 40). The healthy aspects of the dietary shift from PC1, e.g., greater consumption of fruits and vegetables, could possibly represent attempts to eat healthier, however, this is likely to be highly confounded by biased reporting⁴¹. The weak to non-existent effects of cholesterol and other health risks or diseases on these measures of body size suggests that any lifestyle changes upon

disease diagnosis are generally minimal and likely insufficient to have a real impact.

PC3, i.e., predisposition to abdominal fat deposition, was of particular interest as a hypothesis-free emergent component affecting body shape while remaining independent of its size or composition. This is in contrast to WHR which is highly correlated with BMI and body fat percentage, particularly in men⁴². The increased waist circumference at the expense of that of the hips suggests a change in body fat distribution, with a shift from subcutaneous (contributing to both hips and waist) to VAT (contributing to waist alone). This is supported by the enrichment for terms related to energy homeostasis and adipose tissue function, similar to those previously reported for WHR adjusted for BMI⁴³, and the impacts on health which appear specific to lipotoxicity and are similar to what has been found for genetic variants affecting subcutaneous-to-visceral adipose tissue ratio^{12,13}. The implication of genes expressed in adipose tissues, particularly subcutaneous, is in line with the hypothesis that adipose tissue function, in particular its expandability, may be a critical factor in determining the distribution and impact of excess adiposity⁷ and its study may provide pharmaceutical avenues which may reduce health risks associated with obesity. The enrichment of both PCs 3 and 4 (as well as WHR) for genes expressed in several tissues of the digestive tract is not entirely surprising⁴⁴, however, their strong enrichment for all female-specific tissues tested was unexpected. It is possible that these tissues, and the genes expressed therein or associated sexual hormones, contribute to this phenotype, which could explain the increased variance of PC3 (and more generally WHR) observed in women compared to men. Additional studies will be required to determine how these genes and tissues are involved in the regulation of body fat distribution.

The additivity implied by the independence of the PCs' causal effects provides a straightforward estimation of disease risk through linear combination, which proved more accurate than single trait-based prediction. Although causal effects are not ideally suited for within-sample outcome prediction, since they do not take advantage of non-causal correlates, their basis in causality proved reliable even out-of-population. The reason for the drop in accuracy seen for out-of-population PC- and BMI-based prediction of diabetes is unclear, though a number of factors may contribute, such as ethnicity-dependent differences in susceptibility, lifestyle, or ascertainment bias⁴⁵, which exceed the scope of this study.

The systematic application of MR over a broad range of phenotypes carries the risk of violating its assumptions. Despite the steps taken to avoid any such violations or mitigate their impact, some sources of bias may remain inherent to the methods employed. GWAS effect estimates themselves have been shown to be biased due to population effects, such as parental effects, population stratification, and assortative mating⁴⁶. Although such biases may be present in the exposure, they would only bias MR estimates if they affect the outcome in a similar fashion^{34,35}, which is difficult to test in a real setting. This study is also subject to other limitations of the MR approach, such as the estimated univariable causal effects summarizing the consequences of a lifetime exposure of a difference in phenotype at the population level. For example, the effect of 1 SD difference in weight (~14 kg) in the population is not equivalent to a sole gain/loss of 14 kg, since this difference in the population distribution will be accompanied by a difference in average height as well (as illustrated by PC1 loadings). The causal effect being modeled linearly, it also averages what may be different effects in different strata of the population, i.e. if the causal effect is non-linear. For example, the effects of BMI on all cause mortality have been

shown to be non-linear⁴⁷, although the linearity of effects on specific diseases such as diabetes remains unclear⁴⁸. In the case of non-linear effects, the MR causal effect estimates can be interpreted as an average effect at the population mean.

In summary, we established a scale-preserving method for the linear combination of traits at the summary-statistics level, mimicking a GWAS performed on a standardized composite trait but circumventing the need for individual-level data. Furthermore, this allows the combination of traits across different cohorts, since the covariance of effect size estimates between different traits can be estimated from summary statistics via cross-trait LD score regression intercept⁴⁹. We also showed that cross-sex MR can reduce the bias by avoiding sample overlap, while preserving much of the statistical power from a potential one-sample MR using the full sample, still relying only on publicly available summary statistics. Applying these to the PCs of anthropometric traits, body size, adiposity, abdominal fat distribution, and lean mass, we showed that their distinct effects on health- and lifestyle-related outcomes can aid in understanding the etiology of the consequences of obesity. These components can be visualized and their effects on hundreds of outcomes compared through the shiny app which can be downloaded following the instructions at <http://wp.unil.ch/sgg/pca-mr/>. Finally, we showed their effectiveness in predicting obesity-related diseases, confirming the additivity of the conferred risks and emphasizing their relevance to disease etiology. Although we identified four PCs here, larger sample sizes or more detailed anthropometric traits may achieve a finer scale identification of obesity subtypes.

Methods

Data and phenotype selection. We used GWAS summary statistics data derived from the UK Biobank¹⁶, a cohort of ~500,000 participants aged 37–73 (median 58), recruited between 2006 and 2010. UK Biobank obtained ethics approval from the North West Multi-centre Research Ethics Committee (MREC) and obtained informed consent from all participants. Phenome-wide GWAS in UK Biobank data were performed and the summary statistics made available by the Neale Lab (<http://www.nealelab.is/uk-biobank>) for three different sex groups: men, women, and both sexes combined. This includes effect estimates for 13.7 million SNPs which were tested for association with 4203 unique phenotypes across 361,194 unrelated, white British individuals (167,020 men, 194,174 women). All were adjusted for age, age², sex, age * sex, age² * sex, and the top 20 genetic PCs to correct for population stratification. We excluded SNPs from the HLA region (chr6: 28,477,797–33,448,354, www.ncbi.nlm.nih.gov/grc/human/regions/MHC?asm=GRCh37). For continuous phenotypes, we used effect estimates for the inverse rank-normalized trait. Summary statistics for WHR were calculated in the UK Biobank across a similar sample of 378,139 unrelated, white British individuals (175,155 men, 202,984 women) and correcting for the same covariates after inverse rank-normalization. For binary traits, we divided the effect estimates and standard errors by the square root of the variance of the trait (that of the analyzed sample where provided, otherwise across a similar subset of the UK Biobank). As such, all effects are expressed on the SD scale and comparable with continuous traits. The linear models used to estimate the SNP-binary trait association provide well-calibrated *p*-values, as long as rare SNPs (MAF < 0.1%) are not evaluated for a trait with a highly imbalanced case fraction (<10%) or as long as the product of MAF and disease prevalence exceeds 0.0001⁵⁰. Since we analyzed diseases only with >1% prevalence, any SNPs with MAF > 1% are safe to use as instruments. The largest fraction of low-frequency (MAF < 1%) IVs were observed for height (88/1704 SNPs, i.e., 5.2% vs. max 2.8% for other traits/PCs) and the lowest disease prevalence were between 1 and 2%. Still the resulting causal effects from MR analysis between height and these diseases did not change appreciably upon the exclusion of IVs below 1% MAF (Supplementary Note 7).

In total, 232 phenotypes that had at least one GWS SNP ($p < 5 \times 10^{-8}$) were divided into five mutually exclusive categories: body measures (14 phenotypes), continuous measures of health (37 phenotypes), dietary habits (19 phenotypes), diseases (108 phenotypes), and lifestyle factors (54 phenotypes), which are briefly described below. The full list of selected phenotypes can be found in Supplementary Data 1.

Body measures included BMI, height, weight, hip and waist circumference, and WHR, as well as bioimpedance-derived fat and lean mass estimates in arms, legs, trunk, and overall body fat percentage. For arms and legs, summary statistics were available for left and right sides. As these were almost identical, the statistics for the left side alone were used. Basal metabolic rate was also included in body measurements as it is derived from bioimpedance measures.

Continuous health outcomes included the biomarker panel of the UK Biobank, including 34 biomarkers measured in either blood or urine, as well as systolic and diastolic blood pressure (BP), heart rate at rest and during effort, and FVC.

Dietary habits were obtained from a food frequency questionnaire (FFQ). UK Biobank also includes more specific questions such as 24-h recall (food consumed on the last day) which is more reliable but less representative, as well as specific questions such as the type of bread or milk typically consumed, however, these were not included in the present analysis.

Disease summary statistics in the UK Biobank were available for self-reported disease status, ICD-10-classified hospital diagnoses, and diseases curated by the Neale Lab in collaboration with the FinnGen consortium (www.finnngen.fi). We included data from both self-reported answers and the FinnGen curated diseases, excluding the raw ICD10 diagnoses which were considered less informative. We included any diseases which had a prevalence of at least 1% in the analyzed sample.

Lifestyle included both environmental factors and lifestyle choices, mainly relating to physical activity, alcohol consumption, smoking, sleep, work, and SES.

We restricted SNPs to those in common between the summary statistics from Neale and the UK10K reference panel used for LD pruning. We also removed SNPs with a minor allele frequency below 0.001, resulting in 9,675,947 SNPs. For each trait, SNPs were pruned separately using *plink* v1.90b6 with the UK10K European LD panel to obtain independent MR instruments. SNPs were considered independent if separated by more than 10 Mb or the linkage disequilibrium was $r^2 < 0.01$.

High-accuracy body composition measurements. In addition to bioimpedance measurements, the UK Biobank provides other body composition phenotypes derived from more accurate methods, namely DXA and MRI. Unfortunately, the sample sizes for these phenotypes were too low (~5000 individuals) for their inclusion as body phenotypes in the PCA. We did, however, investigate whether approximating these traits through linear combination of other available traits (with regression weights calculated in the UK Biobank) could provide a hypothesis-driven alternative to the hypothesis-free PCA approach. We were able to estimate 57 body composition measurements using the same 14 anthropometric and bioimpedance-based traits with varying accuracy (see Supplementary Data 27). 18 out of 57 had an r^2 above 80%, including abdominal SAT, and another 18 had r^2 above 70%, including VAT. Others, such as bone mass and liver fat percentage could not be accurately approximated using the included traits.

Principal component analysis. PCA was performed on a matrix of effect estimates of independent SNPs on the 14 body traits described above. We selected all SNPs with a GWS effect ($p < 5 \times 10^{-8}$) on at least one of the body traits and pruned them using the same procedure as for single traits (distance > 10 Mb or $r^2 < 0.01$). SNPs were prioritized according to the highest rank within any significantly associated trait, i.e., for each trait, all SNPs significantly associated with it were ranked by *p*-value; then each SNP has attributed a priority based on the highest rank obtained with any significantly associated trait. This was done to avoid the well-powered traits (e.g., height) overshadowing other traits and driving the SNP selection. Any missing effect estimates in the resulting matrix were set to 0. Data were neither centered nor scaled prior to the PCA, as the effect estimates were standardized and should therefore have zero mean and their variance for a given trait is informative (distantly related to trait heritability).

PCA was then performed on the resulting matrix of genetic effects to obtain the PC loadings, which were then used to calculate SNP-PC associations across the entire genome. These were then individually pruned to obtain the final set of SNPs for each PC.

Although the SNP-phenotype associations are standardized (effects are on an SD/SD scale), a linear combination of effects would yield non-standard effects since the resulting traits do not have a variance of 1. To maintain comparable effects between PCs and traits, the PC loadings should be scaled such that the resulting β s would represent SNP effects on composite traits with unit variance, which would typically involve performing the rotation on the phenotypes, standardizing, and rerunning a GWAS on the resulting composite trait with variance 1. Instead, we can consider the linear regression coefficients for trait *j* calculated as:

$$\hat{\beta}_j = (G^T G)^{-1} G^T y_j, \quad (1)$$

where y_j is the vector of values for trait *j* and *G* is the genotype matrix, and $\hat{\beta}_j$ is a weighted sum of *n* variables:

$$y_C = w_1 y_1 + \dots + w_n y_n, \quad (2)$$

where w_j is the weight for trait *j*. The effect size for the composite trait can then be calculated as:

$$\hat{\beta}_C = (G^T G)^{-1} G^T y_C = (G^T G)^{-1} G^T (w_1 y_1 + \dots + w_n y_n) = w_1 \hat{\beta}_1 + \dots + w_n \hat{\beta}_n, \quad (3)$$

The expected variance of variable y_C can be written as⁵¹

$$\text{Var}(y_C) = w^T K w, \quad (4)$$

where *K* is the covariance matrix of the *n* traits composing y_C , which in the case of standardized trait variables, is the correlation matrix. Knowing the variance of the

composite trait, we can simply rescale the vector of weights such that the expected variance in the composite output variable y_C is equal to one, resulting in standardized effect sizes. Applying this to the PC loadings, we rescaled the weights as follows:

$$v_i = \frac{w_i}{\sqrt{w_i^T K w_i}}, \tag{5}$$

where w_i and v_i are the unadjusted and adjusted (trait) loading vectors for PC i , respectively, and K is the pairwise phenotypic correlation matrix of body trait phenotypes, which was calculated on a subset of the UK Biobank similar to that used for the summary statistics (i.e. unrelated white British individuals) and the phenotypes were corrected for the same covariates as used by the Neale lab prior to calculating the correlation. The effect estimates were then calculated using these adjusted loadings, i.e.

$$\hat{\beta}_i = \hat{B} \cdot v_i, \tag{6}$$

where \hat{B} is the matrix of genetic effects on the 14 anthropometric traits. The corresponding standard errors were calculated following the formula for a weighted sum of random variables⁵¹:

$$\hat{\sigma}_{i,j}^2 = w_i^T \cdot \Sigma_j \cdot K \cdot \Sigma_j \cdot w_i, \tag{7}$$

where $\hat{\sigma}_{i,j}$ is the standard error for the association of SNP j with PC i and Σ_j is a diagonal matrix with the standard error of the association of SNP j with each trait. Since we use only UK Biobank summary statistics, the correlation between effect estimates simplifies to the phenotypic correlation between the traits. The advantage of this approach is that we do not need to calculate the composite trait and run a GWAS but can directly compute the association summary statistics.

Tissue specificity and pathway enrichment. We tested PC- and trait-associated SNPs for the enrichment of tissue-specific genes using the MAGMA tool for gene set analysis¹⁷ through the FUMA v 1.3.5e interface¹⁸ using the default parameters, with a Bonferroni-adjusted p -value threshold. We used gene expression data in 54 tissues from GTEx v8¹⁹ as well as brain development and age data from BrainSpan²⁰.

We used the molecular pathway gene sets defined by DEPICT²¹ and tested for enrichment using PASCAL²². As with most gene-based pathway enrichment methods, we assume that most regulatory variants for a given gene are in close (50 kb) physical proximity to the gene body⁵², hence we may ignore more distant regulatory variants.

Mendelian randomization. We used IVW MR to estimate causal effects of body traits and PCs on all non-body traits, as well as the reverse. MR mimics a randomized controlled trial (RCT) where the treatment corresponds to the random allocation of an exposure-associated allele, called an instrumental variable (IV)⁵³. MR relies on three key assumptions to infer causality: (1) the IV is associated with a change in the exposure; (2) the IV is independent of the outcome, except through its association with the exposure; (3) the IV is independent of any confounders of the exposure-outcome association. If these assumptions are verified, MR provides an unbiased estimate of the causal effect of the exposure on the outcome and can be done using summary statistics data alone. We used IVW MR as the default method for all our analyses, but we compared IVW causal effect estimates to those obtained from weighted median-based MR to ensure robustness.

Cross-sex MR. In our case, the use of summary data from the UK Biobank for both the exposure and outcome effect sizes (i.e., full sample overlap) would lead to a bias in MR causal effect estimate in the direction of the observed correlation of the phenotypes⁵⁴. To circumvent this, we used the existing summary statistics for two non-overlapping samples from the Neale Lab, those for men and women separately. Each sex was used as exposure on the other as outcome and then both causal effect estimates were meta-analyzed (using inverse variance weighting). This removed the correlation between the error terms of the effect estimates of IVs on exposure and outcome, considerably reducing the bias from sample overlap, while minimizing loss of power. This would lead to a slight bias away from zero even when there is no true difference in the genetic effects on the exposure in men and women, given that:

$$\left| \alpha \cdot \frac{\beta_m / \beta_f + \beta_f / \beta_m}{2} \right| \geq |\alpha|, \tag{8}$$

where α is the common causal effect (for men and women), β_m and β_f are the coefficients of association for a given IV with the exposure for men and women, respectively. This bias increases if there is a difference between sexes in the strength of association with the exposure, but its magnitude remains small. For example, if the SNP effects on the exposure were consistently 1.37 times stronger in one sex, the relative theoretical bias would only be 5%. In practice, however, the IVs were filtered for genome-wide significance in each sex prior to being used as exposure, resulting in fewer IVs being selected in the sex with weaker effects (on the exposure). This reduced the power and increased the variance in the overestimated causal effect, which was correspondingly down-weighted by the IVW meta-

analysis. For example, we found PC3 to have a strong sex-specificity, with effects on exposure tending to be ~2.2 times stronger in women, which would theoretically lead to a 34% bias away from the null. However, this yielded 242 GWS IVs in women but only 33 in men. This increased the variance in the male exposure-female outcome causal effect estimate by a median of 6.8-fold, resulting in a corresponding down-weighting of the overestimated causal effect. In this example, we expect a meta-analyzed estimate which is biased towards the null by ~32% (since the IVW causal effect estimate is expected to be $((6.1/7.1) * (1/1.7) + (1/7.1) * 1.7) * \alpha = 0.68 * \alpha$). Importantly, since the bias is multiplicative, this introduces no bias in the absence of causal effect, indicating it will not affect the type I error rate. In practice, the causal effect estimates from cross-sex MR were closer to those obtained from a two-sample setting, whereas the standard IVW MR effect estimates were generally overestimated (see Supplementary Notes 8–9).

Although the PCs were constructed using combined-sex summary statistics, the selected IVs were filtered for genome-wide significance in the sex-specific summary statistics prior to MR analysis to avoid weak instrument bias inflating causal effect estimates. This may slightly exacerbate Winner's curse, inflating the SNP-exposure association, which would result in a small bias towards the null. In the presence of a true causal effect, however, the SNP-outcome association (being assessed in the same sample) may be proportionally increased, which would mitigate this bias. Such biases are difficult to correct for without using three independent samples.

For each sex, the GWS IVs associated with the exposure formed the initial set of IVs. Those with significantly larger effects on the outcome than on the exposure were removed, as these would indicate a violation of MR assumptions (likely reverse causality and/or confounding). The effect sizes being on a standardized scale, they were compared directly using a one-sided t -test and removed if the magnitude of the effects in the outcome were significantly greater ($p < 0.05$). To avoid unreliable causal effect estimates, MR was only performed if at least 10 IVs remained. This was done using the *TwoSampleMR* R package v0.5.4⁵⁵.

The MR causal effect estimates from individual IVs were tested for heterogeneity using Cochran's Q test:

$$Q_i = \frac{(\beta_{i,out} - \beta_{MR} \cdot \beta_{i,exp})^2}{\sigma_{out}^2 + \beta_{exp}^2 \cdot \sigma_{MR}^2 + \beta_{MR}^2 \cdot \sigma_{exp}^2 + \sigma_{MR}^2 \cdot \sigma_{exp}^2}, \tag{9}$$

where $\beta_{i,exp}$ and $\beta_{i,out}$ are the coefficients of association of SNP i with the exposure and outcome, respectively, with $\sigma_{i,exp}$ and $\sigma_{i,out}$ the corresponding standard errors, and β_{MR} and σ_{MR} are the MR estimate and standard error of the causal effect of the exposure on the outcome. The test statistic Q_i follows a χ^2 distribution with 1 degree of freedom. If any of the IVs used had an associated $p < 10^{-3}$, the most heterogeneous one (with the lowest p -value) was removed. If at least 5 IVs remained, the MR was then repeated with the remaining IVs. In practice, few SNPs were filtered as outliers (on average <1%), though we also explored alternative methods for dealing with heterogeneity, namely performing no filtering, using weighted median MR with or without filtering, or using exact Q statistics⁵⁶ instead of Cochran's Q (see Supplementary Note 10). All of these produced similar effect estimates, though those from weighted median MR tended to be ~6% lower and lacked the statistical power to find many effects.

Due to the standardized scales of IV effect estimates, the resulting causal effects are on a scale representing the SD change in the outcome for a change of 1 SD in the exposure.

To account for multiple testing, we adjusted the p -value threshold with Bonferroni correction for the number of tests within the exposure-outcome category pair. For example, the p -value threshold for the effect of BMI (body phenotype) on type 2 diabetes (disease) is $0.05/(\text{number of body phenotypes} * \text{number of diseases})$.

Although the forward MR analyses using anthropometric traits (whether individual or composite) as exposures were well-powered to find many associations, the reverse/bi-directional MR analyses often lacked sufficient GWS IVs. This can be seen in Supplementary Data 30, showing the total explained variance by all GWS SNPs for each phenotype in each sex group. Only continuous health outcomes had IVs with a reasonably large (combined) contribution to the phenotypic variance. Those of phenotypes from other categories explained at most 3.5% in the combined-sex group (hypertension), though frequently much less, which was further reduced by the restriction for genome-wide significance within the sex used as exposure (to a maximum explained variance of 2.4% for hypothyroidism in women).

Sex-specific MR. The simplest method to test for sex-specific effects would be to use sex-specific summary statistics for the sex of interest for both exposure and outcome, however, this would bias results away from the null due to correlated errors in the exposure and outcome. In most cases, we were instead able to obtain unbiased causal effect estimates using the exposure summary statistics from the opposite sex on those of the sex of interest in the outcome. This relies on the strength of the association between the IVs and the exposure to be identical (or at least not systematically different) between sexes. We tested this using both paired Wilcoxon signed-rank test on the absolute genetic effects of the IVs and total least squares (TLS) regression. Exposures with non-significant results for both tests were considered identical in both sexes for this purpose and the summary statistics for the opposite sex were used for the exposure.

