View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Explore Bristol Research

Richardson, T. G., Sanderson, E., Elsworth, B. L., Tilling, K. M., & Davey Smith, G. (2020). Use of genetic variation to separate the effects of early and later life adiposity on disease risk: mendelian randomization study. *BMJ*, *369*(8244), [2020;369:m1203]. https://doi.org/10.1136/bmj.m1203

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
10.1136/bmj.m1203

Link to publication record in Explore Bristol Research
PDF-document

This is the final published version of the article (version of record). It first appeared online via BMJ Publishing Group at doi: https://doi.org/10.1136/bmj.m1203 . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/

OPEN ACCESS

Check for updates

Use of genetic variation to separate the effects of early and later life adiposity on disease risk: mendelian randomisation study

Tom G Richardson, Eleanor Sanderson, Benjamin Elsworth, Kate Tilling, George Davey Smith

MRC Integrative Epidemiology Unit (IEU), Population Health Sciences, Bristol Medical School, University of Bristol, Oakfield House, Oakfield Grove, Bristol, BS8 2BN, UK

Correspondence to: T G Richardson Tom.G.Richardson@bristol.ac.uk (or @iTomRichardson on Twitter ORCID 0000-0002-7918-2040)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2020;369:m1203
http://dx.doi.org/10.1136/bmj.m1203

Accepted: 17 March 2020

ABSTRACT

OBJECTIVE
To evaluate whether body size in early life has an independent effect on risk of disease in later life or whether its influence is mediated by body size in adulthood.

DESIGN
Two sample univariable and multivariable mendelian randomisation.

SETTING
The UK Biobank prospective cohort study and four large scale genome-wide association studies (GWAS) consortiums.

PARTICIPANTS
453 169 participants enrolled in UK Biobank and a combined total of more than 700 000 people from different GWAS consortiums.

EXPOSURES
Measured body mass index during adulthood (mean age 56.5) and self-reported perceived body size at age 10.

MAIN OUTCOME MEASURES
Coronary artery disease, type 2 diabetes, breast cancer, and prostate cancer.

RESULTS
Having a larger genetically predicted body size in early life was associated with an increased odds of coronary artery disease (odds ratio 1.49 for each change in body size category unless stated otherwise, 95% confidence interval 1.33 to 1.68) and type 2 diabetes (2.32, 1.76 to 3.05) based on univariable mendelian randomisation analyses. However, little evidence was found of a direct effect (ie, not through adult body size) based on multivariable mendelian randomisation estimates (coronary artery disease: 1.02, 0.86 to 1.22; type 2 diabetes:1.16, 0.74 to 1.82). In the multivariable mendelian randomisation analysis of breast cancer risk, strong evidence was found of a protective direct effect for larger body size in early life (0.59, 0.50 to 0.71), with less evidence of a direct effect of adult body size on this outcome (1.08, 0.93 to 1.27). Including age at menarche as an additional exposure provided weak evidence of a total causal effect (univariable mendelian randomisation odds ratio 0.98, 95% confidence interval 0.91 to 1.06) but strong evidence of a direct causal effect, independent of early life and adult body size (multivariable mendelian randomisation odds ratio 0.90, 0.85 to 0.95). No strong evidence was found of a causal effect of either early or later life measures on prostate cancer (early life body size odds ratio 1.06, 95% confidence interval 0.81 to 1.40; adult body size 0.87, 0.70 to 1.08).

CONCLUSIONS
The findings suggest that the positive association between body size in childhood and risk of coronary artery disease and type 2 diabetes in adulthood can be attributed to individuals remaining large into later life. However, having a smaller body size during childhood might increase the risk of breast cancer regardless of body size in adulthood, with timing of puberty also putatively playing a role.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Obesity in childhood is known to have a detrimental impact on various health conditions and disease risk in later life

Body size in early life has been associated with increased odds of coronary artery disease and type 2 diabetes observationally, but whether this effect is immutable or whether lifestyle modifications can help mitigate it is unclear

The influence of early life body size on risk of other diseases, such as breast cancer and prostate cancer, has been previously reported but the causal relations underlying these observations are complex and require further evaluation

WHAT THIS STUDY ADDS

The influence of genetically predicted early life body size on odds of coronary artery disease and type 2 diabetes was predominantly mediated through later life body size, suggesting that the influence of body size on these outcomes is attributed to a cumulative risk conferred throughout the life course

Findings for body size at age 10 suggested that it had a protective effect on breast cancer risk independent of later life body size, although timing of puberty might play a causal role in this relation

This study focused on cardiometabolic and cancer outcomes, but the analytical approach has potential to investigate how risk factor trajectories across life influence a range of health outcomes

Introduction

Obesity in children is widely recognised as a global public health crisis, yet its prevalence continues to rise.[1 2] Having a high body mass index (BMI) in early life is thought to increase the risk of various health conditions, such as coronary artery disease, type 2 diabetes, and different types of cancer, in later life.[3-7] Whether an individual can reverse the impact of childhood obesity through lifestyle modifications is unclear, particularly as those who are obese in early life tend to remain obese as adults.[8] This makes it challenging to discern whether early life adiposity has an independent and lasting influence on disease risk or if its effect is entirely mediated by later life adiposity. If the latter is the case, then the potential adverse consequences of childhood obesity could be avoided by attaining and maintaining a healthy weight in adulthood.