Using the GIANT consortium as a separate sample²⁵, we were able to confirm that the use of opposite-sex summary statistics for the exposure produced effect estimates which were highly consistent with those using two-sample same-sex MR (see Supplementary Note 11).

Comparison of causal effects. To test the significance of a difference in the effects of two exposures on a given outcome, we used a two-sided Z-test⁵⁷. We accounted for multiple testing using Bonferroni correction, adjusting for the total number of tests, i.e. the number of significant effects of either exposure on the category outcomes. For example, to test whether the effects of BMI on type 2 diabetes were different from those of WHR, the significance threshold would be 0.05/total number of diseases significantly affected by either BMI or WHR.

In some cases, it was useful to compare the causal effects of two exposures on all outcomes of a category (e.g., the effects of weight on disease risk compared to those of PC1) or the same exposure in different experimental settings (comparison of methods or sex-stratified analyses). In these cases, we obtained the slope estimate using TLS regression with no intercept, which considers error in both axes rather than ordinary least squares (OLS) which minimizes only the vertical offset. The standard error on the angle of the regression line (rather than the slope, which is not symmetrical and dependent on the phenotype placed on the y-axis) was computed using a jackknife procedure, which was then used to compare the obtained estimate with the null hypothesis that the true causal effects were identical (i.e., an angle of 45° or a slope of 1). The TLS procedure was performed using the *deming* R package v1.4⁵⁸.

Prediction of disease risk. The accuracy of disease prediction was assessed in-sample in the UK Biobank across 371,523 unrelated, white British individuals (199,699 women, 171,824 men) for diabetes, high cholesterol, and hypertension. PCs were calculated based on the 14 scaled and centered anthropometric traits for each individual. These were then combined in a disease-specific linear combination based on the estimated causal effect on the disease. Rather than ignoring PCs with non-significant effects, the effect estimates were IVW to account for the uncertainty of the effect:

$$d = \frac{\sum_{i=1}^4 \gamma_i p_i}{\sum_{i=1}^4 \sigma_i^2}, \quad (10)$$

where γ_i and σ_i are the estimated causal effect of PC i on the disease of interest and its standard error, respectively, and p_i is the individual's phenotypic realization of that PC. The combined predictor for BMI+WHR was obtained in an analogous manner using the estimated effects of the respective traits on the outcome of interest.

The accuracy of this predictor was assessed using ROC curves by comparing the AUC with those of BMI and WHR. The significance of the difference between the AUCs for each predictor was determined using DeLong's test in the *pROC* R package (version 1.16.2)⁵⁹.

Out-of-sample prediction accuracy was assessed in a sample of 76,756 UK Biobank participants (42,407 women, 34,349 men) who were not flagged as "in.white.British.ancestry.subset" in the sample QC file. This includes individuals who either did not self-report 'white British' or whose genetic ancestry (as determined by the genomic PCs) was dissimilar from other white British individuals. This sample thereby excluded all the individuals included in the original analysis and any relatives in the UK Biobank. The phenotypes were scaled according to the original distribution, i.e. the mean and SD of the original sample were used to scale individuals without relying on the distribution in the second sample. Using the values of the second distribution for scaling changed little (the AUC increased by at most 4.4×10^{-4}).

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

Data availability

UK Biobank summary statistics for WHR can be obtained at wp.unil.ch/sgg/pca-mr/, the others were downloaded from www.nealelab.is/uk-biobank. The external validation summary statistics were downloaded from https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files²⁵. Cross-trait phenotypic correlation and disease risk prediction were performed using data from the UK Biobank (application #16389). L.D. was calculated based on the UK10K data resource (EGAD00001000740, EGAD00001000741).

Code availability

The code for the shiny app and an example pipeline using this method (along with required data files) can be obtained from wp.unil.ch/sgg/pca-mr/.

Received: 3 June 2021; Accepted: 12 August 2021;

Published online: 13 September 2021

References

- Boles, A., Kandimalla, R. & Reddy, P. H. Dynamics of diabetes and obesity: epidemiological perspective. *Biochim. Biophys. Acta Mol. Basis Dis.* **1863**, 1026–1036 (2017).
- Ortega, F. B., Lavie, C. J. & Blair, S. N. Obesity and cardiovascular disease. *Circ. Res.* **118**, 1752–1770 (2016).
- Pischon, T. et al. General and abdominal adiposity and risk of death in Europe. *N. Engl. J. Med.* **359**, 2105–2120 (2008).
- Kaess, B. M. et al. The ratio of visceral to subcutaneous fat, a metric of body fat distribution, is a unique correlate of cardiometabolic risk. *Diabetologia* **55**, 2622–2630 (2012).
- Rosenquist, K. J. et al. Visceral and subcutaneous fat quality and cardiometabolic risk. *JACC Cardiovasc. Imaging* **6**, 762–771 (2013).
- Abraham, T. M., Pedley, A., Massaro, J. M., Hoffmann, U. & Fox, C. S. Association between visceral and subcutaneous adipose depots and incident cardiovascular disease risk factors. *Circulation* **132**, 1639–1647 (2015).
- Virtue, S. & Vidal-Puig, A. Adipose tissue expandability, lipotoxicity and the Metabolic Syndrome—an allostatic perspective. *Biochim. Biophys. Acta* **1801**, 338–349 (2010).
- Heymsfield, S. B. & Wadden, T. A. Mechanisms, pathophysiology, and management of obesity. *N. Engl. J. Med.* **376**, 1492 (2017).
- Zhang, M., Hu, T., Zhang, S. & Zhou, L. Associations of different adipose tissue depots with insulin resistance: a systematic review and meta-analysis of observational studies. *Sci. Rep.* **5**, 18495 (2015).
- Porter, S. A. et al. Abdominal subcutaneous adipose tissue: a protective fat depot? *Diabetes Care* **32**, 1068–1075 (2009).
- Ried, J. S. et al. A principal component meta-analysis on multiple anthropometric traits identifies novel loci for body shape. *Nat. Commun.* **7**, 13357 (2016).
- Yaghootkar, H. et al. Genetic evidence for a normal-weight 'metabolically obese' phenotype linking insulin resistance, hypertension, coronary artery disease, and type 2 diabetes. *Diabetes* **63**, 4369–4377 (2014).
- Yaghootkar, H. et al. Genetic evidence for a link between favorable adiposity and lower risk of type 2 diabetes, hypertension, and heart disease. *Diabetes* **65**, 2448–2460 (2016).
- Lotta, L. A. et al. Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. *Nat. Genet.* **49**, 17–26 (2017).
- Ji, Y. et al. Genome-wide and abdominal MRI data provide evidence that a genetically determined favorable adiposity phenotype is characterized by lower ectopic liver fat and lower risk of type 2 Diabetes, heart disease, and hypertension. *Diabetes* **68**, 207–219 (2019).
- 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).
- de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).
- Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
- Aguet, F. et al. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).
- Miller, J. A. et al. Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199–206 (2014).
- Pers, T. H. et al. Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* **6**, 5820 (2015).
- Lamparter, D., Marbach, D., Rueedi, R., Kutalik, Z. & Bergmann, S. Fast and rigorous computation of gene and pathway scores from SNP-based summary statistics. *PLoS Comput. Biol.* **12**, e1004714 (2016).
- Lang, J. et al. Association of serum albumin levels with kidney function decline and incident chronic kidney disease in elders. *Nephrol. Dial. Transpl.* **33**, 986–992 (2018).
- Marra, M. et al. Assessment of body composition in health and disease using bioelectrical impedance analysis (BIA) and dual energy x-ray absorptiometry (DXA): a critical overview. *Contrast Media Mol. Imaging* **2019**, 3548284 (2019).
- Locke, A. E. et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197–206 (2015).
- van der Klaauw, A. A. & Farooqi, I. S. The hunger genes: pathways to obesity. *Cell* **161**, 119–132 (2015).
- Halban, P. A. et al. β -cell failure in type 2 diabetes: postulated mechanisms and prospects for prevention and treatment. *Diabetes Care* **37**, 1751–1758 (2014).
- Duclos, M. Osteoarthritis, obesity and type 2 diabetes: the weight of waist circumference. *Ann. Phys. Rehabil. Med.* **59**, 157–160 (2016).
- Hyppönen, E., Mulugeta, A., Zhou, A. & Santhanakrishnan, V. K. A data-driven approach for studying the role of body mass in multiple diseases: a

- phenome-wide registry-based case-control study in the UK Biobank. *Lancet Digit Health* **1**, e116–e126 (2019).
30. Schroeder, G. D., Guyre, C. A. & Vaccaro, A. R. The epidemiology and pathophysiology of lumbar disc herniations. *Semin. Spine Surg.* **28**, 2–7 (2016).
 31. Weiler, C. et al. Histological analysis of surgical lumbar intervertebral disc tissue provides evidence for an association between disc degeneration and increased body mass index. *BMC Res. Notes* **4**, 497 (2011).
 32. Wang, Y. & Beydoun, M. A. The obesity epidemic in the united states gender, age, socioeconomic, racial/ethnic, and geographic characteristics: a systematic review and meta-regression analysis. *Epidemiol. Rev.* **29**, 6–28 (2007).
 33. Howe, L. D. et al. Effects of body mass index on relationship status, social contact and socio-economic position: Mendelian randomization and within-sibling study in UK Biobank. *Int. J. Epidemiol.* <https://doi.org/10.1093/ije/dyz240> (2019).
 34. Darrous, L., Mounier, N. & Kutalik, Z. Simultaneous estimation of bi-directional causal effects and heritable confounding from GWAS summary statistics. Preprint at *medRxiv* <https://doi.org/10.1101/2020.01.27.20018929> (2020).
 35. Brumpton, B. et al. Avoiding dynastic, assortative mating, and population stratification biases in Mendelian randomization through within-family analyses. *Nat. Commun.* **11**, 3519 (2020).
 36. Oksanen, A. & Kokkonen, H. Consumption of wine with meals and subjective well-being: a finnish population-based study. *Alcohol Alcohol* **51**, 716–722 (2016).
 37. Mineur, Y. S. et al. Nicotine decreases food intake through activation of POMC neurons. *Science* **332**, 1330–1332 (2011).
 38. Dare, S., Mackay, D. F. & Pell, J. P. Relationship between smoking and obesity: a cross-sectional study of 499,504 middle-aged adults in the UK general population. *PLoS ONE* **10**, e0123579 (2015).
 39. Traversy, G. & Chaput, J.-P. Alcohol consumption and obesity: an update. *Curr. Obes. Rep.* **4**, 122–130 (2015).
 40. Stice, E., Spoor, S., Ng, J. & Zald, D. H. Relation of obesity to consummatory and anticipatory food reward. *Physiol. Behav.* **97**, 551–560 (2009).
 41. Pirastu, N. et al. Using genetics to disentangle the complex relationship between food choices and health status. In *PLoS Genetics* <https://doi.org/10.1101/829952> (2019).
 42. Sulc, J., Winkler, T. W., Heid, I. M. & Kutalik, Z. Heterogeneity in obesity: genetic basis and metabolic consequences. *Curr. Diab. Rep.* **20**, 1 (2020).
 43. Shungin, D. et al. New genetic loci link adipose and insulin biology to body fat distribution. *Nature* **518**, 187–196 (2015).
 44. Winkler, T. W. et al. A joint view on genetic variants for adiposity differentiates subtypes with distinct metabolic implications. *Nat. Commun.* **9**, 1946 (2018).
 45. Goff, L. M. Ethnicity and type 2 diabetes in the UK. *Diabet. Med.* **36**, 927–938 (2019).
 46. Young, A. I. et al. Mendelian imputation of parental genotypes for genome-wide estimation of direct and indirect genetic effects. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.07.02.185199> (2020).
 47. Sun, Y.-Q. et al. Body mass index and all cause mortality in HUNT and UK Biobank studies: linear and non-linear mendelian randomisation analyses. *BMJ* **364**, 11042 (2019).
 48. Wainberg, M. et al. Homogeneity in the association of body mass index with type 2 diabetes across the UK Biobank: a Mendelian randomization study. *PLoS Med.* **16**, e1002982 (2019).
 49. Bulik-Sullivan, B. et al. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
 50. Loh, P.-R., Kichaev, G., Gazal, S., Schoech, A. P. & Price, A. L. Mixed-model association for biobank-scale datasets. *Nat. Genet.* **50**, 906–908 (2018).
 51. Johnson, A. R. & Wichern, D. W. Applied multivariate statistical analysis. *Biometrics* **44**, 920 (1988).
 52. Boix, C. A., James, B. T., Park, Y. P., Meuleman, W. & Kellis, M. Regulatory genomic circuitry of human disease loci by integrative epigenomics. *Nature* **590**, 300–307 (2021).
 53. Burgess, S., Butterworth, A. & Thompson, S. G. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet. Epidemiol.* **37**, 658–665 (2013).
 54. Burgess, S., Davies, N. M. & Thompson, S. G. Bias due to participant overlap in two-sample Mendelian randomization. *Genet. Epidemiol.* **40**, 597–608 (2016).
 55. Hemani, G. et al. The MR-Base platform supports systematic causal inference across the human phenome. *eLife* **7**, e34408 (2018).
 56. Bowden, J. et al. Improving the accuracy of two-sample summary-data Mendelian randomization: moving beyond the NOME assumption. *Int. J. Epidemiol.* **48**, 728–742 (2019).
 57. Paternoster, R., Brame, R., Mazerolle, P. & Piquero, A. Using the correct statistical test for the equality of regression coefficients. *Criminology* **36**, 859–866 (1998).
 58. Therneau, T. *deming: Deming, Theil-Sen, Passing-Bablok and Total Least Squares Regression, R package version 1.4.* <https://CRAN.R-project.org/package=deming> (2018).
 59. Robin, X. et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinforma.* **12**, 77 (2011).

Acknowledgements

This research has been conducted using the UK Biobank resource (#16389), which has been approved by the National Research Ethics Service Committee. The computations have been carried out on the HPC server of the Lausanne University Hospital. L.D. was calculated based on the UK10K data resource (EGAD00001000740, EGAD00001000741). Z.K. was funded by the Swiss National Science Foundation (31003A-143914 and 310030-189147). B.D. is supported by the Swiss National Science Foundation (NCCR Synapsy, project grant Nr. 32003B_135679, 32003B_159780, 324730_192755, and CRSK-3_190185) and the Leenaards Foundation. LREN is very grateful to the Roger De Spoelberch and Partridge Foundations for their generous financial support. T.O.K. was funded by the Novo Nordisk Foundation (NNF17OC0026848, NNF18CC0034900).

Author contributions

A.S. and J.S. preprocessed the data and performed the initial analyses. N.M. assisted with initial data preprocessing. B.D. contributed to the initial design. J.S. performed the remaining analyses. Z.K. and J.S. developed the methods. Z.K. designed and supervised the project. J.S. wrote the initial manuscript. A.S., N.M., C.A., E.M., L.D., B.D., T.O., P.J., R.F., and Z.K. provided critical feedback, helped interpret the findings and revise the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-021-02550-y>.

Correspondence and requests for materials should be addressed to Zoltán Kutalik.

Peer review information *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editor: George Inglis. Peer reviewer reports are available.

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

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



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

© The Author(s) 2021

2.2.1 Choice of methods

In addition to PCA, we tested several methods to define meaningful composite phenotypes or, alternatively, *obesity subtypes*².

NEURAL NETWORKS were investigated for their demonstrated power to produce meaningful classification and draw reliable conclusions from numerous variables with (individually) low information content, as has been shown in particular for image processing. Their performance in our initial attempts was less than stellar and did not exceed that of, e.g., random forests.

I suspect that the lack of effectiveness in our application was due to two specific reasons, the linearity of genetic effects and the low level of interactions across variants. Note that I do not imply that these limitations are intrinsic to genetic variants and their effects on biological systems, rather that they stem from the standard approach we applied to overcome the overwhelming nature of genetic data: the preliminary selection of independent and significantly associated variants as a first step. Indeed, using the entire genome without filtering would have been excessively costly in terms of computational resources and posed the risk of over-fitting. Restrictions can be imposed through adapted architecture (e.g. imposing a bottleneck layer would reduce over-fitting) however the computational intensity of full data usage nevertheless rendered this impractical. Inclusion of non-independent neighboring variants could have conceivably improved prediction/classification accuracy, however preliminary tests showed little gain.

The minimal advantages, combined with the drawbacks of the ‘black box’ nature³ of neural networks, led us to drop them after these initial attempts.

THE ITERATIVE SIGNATURE ALGORITHM (ISA) [64] was considered as an alternative to PCA. This method has the advantage of bi-clustering on both

² *Obesity subtypes* were intended to circumscribe groups of individuals who would possess (or be predisposed to) certain body shapes associated with specific health risks or benefits.

³ Neural networks are notoriously effective in classification or prediction, but understanding what determines a result is far from trivial.

dimensions of an input data matrix, i.e. it would provide a combination of traits which meaningfully distinguished a subset of individuals or variants, which could have led to the informed creation of obesity subtypes. In the end, ISA did not provide substantial improvements over alternative methods (within the context of body shape and obesity), whereas its stochastic nature requiring the pooling/averaging of similar bi-clusters and the semi-arbitrary threshold proved impractical.

FACTOR ANALYSIS (FA) was considered as a viable alternative to PCA. Once we had found that four components would likely suffice to describe most of the variance, FA could be used to create independent components in the same way as PCA. The explicit assumption of latent variables in FA was consistent with our interpretation of PCs. The main difference between FA and PCA, that individual variance which is not dependent on the shared factors is ignored by FA, ultimately made it undesirable for our application since a component of interest to us (PC3, predisposition to abdominal fat distribution) was discarded by FA. Although this might make sense in other applications where FA is more typically used and individual elements are of no particular interest on their own (e.g. making sense of the main drivers underlying patterns of answers in large questionnaires), we aimed to understand the driving factors defining the covariance structure underlying body morphology, regardless of whether we included redundant metrics for certain aspects.

CANONICAL CORRELATION ANALYSIS (CCA) AND THE PING-PONG ALGORITHM (PPA) were both tested as well. These methods are similar in that they are extensions of PCA and ISA, respectively, but make use of a secondary dataset in the construction of the composite traits or bi-clusters. For example, CCA finds linear combinations within one dataset which are maximally correlated with

corresponding linear combinations in the second dataset. Although these are powerful approaches, we encountered two issues.

The first derives from the hypothesis-driven nature of these approaches: the inclusion of outcome data in the creation of the composite traits would complicate any reliable conclusion as to the *causality* of a discovered association. Indeed, these methods would favor the selection of any pairs of correlated composite traits, including those whose correlation is due to pleiotropy, confounding, or reverse causation. This would increase the likelihood of violating MR assumptions and inflate the type I error rate (false positives).

The second problem came from our trait selection which aimed to include as many health- and lifestyle-related outcomes as possible. The inclusion of phenotypes with overlapping definitions led to very high correlation of their genetic effect estimates⁴ which in turn led to the methods attributing opposing weights to these traits to drive diminutive gains in a process similar to overfitting. This could have been overcome by the intentional selection of meaningfully distinct and (mostly) uncorrelated traits⁵.

⁴ For example, the genetic effects for ‘diseases of the respiratory system’ and ‘chronic obstructive pulmonary disorder’ had a correlation of $r = 0.97$ genome-wide.

⁵ In our hypothesis-free, data-driven approach, we simply included all diseases with a prevalence above 1%.

THE SELECTION OF PCA as our method of choice was due to the above-mentioned drawbacks of the alternatives, as well as other advantages. The independence of PCs enabled simple and meaningful comparisons between their effects, allowing us to disentangle the etiology of obesity-associated diseases and identify specific aspects of body shape which were driving the causal relationship. This also ensured that their effects were additive and could easily be combined to predict the risk of disease.

PCA is also widely used, making the results more easily interpretable and understandable by others. The non-stochastic nature of PCA and the absence of (potentially arbitrary) parameters both ensured the replicability of our results.

2.2.2 Choice of variance-covariance matrix

In addition to the choice of method, we considered alternatives for the variance-covariance matrix to analyze. The main concern was whether to use *phenotypic* or *genetic* information.

PHENOTYPIC DATA would provide several advantages, the main one being that it would ensure the independence of the composite traits at the phenotypic level, i.e. upon calculating the resulting values in the individuals. This approach would also provide insight into *all* the shared mechanisms leading to correlation in anthropometric traits, rather than simply genetic ones.

THE CHOICE TO USE GENETIC DATA was driven by its focus on the shared biological mechanisms underlying the correlation in these traits, rather than what was, from our point of view, environmental confounding. The non-independence of the resulting phenotypic PCs was less problematic because of the genetic basis of MR, ensuring that the estimated effects we obtained were independent even if the phenotypic values in the individuals were not.

We calculated the variance-covariance matrix based on the independent, GWS SNPs associated with the traits, which focused on the correlation of effects of genomic loci which were detectable in the sample we had. We could have instead used the genetic correlation between traits, which might have reduced the influence of statistical power within each trait (based on its heritability and genetic architecture, affecting the number of SNPs detected), although we would expect these two alternatives to converge with larger samples sizes. Given the hundreds of SNPs associated with each trait, the results would likely have been quite similar.