Mendelian randomisation is an approach that can help deal with these challenges[9] through the use of genetic variants as instrumental variables

BMJ: first published as 10.1136/bmj.m1203 on 6 May 2020. Downloaded from http://www.bmj.com/ on 25 June 2020 at BVA. Protected by copyright.

to infer causality among correlated traits.[10 11] As an individual's genotype is established at zygote formation then genetic variation is robust to reverse causation and confounding is considerably less evident than in conventional observational studies.[12] Furthermore, recent methods have been developed to determine whether several exposures influence an outcome along the same causal pathway or whether the exposures have independent effects.[13 14] One of these developments is known as multivariable mendelian randomisation.[13 15]

In this mendelian randomisation study we evaluated whether genetically predicted early life body size has an effect on four disease outcomes that have been linked to childhood adiposity: coronary artery disease, type 2 diabetes, breast cancer, and prostate cancer.[3 4 7 16] We identified genetic instruments by undertaking a genome-wide association study of 453 169 participants in the UK Biobank study with measures of BMI in adulthood (mean age 56.5) and who self-reported perceived body size at age 10 years. This allowed us to conduct both univariable and multivariable mendelian randomisation analyses to discern whether the predicted causal influence of early life body size has an independent effect on disease risk, or whether the effect is mediated through later life body size.

We postulated that if body size in early life has a causal influence on disease risk, then there are likely three main categories that such effects could be grouped into in a multivariable framework after accounting for body size during adulthood (fig 1). Firstly, the influence of early life body size on disease risk could be mediated by later life body size. As such, early life body size only has an indirect effect on disease risk through adult body size, which, for example, could be attributed to those with a large body size at age 10 remaining overweight into adulthood (fig 1 top panel). Secondly, early life body size might have only an independent (ie, direct) effect on disease risk, which is not mediated through adult body size (fig 1 middle panel). Thirdly, body size at separate time points might influence disease risk through alternate causal pathways (ie, early life body size has both a direct and an indirect effect on outcome; fig 1 bottom panel). These scenarios are possible to investigate using multivariable mendelian randomisation given that a genome-wide association study has recently shown that genetic variation might have varying effects on body size at different stages in the life course.[17] Supplementary figure 1 provides additional descriptions of direct, indirect, and total effects within a multivariable framework.

Methods

UK Biobank and disease outcome datasets

Between 2006 and 2010 the UK Biobank study enrolled 500 000 adults aged between 40 and 69 years at baseline across 22 assessments centres in the United Kingdom.[18] Data were collected based on clinical examinations, assays of biological samples, detailed information on self-reported health characteristics, and genome-wide genotyping.[19] BMI was derived using height (measured in whole centimetres) and weight (to the nearest 0.1 kg) measured at baseline. Participants were also asked "When you were 10 years old, compared to average would you describe yourself as thinner, plumper, or about average?" This measure is referred to as early life body size. We performed validation and simulation analyses to account for possible limitations of using perceived body size rather than a measured variable. Only those with measures in both early life and later life were included in analyses. BMI in adults was converted into a categorical variable with three groups based on the same proportions as the early life body size variable (ie, thinner, plumper, and about average). This was to ensure that both measures were as comparable as possible. This measure is referred to as adult body size. Effect estimates from our results can be interpreted as the increase in odds conferred for each additive change in body size category. All individual participant data used in this study were obtained from the UK Biobank study. Participants enrolled in UK Biobank have signed consent forms.

In total, 12 370 749 genetic variants in 463 005 people were available for analysis, as described previously.[20] Briefly, UK Biobank participants were selected based on those of European descent (using K means clustering (K=4)) after standard exclusions, including withdrawn consent, mismatch between genetic and reported sex, and putative sex chromosome aneuploidy.[21] Effect estimates for genome-wide genetic variants on coronary artery disease, type 2 diabetes, breast cancer, and prostate cancer were obtained using findings from large scale consortiums, which did not include data from the UK Biobank (supplementary table 1).[22-25] Since earlier age at menarche is an established risk factor for breast cancer we also obtained summary statistics from a genome-wide association study of age at menarche that did not include individual's from the UK Biobank, which was used in additional analyses relating to this outcome.[26]

Avon Longitudinal Study of Parents and Children

The Avon Longitudinal Study of Parents and Children (ALSPAC) is a population based cohort study investigating genetic and environmental factors that affect the health and development of children. The study methods are described in detail elsewhere.[27 28] Briefly, 14 541 pregnant women residing in the former region of Avon, UK, with an expected delivery date between 1 April 1991 and 31 December 1992 were eligible to take part in ALSPAC. Detailed phenotypic information, biological samples, and genetic data have been collected from the participants, which are available through a searchable data dictionary (http://www.bristol.ac.uk/alspac/researchers/our-data/). Written informed consent was obtained for all study participants. Ethical approval for the study was obtained from the ALSPAC ethics and law committee and the local research ethics committees. Supplementary note 1 provides further details on genotyping and trait measurements in the ALSPAC cohort.