2.2.3 Obesity and the brain

My PhD originally aimed to assess and investigate the possible link between obesity and the brain which has been hinted at through various results, most prominent of which is the enrichment of BMI-associated SNPs for genes expressed in brain tissues [59, 60]. Unfortunately, the summary statistics we were relying on [88] for our approach proved unsatisfactory. Indeed, all brain *imaging-derived phenotypes* (IDPs)⁶ of the brain were adjusted for using the *head scaling factor*, a standard procedure to reduce head size-related variability between individuals which enables the estimation of, e.g., atrophy from cross-sectional data [89]. We found that the inclusion of this head scaling factor as a covariate led to overcorrection, inducing spurious associations with other anthropometric traits. For example, ‘uncorrected’ brain volume is expected to increase with greater standing height (linear regression: $slope = 0.55$, $p < 10^{-300}$), however they become negatively associated after ‘correction’ ($slope = -0.14$, $p = 10^{-51}$). It is currently unclear whether this is due to the assumed relationship between skull size and brain volume [90], or whether the SIENAX method for head scaling [89] will require refining for application to large datasets where small biases become relevant.⁷ Rather than risk an unknown amount of bias being introduced into our downstream analyses, we chose to drop the brain-related aspect of the project.

This decision was due to the time we estimated might be necessary to overcome these issues, the low statistical power from the small sample size available⁸, and the phenotypic proximity of brain morphometry and anthropometric traits which greatly increased the likelihood of violating MR assumptions. Had we chosen to pursue this aspect of the project, I might have attempted to model brain volume as a function of, e.g., height (and other covariates) to obtain a scaling factor which could provide a *normalized brain volume* which would be

⁶ *Imaging-derived phenotypes* refer to the processed phenotypes, e.g. total gray matter volume, which were derived from the raw scans.

⁷ For context, most brain imaging studies are comparatively small and the SIENAX method was described based on a dataset of 16 healthy individuals, resulting in scaling factors ranging from 1.05 to 1.50. The UK Biobank contains thousands of individuals, inevitably leading to scaling factors considerably beyond this range (0.80–1.94), for which the accuracy of the method has, perhaps, not been fully investigated. I suspect that the association between skull size and brain volume is stronger at the lower end of the range, i.e. the brain occupies closer to maximum capacity if the total available space is more limiting, though interrogating this hypothesis would require a healthy cohort of younger individuals (to reduce the impact of age-related atrophy).

⁸ The original GWAS was based on brain imaging from the 4574 unrelated white European individuals which were available then. At the time of this writing 42945 individuals have been *scanned* (some of which have already been subject to a follow-up scan), though it is unclear when the scans and the corresponding IDPs will be made available.

height-/body size-neutral for use as a covariate in the GWAS. This would of course pose additional challenges, especially for any use of these summary statistics in downstream MR analyses with respect to anthropometric traits. Once more individuals have received a follow-up scan, the (relative) change in brain morphometry may provide an interesting alternative avenue of investigation.

The mention of the partial results from brain IDPs was eventually removed from the manuscript [87] as their inclusion brought mostly confusion and lengthened the already sizable paper.

2.2.4 Contributions

Although I performed the majority of the work on this project, Anthony Sonrel made significant contributions to the early stages (much of the data preprocessing and early implementation and testing).

3

Non-linear causal inference

3.1 Polynomial MR, simulations and application

This project aims to extend MR to allow non-linear causal inference with a fully parametric approach to increase the statistical power over alternative methods, namely LACE [83]. The manuscript is still in preparation but is included below in its current form.

We developed PolyMR, a method which allows for the polynomial approximation of an arbitrary *causal function*¹ of the effect of an exposure on an outcome. We performed extensive simulations to test the validity of the inferred functions and compare its performance with that of one of the only currently available alternatives², LACE [83]. The results showed very good agreement between the inferred and true causal functions for both methods, however PolyMR provided greater power and accuracy in all tested situations.

We also tested an alternative method to the two OLS regressions using *maximum likelihood estimation*, however we did not observe any reduction in bias or error in the simulations and runtime was considerably increased. This alternative was therefore not pursued further (nor is it described in the paper).

Applying this method to data from the UK Biobank, we showed that non-linear causal effects underlie the dependence³ between anthropometric and continuous health traits.

The original method was developed by Zoltán and implemented in Matlab. I translated the orig-

¹ We define the causal function as the relationship which determines the part of the outcome which is (causally) dependent on the exposure, i.e. $y = f(x) + \epsilon$, where y is the outcome, x the exposure, and $f(\cdot)$ is the causal function.

² The other alternative is another control function-based approach termed SpotIV [84] which detects non-linearity based on differences of effects at different exposure levels rather than the inference of a causal function. This difference would have made comparisons difficult at best, but the implementation available at the time was also computationally demanding, slow, and not robust to changes in the settings. After initial testing, we decided against systematic comparison with it.

³ I deliberately use the term ‘dependence’ rather than ‘correlation’ here, as many trait pairs may be clearly dependent with very low correlation. For example, cholesterol levels show a pronounced inverted U-shaped association with BMI, which leads to close to zero correlation

inal code to R before investigating the limits, testing alternative approaches, refining the method, and eventually applying it to data from the UK Biobank. I have also written drafts of the Methods (based on Zoltán’s original draft), Results, and Discussion sections for the manuscript. Finally, I compiled the method and code into a preliminary R package currently on GitHub: <https://github.com/JonSulc/PolyMR>.

3.1.1 Methods

Let X and Y denote two random variables representing complex traits. We intend to use Mendelian Randomization (MR) to estimate a non-linear causal effect of X on Y . The genotype data of the SNPs to be used as instrumental variables (IVs) is denoted by G . To simplify notation we assume that $E(X) = E(Y) = E(G) = 0$ and $Var(X) = Var(Y) = Var(G) = 1$. The effect sizes of the instruments on X are denoted by β . Let us assume the following models

$$X = G \cdot \beta + \epsilon_x \quad (3.1)$$

$$Y = f_\alpha(X) + \epsilon_y \quad (3.2)$$

where the parametric function $f_\alpha(\cdot)$ denotes the shape of the causal relationship between X and Y and ϵ_x and ϵ_y are zero-mean errors. For simplicity, we assume that $f_\alpha(\cdot)$ is a polynomial—even if it is not it can be approximated by one with arbitrary precision over the range of the majority of values X can take. For example, if we intend to test a quadratic causal relationship, $f_\alpha(x) = \alpha_0 + \alpha_1 \cdot x + \alpha_2 \cdot x^2$. Thus, the model can be rewritten as

$$Y = \sum_{i=0}^k \alpha_i \cdot X^i + \epsilon_y \quad (3.3)$$

The above equation can be expanded to

$$Y = \sum_{i=0}^k \alpha_i \left(\sum_{j=0}^i \binom{i}{j} (G\beta)^j \cdot \epsilon_x^{i-j} \right) + \epsilon_y \quad (3.4)$$

We rely on the INSIDE assumption [71], which ensures that $G\boldsymbol{\beta} \perp \epsilon_y$. The error terms ϵ_x and ϵ_y can nevertheless be correlated because of a potential causal effect of Y on X (reverse causation) and/or due to confounders. Let us split ϵ_y into ϵ_x -dependent and ϵ_x -independent parts.

$$\epsilon_y = (\epsilon_y | \epsilon_x, \epsilon_x^2, \dots, \epsilon_x^l) + \tau_y = \sum_{i=0}^l r_i \cdot \epsilon_x^i + \tau_y \quad (3.5)$$

Since $\text{cov}(G\boldsymbol{\beta}, \epsilon_x) = \text{cov}(G\boldsymbol{\beta}, \epsilon_y) = 0$, the residual noise τ_y is independent of both $G\boldsymbol{\beta}$ and ϵ_x . As a consequence $\text{cov}(X, \tau_y) = \text{cov}(X - G\boldsymbol{\beta}, \tau_y) = 0$. This allows us to rewrite the main model equations as

The $X - G\boldsymbol{\beta}$ term used here is the previously discussed control function [85], see Section 1.4.

$$X = G \cdot \boldsymbol{\beta} + \epsilon_x$$

$$Y = \sum_{i=1}^k \alpha_i X^i + \sum_{i=0}^l r_i \epsilon_x^i + \tau_y = \sum_{i=1}^k \alpha_i X^i + \sum_{i=0}^l r_i (X - G\boldsymbol{\beta})^i + \tau_y \quad (3.6)$$

The advantage of this equation system is that the error terms (ϵ_x and τ_y) are uncorrelated and independent of the respective explanatory variables.

Let the realizations of the random variables X, Y, G be denoted by $\mathbf{x}, \mathbf{y}, G$, observed in a sample of size n . The parameters $\{\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{r}\}$ can be estimated by computing two ordinary least squares estimates, first estimating $\boldsymbol{\beta}$ using the first equation and substituting this into the second equation:

$$\hat{\boldsymbol{\beta}} = (G'G)^{-1} \cdot G'\mathbf{x} \quad (3.7)$$

$$\begin{pmatrix} \hat{\boldsymbol{\alpha}} \\ \hat{\mathbf{r}} \end{pmatrix} = (M'M)^{-1} \cdot M'\mathbf{y}, \quad (3.8)$$

where

$$M := [\mathbf{x}, \dots, \mathbf{x}^k, \mathbf{1}, (\mathbf{x} - G\hat{\boldsymbol{\beta}}), \dots, (\mathbf{x} - G\hat{\boldsymbol{\beta}})^l]$$

The special case of this approach when $k = l = 1$ is equivalent to the standard control function approach. For higher orders, the terms k and l are not required to be equal, as k represents the powers of

the $f_\alpha(\cdot)$ function describing the causal relationship between X and Y , whereas l concerns the order of the confounding/reverse causation.

Implementation

We implemented this method in R [91]. For the polynomial approximation of $f_\alpha(\cdot)$ function of the causal relationship, we used $k = 10$, iteratively eliminating coefficients which are not significant at a Bonferroni-corrected level (i.e. $0.05/a$, where a is the number of coefficients remaining), setting them to zero. We set l to be equal to the polynomial order of the function, retaining all terms up to l such that any association with higher orders of the exposure driven by confounding are properly accounted for. Once all remaining coefficients were significant, the non-linearity p-value is obtained with a likelihood ratio test (LRT), comparing the full model with that including only the linear effect but retaining all residual correction terms (r_j). The causally explained variance was determined as the difference in explained variance (r^2) between the full model and that excluding all $\alpha_j \cdot x_j$ terms (i.e. accounting only for potential confounding).

The polynomial approximation is the result of a multivariable regression, which also provides us with the variance-covariance matrix of the polynomial coefficients. From these, we can generate functions drawn from the same multivariable distribution to obtain the 95% confidence hull.

For comparison, LACE [83] was also implemented and polynomial approximation was obtained in an analogous fashion. The piecewise linear LACE approach was not tested here. We examined the limitations of the standard (1st order) control function approach by running PolyMR with $l := 1$, hereafter referred to as PolyMR-L1, in specific settings.

Simulations

We simulated data according to the following model

$$X = G \cdot \beta + q_x \cdot U + \epsilon_x \quad (3.9)$$

$$Y = f_\alpha(X) + q_y \cdot U + \epsilon_y \quad (3.10)$$

where U is a confounder drawn from a standard normal distribution. Columns of G were drawn from a binomial distribution with minor allele frequencies following a beta distribution with shape parameters equal to 1 and 3, and then normalized each column to have zero mean and unit variance. The genetic effects β_i were drawn from normal distributions based on the minor allele frequencies, specifically $\beta_i \sim \mathcal{N}(0, (p_i * (1 - p_i))^{-0.25})$, where p_i is the minor allele frequency of SNP i , and scaled such that the total explained variance matches the predefined heritability, i.e. $\sum \beta_i^2 = h^2$. These effect sizes are realistic and are according to the baseline LDAK heritability model (without functional categories) with a selection strength of -0.25 [92, 93]. For the basic settings, we included moderate confounding ($q_x = 0.2$ and $q_y = 0.5$) and a quadratic causal function ($f_\alpha(X) = 0.1X + 0.05X^2$). The heritability h^2 was set to 0.5, explained by $m = 100$ causal SNPs, and sample size was set to 100,000 individuals. Causal SNPs were filtered for genome-wide significance of their univariate effects in the simulated data prior to their use as IVs. Variations on these settings were tested, as shown in Table 3.1. Each combination of parameters was used to generate 1000 sets of data, to which we applied both PolyMR and LACE. We also tested PolyMR-L1 in the base settings, in the presence of weak quadratic confounding ($q_x \times q_{y2} = 0.04$), as well as in the absence of quadratic causal effect but with quadratic confounding ($q_x \times q_{y2} = 0.1$) creating a similar observed association between traits as in the base settings.

The theoretical 95% confidence hulls were compared to the empirical distribution of estimated models across simulations. At each percentile of

Causal functions	$f_{\alpha}(X)$
*Base settings	$0.1 \cdot X + 0.05 \cdot X^2$
Null	0
Linear effect	$0.1 \cdot X + \mathbf{0} \cdot X^2$
Stronger effect	0.3 · X + 0.1 · X ²
Weak quadratic effect	$0.1 \cdot X + \mathbf{0.01} \cdot X^2$
Cubic effect	$0.1 \cdot X + 0.05 \cdot X^2 + \mathbf{0.05} \cdot X^3$
Fourth order effect	$0.1 \cdot X + 0.05 \cdot X^2 + \mathbf{0.05} \cdot X^4$
3rd & 4th order effects	$0.1 \cdot X + 0.05 \cdot X^2 + \mathbf{0.03} \cdot X^3 + \mathbf{0.01} \cdot X^4$
Exponential effect	0.1 · e ^X
Square root effect	0.1 · <i>sgn</i> (X) · √ X
Sigmoid effect (1)	0.1 · $\frac{1}{1+e^{-X}}$
Sigmoid effect (2)	0.1 · $\frac{1}{1+e^{-2 \cdot X}}$
Sigmoid effect (3)	0.1 · $\frac{1}{1+e^{-3 \cdot X}}$
Other settings	
Strong confounding	$q_x = 0.5, q_y = 0.8$
Negative confounding	$q_y = -0.5$
*Quadratic confounding	$q_{y2} \cdot U^2$ term added to Y, where $q_{y2} \in \{0.1, 0.2\}$
*Quadratic confounding, linear effect	$0.2 \cdot U^2$ term added to Y, $f_{\alpha}(X) = 0.1 \cdot X$
Heritability & polygenicity	$h^2 \in \{0.2, 0.3, 0.5, 0.8\}$, $m \in \{20, 100, 1K, 5K, 10K\}$

Table 3.1: Causal functions and setting parameter combinations simulated for PolyMR. Modifications to the base setting’s causal function are indicated in **bold**. Asterisks (*) denote settings where PolyMR-L1 was also applied for comparison.

the exposure distribution, the size of the predicted 95% confidence intervals (CIs) was compared to the empirical one.

Application to UK Biobank data

The UK Biobank is a prospective cohort of over 500’000 participants recruited in 2006–2010 and aged 40–69 [54]. We tested for non-linear causal effects of anthropometric traits (body mass index [BMI], weight, body fat percentage [BFP], and waist-to-hip ratio [WHR]) on continuous health outcomes (pulse rate [PR], systolic blood pressure [SBP], diastolic blood pressure [DBP], and glucose, and LDL, HDL, and total cholesterol [TC] levels in blood), as well as the reverse. We also tested for effects of both BMI and age completed full time education on life expectancy. Since most participants in the UK Biobank are still alive, we scaled

the mother’s and father’s age of death (separately) and used the mean of these standardized phenotypes as a proxy for the individual’s life expectancy.

We selected 377’607 unrelated white British participants and all phenotypes were corrected for age, age², sex, age × sex, age² × sex, as well as the top 10 genetic principal components. With the exception of WHR, IVs were selected using the `TwoSampleMR` R package [version 0.5.5, 94] with default settings ($p < 5 * 10^{-8}$ and $r^2 < 10^{-3}$ or $d > 10^4$ kb) from the GWAS in the `ieugwasr` R package [version 0.1.5, 95] with the largest number of instruments overlapping our dataset. For WHR, we used a previously-performed GWAS on the aforementioned sample from the UK Biobank, adjusting for covariates as above [96].

For the purpose of comparison, we also used inverse-variance weighted MR and MR Egger on each of these exposure–outcome pairs. These were performed with the same IVs and (in-sample) association statistics using the `TwoSampleMR` R package [version 0.5.5, 94]. We also compared the results of standard PolyMR with those PolyMR-L1 to determine whether accounting for higher order confounding is necessary in real data applications.

3.1.2 Results

Simulations

We simulated a variety of settings, including many combinations of heritability and polygenicity in the exposure, sample size, and shape of the causal function $f_{\alpha}(\cdot)$ and confounding. Where the true underlying function was polynomial, our approach correctly captured its shape (Fig. 3.1) although a slight bias from confounding was introduced in certain settings with high polygenicity (> 1000 causal SNPs), high heritability ($h^2 = 0.8$), or strong confounding (where linear confounding explained $\sim 40\%$ of the exposure–outcome association) (Fig. 3.2). The distribution of this bias was affected by the

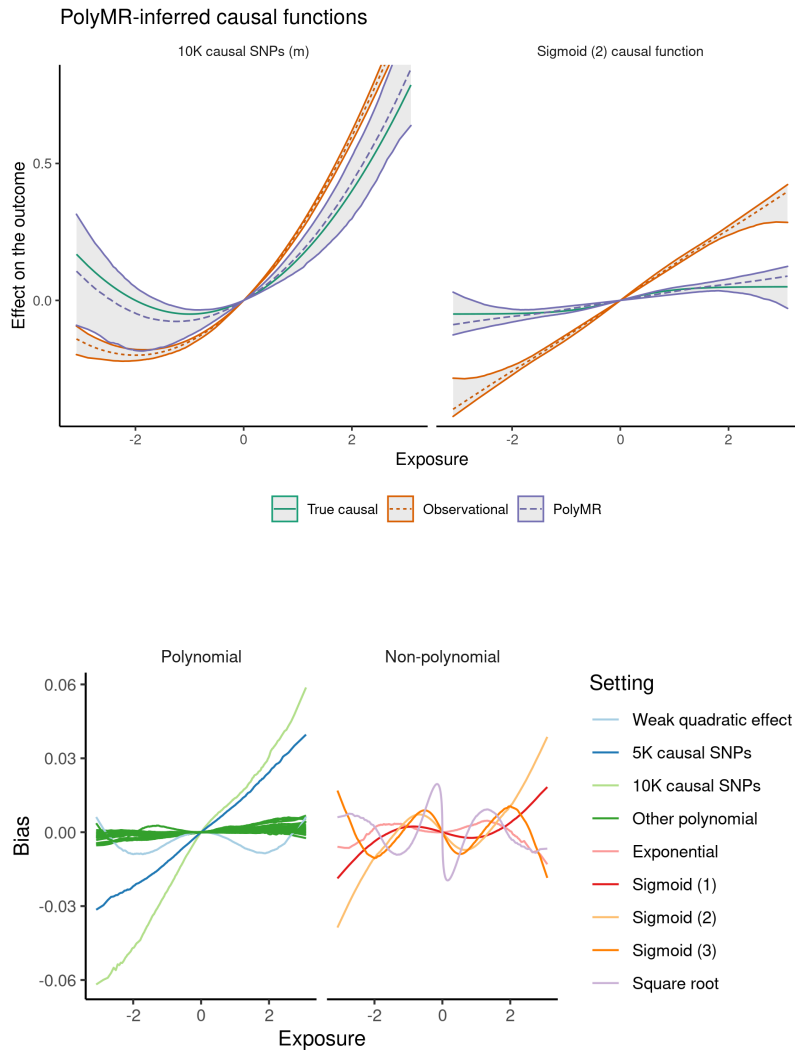


Figure 3.1: **PolyMR is able to recover the shape of the causal function.** The true causal function is shown in green (solid line). The observed association model is shown in orange (short-dashed) while that obtained using PolyMR is shown in purple (long-dashed). The hulls around the model curves show the 95% coverage hull across 1000 simulations. Shown here are (a) high polygenicity (10'000 causal SNPs accounting for the heritability of 0.3); and (b) a sigmoid causal effect ($f_{\alpha}(X) = 0.1 \cdot \frac{1}{1+e^{-2X}}$). The Y-axis shows the expected association with/effect of the exposure on the outcome, relative to the outcome level at the mean population exposure.

Figure 3.2: **Bias as a function of exposure across settings.** In settings with a polynomial causal function $f_{\alpha}(\cdot)$, slight bias from non-genetic confounding was induced under certain combinations of high polygenicity, high heritability, or strong confounding. The bias found in non-polynomial settings was expected due to the polynomial approximation approach.

shape of the confounding, i.e. in situations with quadratic confounding, the bias was quadratic with respect to exposure. In all simulation settings, this bias was nevertheless orders of magnitude smaller than both the causal effect and the confounding (e.g. Fig. 3.1 A). Although this bias was minimal with the standard PolyMR settings ($l = k$), quadratic confounding produced significant bias when higher orders of the control function term $(x - G\hat{\beta})$ were ignored (i.e. $l = 1$), which is the standard approach for control function use.

In the case of non-polynomial functions, PolyMR nevertheless provided reasonable estimates of the true shape of the causal function (Fig. 3.1 B). The bias introduced in these cases (Fig. 3.1 B, 3.2 B) is

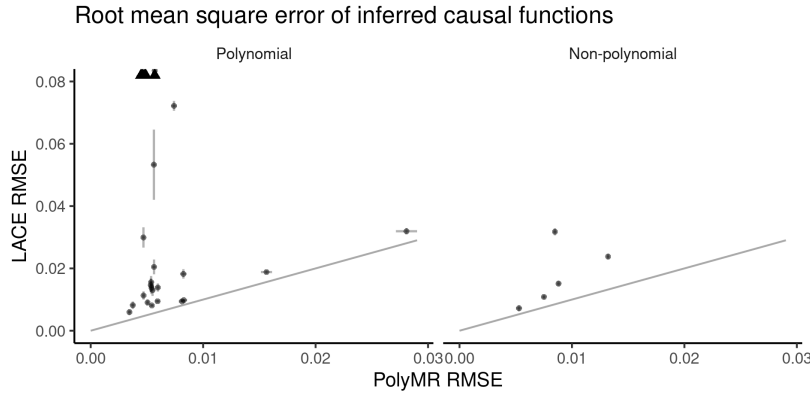


Figure 3.3: **PolyMR provided greater accuracy in the estimation of causal functions.** The root mean square errors (RMSEs) are shown for both PolyMR and LACE. Each point is the mean RMSE for a given setting with the error bars showing the 95% confidence interval (CI) of the true mean across simulations. Settings were split into polynomial and non-polynomial causal functions. Arrows in the polynomial plot indicate RMSEs which exceed the bounds of the plot.

consistent with expectations of polynomial approximation with limited power and dependent on the shape of the non-polynomial function.