BMJ: first published as 10.1136/bmj.m1203 on 6 May 2020. Downloaded from http://www.bmj.com/ on 25 June 2020 at BVA. Protected by copyright.

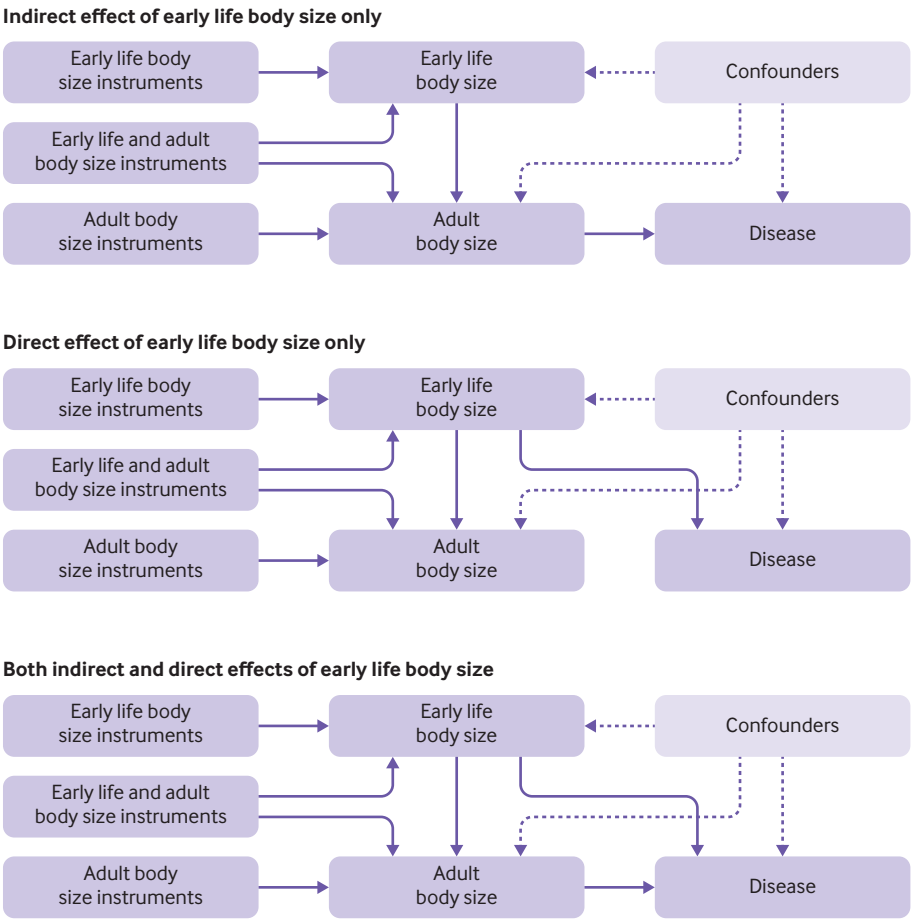

Fig 1 | Directed acyclic graphs depicting three possible scenarios that could explain a causal effect between body size at age 10 years and disease outcomes in adulthood. (Top) Early life body size has an indirect effect on disease risk only through body size in adulthood, (middle) early life body size has a direct effect on disease risk independent of body size in adulthood, and (bottom) early life body size exerts both direct and indirect effects on disease risk in adulthood

Statistical analysis

Identifying instruments

We used the software BOLT-LMM to assess the association between genetic variants across the human genome and both measures of body size.[20 29] This approach applies a bayesian linear mixed model to evaluate the association between each genetic variant with each measure of body size in turn, while accounting for both relatedness and population stratification. We added age at baseline and type of genotyping array as covariates in the model. Analyses were undertaken three times, once in all eligible participants after additionally adjusting for sex, and then stratifying by sex and analysing each separately. This allowed us to identify genetic variants that could be used as instrumental variables for outcomes based on populations of women only (ie, breast cancer), populations of men only (ie, prostate cancer), and mixed populations (ie, coronary artery disease and type 2 diabetes). Analyses of early life body size were additionally adjusted for month of birth as we hypothesised that peoples' relative age within their school year might bias self-reporting. Effect estimates from this analysis were used in the subsequent mendelian randomisation analyses. We used linear regression—that is, assuming that the effect of a given single nucleotide polymorphism on moving from the lowest to the middle category of the body size variables is the same as its effect on moving from the middle to the highest.