LACE also produced some bias in estimating the causal function. The magnitude of the bias introduced by either method was dependent on the settings used. PolyMR yielded less biased estimates in the cases of high order causal functions whereas LACE was overall slightly less biased by high levels of confounding. However, in all settings tested, PolyMR provided lower root mean square errors (RMSEs) than LACE on average (Fig. 3.3), partly driven by smaller SEs.

To ensure that the variance estimated from the variance-covariance matrix of the model was correctly calibrated, we assessed the coverage of the 95% confidence interval. We did so by comparing the predicted 95% confidence intervals (CIs) of the curves with those derived empirically from repeated simulations. We found that in the case of most polynomial functions, the CIs were properly calibrated with the theoretical and empirical CIs being almost equal across most of the exposure distribution. Note that under some simulation settings (e.g. weak quadratic effects, where $\alpha_2 = 0.01$) allowing the polynomial degree to vary led to underestimation of the variance of the model due to post-selection inference (see [Supplementary Fig. 3.6 A](#)). If we consider only those simulations with the same polynomial order selected, the em-

pirical CIs were properly calibrated and very close to those predicted by the method (Supplementary Fig. 3.6 B).

UK Biobank

Given its favorable performance throughout all simulation settings, we applied the PolyMR method to data from the UK Biobank. We set out to estimate the causal effects of four anthropometric traits (BMI, weight, BFP, and WHR) on each of seven continuous traits commonly used as health biomarkers (PR, SBP, DBP, and the levels of glucose, LDL, HDL, and TC in blood). We also tested for reverse causal effects for these trait pairs, as well as any effects of BMI or education on life expectancy.

The effects of the anthropometric traits were qualitatively similar to one another and significant against all tested outcomes with significant non-linearity in most cases (~86%). Those of WHR and BFP tended to be more similar to one another, monotonically increasing DBP, SBP, and PR with linear to slightly non-linear effects. BMI also increased these traits overall, though the effects of BMI on DBP and SBP plateaued at around 2 SD above the population mean (~36.9 kg/m²) and the causal function for BMI on PR showed a positive slope for values between approximately 1 SD below to 2 SDs above the population mean (~22.7–36.9 kg/m²) with negative slopes beyond these. The effects of weight were weaker, but qualitatively similar to those of BMI. Glucose was increased by all of these, though the effects of a change in exposure were negligible below ~-1 SD for all traits and intensified at higher values. For example, the estimated slope of the standardized effect of BMI was 0.17 at the population mean but increased to 0.31 at +2 SD. The strongest non-linearity in the effects of anthropometric traits was found for total cholesterol, mainly driven by the LDL fraction (Fig. 3.4 A), where the causal function took a

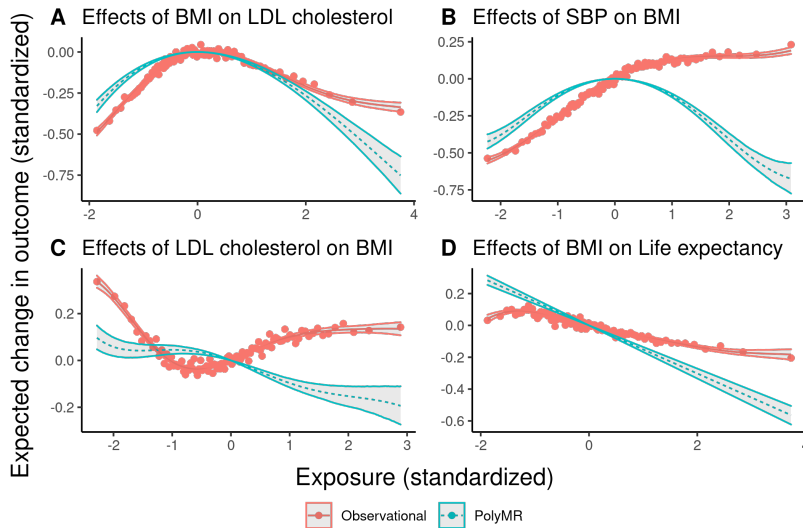


Figure 3.4: **Most tested causal effects have strong non-linear components in the UK Biobank.** The red points show the mean outcome plotted against the median exposure for each of 100 bins, split by covariate-adjusted exposure level. The red curve (solid) is the multi-variable regression model whereas the teal one (dashed) corresponds to the estimated causal function obtained using PolyMR. The hulls around both curves correspond to the 95% confidence interval.

strong inverted U-shape. In contrast to this, their effects on HDL were all monotonic decreasing.

The impact of PR was limited to weak linear effects on BFP and WHR. TC linearly decreased BMI, weight, and BFP, with no detectable effect on WHR. SBP and DBP both had inverted U-shaped causal functions for their effects on all anthropometric traits ($p < 1.6 * 10^{-47}$), with the effects on WHR being slightly weaker. Glucose levels had nearly no effect on most traits across most of the distribution, but drove strong reductions at higher values, with the exception of WHR, which was in fact slightly increased by glucose levels up to ~ 3 SD before being decreased at higher levels. HDL had a slight U-shaped effect on these traits, with a stronger increase for high values of the exposure on BFP and no increase in WHR.

Although the observational association of SBP/DBP and the anthropometric traits was mostly monotonic increasing, the estimated causal function on these had an inverted U-shape (Fig. 3.4 B), with slightly weaker effects on WHR. The causal effects of PR on anthropometric traits show a slight positive slope close to the population median, but the directionality switches at either extreme of the distribution. LDL cholesterol decreased the outcomes near monotonically, though the effect close to the

population median was weak to null (Fig. 3.4 C). The impact of glucose was slightly different across the anthropometric traits. Both BMI and weight were overall negatively affected by glucose levels (with weaker effects around zero). BFP was also decreased, although the effect was much weaker. WHR, however, was slightly increased by glucose levels up to ~ 3 SD before being decreased at higher levels. Note that the effect close to the population mean is likely driven by a decrease in hip circumference rather than an increase in the waist's, similar to what we've shown previously for the effects of diabetes risk and triglyceride levels on WHR-related metrics [87].

The effects of BMI and education on life expectancy are directionally as expected but we found no evidence of non-linearity. The BMI-life expectancy (causal) relationship was decreasing, though the intensity of the effect was greater than the observed association (Fig. 3.4 D), whereas higher education increased life expectancy.

The exclusion of higher order control function terms in PolyMR-L1 produced somewhat different inferred causal functions, with generally stronger non-linear components, resulting in inferred causal functions which were closer to the observed associations.

3.1.3 Discussion

We've developed a Mendelian randomization (MR)-based method, PolyMR, for the inference of (potentially) non-linear causal effect of an exposure on an outcome. This method accurately recovers the parameters where the causal function is polynomial in nature, otherwise providing a suitable polynomial approximation of the function shape. Although a very slight bias ($\sim 2\%$) was introduced in the most extreme cases of confounding (accounting for $\sim 40\%$ of the variance in the outcome), PolyMR significantly improve the power and accuracy of non-linear causal inference over

the semi-parametric approach employed by LACE [83]. Applying PolyMR to a selection of anthropometric and health-related traits in the UK Biobank [54], we found most of these to be significantly non-linear.

Assuming linearity in the presence of non-linear causal effects may have different consequences depending on the context, as well as the method used. Standard IVW MR [70] will provide an estimate of the slope at the population mean, while the inclusion of an intercept in MR-Egger [71] yields an effect closer to the average slope over the distribution, i.e. the non-linearity is treated as directional pleiotropy. Other methods designed to be robust to outliers or pleiotropy are liable to attribute some of the non-linearity (particularly of strong IVs) to these effects and “account” for this by excluding or down-weighting them in their respective ways. Even the “best” linear approximation will introduce bias in the presence of underlying non-linearity, the extent of which will depend on many factors such as the exposure distribution and the amplitude of the deviation from linearity. There is one factor worth discussing in this respect: the *monotonicity* of the causal function.

In the case of a monotonic causal function, the linear effect estimate can be considered an average causal effect over the range of the exposure and could be used as an adequate approximation in some cases. This would not affect the qualitative conclusion of the analysis, i.e. exposure X increases outcome Y , however the bias introduced may impact downstream analyses (e.g. mediation analysis) and affect predictions made for certain strata of the population. For example, BMI monotonically increased glucose levels in blood and linear effect estimates from IVW and MR-Egger (~ 0.16 – 0.17) are consistent with the slope of the effect at the population mean. However, the effect was nearly twice as large for individuals with severe obesity ($>35 \text{ kg/m}^2$) and close to zero for normal weight individuals, which is in line with what we would

expect biologically: individuals who are not overweight are able to maintain normoglycemia. A biologist (or a medical professional) might know not to apply the linear effect estimate to normal-weight individuals *in this specific example*, however this is not reflected in the statistical estimate and might result in poor public health guidelines. Just as importantly, the effect in people with extreme obesity may be dangerously underestimated. Conversely, the increasing monotonic function for the effects of BMI on SBP shows that weight loss (or gain) in extremely obese individuals will not have much of an effect on their blood pressure.

The case of non-monotonic causal functions is, predictably, worse. For example, the effect of BMI on LDL cholesterol has previously been estimated to be either null [97, 98] or weakly negative [87, 99], however PolyMR revealed a strong non-linear effect with an inverted U-shape. The close-to-null effect size estimated by linear MR methods likely occurs due to the proximity of the population mean to the maximum of the causal function, where the slope (i.e. linear effect point estimate) is close to zero. Although the conclusions of these studies are not contradictory (the absence of evidence is not evidence of absence), none of them accurately portray the effect BMI has on LDL cholesterol. In other instances, such non-linearity could lead to contradictory results if the mean exposure level differs between studies, even in the absence of any meaningful biological or methodological differences.

NON-LINEARITY MAY ARISE FOR A NUMBER OF REASONS (Fig. 3.5), the foremost of which is the one generally assumed in this type of modeling; that the biological mechanism through which the exposure affects the outcome has additional downstream constraints which modulate the effect (Fig. 3.5 A). This could be considered a *true* non-linear effect and often occurs in biological systems through the presence of feedback loops or other

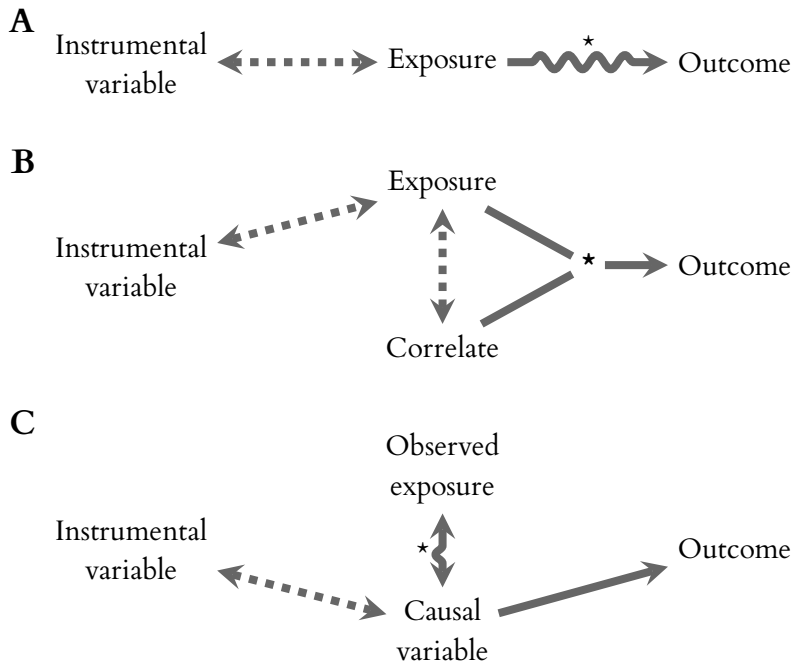


Figure 3.5: **Non-linear causal effects may be observed in several situations.** (A) It may occur biologically in what could be considered *true non-linearity*, (B) as an interaction between the exposure and a correlated variable, or (C) due to the scale on which the exposure is measured. The asterisks (*) and/or wavy arrows mark the source of non-linearity in each case.

controls which modulate the effect of the exposure on the outcome (e.g. limiting the bio-availability of other required components). This is most the case in the effects of BMI on glucose homeostasis: normal weight individuals are able to maintain normoglycemia, and the effect of BMI increases as their capacity to do so breaks down. Similarly, the plateauing of BP for high BMI could represent a homeostatic feedback loop or it may simply indicate the physiological bounds of BP. This could conceivably represent a behavioral feedback loop as well, such as the use of BP-reducing medication, however Staley and Burgess [83] found a similar plateauing even after adjusting for the use of medication. A similar effect might be observed in the case of selection bias [100], whereby individuals with more extreme outcome values may have been selected out of the observed sample.

Traits interacting to affect an outcome may likewise yield non-linear effects in certain conditions, particularly if they are not independent (Fig. 3.5 B). For example, if two traits are positively correlated and the effect on the outcome scales with their product, then the magnitude of the effect will scale

quadratically with either trait used as exposure.

The simplest (and possibly least interesting) cause for non-linearity we will consider here is that of the scale of observation (Fig. 3.5 C). Even in the presence of a purely linear effect, it may appear non-linear if either the exposure or outcome is observed on a scale which is non-linear with respect to the true causal trait. For example, adiposity can be studied using any of a number of different metrics, such as body fat percentage or body fat mass. However, if the effects of adiposity on a particular outcome are proportional to, e.g., the *surface area* of the adipose tissue, then we can expect any of the usual metrics of adiposity to show non-linear effects on this outcome simply due to the non-linearity of their association with the true proximal cause involved in the etiology of the effect on the outcome. This is conceptually related to the misidentification problem which can never be fully excluded in observational studies. In some cases, the true cause can be considered a confounder of the (tested) exposure–outcome association and violate MR assumptions. This may introduce bias which will increase the more dissimilar the true cause and observed exposure are.

POLYMR HAS BEEN DESIGNED USING POLYNOMIAL REGRESSION but could easily be extended to include other terms such as logarithmic or sigmoid. Coefficients for these terms could be estimated in addition to or instead of the proposed polynomial coefficients. The inclusion of more complex flexibility including, e.g., parameters to be fitted *within* a term (e.g. finding α_e in $f_\alpha(x) = c \cdot e^{\alpha_e x}$) might be possible with a maximum likelihood approach.

PolyMR can be applied as is with binary outcomes such as diseases, though an approach based on logistic regression, which already includes non-linearity in the link function, might prove more advantageous.

THE LIMITATIONS OF POLYMR are partly dependent on the statistical power available. In an infinite

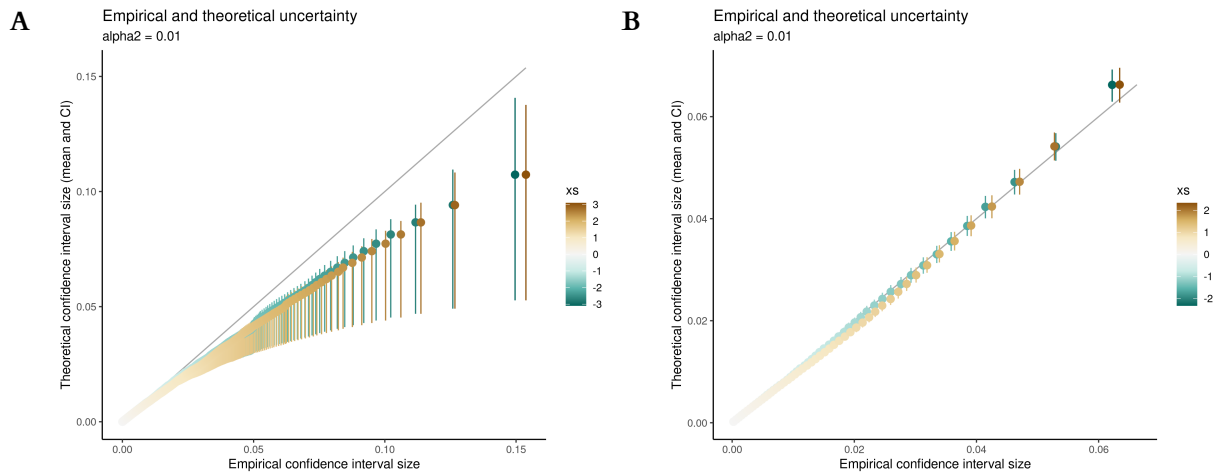
sample size any function could be modeled flawlessly, however in finite sample sizes the approximation of the causal function in the central portion of the exposure distribution affects its behavior at the tails. In our simulations, this is accounted for by the considerably wider confidence interval in these areas. It is nevertheless possible that a curve whose shape deviates strongly from polynomial near the extremes will be poorly represented by PolyMR.

In cases where the modeling assumptions are violated, the inferred causal function may be biased. In particular, the possible influence of covariates on the approximation of causal functions has not been fully investigated.

The post-selection inference in PolyMR (i.e. the selection and elimination of coefficients) can lead to a slight under-coverage of the confidence intervals in certain situations. This can be avoided if a model is postulated beforehand, such as by selecting coefficients in a separate training sample, and kept throughout the PolyMR modeling.

IN SUMMARY, we have developed a method for the polynomial approximation of non-linear causal functions. Applying it to data from the UK Biobank, we have shown that non-linearity is pervasive in the effects between anthropometric and metabolic or other continuous health traits.

3.1.4 Supplementary Figures



Supplementary Figure 3.6: **The confidence intervals (CIs) of PolyMR are correctly calibrated for the returned function order.** The size of the theoretical 95% CI was calculated for each simulation at each percentile of the exposure, yielding a distribution of CI sizes. The empirical CIs come from the distribution of the estimated causal function. Both panels show the CIs for the weak quadratic effect setting, with either (A) all results or (B) only those results where the correct order of the causal function was determined (920 out of 1000 simulations).

3.2 Further applications and extensions

In addition to the basic use of PolyMR method described in the section above, there are other possible applications which exceeded the scope of the manuscript in preparation. Here I describe two of these, namely the use of composite traits as exposures and that of binary traits (e.g. diseases) as outcomes. In addition, I describe an additional extension to PolyMR which enables the modeling of interactions within the context of non-linear effects.

3.2.1 Non-linear effects of composite traits

I applied the same method as described in [Sec. 3.1](#) to the phenotypic values of the four body PCs previously described [87], testing for non-linear effects on the same continuous health traits (pulse rate [PR], systolic blood pressure [SBP], diastolic blood pressure [DBP], and glucose, and LDL, HDL, and total cholesterol [TC] levels in blood). Most (19/28) were significantly non-linear (with an additional

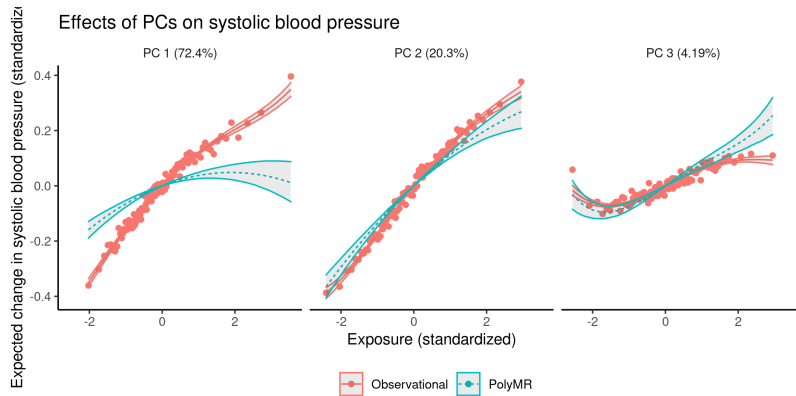


Figure 3.7: PCs 1–3 mostly increased SBP, though with qualitatively different effects (teal, dashed). The values on the y-axis are given relative to the expected outcome for an exposure level equal to that of the population mean. The points represent the mean difference in observed outcome for each centile of the population ranked by exposure level, with the red line (solid) showing the observed association. The values in parentheses indicate the variance (in body shape) explained by each PC.

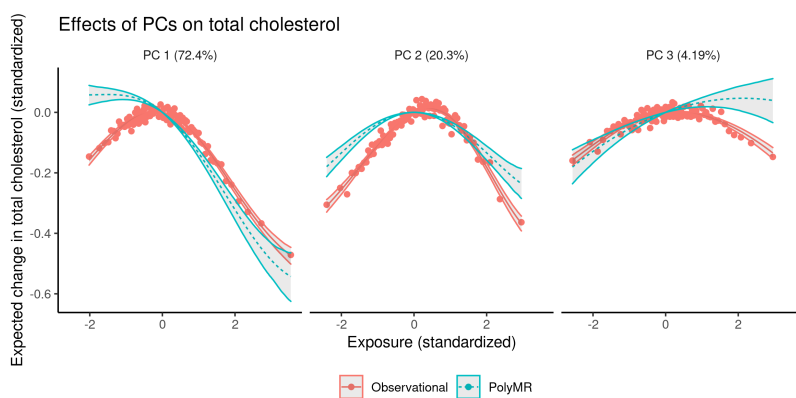


Figure 3.8: PCs 1–3 had distinct effects on total cholesterol levels in blood (blue, dashed), despite showing similar inverted-U shaped observational associations (red, solid). The values on the y-axis are given relative to the expected outcome for an exposure level equal to that of the population mean. The points represent the mean difference in observed outcome for each centile of the population ranked by exposure level. The values in parentheses indicate the variance (in body shape) explained by each PC.

four linear effects).⁴ The effects of PCs 1–3 on glucose were similar to those of other obesity traits, i.e. no effect in normal-weight individuals but increasingly strong ones for higher values. This is consistent with the somewhat similar (linear) effect sizes we found for these PCs on glucose, glycated hemoglobin, and diabetes in the original paper [87]. Their effects on the other traits were not so similar.

PCs 1–3 all broadly increased SBP, but the shape of the inferred causal function and the intensity of the effect differed considerably (Fig. 3.7). PCs 1–2 were similarly associated with SBP *observationally*, however the effects of PC1 were much weaker and rapidly plateau, while those of PC2 explain the observed association nearly entirely. Interestingly, although the observed PC3–SBP association was much weaker, the amplitude of effects shown by PC3 exceeds that of PC1.

⁴ All five null effects (and two of the linear ones) were predicted for PC4 (lean mass), which is likely due, in part, to low statistical power.

These 3 PCs had qualitatively similar, inverted U-shaped observed associations with total cholesterol levels in blood, but different effects (Fig. 3.8): PC1 decreased it, PC2 had a strong inverted-U shape, and PC3 increased it. This is driven in large part by their effects on LDL-c which is expected to contribute a larger fraction of the total [113], which is dominated by stronger negative effects for higher values of PC1, balanced for PC2, but linearly increasing for PC3. All three PCs decreased HDL-c, though the effects were weaker for PC2.

Even across such a small subset of continuous health outcomes, the shapes of the inferred causal functions of PCs 1–3 provided interesting comparisons. The similarity and monotonicity of their effects on glucose are consistent with those (linearly estimated) we reported originally [87], not only on glucose but glycated hemoglobin and diabetes as well. On the other hand, the shapes of the causal functions estimated for total and LDL cholesterol provide insight as to why PCs 2 and 3 produced non-significant linear effect estimates; not because they have no effect but because the slope of the causal function at the mean population exposure value is close to zero.

3.2.2 *Non-linear increase in disease risk*

As mentioned in the PolyMR manuscript (Sec. 3.1), the method itself does not preclude the use of binary variables as outcomes. As a proof of concept, I examined the shape of the effects of anthropometric traits (namely BMI and BFP) on self-reported diseases in the UK Biobank, selected for their relevance to obesity (diabetes, heart attack, osteoarthritis) or arbitrary curiosity (depression).

Both BMI and BFP increased the risk of all investigated diseases. Their effects on the risk of diabetes were (unsurprisingly) non-linear and comparable to those we previously found on glucose: no effect in normal-weight individuals with increasing effects for higher values. Their effects on other

diseases were closer to linear, likely due in part to weaker causality which reduced statistical power. The effects of BFP on the risk of osteoarthritis were similar in shape to those on diabetes, with little to no effect at low values but increasingly raised the risk of disease at higher values.

BMI showed a J-shaped association with depression but linearly increased risk, while BFP showed an exponential-shaped association but a J-shaped effect on the risk.

3.2.3 Extending PolyMR to model interactions

The inclusion of non-linear terms in the causal function adds further complexity, beyond the simple shape of the causal effect, in particular in the form of *interactions*.

In traditional (linear) MR, confounders of the exposure–outcome relationship can safely be ignored, provided they don’t violate MR assumptions through association with an IV. We adjust for genetic PCs in many analyses precisely because population stratification would violate MR assumptions (and GWAS requirements for that matter). Other potential confounders are typically also regressed out of both traits to reduce noise in the IV–trait association (e.g. sex and age). In reality, these covariates may alter the effect of the exposure, i.e. the covariate may *interact* with the exposure to produce a certain effect on the outcome (Fig. 3.9). For example, the exposure might have a greater effect in men than in women, in which case the exposure *interacts* with sex to produce an effect on the outcome. Adjusting for this covariate in traditional MR will produce the expected (and arguably desirable) result of providing the causal effect estimated for an average exposure level (or the weighted average effect). With non-linear causality, it’s a bit more complicated.

IN THE INFERENCE OF NON-LINEAR CAUSAL EFFECTS,⁵ adjusting for covariates has a similar effect

⁵ Although I will discuss this mainly concerning PolyMR, this applies to other methods as well.

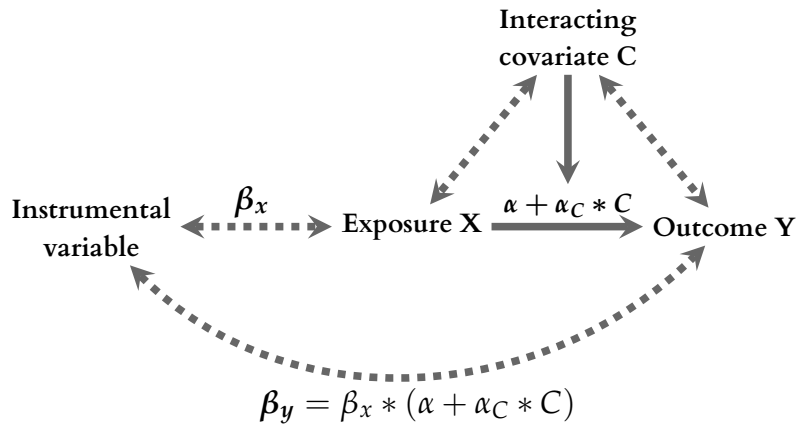


Figure 3.9: **Example interaction scenario** in (linear) summary statistics-Mendelian randomization, where the effect of exposure X is altered by the presence of covariate C. This covariate could also be a confounder, i.e. affecting both exposure and outcome, in addition to interacting with the exposure. Solid lines represent causal effects, dashes show associations.

but different interpretation: the non-linear causal effects are expressed as function of the *covariate-corrected variables*. In other words, the causal function predicts the magnitude of the effect based on *the difference between the actual exposure value compared to what we would expect based on the covariates*. Conversely, choosing not to adjust for a given covariate implies that the effect depends on the exposure level, independent of that specific covariate. Biologically, it might seem to make sense that it is the absolute level of the exposure that matters, regardless of what covariates caused it to take on that value, and in this case we might favor unadjusted variables. But this isn't necessarily the case, for example, women tend to have proportionally more body fat than men but are not necessarily more at risk of obesity-related diseases [104].

In the case of sex, a straightforward solution is stratification: the non-linear causal function can be inferred in men and women separately, avoiding the issue entirely and providing unbiased estimates (with respect to the interaction effect of sex). The resulting curves can be compared to see whether adjusting for sex is advisable (thereby enabling the use of the full dataset), although the individual causal functions can be used as well. Applying PolyMR to stratified data from the UK Biobank, we can observe the two scenarios in the effects of BFP and WHR on LDL cholesterol (Fig. 3.10).

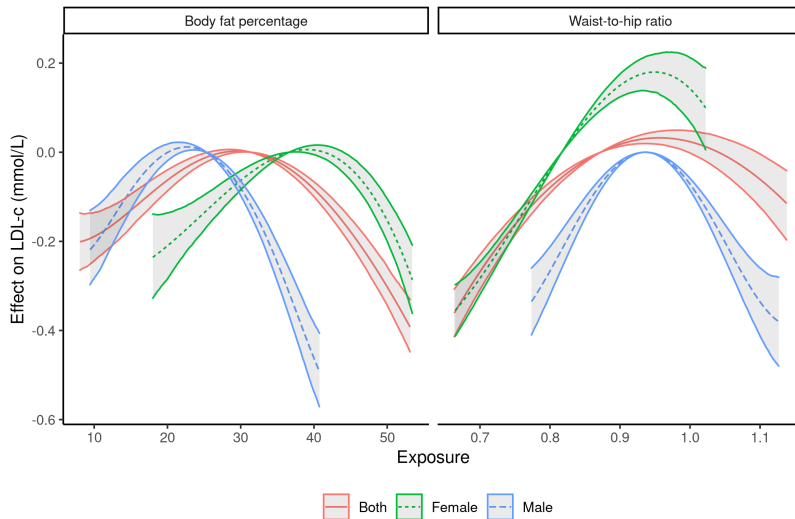


Figure 3.10: **Adjusting for sex can have different consequences** based on the exposure-outcome relationship. The sex-specific and -combined effects of body fat percentage (BFP, left) and waist-to-hip ratio (WHR, right) on LDL cholesterol illustrate the two extremes. The horizontal shift between the sex-specific curves (green and blue) for BFP indicates an interaction with sex, and, in this case, adjusting produces a relatively similar sex-combined curve (red). The absence of such a *horizontal* shift for WHR suggests no interaction with sex, in which case adjusting for sex produces a biased (flattened) sex-combined causal function. Note that the vertical (mis)alignment is simply due to the y-axis being defined relative to the expected effect at the (respective) population mean exposure values.

In all cases, the inferred causal function had an inverse-U shape with roughly corresponding intensities between sexes, however it is the relative positions of these curves that is interesting: they are horizontally shifted for BFP but not WHR.⁶ From this, we can expect that adjusting for sex will produce better results for BFP but worse for WHR and indeed, we can see that the sex-combined causal function for BFP is similar to the sex-specific ones, whereas that of WHR is much more flattened.

This much we can evaluate from a quick visual inspection of the sex-stratified causal functions, but this is hardly ideal. It's simple enough to stratify by sex, but this is less trivial for continuous variables such as age. Furthermore, a closer inspection reveals that the sex-specific curves of BFP are *not*, in fact, identical, nor can we reasonably expect adjusting for sex to necessarily align the exposures exactly, such that their effects on a specific outcome are equivalent. For example, adjusting BFP for sex is likely to improve the inferred causal function to some extent (i.e. the sex-adjusted curves will be similar to the sex-specific ones), however we can see that the population means do not sit in the same portion of their respective curves: the slope at the (sex-specific) mean exposure level is positive in females but negative in males. We can

⁶ Note that the vertical displacement in the case of WHR is not relevant here, as the y-axis is defined as the expected effect relative to the sex-specific sample mean rather than any absolute value.

therefore expect adjusting for sex to dilute the signal somewhat, which is what we observe in the sex-combined PolyMR, where the variables were adjusted for sex and the causal function (red) is slightly flattened. A more accurate approach would require modeling the interaction effect simultaneously, for which I propose a method below.

Methods

Including sex as a possibly interacting covariate can be achieved with minimal changes to the PolyMR equations:

$$X = G \cdot \beta + \epsilon_x \quad (3.11)$$

$$Y = \sum_{i=1}^k \alpha_i \cdot X^i + S \odot \sum_{i=1}^l \alpha_{si} \cdot X^i + \epsilon_y, \quad (3.12)$$

where X and Y are the covariate-adjusted exposure and outcome, respectively, S is the sex (with $E(S) = 0$ and $Var(S) = 1$), and ϵ_x and ϵ_y are the error terms as described in [Section 3.1](#). Similar to the original procedure, ϵ_y can be split into ϵ_x -dependent and -independent terms:

$$\begin{aligned} \epsilon_y &= (\epsilon_y | \epsilon_x, \dots, \epsilon_x^m, S \odot \epsilon_x, \dots, S \odot \epsilon_x^n) + \tau_y \\ &= \sum_{i=0}^m r_i \cdot \epsilon_x^i + S \odot \sum_{i=0}^n r_{si} \cdot \epsilon_x^i + \tau_y \end{aligned} \quad (3.13)$$

The *exclusion-restriction* assumption requires that IVs be independent of any confounders of the exposure-outcome relationship, ensuring that $cov(G, S) = cov(G\beta, S) = 0$, which we would anyway expect to be the case for autosomal IVs. The exposure and outcome are adjusted for sex, ensuring that $cov(X, S) = 0$. We further require that $cov(S \odot G\beta, \epsilon_y) = 0$, which is no more constraining than that the InSIDE assumption [71] be verified for the full dataset, as well as within each stratum (here sex), which is the same as would be required to use MR in each of these settings separately. Substituting [Equation 3.13](#) into [3.12](#), we obtain

$$\begin{aligned}
Y &= \sum_{i=1}^k \alpha_i \cdot X^i + S \odot \sum_{i=1}^l \alpha_{si} \cdot X^i \\
&+ \sum_{i=0}^m r_i \cdot (X - G\beta)^i + S \odot \sum_{i=0}^n r_{si} \cdot (X - G\beta)^i + \tau_y
\end{aligned} \tag{3.14}$$

Here, the error terms (ϵ_x and τ_y) are uncorrelated and independent of the respective explanatory variables.

Let the realizations of the random variables X, Y, G, S be denoted by $\mathbf{x}, \mathbf{y}, G, \mathbf{s}$, observed in a sample of size n . The parameters $\{\beta, \alpha, \mathbf{r}\}$ can be estimated by computing two ordinary least squares estimates, first estimating β using the first equation and substituting this into the second equation:

$$\hat{\beta} = (G'G)^{-1} \cdot G'\mathbf{x} \tag{3.15}$$

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\mathbf{r}} \end{pmatrix} = (M'M)^{-1} \cdot M'\mathbf{y}, \tag{3.16}$$

where

$$\begin{aligned}
M &:= [\mathbf{x}, \dots, \mathbf{x}^k, \mathbf{s} \odot \mathbf{x}, \dots, \mathbf{s} \odot \mathbf{x}^m, \\
&\mathbf{1}, (\mathbf{x} - G\hat{\beta}), \dots, (\mathbf{x} - G\hat{\beta})^l, \\
&\mathbf{s}, \mathbf{s} \odot (\mathbf{x} - G\hat{\beta}), \dots, \mathbf{s} \odot (\mathbf{x} - G\hat{\beta})^n]
\end{aligned}$$

I used a similar approach to that described in [Section 3.1](#) to implement this proposed extension. The exposure and outcome variables are adjusted for covariates as normal, independent IVs are selected using the `ieugwasr` R package [version 0.1.5, 95], and the initial regression (Eq. 3.15) is performed as previously described. For the second regression (Eq. 3.16), all polynomial exponents up to 10 are included, i.e. $k = m = 10$, and then the least significant coefficient (α_i or α_{si}) is iteratively removed until all remaining ones are significant (at a Bonferroni-corrected threshold). The ϵ_x -dependent coefficients (r_i and r_{si}) were kept up to the highest included corresponding x -dependent term, i.e. $l := k$ and $n := m$.

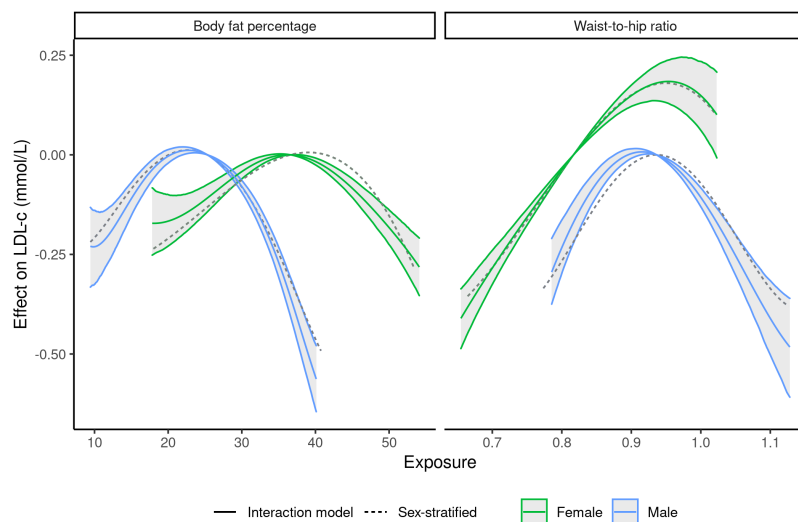


Figure 3.11: **Modeling the interaction** directly recovers the respective sex-specific curves (solid lines), which are highly similar to those obtained from stratification (dashed). The hulls shown are those of the interaction modeling.

I applied this method to data from the UK Biobank using the same anthropometric traits as previously (BMI, BFP, WHR, and weight) as exposures on both the continuous (Sec. 3.1) and binary (Sec. 3.2.2) outcomes, in addition to which I added hypothyroidism as we had previously noted sex-specificity in the effects of obesity on it [87].

Results

The curves from this extended method closely match those from the sex-stratified analysis (Fig. 3.11). Where a distinction is required, I will generally refer to the unadjusted exposure level as the *absolute* value and the covariate-corrected value (which is used in the modeling) as the *relative* value, which is expressed relative to the sample distribution for a given sex. Unless otherwise noted, the *shape* of the sex-specific causal functions were comparable to those of the sex-combined PolyMR (Sec. 3.1).

In most cases, the effects were altered by sex, often resulting in a change of intensity and/or a shift in the range of exposure values (absolute and/or relative) over which it occurred, but usually without much alteration in the overall shape. The differences can nevertheless be considerable. For example, an increase in absolute BFP from

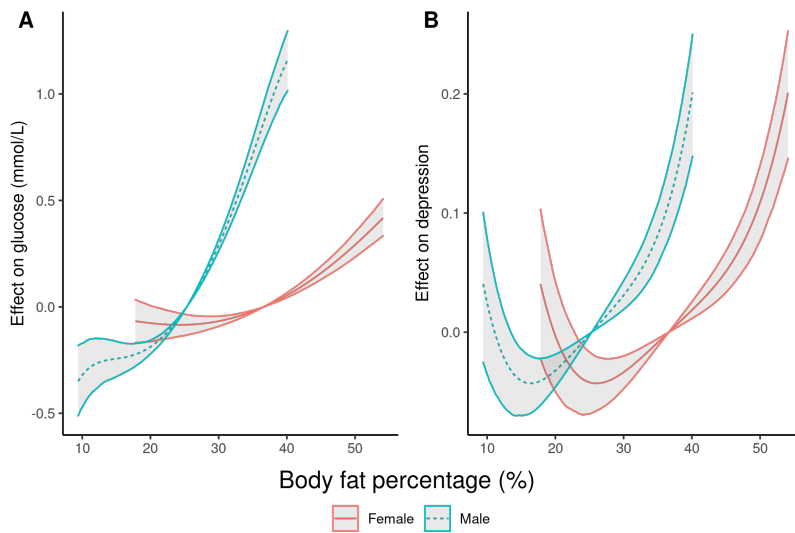


Figure 3.12: The effects of body fat percentage (BFP) on (A) glucose levels in blood and (B) risk of depression show very different sex-specificity. The increase in blood glucose from higher BFP in men (teal, dashed) is much greater than that inferred in women (red, solid), whereas the effects on depression were identical across the respective sex-specific exposure distributions (i.e. no significant sex-interaction coefficient). The estimated effect (y-axis) is here shown relative to that estimated for the sex-specific population mean exposure.

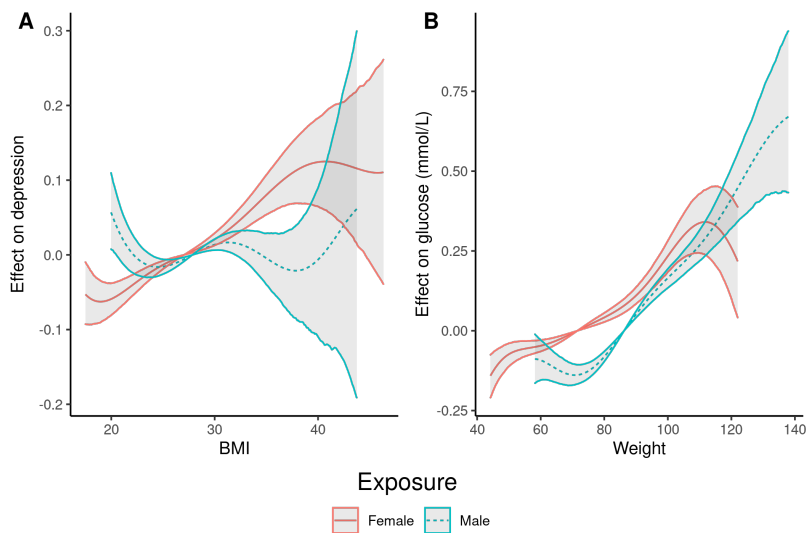


Figure 3.13: Some inferred causal functions showed considerable irregularity, with many inflection points, as shown here for the effects of (A) BMI on the risk of depression and (B) weight on glucose. The inferred (sex-interacting) effect is plotted (y-axis) as a function of the absolute exposure value for each sex, compared to the expected effect at the sex-specific population mean exposure.

20% to 30% in men increased blood glucose levels by 0.48 mmol/L but only 0.01 mmol/L in women. Even increasing BFP from 20% to 50% in women⁷ is only estimated to increase blood glucose by 0.37 mmol/L (Fig. 3.12 A).

Overall, the effects of BMI changed mostly in intensity, where it affected PR, glucose, and risk of diabetes and heart attack more strongly in men but hypothyroidism more in women. Its effects on LDL-c were shifted (in both relative and absolute terms), where the effect in women occurred at slightly higher BMI than would be expected from

⁷ In terms of relative exposure, this corresponds to a shift from the 1st percentile to the 97th percentile. For comparison, the increase from 20% to 30% in men corresponds to a shift from 17th to 78th percentile.

the effect in men. BMI tended to increase the risk of depression, especially in women, though the inferred model was irregular (Fig. 3.13 A).