We used bivariate GREML analysis with GCTA software to calculate the genetic correlation between the early life and adult body size results from a genome-wide association study.[30] A genetic relation matrix was derived from 10 000 randomly selected UK Biobank participants who were unrelated and of European descent. To identify independent variants from our genome-wide association study we undertook linkage disequilibrium clumping using the software PLINK.[31] This was based on a threshold of $r^2<0.001$ using genotype data from European individuals from phase 3 (version 5) enrolled in the 1000 genomes project as a reference panel.[32] For independent variants associated with either early life or adult body size in our genome-wide association study of all participants (based on $P<5\times10^{-08}$), we compared the effect estimates of these variants between time points. This was undertaken by stacking observations (ie, so that each individual

BMJ: first published as 10.1136/bmj.m1203 on 6 May 2020. Downloaded from http://www.bmj.com/ on 25 June 2020 at BVA. Protected by copyright.

had two body size measurements) and then regressing the categorical body size outcomes on each variant in turn adjusting for age, sex, and time (ie, early life or later life). This analysis was then repeated along with adjustment for the interaction between genetic variant and time period. To account for multiple testing, we applied a Bonferroni correction to the P values of the interaction terms (ie, P<0.05/number of individual variants assessed) as a heuristic to highlight genetic loci with the strongest evidence of an interaction with time. This analysis was undertaken to show that various genetic variants seemingly appear to have a stronger influence on body size at different time points in the life course.

Validation of genetic scores using external datasets

Data on clinically measured BMI from the ALSPAC cohort was used to validate the scores we generated in the UK Biobank study. BMI data were obtained at three time points; two based on the ALSPAC offspring (at mean ages 9.9 and 17.9 years) and one in the ALSPAC mothers (mean age 50.8 years). Measures of BMI were dichotomised to classify individuals higher than the 85th centiles as overweight. Receiver operator characteristic curves were generated using the R package plotROC to compare the predictive ability of the scores for both early life and adult body size at all three time points.

We also undertook linkage disequilibrium score regression[33] to evaluate the genetic correlation between the two measures of body size from our analysis with directly measured BMI in adulthood,[34] and childhood obesity[35] using large scale external populations based on data from the linkage disequilibrium hub.[36] This was undertaken to further validate our early life measure of body size by showing that it is more strongly correlated with measured obesity in childhood than with BMI in adulthood. This was particularly important to show given that the early life body size score in UK Biobank was derived using questionnaire recall data, which could lead to bias if not appropriately validated using external datasets.

Mendelian randomisation

Univariable mendelian randomisation analyses were conducted using the MR-Base platform[37] to investigate total effect estimates between genetically predicted early life and adult body size individually with each disease outcome in turn. This was based on the inverse variance weighted method, which estimates the causal effect of an exposure on an outcome by combining ratio estimates using each variant in a fixed effect meta-analysis model.[38] Effect estimates on early life body size were used for the analysis, with additional adjustment for month of birth. However, the "total" effect of adult body size should be interpreted with caution. A strong correlation is likely to exist between genetic determinants of early and adult body size, and thus the total effect of adult body size is likely to be driven by pleiotropy if there is also a direct effect of early body size on the outcome. For this reason we only discuss the total effect of early body size and the direct effects of early and adult body size. Supplementary figure 1 shows a direct acyclic graph illustrating direct, indirect, and total effects as investigated within a multivariable mendelian randomisation framework.

We undertook multivariable mendelian randomisation analyses to estimate the direct effect of early life body size on each outcome in turn. Variants from the univariable analysis were used again here after undertaking further linkage disequilibrium clumping to account for instrument correlation between the two sets. We also undertook a negative control analysis with early life and adult body size as exposures in a multivariable framework and age at menarche as an outcome using findings from a genome-wide association study published before the release of UK Biobank.[26] We hypothesised that early life body size should only influence age at menarche directly in this framework as an individual's adult body size cannot influence timing of puberty (supplementary figure 2). As such, direct and total effects derived from a univariable framework should be comparable for early life body size on this outcome, whereas evidence of an indirect effect (ie, through adult body size) should be weak.

Furthermore, we undertook a sensitivity analysis for the breast cancer outcome to discern whether timing of puberty might also play a role in disease risk. This involved repeating analyses on breast cancer with age at menarche as a third exposure (along with early life and adult body size). We also repeated the univariable and multivariable mendelian randomisation analyses using findings derived from women with oestrogen receptor positive and oestrogen receptor negative breast cancer. Moreover, we undertook various sensitivity analyses to investigate results from the univariable and multivariable analyses, such as assessing instrument strength by deriving F statistics,[39] examining heterogeneity using Cochran's Q,[40] and investigating horizontal pleiotropy using the MR-Egger method.[41 42] Finally, we undertook a simulation study to investigate how misclassification of early life body size could affect our results (supplementary note 2).

All analyses were undertaken using R (version 3.5.1). Plots were created using the packages ggplot2[43] and metaphor.[44]

Patient and public involvement

This research did not involve patients or the public as it uses data from the UK Biobank study that were previously obtained from a cohort of people who had already been recruited. As such, no patients or member of the public were involved in the design or implementation of this study or the research questions addressed.

Results

In a genome-wide association study of 453 169 UK Biobank participants, 295 and 557 independent associations were detected with early life and adult body size, respectively, based on conventional genome-wide corrections (ie, $P<5\times10^{-08}$; supplementary tables

BMJ: first published as 10.1136/bmj.m1203 on 6 May 2020. Downloaded from http://www.bmj.com/ on 25 June 2020 at BVA. Protected by copyright.

2 and 3). A genetic correlation coefficient of rG=0.61 was calculated between these two sets of results. Using individual level data from the ALSPAC cohort, a genetic score based on the early life body size variants was found to be a stronger predictor of childhood BMI (mean age 9.9 years) compared with a score based on the adult variants (fig 2 left panel). Investigating the prediction of both early life and adult body size scores during later adolescence (mean age 17.8 years) suggested that neither score was optimal at this time point (fig 2 middle panel). During the adulthood time point (mean age 50.8 years), the adult body size score was a better predictor than using the early life body size variants (fig 2 right panel). These findings show the ability of these genetic instruments to separate early life size from adult body size.

Linkage disequilibrium score regression analyses identified strong genetic correlation between our derived early life measure of body size and a previous genome-wide association study of childhood obesity (rg=0.85) (supplementary figure 3). In contrast, our adult body size measure was considerably more strongly genetically correlated with a previous genome-wide association study of BMI in adulthood (rg=0.96) than to obesity in childhood (rg=0.64). Evaluating the difference between the genetic variant associations at the two different time points suggested that 75 genetic variants had stronger effect estimates on body size in adulthood compared with body size in early life, whereas 23 genetic variants had a stronger effect in early life compared with adult body size estimates (at $P<7.19\times10^{-05}$ (ie, 0.05/694 tests), supplementary table 4). Demonstrating that the magnitude of effect for these genetic variants differs with respect to early life and adult body size further suggests that they can be separated as two exposures in a multivariable framework. Plotting the relation between early life body size and adult body size in UK Biobank did not indicate a non-linear relation between them in line with the assumptions of multivariable mendelian randomisation (supplementary figure 4). Linear trajectories between childhood and adulthood have also been presented previously in cohorts, such as the Fels Longitudinal Study.[45]

Figure 3 illustrates findings from both early life and adult body size genome-wide association studies using a bidirectional Manhattan plot (sometimes referred to as a Miami plot) where some examples of time point specific effects have been highlighted. Repeating our genome-wide association study stratified by sex identified 135 and 215 variants that survived genome-wide association study corrections (ie, $P<5\times10^{-08}$) in women only with early life and adult body size, respectively (n=246 511) (supplementary tables 5 and 6). In the analysis restricted to men only, 69 genetic variants were associated with early life body size and 159 genetic variants were associated with adult body size (n=206 658) (supplementary tables 7 and 8). Repeating early life body size analyses with additional adjustment for month of birth did not seem to materially influence overall findings (supplementary tables 9-11).

Univariable and multivariable mendelian randomisation analyses

Odds ratios from these analyses reflect the change in odds for each change in category for our derived early life and adult body size variables. Univariable analyses provided strong evidence that early life body size is associated with risk of coronary artery disease (univariable mendelian randomisation odds ratio 1.49, 95% confidence interval 1.33 to 1.68). However, the direct effect of early life body size (ie, not mediated through adult body size) was much smaller than the total effect (multivariable mendelian randomisation odds ratio 1.02, 95% confidence interval 0.86 to 1.22), whereas strong evidence of a direct effect was identified for adult body size (multivariable mendelian randomisation odds ratio 1.82, 1.59 to 2.09). Similar findings were identified for analyses of type 2 diabetes, as the magnitude of the direct causal effect for early life body size was much smaller than the total effect (multivariable mendelian randomisation odds ratio: early life body size 1.16, 0.74 to 1.82, adult body size 2.80, 1.89 to 4.15). Figure 4 shows effect estimates from both univariable and multivariable mendelian randomisation analyses (supplementary table 12 also provides a full list). The results from this analysis were robust to the various sensitivity analyses applied in this study (supplementary tables 13-15).

The univariable analysis evaluating the effect of early life body size on age at menarche as an outcome provided strong evidence that having a larger body size at age 10 might result in earlier puberty (β −0.93 standard deviation change in age at menarche for each change in body size category, 95% confidence interval −0.66 to −1.20). When adult body size was included in the multivariable framework, along with early life body size as an exposure, the direct effect estimate for early life body size was found to be essentially unchanged (table 1). In contrast, evidence that adult body size has an effect on timing of puberty was weak. These findings reflect the scenario depicted in supplementary figure 2, suggesting that early life body size can only affect timing of puberty directly and not indirectly through adult body size. This analysis serves as a proof of concept for our multivariable framework, given that adult body size cannot influence timing of puberty as it occurs at a later time point in the life course.

In the univariable analysis strong evidence was found that higher genetically predicted early life body size was associated with a reduction in risk of breast cancer (univariable mendelian randomisation odds ratio 0.63, 0.55 to 0.72). When both measures of body size were analysed in a multivariable framework, the direct effect of adult body size (multivariable mendelian randomisation odds ratio 1.08, 0.93 to 1.27) was smaller in magnitude than the direct protective effect of early life body size (0.59, 0.50 to 0.71). Similar findings were identified in the analysis of women with oestrogen receptor positive and oestrogen receptor negative breast cancer (supplementary table 16).

Evaluating the relation between genetically predicted age at menarche and risk of breast cancer

BMJ: first published as 10.1136/bmj.m1203 on 6 May 2020. Downloaded from http://www.bmj.com/ on 25 June 2020 at BVA. Protected by copyright.