The effects of BFP were comparable between men and women, though usually on the relative exposure scale. The sex-specific intensity of the effects varied with the same phenotypes as for BMI, i.e. stronger in men for PR, glucose, diabetes, and heart attack, weaker for hypothyroidism. Interestingly, the modeling revealed no sex-specific coefficients for the effects of BFP on depression, with the same J-shaped association as previously noted (Sec. 3.2.2) aligned to the relative exposure (Fig. 3.12 B).

Modeling the effects of WHR revealed many sex-specific coefficients, with more of a tendency to align with the absolute exposure level, as previously shown through stratification. The effects in women were still slightly attenuated for PR, glucose, and total cholesterol, while slightly stronger for SBP. They were shifted in DBP, lining up more closely with the relative than absolute exposures.

The effects of weight were generally comparable on the relative exposure scale. The intensity was stronger in men for the risk of heart attack and glucose (though the model was slightly irregular, Fig. 3.13 B), but weaker for HDL-c, hypothyroidism, and osteoarthritis. The effects of weight on LDL-c were close to identical on the *absolute* scale. Interestingly, the effects of weight on the risk of depression were ~zero close to the (relative) population mean, but affected both in opposite directions for higher values, increasing the risk in women but decreasing it in men.

Discussion

Overall, the results are as expected, with the interaction terms allowing sex-specific effects to be modeled. The shape of most effects were comparable between sexes but often different in magnitude

and/or the exposure range where it occurred. Interestingly, sometimes the effects were similar on an *absolute* exposure level (i.e. would be comparable on untransformed, unadjusted values), while others were more similar on the *relative* scale (i.e. dependent on where the values lie in the sex-specific exposure distribution).

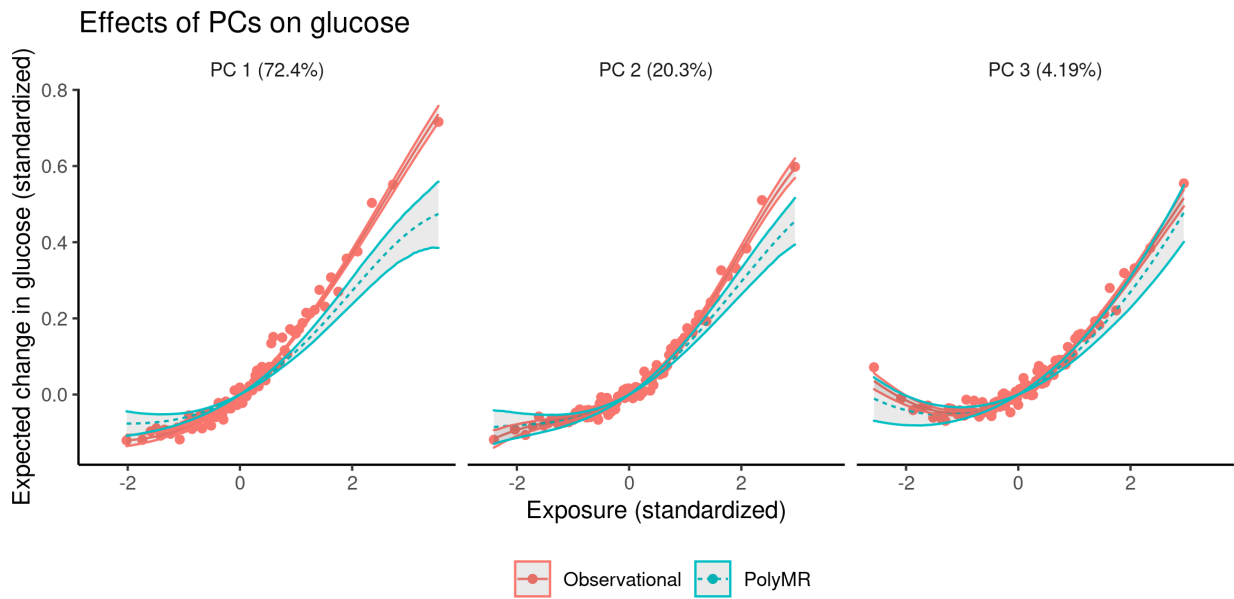
The sex-specificity generally manifested as a difference in intensity (particularly for BMI) or relevant exposure range.

In the absence of any sex-specificity of the causal function, the method provides the exact same model as would PolyMR on the full dataset. In the absence of any joint effect, i.e.

$$\forall i, \alpha_i \neq 0 \implies \alpha_{si} \neq 0,$$

then the result will be similar to the sex-stratified approach. In the other cases, the shared variance will improve the statistical power while avoiding the misalignment that may occur when naively adjusting for sex and allowing some sex-specific variation in causal function shape. The ‘additional’ assumptions clarified here would be required for any MR analysis involving possibly sex-specific effects.

Supplementary figures



Supplementary Figure 3.14: PCs 1–3 show similar effects on glucose levels in blood, with no effect in normal-weight individuals but monotonic increasing effects for larger values.

4

Minor contributions

4.1 Quantifying gene by environment interactions

This manuscript [101] describes a method to quantify the gene by environment interaction, largely developed and written by Zoltán. My contributions to this are mainly the creation of a streamlined R package based on Zoltán's Matlab implementation, as well as a few minor corrections and suggestions.

Discussion

The aim of this thesis was to improve the characterization of obesity, identify its different components, and disentangle their specific contributions to the consequences typically attributed to “obesity.” In the review [1], we described the current understanding of body fat distribution, its genetic basis, sexual dimorphism, and its metabolic consequences. In the first project [87], we decomposed body shape into four genetically independent components and showed that these differ in both their genetic basis and consequences, the comparison of which provided insight into the etiology of obesity-related diseases. In the second project (Sec. 3.1), we developed a method to approximate non-linear causal effects and used it to show that most anthropometric traits affect and are affected by metabolic traits in a non-linear fashion.

In this final chapter, I provide a broader context for the non-linearity of obesity-related effects and considerations for the wider application of PolyMR and similar methods to composite or binary traits. I also discuss the impact of interactions on non-linearity and its modeling. I conclude with some remarks on the study of obesity from the perspective of public health.

5.0.1 Non-linear effects of composite traits

The four body PCs being constructed from the linear combination of phenotypes with non-linear effects, it is unsurprising that most of the inferred

causal functions were significantly non-linear. The linear independence of these PCs with each other, which enables the dissection of obesity-related consequences, ensures sufficient dissimilarity between them to provide causal functions which were often both quantitatively and qualitatively dissimilar. The comparison of these differences in shape, as well as the directionality and intensity described in the original paper [87], can here as well provide further insight into the specific biological mechanisms underlying obesity-related consequences.

However there is one caveat in this application: while the body PCs are (genetically) *uncorrelated*, they are not *statistically independent*. PCA ensures that the eigenvectors are orthogonal and the PCs uncorrelated, but this does not imply that higher orders of these PCs are also uncorrelated (as would be the case if they were statistically independent). This implies that, unlike with linear MR methods, the effects inferred by PolyMR for each PC are not fully independent and therefore not additive. The orthogonality of PCs nevertheless results in causal functions which were more dissimilar than those of individuals obesity-related traits.

5.0.2 Obesity and disease risk

Although application of PolyMR for the inference of disease risk has not been extensively validated through simulations, the causal functions are largely as expected. In particular, the effects of obesity-related exposures on diabetes were nearly identical to those on glucose levels in blood, as is expected for one of the primary biomarkers for the disease. The near-linear effects of BMI were somewhat unexpected, though this may be due to covariate-dependent interactions, leading to the flattening of the effect (see [Sec. 5.1](#)). The non-linear effects of BFP may nevertheless contribute to the non-linearity of the observed association between obesity and depression [102, 103].

Given that most obesity-related increases in dis-

ease risk (tested here) were monotonic (with the exception of BFP on depression), we can nevertheless expect linear causal effect estimates to correctly assess the directionality of the effect, however its magnitude will be misestimated for individuals away from the population mean.

In addition to the use of binary traits as outcomes, it would be possible to include them as exposures as well. Although modeling higher order functions of a binary trait seems pointless at first glance, the *covariate-adjusted* variable included in the model would not, in fact, be binary. This adjusted binary variable is essentially a measure of the difference between the predicted disease state and the diagnosis, which can be likened to a disease intensity after adjusting for risk factors. For example, early onset Alzheimer's disease might receive a higher covariate-adjusted value than late-onset. The usefulness of such an approach would require additional investigation.

5.1 *Non-linear effects and interactions*

The differences in sex-specificity between BMI, WHR, and BFP illustrate the importance of more advanced methods for causal inference and the careful consideration of the assumptions underlying not just the method but the preprocessing of data as well. BMI is perhaps the least sex-specific in its distribution and only the magnitude of the effect tended to be different between sexes. Both BFP and WHR show much greater differences in range between men and women, but it is interesting to note the difference in interpretation with respect to health risks: women tend to have higher BFP without this affecting their overall risk of disease (at the population level), whereas the tendency for men to have higher WHR leads to greater risk of adverse health conditions.

Given the similarity between the sex-stratified curves and those from including sex as an interact-

ing factor, it is worth considering why we might want to model interactions rather than simply stratifying. There are two main advantages to the joint modeling.

First, the use of *sex* as an interactor is more of a proof of concept, where the stratification is simple and provides us with a reference for easy comparison, however this method could be extended to continuous covariates, such as age, where stratification is likely to introduce bias from arbitrary thresholding and simply not an ideal option. Including such factors may require considering the inclusion of higher order terms for the interactors as well, something which was not necessary in the case of *sex*.

Second, the joint modeling of the interaction effect can provide a considerable boost in statistical power over not only the sex-stratified analysis, but also the sex-combined one due to the issue of mis-alignment of relative exposures with respect to their effects on the outcome.

Although the variables were adjusted and controlled where possible and steps were taken to avoid violating MR assumptions, it is possible that some bias from these remains. I have investigated some more stringent variations of the method as well,¹ and results suggest that even *if* the exact magnitude and range of the effects are biased, the conclusion concerning the shape and magnitude of the effect remains relevant.

THIS EXTENSION IS IN A PRELIMINARY FORM and may require additional investigation to ensure the validity of the inferences made, although the generally-good agreement between the sex-stratified results and the interaction modeling suggests that the risk of significant bias introduced by these modifications is minimal. The main advantage of this method is that it would lend itself to the modeling of other non-genetic² interactors such as age with minimal adjustments. Future extensions could provide opportunities to model many other complex relationships.

¹ In particular, I investigated the possible necessity for sex-specific *scaling* rather than simply adjusting. This eliminates the statistical dependency between S and ϵ and ensures that $\forall i, cov(S, \epsilon^i) = 0$. The relevance of this approach, the bias it may avoid versus that which it could introduce may require further investigation.

² Although *sex* is *technically* a genetic factor, its definition purely based on the sexual chromosomes makes it mathematically independent of autosomal IVs, which is equivalent for the purposes of this extension.

I described a minimal extension for the modeling of interactions with the *covariate-adjusted* variables, but a more interesting question (biologically) is whether the *absolute* exposure interacts with sex (or another interactor). This would provide two main advantages for the PolyMR method. Firstly, the selection/elimination of sex-specific coefficients would use the null hypothesis equivalent to “the absolute exposure has the same effect in both sexes” rather than, essentially, testing whether the sex-specificity is exactly equal to the sex-adjustment performed based on the sample distributions.³ Secondly, we could use a likelihood ratio test to compare the sex-interacting model is better than the combined model, effectively testing the significance of the biological interaction overall.

Another possible avenue of research would be the inclusion of sex-specificity in the genetic effects. For example, we know that many SNPs have a stronger effect on WHR in women [52], however in PolyMR we disregard any such specificity. The inclusion of gene-sex interactions in the first equation (Sec. 3.2.3, Eq. 3.11) could provide more accurate estimation of sex-specific causal effects (or even possibly overall effects) by reducing the bias in the IV-exposure association (which could result in bias in the control function term).

Many other factors are also known to affect both obesity and its impact on health, though modeling their interactions is not always so straightforward, especially if they are not independent of genetics. For example, ethnicity is well-known to be relevant to obesity and its consequences [13], but population stratification is something we usually avoid (with good reason) by restricting our samples to a homogeneous population (e.g. here white British) and adjusting for genetic PCs. Unfortunately, this approach limits how generalizable our conclusions may be (and reduces statistical power). Recent methodological developments using generalized mixed models have overcome this hurdle in GWAS

³ Note that this may be a valid null hypothesis for the societal impact of obesity (e.g. depression), where its perception may rely on the comparison with others.

[105], although it is unclear whether/how any such information could be incorporated in PolyMR.

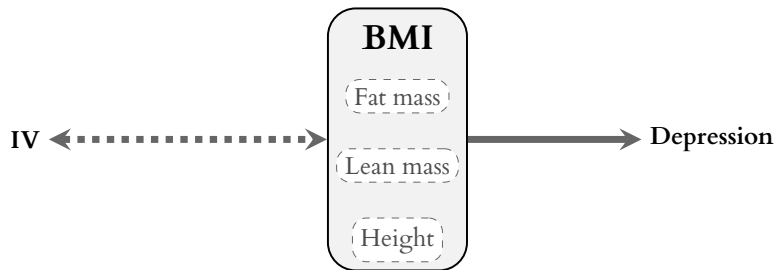
The inclusion of additional *exposures* (similar to what is modeled in Factorial MR [76]), possibly interacting with one another, could also prove an interesting extension to PolyMR. In the context of obesity, we would naturally expect many such interactions, e.g. between BFP and body fat distribution: the latter is only expected to become deleterious upon reaching certain levels of the former, and investigating such interactions may provide further insight into the etiology of obesity-related diseases and aid in their management.

THE VIOLATION OF ASSUMPTIONS in PolyMR (or indeed any MR-based method) can never be excluded and the risk of this increases considerably with certain exposures/outcomes. This does not preclude its use and the results may still be interesting and informative, but their interpretation warrants some caution. I will mainly consider two examples to illustrate this, both of which produced fairly irregular causal functions: the sex-interacting effects of BMI on the risk of depression and those of weight on glucose levels in blood (Fig. 3.13). While the apparent irregularity⁴ may be exacerbated by the polynomial approximation of an arbitrary function, there are reasons to consider the possible violation of assumptions.

As described in the introduction (Sec. 1.1), BMI itself is already less than ideal as an exposure. It approximates excess body mass (poorly) and depends on total mass (lean and fat) and height. Strictly speaking, each of these constituents acts as a *confounder* in the MR graph (Sec. 1.4, Fig. 1.3). Nevertheless, if they exert concerted effects (relative to the scale of BMI), then the interpretation remains valid. If, however, these constituents act differently on the outcome, as has been shown to be the case for depression [106] (illustrated in Fig. 5.1), then the modeled causal function will be an amalgama-

⁴ I.e. There are comparatively many inflection points, resulting in a less smooth function.

A. Intended model



B. Underlying model

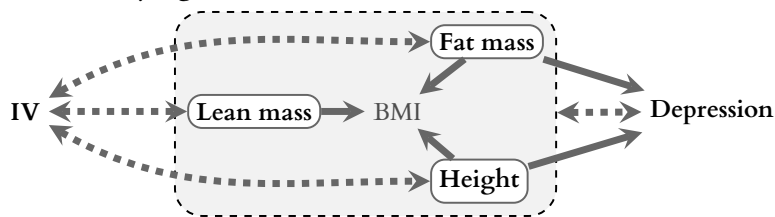


Figure 5.1: Graph of the constituents of BMI and their effects on depression (A) as intended when using BMI as exposure for MR and (B) the more likely underlying model of causality. Solid arrows represent causal effects, while dashed lines illustrate associations.

tion of the disparate effects and is likely to result in a multi-modal shape. This heterogeneity of effects is reflected in the large confidence hulls shown in Fig. 3.13. Bearing these limitations in mind, the resulting model can nevertheless be useful and interpreted qualitatively. Acknowledging that these effects are likely affected by a combination of fat mass and height [106], we can see that the impact is close to monotonic increasing in women and much weaker in men.

The “irregularity” in the weight-glucose causal function is less severe. Excluding the behavior of the curve at the extremes of the distribution (where polynomial approximation may be less reliable), we can consider glucose to increase (near-)monotonically with weight, which is in line with expectations. It is worth noting that the effects inferred here correspond to the lifelong impact of *genetically*-predicted weight. Genetic weight is not necessarily predictive of *obesity per se* (although it is associated) and we have found it to be closer to an overall *body size*, correlated with height and all types of body mass [87]. This may reduce the specificity of the effects detected and, even assum-

ing they don't directly violate MR assumptions, would likely introduce heterogeneity in the form of a sum of multiple, possibly non-linear and interacting, effects.

This type of heterogeneity is likely due to mechanisms similar to what is modeled in MR-Clust [75], where IVs are grouped according to their MR effect estimates, under the assumption that these represent either (1) different components of a risk factor (e.g. lean mass, fat mass, and height for BMI), (2) different pathways through which the exposure affects the outcome, or (3) pleiotropic pathways. Although all three of these scenarios violate MR assumptions, the information provided by the analysis is not useless requires care in its interpretation.

"All models are wrong but some are useful." —GEORGE BOX

5.2 Limitations of polynomial regression

We developed PolyMR as a general method for the inference of non-linear effects through polynomial regression, justifying the use of a polynomial function by the arbitrary precision with which a non-polynomial function could be modeled. In real-world applications with finite data, however, there are limitations to the precision which can be attained with this approach. The two main limitations I will consider here are the *post-selection inference bias* mentioned in the discussion (Sec. 3.1.3) and the *global* nature of the polynomial approximation.

POST-SELECTION INFERENCE BIAS is quite simple to avoid: the model simply need be fixed beforehand, such that only the coefficients are estimated. This seems simple enough, however the selection of terms to include is not trivial without observing the data (which, again, would lead to post-selection inference bias). A straightforward solution would be to use a separate dataset to select the model,

however this requires more data (or equivalently reduces statistical power by not allowing the full dataset to be used in coefficient estimation). Model misspecification may drastically reduce power if too many terms are left in or introduce bias if too few are included. The inclusion of multiple high order terms is required to provide polynomial regression with the appropriate flexibility, particularly to model the extremes of the exposure distribution. This brings us to the second limitation: the modeling of the effect as a global function.

GLOBAL APPROXIMATION of the causal function is a defining characteristic of polynomial regression. In the case of PolyMR, this implies that the effect of the exposure is expected to follow a single function across the entire distribution. The biological validity of this is debatable in some cases (e.g. the effect of body fat on glucose is likely different in normal weight or obese individuals) but more importantly this has consequences for the inferred model. Specifically, the inclusion of higher order terms to allow the model to conform to the “true” function at one extremity of the distribution inevitably affects the other end as well (and to some extent the entire range). Although this can be corrected to an arbitrary precision in theory, in practice this is dependent on the amount of data (and the strength of the effect) and inevitably leads to bias in the inferred model.

Both of these limitations are (to some extent) caused by the requirements and specificities of polynomial regression as a means to approximate the true function. There are, however, alternatives such as *spline regression* which can reduce or eliminate these limitations.

SPLINE REGRESSION is a well-established method, often used for interpolation as well as approximation. In essence, it is a piecewise polynomial approach with some added constraints. The exposure range is subdivided into intervals by *knots*, essen-

⁵ I use the terms exposure and outcome in relation to its potential application for causal inference.

tially inflection points at which the function can change shape. The causal relationship is then approximated by a separate polynomial (usually cubic) function for each interval. The value of both functions around a knot are constrained to be equal at that point, as are the first and second derivatives, ensuring a continuous and smooth function. Additional constraints are often added at the extremities of the distribution, forcing the second and third derivatives to be zero beyond the endpoints, resulting in so-called *natural splines*. This additional constraint avoids unnatural sharp turns at the ends of the distribution to fit outliers.

The use of spline instead of polynomial regression would enable the *local* adjustment of the function without unduly affecting the rest of the model. Although spline regression might *seem* to include more coefficients (by modeling multiple polynomial functions), the added constraints reduce the number of degrees of freedom such that complex functions can be modeled using only 4–5 coefficients (a natural spline with k knots only has k degrees of freedom). This also implies that we can avoid selecting the terms when modeling and avoid post-selection inference bias. Preliminary tests do indeed show differences in the extremities of the causal function, suggesting bias introduced by the polynomial model. Furthermore, I expect that this approach could improve the accuracy of the inferred models. For these preliminary results, I implemented this method in `R` and included it in a separate branch of the `POLYMR` package on github [91]. Not requiring iterative coefficient selection, this alternative method was also considerably faster.

5.3 Towards a more comprehensive view of obesity

The complexities of obesity are now well-recognized and many of its aspects have been described, and their causes and consequences continue to be investigated. The traditional scientific method of iso-

lating individual components of the system (here obesity) to understand its role is gradually reaching the limits of the information it can provide. Luckily, the combination of methodological advances and the ever-increasing size of datasets now enables the interrogation of more intricate hypotheses than ever before.

Here I presented two such investigations: the first employing an integrative approach which combines multiple related traits to uncover the basis of their interdependency and disentangle their specific consequences; the second extending Mendelian randomization to infer the non-linearity which characterizes many consequences of obesity. I further discussed other possible sources of complexity in the underlying causality and supplied a proof-of-concept approach whereby an extension to the PolyMR method could be employed to test such hypotheses. To conclude, I would like discuss two topics which are central to the global obesity crisis: the impact of environment on obesity and the practical implications of scientific findings.

THE IMPACT OF THE ENVIRONMENT on the body is undeniable and well established: obesity, or the accumulation of excess fat mass, occurs when excessive eating and/or insufficient energy expenditure result in a positive caloric balance, which is then stored in the form of adipose tissue. While the genetics which have occupied much of this thesis are by no means inconsequential, it is the *environment* which has the dominant role in the *emergence* of obesity. This is, of course, one of the reasons obesity has become a problem worldwide, not because of a genetic change in the human population but because our environment has changed in a way that the human body is ill-suited to handle.