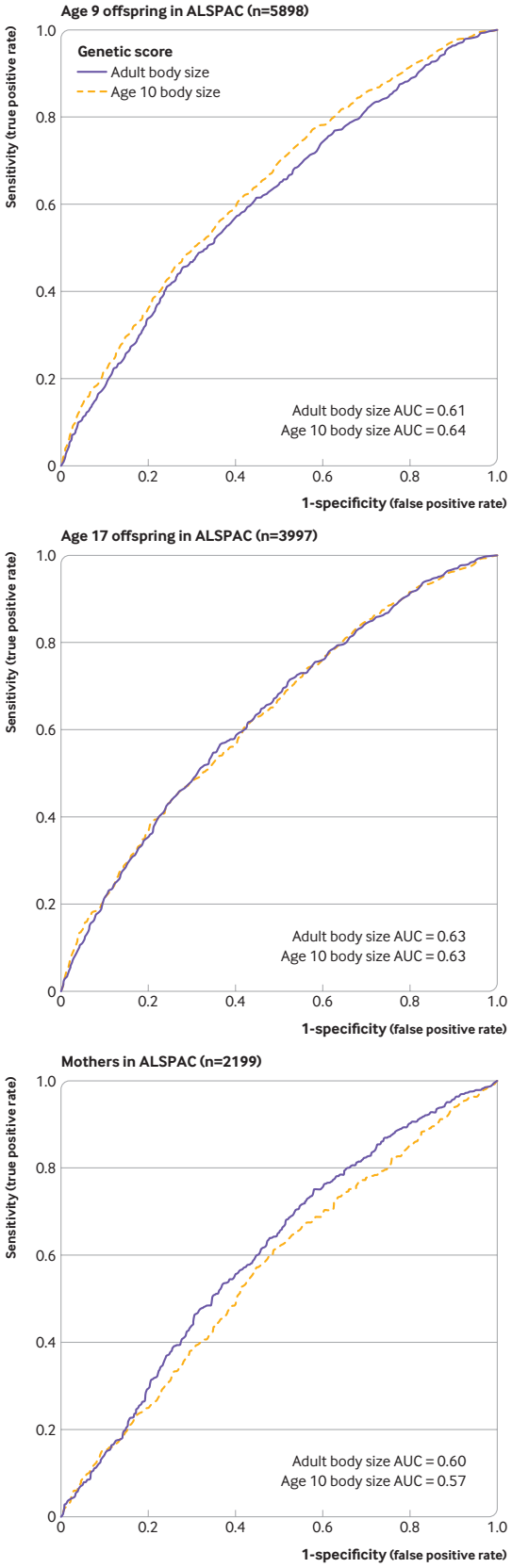

Fig 2 | Receiver operator characteristic curves to compare the predictive capability of early life and adult body size scores across three time points in Avon Longitudinal Study of Children and Parents (ALSPAC). (Top) Mean age 9.9 years in the ALSPSAC offspring; (middle) mean age 17.9 years in the ALSPAC offspring; and (bottom) mean age 50.8 years in the ALSPAC mothers. Body mass index in the ALSPAC cohort was dichotomised based on the 85th centile in all analyses. AUC=area under curve

in a univariable analysis provided weak evidence of a total effect (univariable mendelian randomisation odds ratio 0.98, 0.90 to 1.06). However, when modelled together with both early and adult body size (ie, all three exposures in the multivariable model) there was strong evidence that later age at menarche might lower the risk of breast cancer (multivariable mendelian randomisation odds ratio 0.90, 0.85 to 0.95). Effect estimates for predicted early life body size had a larger effect in this multivariable analysis compared with the direct effect of adult body size (table 1). Evidence that early life body size has a predicted causal effect on prostate cancer directly was weak (multivariable mendelian randomisation odds ratio 1.06, 0.81 to 1.40), which was also the case for adult body size (0.87, 0.70 to 1.08).

Early life body size was self-reported later in life and so is prone to misclassification, whereas adult body size is directly measured and is likely to be subject only to random error. A simulation was therefore conducted to identify how such misclassification, affecting only one exposure, could influence the effect estimation in the multivariable mendelian randomisation analyses. The simulation study (see supplementary note 2) shows that misclassification of early life body size is associated with a weakening in strength of association between the single nucleotide polymorphisms and our measure of early life body size and so potentially biases the estimated effect of early life body size in the multivariable mendelian randomisation analysis. These simulations do not cover all of the potential scenarios for measurement error. In the scenarios considered, however, misclassification of early life body size only introduces bias in the estimated effect of adult body size on the outcome when the misclassification in early life body size is dependent on actual adult body size and there is an effect of early life body size on the outcome. The simulations also show that the direction of this bias on the estimated effect of the adult variable depends on the direction of effect of the adult body size relative to the early life measure. Furthermore, the effect of adult body size is unbiased when early life body size has no associated effect on the outcome.

The simulation study also shows that the potential measurement error in self-reported body size at age 10 has the potential to bias adult effects in the multivariable mendelian randomisation when the measurement error depends on observed adult BMI and early life body size has a causal effect on the outcome. This therefore has the potential to affect the results we obtain for adult body size and risk of breast cancer. These simulations show that the bias from misclassification depends on both the type of misclassification and the size and direction of the effects of both exposures on the outcome. However, they suggest that this misclassification only masks the adult effect when it acts in the same direction as the early life effect on the outcome. We therefore do not believe that this measurement error is hiding a risk increasing effect of adult body size on breast cancer,

BMJ: first published as 10.1136/bmj.m1203 on 6 May 2020. Downloaded from http://www.bmj.com/ on 25 June 2020 at BVA. Protected by copyright.

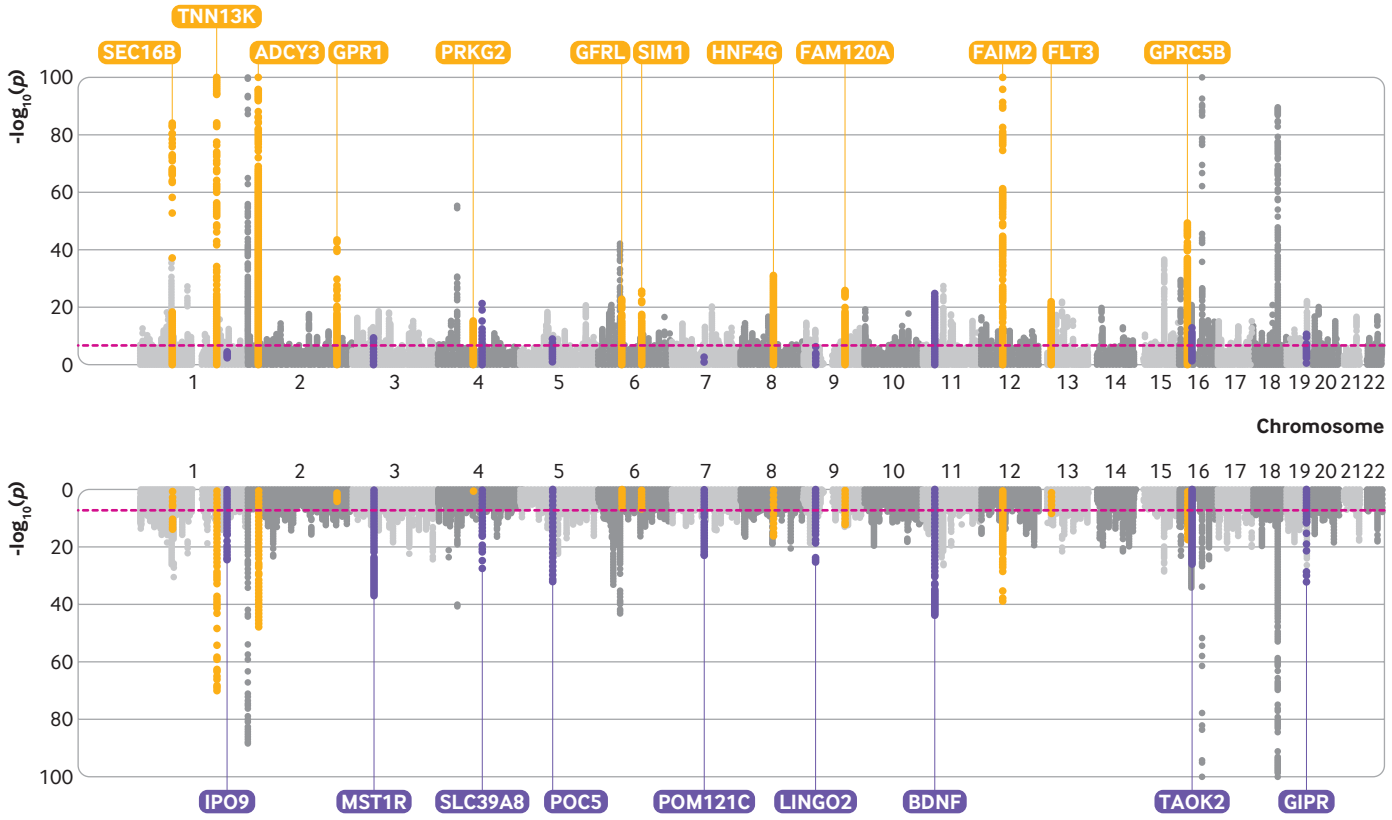

Fig 3 | Bidirectional Manhattan plot to illustrate genetic variants across the genome associated with (top) body size in early life (age 10 years) and body size in adulthood (bottom). Yellow points highlight effects that provided stronger evidence of association with early life body size, whereas purple points highlight effects more strongly associated with adult body size. These genetic loci have been annotated based on the nearest protein coding gene to the top associated variant

although it is possible that it is masking a larger protective effect.