Although we have a clear idea of the basics of how the environment can give rise to obesity, very little is known about the specific effects of individual dietary and lifestyle choices, how they interact with genetics to shape the body and maintain a

healthy metabolism or, failing to do so, lead to disease. Our limited understanding of these specifics is largely due to the challenges related to its study. Observational analyses on the topic suffer from the usual problems of causal inference from correlation, compounded by strong population stratification and weak genetic associations. Added to this, the data that we do have available is generally from questionnaires, which can suffer considerable bias [107]. Experimental studies could provide more reliable causal inferences, however these are challenging to put in practice in humans. The required diet/lifestyle interventions for even short-term effects is likely to result in low compliance of test subjects. For example, several studies comparing the effectiveness of various diets on short-term weight loss have found that adherence to the diet was the most important factor in determining weight loss [108]. The interventions and sample sizes which might be required to investigate the specific effects of individual elements would likely be unrealistic.

Solutions to some of these problems already exist or have been proposed. Accelerometry data from wearable devices can accurately relay a person's physical activity throughout the day, though the analysis of this data remains challenging and it is unknown whether the knowledge of the device biases the individual's short-term behavior (i.e. a placebo-like effect). The use of food *preference* rather than food *frequency* questionnaires has the potential for greater reliability, since a person is more likely to accurately recall their preferences rather than their estimated consumption, and this relates to their overall diet [107]. A possible strategy for experimental approaches is to adopt various strategies which increase the probability of compliance in interventional studies [108]. The implementation of such approaches integrating environmental factors will be key to improving our understanding of obesity as a whole and developing effective interventions to improve public health.

After all, genetics may provide insight into the specific mechanisms underlying body shape and the etiology of diseases, but it is through the environment that obesity must be addressed at the population level.

PREVENTING OBESITY AND ITS DELETERIOUS CONSEQUENCES is the overarching aim of the intensive research on the topic and our ability to affect undesirable outcomes improves with our understanding of the intricacies underlying their etiology enabling a more personalized approach to health. Or does it?

Despite the popularity of the term *personalized medicine*, the adoption of meaningful changes in clinical practice to reflect this have, so far, been less than stellar. For example, abdominal obesity, measured by simple waist circumference, has been known to provide valuable information about a person's health risks since the 1980s [109], yet its assessment in clinical practice is still far from standard procedure [110]. The delay between a scientific finding and its application in the field is well known [111] and efforts have been/are being made to bridge this gap [112]. With the ever-increasing prevalence of obesity in the population, it is critical that we not only deepen our understanding of the complexities of obesity but expedite their translation into meaningful changes to improve public health and healthcare.

Bibliography

- [1] Jonathan Sulc et al. “Heterogeneity in Obesity: Genetic Basis and Metabolic Consequences”. In: *Current Diabetes Reports* 20.1 (2020), pp. 1–8. ISSN: 15390829. DOI: [10.1007/s11892-020-1285-4](https://doi.org/10.1007/s11892-020-1285-4). URL: <https://doi.org/10.1007/s11892-020-1285-4>.
- [2] Annette Boles, Ramesh Kandimalla, and P. Hemachandra Reddy. *Dynamics of diabetes and obesity: Epidemiological perspective*. 2017. DOI: [10.1016/j.bbadis.2017.01.016](https://doi.org/10.1016/j.bbadis.2017.01.016). URL: <https://pubmed.ncbi.nlm.nih.gov/28130199/>.
- [3] Francisco B. Ortega, Carl J. Lavie, and Steven N. Blair. *Obesity and cardiovascular disease*. 2016. DOI: [10.1161/CIRCRESAHA.115.306883](https://doi.org/10.1161/CIRCRESAHA.115.306883). URL: <http://circres.ahajournals.org>.
- [4] Barry M. Popkin et al. “Individuals with obesity and COVID-19: A global perspective on the epidemiology and biological relationships”. In: *Obesity Reviews* 21.11 (2020), e13128. ISSN: 1467789X. DOI: [10.1111/obr.13128](https://doi.org/10.1111/obr.13128). URL: <https://doi.org/10.1111/obr.13128>.
- [5] Jun Yang, Jiahui Hu, and Chunyan Zhu. *Obesity aggravates COVID-19: A systematic review and meta-analysis*. 2021. DOI: [10.1002/jmv.26237](https://doi.org/10.1002/jmv.26237). URL: <https://onlinelibrary.wiley.com/doi/full/10.1002/jmv.26237>.
- [6] Ho Namkoong et al. “Japan COVID-19 Task Force: a nationwide consortium to elucidate host genetics of COVID-19 pandemic in Japan”. In: *medRxiv* (2021), p. 2021.05.17.21256513. DOI: [10.1101/2021.05.17.21256513](https://doi.org/10.1101/2021.05.17.21256513). URL: <https://www.medrxiv.org/content/10.1101/2021.05.17.21256513v1>.
- [7] World Health Organization. *Obesity and Overweight*. URL: <https://www.who.int/en/news-room/fact-sheets/detail/obesity-and-overweight> (visited on 06/23/2021).
- [8] Mohammad Ali Mansournia and Douglas G. Altman. *Population attributable fraction*. 2018. DOI: [10.1136/bmj.k757](https://doi.org/10.1136/bmj.k757). URL: <http://www.bmj.com/permissionsSubscribe>:<http://www.bmj.com/subscribeBMJ2018;360:k757doi:10.1136/bmj.k757>.

- [9] Cristiana Abbafati et al. “Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019”. In: *The Lancet* 396.10258 (2020), pp. 1223–1249. ISSN: 1474547X. DOI: [10.1016/S0140-6736\(20\)30752-2](https://doi.org/10.1016/S0140-6736(20)30752-2). URL: www.thelancet.com.
- [10] Eric A. Finkelstein et al. “Annual medical spending attributable to obesity: Payer- and service-specific estimates”. In: *Health Affairs* 28.5 (2009). ISSN: 02782715. DOI: [10.1377/hlthaff.28.5.w822](https://doi.org/10.1377/hlthaff.28.5.w822).
- [11] John Cawley and Chad Meyerhoefer. “The medical care costs of obesity: An instrumental variables approach”. In: *Journal of Health Economics* 31.1 (2012), pp. 219–230. ISSN: 01676296. DOI: [10.1016/j.jhealeco.2011.10.003](https://doi.org/10.1016/j.jhealeco.2011.10.003).
- [12] R. Sturm. “Increases in morbid obesity in the USA: 2000–2005”. In: *Public Health* 121.7 (2007), pp. 492–496. ISSN: 00333506. DOI: [10.1016/j.puhe.2007.01.006](https://doi.org/10.1016/j.puhe.2007.01.006). URL: [/pmc/articles/PMC2864630/](https://pubmed.ncbi.nlm.nih.gov/PMC2864630/).
- [13] Maria Chiu et al. “Deriving ethnic-specific BMI cutoff points for assessing diabetes risk”. In: *Diabetes Care* 34.8 (2011), pp. 1741–1748. ISSN: 01495992. DOI: [10.2337/dc10-2300](https://doi.org/10.2337/dc10-2300).
- [14] Ali H. Mokdad et al. “Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001”. In: *Journal of the American Medical Association* 289.1 (2003), pp. 76–79. ISSN: 00987484. DOI: [10.1001/jama.289.1.76](https://doi.org/10.1001/jama.289.1.76). URL: <https://jamanetwork.com/>.
- [15] T. Pischon et al. “General and Abdominal Adiposity and Risk of Death in Europe”. In: *New England Journal of Medicine* 359.20 (2008), pp. 2105–2120. ISSN: 0028-4793. DOI: [10.1056/NEJMoa0801891](https://doi.org/10.1056/NEJMoa0801891). URL: <http://www.nejm.org/doi/abs/10.1056/NEJMoa0801891>.
- [16] Adolphe Quetelet. *A treatise on man and the development of his faculties*. 1842.
- [17] Garabed Eknoyan. “Adolphe Quetelet (1796–1874) – The average man and indices of obesity”. In: *Nephrology Dialysis Transplantation* 23.1 (2008). ISSN: 09310509. DOI: [10.1093/ndt/gfm517](https://doi.org/10.1093/ndt/gfm517).
- [18] Frank Q. Nuttall. *Body mass index: Obesity, BMI, and health: A critical review*. 2015. DOI: [10.1097/NT.0000000000000092](https://doi.org/10.1097/NT.0000000000000092). URL: [/pmc/articles/PMC4890841/](https://pubmed.ncbi.nlm.nih.gov/PMC4890841/).
- [19] Richard S. Taylor. *Use of body mass index for monitoring growth and obesity*. 2010. DOI: [10.1093/pch/15.5.258](https://doi.org/10.1093/pch/15.5.258). URL: [/pmc/articles/PMC2912631/](https://pubmed.ncbi.nlm.nih.gov/PMC2912631/).
- [20] Nick Trefethen. *New BMI (New Body Mass Index)*. URL: <https://people.maths.ox.ac.uk/trefethen/bmi.html> (visited on 06/25/2021).

- [21] T B VanItallie et al. “Height-normalized indices of the body’s fat-free mass and fat mass: potentially useful indicators of nutritional status”. In: *The American Journal of Clinical Nutrition* 52.6 (1990), pp. 953–959. ISSN: 0002-9165. DOI: [10.1093/AJCN/52.6.953](https://doi.org/10.1093/AJCN/52.6.953). URL: <https://academic.oup.com/ajcn/article/52/6/953/4651256>.
- [22] Y Schutz, UUG Kyle, and C Pichard. “Fat-free mass index and fat mass index percentiles in Caucasians aged 18–98 y”. In: *International Journal of Obesity* 26:7 26.7 (2002), pp. 953–960. ISSN: 1476-5497. DOI: [10.1038/sj.ijo.0802037](https://doi.org/10.1038/sj.ijo.0802037). URL: <https://www.nature.com/articles/0802037>.
- [23] Wikipedia. *Arnold Schwarzenegger*. URL: https://en.wikipedia.org/wiki/Arnold_Schwarzenegger (visited on 06/28/2021).
- [24] Center for Disease Control and Prevention. *About Adult BMI | Healthy Weight, Nutrition, and Physical Activity*. URL: https://www.cdc.gov/healthyweight/assessing/bmi/adult_bmi/index.html (visited on 06/28/2021).
- [25] A. Romero-Corral et al. “Accuracy of body mass index in diagnosing obesity in the adult general population”. In: *International Journal of Obesity* 32.6 (2008), pp. 959–966. ISSN: 03070565. DOI: [10.1038/ijo.2008.11](https://doi.org/10.1038/ijo.2008.11). URL: www.nature.com/ijo.
- [26] D. O. Okorodudu et al. *Diagnostic performance of body mass index to identify obesity as defined by body adiposity: A systematic review and meta-analysis*. 2010. DOI: [10.1038/ijo.2010.5](https://doi.org/10.1038/ijo.2010.5). URL: www.nature.com/ijo.
- [27] Marc Andre Cornier et al. “Assessing adiposity: A scientific statement from the american heart association”. In: *Circulation* 124.18 (2011), pp. 1996–2019. ISSN: 00097322. DOI: [10.1161/CIR.0b013e318233bc6a](https://doi.org/10.1161/CIR.0b013e318233bc6a). URL: <http://circ.ahajournals.org>.
- [28] Moonseong Heo et al. “Percentage of body fat cutoffs by sex, age, and race-ethnicity in the US adult population from NHANES 1999–2004”. In: *American Journal of Clinical Nutrition* 95.3 (2012), pp. 594–602. ISSN: 00029165. DOI: [10.3945/ajcn.111.025171](https://doi.org/10.3945/ajcn.111.025171). URL: <https://academic.oup.com/ajcn/article/95/3/594/4578297>.
- [29] Amy M. Goss et al. “Effects of diet macronutrient composition on body composition and fat distribution during weight maintenance and weight loss”. In: *Obesity* 21.6 (2013), pp. 1139–1142. ISSN: 1930-739X. DOI: [10.1002/OBY.20191](https://doi.org/10.1002/OBY.20191). URL: <https://onlinelibrary.wiley.com/doi/full/10.1002/oby.20191>.
- [30] Clemens Drenowatz et al. “The independent association between diet quality and body composition”. In: *Scientific Reports* 2014 4:1 4.1 (2014), pp. 1–6. ISSN: 2045-2322. DOI: [10.1038/srep04928](https://doi.org/10.1038/srep04928). URL: <https://www.nature.com/articles/srep04928>.

- [31] Frank B. Hu. “Physical Activity, Sedentary Behaviors, and Obesity”. In: *Obesity Epidemiology* (2009), pp. 301–319. DOI: [10.1093/ACPROF:050/9780195312911.003.0015](https://doi.org/10.1093/ACPROF:050/9780195312911.003.0015).
- [32] Neale lab. *UK Biobank — Neale lab*. 2018. URL: <http://www.nealelab.is/uk-biobank> (visited on 08/23/2021).
- [33] Cindy D. Davis. “The gut microbiome and its role in obesity”. In: *Nutrition Today* 51.4 (2016), pp. 167–174. ISSN: 15389839. DOI: [10.1097/NT.000000000000167](https://doi.org/10.1097/NT.000000000000167). URL: [/pmc/articles/PMC5082693/](https://pubmed.ncbi.nlm.nih.gov/3402693/).
- [34] Giovanna Muscogiuri et al. “Gut microbiota: a new path to treat obesity”. In: *International Journal of Obesity Supplements* 9.1 (2019), pp. 10–19. ISSN: 2046-2166. DOI: [10.1038/s41367-019-0011-7](https://doi.org/10.1038/s41367-019-0011-7). URL: <https://doi.org/10.1038/s41367-019-0011-7>.
- [35] Bei Fan Zhou. “Predictive values of body mass index and waist circumference for risk factors of certain related diseases in Chinese adults—study on optimal cut-off points of body mass index and waist circumference in Chinese adults.” In: *Biomedical and environmental sciences : BES* 15.1 (2002), pp. 83–96. ISSN: 08953988. DOI: [10.1046/j.1440-6047.11.s8.9.x](https://doi.org/10.1046/j.1440-6047.11.s8.9.x). URL: <https://onlinelibrary.wiley.com/doi/full/10.1046/j.1440-6047.11.s8.9.x>.
- [36] Yuji Matsuzawa et al. “New criteria for ‘obesity disease’ in Japan”. In: *Circulation Journal* 66.11 (2002), pp. 987–992. ISSN: 13469843. DOI: [10.1253/circj.66.987](https://doi.org/10.1253/circj.66.987). URL: <https://pubmed.ncbi.nlm.nih.gov/12419927/>.
- [37] J. Kullberg et al. “Whole-body adipose tissue analysis: Comparison of MRI, CT and dual energy X-ray absorptiometry”. In: *British Journal of Radiology* 82.974 (2009), pp. 123–130. ISSN: 00071285. DOI: [10.1259/bjr/80083156](https://doi.org/10.1259/bjr/80083156). URL: <https://pubmed.ncbi.nlm.nih.gov/19168691/>.
- [38] Magnus Borga et al. *Advanced body composition assessment: From body mass index to body composition profiling*. 2018. DOI: [10.1136/jim-2018-000722](https://doi.org/10.1136/jim-2018-000722). URL: <http://jim.bmj.com/>.
- [39] Guang Sun et al. “Comparison of multifrequency bioelectrical impedance analysis with dual-energy X-ray absorptiometry for assessment of percentage body fat in a large, healthy population”. In: *The American Journal of Clinical Nutrition* 81.1 (2005), pp. 74–78. DOI: [10.1093/ajcn/81.1.74](https://doi.org/10.1093/ajcn/81.1.74). URL: <https://academic.oup.com/ajcn/article/81/1/74/4607682>.
- [40] Najate Achamrah et al. “Comparison of body composition assessment by DXA and BIA according to the body mass index: A retrospective study on 3655 measures.” In: *PloS one* 13.7 (2018), e0200465. ISSN: 1932-6203. DOI: [10.1371/journal.pone.0200465](https://doi.org/10.1371/journal.pone.0200465). URL: <http://www.ncbi.nlm.nih.gov/pubmed/30001381>.

- [41] Maurizio Marra et al. *Assessment of body composition in health and disease using bioelectrical impedance analysis (bia) and dual energy x-ray absorptiometry (dxa): A critical overview*. 2019. DOI: [10.1155/2019/3548284](https://doi.org/10.1155/2019/3548284).
- [42] Carolina H.Y. Ling et al. “Accuracy of direct segmental multi-frequency bioimpedance analysis in the assessment of total body and segmental body composition in middle-aged adult population”. In: *Clinical Nutrition* 30.5 (2011), pp. 610–615. ISSN: 02615614. DOI: [10.1016/j.clnu.2011.04.001](https://doi.org/10.1016/j.clnu.2011.04.001). URL: <http://www.clinicalnutritionjournal.com/article/S0261561411000665/fulltext>.
- [43] Tommy Boone et al. “JEPonline Validating InBody® 570 Multi-frequency Bioelectrical Impedance Analyzer versus DXA for Body Fat Percentage Analysis”. In: *Journal of Exercise Physiologyonline* October 19 (2016). ISSN: 1097-9751.
- [44] Geesje H. Hofsteenge, Mai J.M. Chinapaw, and Peter J.M. Weijs. “Fat-free mass prediction equations for bioelectric impedance analysis compared to dual energy X-ray absorptiometry in obese adolescents: A validation study”. In: *BMC Pediatrics* 15.1 (2015). ISSN: 14712431. DOI: [10.1186/s12887-015-0476-7](https://doi.org/10.1186/s12887-015-0476-7). URL: [/pmc/articles/PMC4608267/](http://pmc/articles/PMC4608267/).
- [45] B. M. Kaess et al. “The ratio of visceral to subcutaneous fat, a metric of body fat distribution, is a unique correlate of cardiometabolic risk”. In: *Diabetologia* 55.10 (2012), pp. 2622–2630. ISSN: 0012-186X. DOI: [10.1007/s00125-012-2639-5](https://doi.org/10.1007/s00125-012-2639-5). URL: <http://link.springer.com/10.1007/s00125-012-2639-5>.
- [46] Klara J. Rosenquist et al. “Visceral and Subcutaneous Fat Quality and Cardiometabolic Risk”. In: *JACC: Cardiovascular Imaging* 6.7 (2013), pp. 762–771. ISSN: 1936-878X. DOI: [10.1016/J.JCMG.2012.11.021](https://doi.org/10.1016/J.JCMG.2012.11.021). URL: <https://www.sciencedirect.com/science/article/pii/S1936878X13002696?via%3Dihub>.
- [47] Tobin M Abraham et al. “Association between visceral and subcutaneous adipose depots and incident cardiovascular disease risk factors.” In: *Circulation* 132.17 (2015), pp. 1639–47. ISSN: 1524-4539. DOI: [10.1161/CIRCULATIONAHA.114.015000](https://doi.org/10.1161/CIRCULATIONAHA.114.015000). URL: <https://www.ahajournals.org/doi/10.1161/CIRCULATIONAHA.114.015000>.
- [48] Sam Virtue and Antonio Vidal-Puig. “Adipose tissue expandability, lipotoxicity and the Metabolic Syndrome — An allostatic perspective”. In: *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* 1801.3 (2010), pp. 338–349. ISSN: 1388-1981. DOI: [10.1016/J.BBALIP.2009.12.006](https://doi.org/10.1016/J.BBALIP.2009.12.006). URL: <https://www.sciencedirect.com/science/article/pii/S1388198109002868?via%3Dihub>.
- [49] Satoshi Nishimura, Ichiro Manabe, and Ryoza Nagai. “Adipose Tissue Inflammation in Obesity and Metabolic Syndrome”. In: *Discovery Medicine* (2009).

- [50] Shannon M. Reilly and Alan R. Saltiel. “Adapting to obesity with adipose tissue inflammation”. In: *Nature Reviews Endocrinology* 2017 13:11 13.11 (2017), pp. 633–643. ISSN: 1759–5037. DOI: [10.1038/nrendo.2017.90](https://doi.org/10.1038/nrendo.2017.90). URL: <https://www.nature.com/articles/nrendo.2017.90>.
- [51] Hanieh Yaghoobkar et al. “Genetic evidence for a normal-weight “metabolically obese” phenotype linking insulin resistance, hypertension, coronary artery disease, and type 2 diabetes.” In: *Diabetes* 63.12 (2014), pp. 4369–77. ISSN: 1939–327X. DOI: [10.2337/db14-0318](https://doi.org/10.2337/db14-0318). URL: <http://www.ncbi.nlm.nih.gov/pubmed/25048195>.
- [52] Thomas W Winkler et al. “A joint view on genetic variants for adiposity differentiates subtypes with distinct metabolic implications”. In: *Nature Communications* 9.1 (2018), p. 1946. ISSN: 2041–1723. DOI: [10.1038/s41467-018-04124-9](https://doi.org/10.1038/s41467-018-04124-9). URL: <http://www.nature.com/articles/s41467-018-04124-9>.
- [53] Vivian Tam et al. “Benefits and limitations of genome-wide association studies”. In: *Nature Reviews Genetics* 2019 20:8 20.8 (2019), pp. 467–484. ISSN: 1471–0064. DOI: [10.1038/s41576-019-0127-1](https://doi.org/10.1038/s41576-019-0127-1). URL: <https://www.nature.com/articles/s41576-019-0127-1>.
- [54] Cathie Sudlow et al. “UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age”. In: *PLOS Medicine* 12.3 (2015), e1001779. ISSN: 1549–1676. DOI: [10.1371/journal.pmed.1001779](https://doi.org/10.1371/journal.pmed.1001779). URL: <https://dx.plos.org/10.1371/journal.pmed.1001779>.
- [55] Ruth J. F. Loos. “15 years of genome-wide association studies and no signs of slowing down”. In: *Nature Communications* 2020 11:1 11.1 (2020), pp. 1–3. ISSN: 2041–1723. DOI: [10.1038/s41467-020-19653-5](https://doi.org/10.1038/s41467-020-19653-5). URL: <https://www.nature.com/articles/s41467-020-19653-5>.
- [56] J. Graham Ruby et al. “Estimates of the Heritability of Human Longevity Are Substantially Inflated due to Assortative Mating”. In: *Genetics* 210.3 (2018), pp. 1109–1124. ISSN: 0016–6731. DOI: [10.1534/GENETICS.118.301613](https://doi.org/10.1534/GENETICS.118.301613). URL: <https://www.genetics.org/content/210/3/1109>.
- [57] Loic Yengo et al. “Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry”. In: *Human molecular genetics* 27.20 (2018). ISSN: 14602083. DOI: [10.1093/hmg/ddy271](https://doi.org/10.1093/hmg/ddy271). arXiv: [1105](https://arxiv.org/abs/1105).
- [58] Sara L Pulit et al. “Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry”. In: *Human Molecular Genetics* 28.1 (2019), pp. 166–174. ISSN: 0964–6906. DOI: [10.1093/hmg/ddy327](https://doi.org/10.1093/hmg/ddy327). URL:

- <https://academic.oup.com/hmg/article/28/1/166/5098227>.
- [59] Adam E. Locke et al. “Genetic studies of body mass index yield new insights for obesity biology”. In: *Nature* 518.7538 (2015), pp. 197–206. ISSN: 0028–0836. DOI: [10.1038/nature14177](https://doi.org/10.1038/nature14177). URL: <http://www.nature.com/doifinder/10.1038/nature14177>.
- [60] Valérie Turcot et al. “Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure underpinning obesity”. In: *Nature genetics* 50.1 (2018), p. 26. DOI: [10.1038/S41588-017-0011-X](https://doi.org/10.1038/S41588-017-0011-X). URL: [/pmc/articles/PMC5945951/](https://pmc/articles/PMC5945951/).
- [61] Anne E. Justice et al. “Protein-coding variants implicate novel genes related to lipid homeostasis contributing to body-fat distribution”. In: *Nature Genetics* 51.3 (2019), pp. 452–469. ISSN: 15461718. DOI: [10.1038/s41588-018-0334-2](https://doi.org/10.1038/s41588-018-0334-2).
- [62] Hanieh Yaghoobkar et al. “Genetic Evidence for a Link Between Favorable Adiposity and Lower Risk of Type 2 Diabetes, Hypertension, and Heart Disease.” In: *Diabetes* 65.8 (2016), pp. 2448–60. ISSN: 1939–327X. DOI: [10.2337/db15-1671](https://doi.org/10.2337/db15-1671). URL: <http://www.ncbi.nlm.nih.gov/pubmed/27207519>.
- [63] Yingjie Ji et al. “Genome-Wide and Abdominal MRI Data Provide Evidence That a Genetically Determined Favorable Adiposity Phenotype Is Characterized by Lower Ectopic Liver Fat and Lower Risk of Type 2 Diabetes, Heart Disease, and Hypertension.” In: *Diabetes* 68.1 (2019), pp. 207–219. ISSN: 1939–327X. DOI: [10.2337/db18-0708](https://doi.org/10.2337/db18-0708). URL: <http://www.ncbi.nlm.nih.gov/pubmed/30352878>.
- [64] Sven Bergmann, Jan Ihmels, and Naama Barkai. “Iterative signature algorithm for the analysis of large-scale gene expression data”. In: *Physical Review E* 67.3 (2003), p. 031902. ISSN: 1063–651X. DOI: [10.1103/PhysRevE.67.031902](https://doi.org/10.1103/PhysRevE.67.031902). URL: <http://www.ncbi.nlm.nih.gov/pubmed/12689096>.
- [65] Zoltán Kutalik, Jacques S Beckmann, and Sven Bergmann. “A modular approach for integrative analysis of large-scale gene-expression and drug-response data”. In: *Nature Biotechnology* 26:5 26.5 (2008), pp. 531–539. ISSN: 1546–1696. DOI: [10.1038/nbt1397](https://doi.org/10.1038/nbt1397). URL: <https://www.nature.com/articles/nbt1397>.
- [66] Janina S. Ried et al. “A principal component meta-analysis on multiple anthropometric traits identifies novel loci for body shape”. In: *Nature Communications* 7.1 (2016), p. 13357. ISSN: 2041–1723. DOI: [10.1038/ncomms13357](https://doi.org/10.1038/ncomms13357). URL: <http://www.nature.com/articles/ncomms13357>.

- [67] Martijn B. Katan. “Apolipoprotein E isoforms, serum cholesterol, and cancer”. In: *The Lancet* 327.8479 (1986), pp. 507–508. DOI: [10.1016/S0140-6736\(86\)92972-7](https://doi.org/10.1016/S0140-6736(86)92972-7). URL: <https://www.sciencedirect.com/science/article/pii/S0140673686929727?via%3Dihub>.
- [68] R Gray and K Wheatley. “How to avoid bias when comparing bone marrow transplantation with chemotherapy.” In: *Bone marrow transplantation* 7 Suppl 3 (1991), pp. 9–12. URL: <http://www.ncbi.nlm.nih.gov/pubmed/1855097>.
- [69] D. H. Hamer and L. Sirota. “Beware the chopsticks gene”. In: *Molecular Psychiatry* 5.1 (2000), pp. 11–13. DOI: [10.1038/SJ.MP.4000662](https://doi.org/10.1038/SJ.MP.4000662).
- [70] Stephen Burgess, Adam Butterworth, and Simon G. Thompson. “Mendelian Randomization Analysis With Multiple Genetic Variants Using Summarized Data”. In: *Genetic Epidemiology* 37.7 (2013), pp. 658–665. DOI: [10.1002/gepi.21758](https://doi.org/10.1002/gepi.21758). URL: <http://doi.wiley.com/10.1002/gepi.21758>.
- [71] Jack Bowden, George Davey Smith, and Stephen Burgess. “Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression”. In: *International Journal of Epidemiology* 44.2 (2015), pp. 512–525. ISSN: 0300-5771. DOI: [10.1093/ije/dyv080](https://doi.org/10.1093/ije/dyv080). URL: <https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dyv080>.
- [72] Jack Bowden et al. “Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator.” In: *Genetic epidemiology* 40.4 (2016), pp. 304–314. ISSN: 1098-2272. DOI: [10.1002/gepi.21965](https://doi.org/10.1002/gepi.21965). URL: <http://www.ncbi.nlm.nih.gov/pubmed/27061298>.
- [73] Fernando Pires Hartwig, George Davey Smith, and Jack Bowden. “Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption”. In: *International Journal of Epidemiology* 46.6 (2017), p. 1985. DOI: [10.1093/IJE/DYX102](https://doi.org/10.1093/IJE/DYX102). URL: [/pmc/articles/PMC5837715/](https://pubmed.ncbi.nlm.nih.gov/35837715/).
- [74] Marie Verbanck et al. “Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases”. In: *Nature Genetics* 50.5 (2018), pp. 693–698. ISSN: 1061-4036. DOI: [10.1038/s41588-018-0099-7](https://doi.org/10.1038/s41588-018-0099-7). URL: <http://www.nature.com/articles/s41588-018-0099-7>.
- [75] Christopher N Foley et al. “MR-Clust: clustering of genetic variants in Mendelian randomization with similar causal estimates”. In: *Bioinformatics* 37.4 (2021), pp. 531–541. ISSN: 1367-4803. DOI: [10.1093/BIOINFORMATICS/BTAA778](https://doi.org/10.1093/BIOINFORMATICS/BTAA778). URL: <https://academic.oup.com/bioinformatics/article/37/4/531/5904264>.

- [76] Jessica M B Rees, Christopher N Foley, and Stephen Burgess. “Factorial Mendelian randomization: using genetic variants to assess interactions”. In: *International Journal of Epidemiology* 49.4 (2020), pp. 1147–1158. ISSN: 0300-5771. DOI: [10.1093/ije/dyz161](https://doi.org/10.1093/ije/dyz161). URL: <https://academic.oup.com/ije/article/49/4/1147/5542600>.
- [77] Jean Morrison et al. “Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics”. In: *Nature Genetics* 2020 52:7 52.7 (2020), pp. 740–747. ISSN: 1546-1718. DOI: [10.1038/s41588-020-0631-4](https://doi.org/10.1038/s41588-020-0631-4). URL: <https://www.nature.com/articles/s41588-020-0631-4>.
- [78] Liza Darrous, Ninon Mounier, and Zoltán Kutalik. “Simultaneous estimation of bi-directional causal effects and heritable confounding from GWAS summary statistics”. In: *medRxiv* (2020), p. 2020.01.27.20018929. DOI: [10.1101/2020.01.27.20018929](https://doi.org/10.1101/2020.01.27.20018929). URL: <https://www.medrxiv.org/content/10.1101/2020.01.27.20018929v1>.
- [79] Luke J. O’Connor and Alkes L. Price. “Distinguishing genetic correlation from causation across 52 diseases and complex traits”. In: *Nature Genetics* (2018), p. 1. ISSN: 1061-4036. DOI: [10.1038/s41588-018-0255-0](https://doi.org/10.1038/s41588-018-0255-0). URL: <http://www.nature.com/articles/s41588-018-0255-0>.
- [80] S. Burgess and S. G. Thompson. “Multivariable Mendelian Randomization: The Use of Pleiotropic Genetic Variants to Estimate Causal Effects”. In: *American Journal of Epidemiology* 181.4 (2015), pp. 251–260. DOI: [10.1093/aje/kwu283](https://doi.org/10.1093/aje/kwu283). URL: <https://academic.oup.com/aje/article-lookup/doi/10.1093/aje/kwu283>.
- [81] Eric A. W. Slob and Stephen Burgess. “A comparison of robust Mendelian randomization methods using summary data”. In: *Genetic Epidemiology* 44.4 (2020), pp. 313–329. ISSN: 1098-2272. DOI: [10.1002/GEPI.22295](https://doi.org/10.1002/GEPI.22295). URL: <https://onlinelibrary.wiley.com/doi/full/10.1002/gepi.22295>.
- [82] Andrew J. Grant and Stephen Burgess. “Pleiotropy robust methods for multivariable Mendelian randomization”. In: *Statistics in Medicine* (2021), sim.9156. ISSN: 1097-0258. DOI: [10.1002/SIM.9156](https://doi.org/10.1002/SIM.9156). URL: <https://onlinelibrary.wiley.com/doi/full/10.1002/sim.9156>.
- [83] James R. Staley and Stephen Burgess. “Semiparametric methods for estimation of a nonlinear exposure-outcome relationship using instrumental variables with application to Mendelian randomization”. In: *Genetic Epidemiology* 41.4 (2017), pp. 341–352. ISSN: 07410395. DOI: [10.1002/gepi.22041](https://doi.org/10.1002/gepi.22041). URL: <http://doi.wiley.com/10.1002/gepi.22041>.

- [84] Sai Li and Zijian Guo. “Causal Inference for Nonlinear Outcome Models with Possibly Invalid Instrumental Variables”. In: (). arXiv: [2010.09922v1](https://arxiv.org/abs/2010.09922v1).
- [85] James J. Heckman and Richard Robb. “Alternative methods for evaluating the impact of interventions: An overview”. In: *Journal of Econometrics* 30.1-2 (1985), pp. 239–267. ISSN: 0304-4076. DOI: [10.1016/0304-4076\(85\)90139-3](https://doi.org/10.1016/0304-4076(85)90139-3).
- [86] Sai Li. *SpotIV*. URL: <https://github.com/saili0103/SpotIV> (visited on 08/04/2021).
- [87] Jonathan Sulc et al. “Composite trait Mendelian randomization reveals distinct metabolic and lifestyle consequences of differences in body shape”. In: *Communications Biology* 2021 4:1 4.1 (2021), pp. 1–13. ISSN: 2399-3642. DOI: [10.1038/s42003-021-02550-y](https://doi.org/10.1038/s42003-021-02550-y). URL: <https://www.nature.com/articles/s42003-021-02550-y>.
- [88] Lloyd T. Elliott et al. “Genome-wide association studies of brain imaging phenotypes in UK Biobank”. In: *Nature* 562.7726 (2018), pp. 210–216. ISSN: 0028-0836. DOI: [10.1038/s41586-018-0571-7](https://doi.org/10.1038/s41586-018-0571-7). URL: <http://www.nature.com/articles/s41586-018-0571-7>.
- [89] Stephen M. Smith et al. “Accurate, Robust, and Automated Longitudinal and Cross-Sectional Brain Change Analysis”. In: *NeuroImage* 17.1 (2002), pp. 479–489. ISSN: 1053-8119. DOI: [10.1006/NIMG.2002.1040](https://doi.org/10.1006/NIMG.2002.1040). URL: <https://www.sciencedirect.com/science/article/pii/S1053811902910402?via%3Dihub>.
- [90] S. Sgouros et al. “Skull Base Growth in Childhood”. In: *Pediatric Neurosurgery* 31.5 (1999), pp. 259–268. ISSN: 1016-2291. DOI: [10.1159/000028873](https://doi.org/10.1159/000028873). URL: <https://www.karger.com/Article/FullText/28873>.
- [91] Jonathan Sulc and Zoltán Kutalik. “PolyMR: Polynomial Mendelian randomization for the inference of non-linear causal effects”. 2021. URL: <https://github.com/JonSulc/PolyMR>.
- [92] Doug Speed et al. “Reevaluation of SNP heritability in complex human traits”. In: *Nature Genetics* 2017 49:7 49.7 (2017), pp. 986–992. ISSN: 1546-1718. DOI: [10.1038/ng.3865](https://doi.org/10.1038/ng.3865). URL: <https://www.nature.com/articles/ng.3865>.
- [93] Doug Speed, John Holmes, and David J. Balding. “Evaluating and improving heritability models using summary statistics”. In: *Nature Genetics* 2020 52:4 52.4 (2020), pp. 458–462. ISSN: 1546-1718. DOI: [10.1038/s41588-020-0600-y](https://doi.org/10.1038/s41588-020-0600-y). URL: <https://www.nature.com/articles/s41588-020-0600-y>.
- [94] Gibran Hemani et al. “The MR-Base platform supports systematic causal inference across the human phenome”. In: *eLife* 7 (2018). DOI: [10.7554/eLife.34408](https://doi.org/10.7554/eLife.34408). URL: <https://elifesciences.org/articles/34408>.

- [95] Gibran Hemani. *MRCIEU/ieugwasr: R interface to the IEU GWAS database API*. URL: <https://github.com/MRCIEU/ieugwasr>.
- [96] Jonathan Sulc et al. *GWAS summary statistics for waist-to-hip ratio and body principal components*. 2021. DOI: [10.5281/ZENODO.5171807](https://doi.org/10.5281/ZENODO.5171807). URL: <https://zenodo.org/record/5171807>.
- [97] Lin Xu et al. “The role of glycaemic and lipid risk factors in mediating the effect of BMI on coronary heart disease: a two-step, two-sample Mendelian randomisation study”. In: *Diabetologia* 60.11 (2017), pp. 2210–2220. ISSN: 0012-186X. DOI: [10.1007/s00125-017-4396-y](https://doi.org/10.1007/s00125-017-4396-y). URL: <http://link.springer.com/10.1007/s00125-017-4396-y>.
- [98] Xun Hu et al. “Exploring the causal pathway from body mass index to coronary heart disease: a network Mendelian randomization study”. In: *Therapeutic Advances in Chronic Disease* 11 (2020), p. 204062232090904. ISSN: 2040-6223. DOI: [10.1177/2040622320909040](https://doi.org/10.1177/2040622320909040). URL: <http://journals.sagepub.com/doi/10.1177/2040622320909040>.
- [99] Michael V. Holmes et al. “Causal effects of body mass index on cardiometabolic traits and events: A Mendelian randomization analysis”. In: *American Journal of Human Genetics* 94.2 (2014), pp. 198–208. ISSN: 00029297. DOI: [10.1016/j.ajhg.2013.12.014](https://doi.org/10.1016/j.ajhg.2013.12.014). URL: [/pmc/articles/PMC3928659/?report=abstract](https://pubmed.ncbi.nlm.nih.gov/24811904/).
- [100] Felix Elwert and Christopher Winship. “Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable.” In: *Annual review of sociology* 40 (2014), pp. 31–53. ISSN: 0360-0572. DOI: [10.1146/annurev-soc-071913-043455](https://doi.org/10.1146/annurev-soc-071913-043455). URL: <http://www.ncbi.nlm.nih.gov/pubmed/30111904>.
- [101] Jonathan Sulc et al. “Quantification of the overall contribution of gene–environment interaction for obesity–related traits”. In: *Nature Communications* 2020 11:1 11.1 (2020), pp. 1–13. ISSN: 2041-1723. DOI: [10.1038/s41467-020-15107-0](https://doi.org/10.1038/s41467-020-15107-0). URL: <https://www.nature.com/articles/s41467-020-15107-0>.
- [102] Leonore M de Wit et al. “Depression and body mass index, a u-shaped association”. In: *BMC Public Health* 2009 9:1 9.1 (2009), pp. 1–6. ISSN: 1471-2458. DOI: [10.1186/1471-2458-9-14](https://doi.org/10.1186/1471-2458-9-14). URL: <https://bmcpublihealth.biomedcentral.com/articles/10.1186/1471-2458-9-14>.
- [103] Bhautesh Dinesh Jani et al. “Revisiting the J shaped curve, exploring the association between cardiovascular risk factors and concurrent depressive symptoms in patients with cardiometabolic disease: Findings from a large cross-sectional study”. In: *BMC Cardiovascular Disorders* 2014 14:1 14.1 (2014), pp. 1–8. ISSN: 1471-2261. DOI: [10.1186/1471-2261-14-139](https://doi.org/10.1186/1471-2261-14-139). URL: <https://bmccardiovascdisord.biomedcentral.com/articles/10.1186/1471-2261-14-139>.

- [104] The Emerging Risk Factors Collaboration. “Separate and combined associations of body-mass index and abdominal adiposity with cardiovascular disease: collaborative analysis of 58 prospective studies”. In: *The Lancet* 377.9771 (2011), pp. 1085–1095. ISSN: 0140-6736. DOI: [10.1016/S0140-6736\(11\)60105-0](https://doi.org/10.1016/S0140-6736(11)60105-0). URL: <http://www.thelancet.com/article/S0140673611601050/fulltext>.
- [105] Wei Zhou et al. “Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies”. In: *Nature Genetics* 2018 50:9 50.9 (2018), pp. 1335–1341. ISSN: 1546-1718. DOI: [10.1038/s41588-018-0184-y](https://doi.org/10.1038/s41588-018-0184-y). URL: <https://www.nature.com/articles/s41588-018-0184-y>.
- [106] Maria S. Speed et al. “Investigating the association between body fat and depression via Mendelian randomization”. In: *Translational Psychiatry* 2019 9:1 9.1 (2019), pp. 1–9. ISSN: 2158-3188. DOI: [10.1038/s41398-019-0516-4](https://doi.org/10.1038/s41398-019-0516-4). URL: <https://www.nature.com/articles/s41398-019-0516-4>.
- [107] Nicola Pirastu et al. “Using genetics to disentangle the complex relationship between food choices and health status”. 2019. DOI: [10.1101/829952](https://doi.org/10.1101/829952). URL: <https://doi.org/10.1101/829952>.
- [108] Alice A. Gibson and Amanda Sainsbury. “Strategies to Improve Adherence to Dietary Weight Loss Interventions in Research and Real-World Settings”. In: *Behavioral Sciences* 7.3 (2017). DOI: [10.3390/BS7030044](https://doi.org/10.3390/BS7030044). URL: [/pmc/articles/PMC5618052/](https://pmc/articles/PMC5618052/).
- [109] Eric T. Poehlman. “Abdominal obesity: the metabolic multi-risk factor”. In: *Coronary Artery Disease* 9.8 (1998), pp. 469–472. URL: https://journals.lww.com/coronary-artery/Citation/1998/09080/Abdominal_obesity_the_metabolic_multi_risk_factor.1.aspx.
- [110] Robert Ross et al. “Waist circumference as a vital sign in clinical practice: a Consensus Statement from the IAS and ICCR Working Group on Visceral Obesity”. In: *Nature Reviews Endocrinology* 2020 16:3 16.3 (2020), pp. 177–189. ISSN: 1759-5037. DOI: [10.1038/s41574-019-0310-7](https://doi.org/10.1038/s41574-019-0310-7). URL: <https://www.nature.com/articles/s41574-019-0310-7>.
- [111] Declan Butler. “Translational research: Crossing the valley of death”. In: *Nature* 453.7197 (2008), pp. 840–842. DOI: [10.1038/453840A](https://doi.org/10.1038/453840A).
- [112] Nina Fudge et al. “Optimising Translational Research Opportunities: A Systematic Review and Narrative Synthesis of Basic and Clinician Scientists’ Perspectives of Factors Which Enable or Hinder Translational Research”. In: *PLOS ONE* 11.8 (2016), e0160475. ISSN: 1932-6203. DOI: [10.1371/JOURNAL.PONE.0160475](https://doi.org/10.1371/JOURNAL.PONE.0160475). URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0160475>.

- [113] G R Warnick et al. "Estimating low-density lipoprotein cholesterol by the Friedewald equation is adequate for classifying patients on the basis of nationally recommended cutpoints." eng. In: *Clinical chemistry* 36.1 (1990), pp. 15–19. ISSN: 0009-9147 (Print).