Discussion

In this study we examined the influence of body size in early life (age 10 years) on risk of disease in later life and whether this putative causal effect occurs independently (direct effect) or through the same causal pathway (indirect effect) as later life body size. Our univariable mendelian randomisation analysis suggested that genetically predicted early life body size is associated with an increased risk of coronary artery disease and type 2 diabetes. When early life body size was analysed together with adult body size in a multivariable framework, however, the direct effect estimates for early life body size were considerably attenuated and fully compatible with, and close to, the null compared with the estimates of the total effects, suggesting that the influences of early life adiposity on these outcomes are mediated by body size in later life (see fig 1). These findings imply that observed associations between early life obesity and increased risk of coronary artery disease and type 2 diabetes are likely attributed to those with a large body size in childhood that persists into later life. This suggests that a window of opportunity exists to mitigate the detrimental impact of early life body size on risk of these disease outcomes.[46 47] These findings corroborate those from previous studies suggesting that there is no persistent influence of childhood obesity on risk of type 2 diabetes and cardiovascular disease unless adiposity is sustained.[48 49] Furthermore, these findings highlight the importance of taking into account adult body size to assess whether childhood body size has a direct or persistent effect on disease risk over the life course.[50 51]

Our results also provide strong evidence that early life body size has a protective effect on risk of breast cancer, as has been previously reported from both observational and mendelian randomisation studies that have not taken into account later life body size.[52-55] This was identified as both a total and a direct effect using univariable and multivariable mendelian randomisation analyses, respectively, suggesting that the effect of early life body size might persist into later life regardless of interventions that influence variation in body size. Furthermore, reported protective effect estimates from mendelian randomisation studies between later life body size and breast cancer risk could be attributed to effects from childhood. This is in contrast to observational estimates suggesting that higher BMI in adulthood might increase the risk of breast cancer, which could be attributed to confounding factors to which mendelian randomisation analyses are more robust.[55] This effect requires further investigation in subpopulations of

BMJ: first published as 10.1136/bmj.m1203 on 6 May 2020. Downloaded from http://www.bmj.com/ on 25 June 2020 at BVA. Protected by copyright.

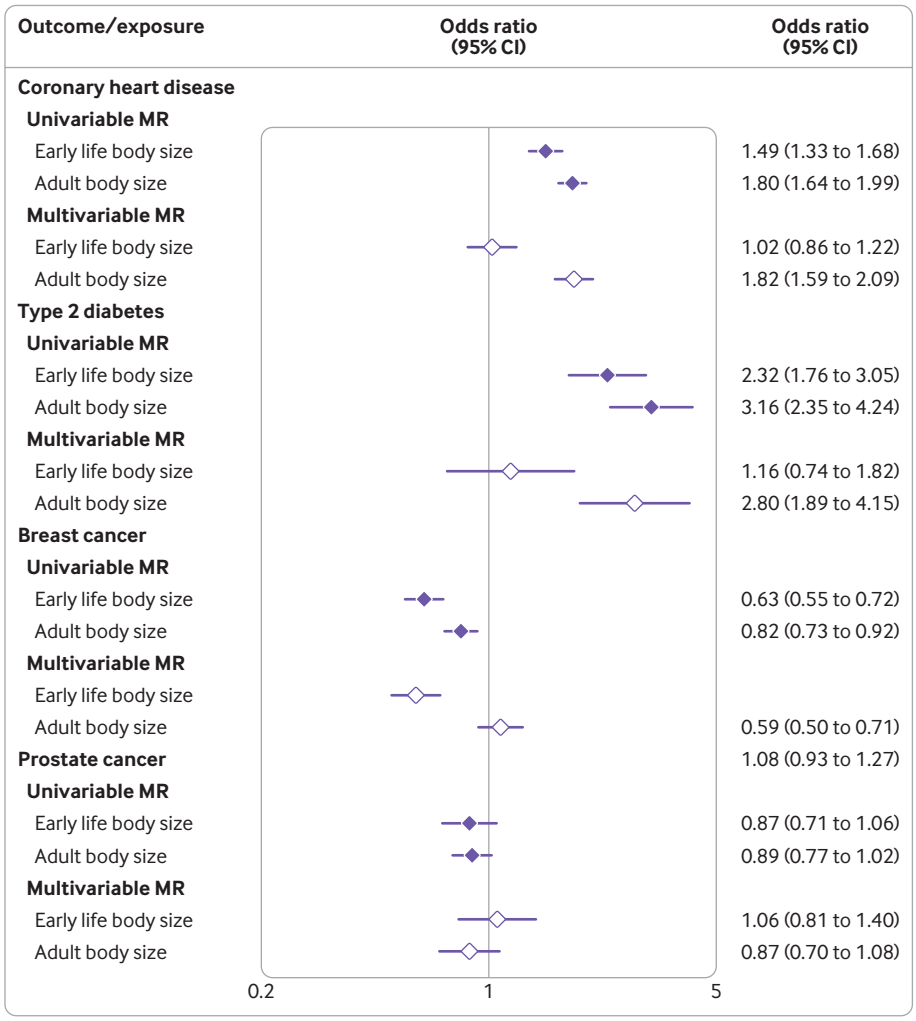

Fig 4 | Forest plot illustrating direct and indirect effects for genetically predicted body size in early life (age 10 years) on four different disease outcomes (coronary artery disease, type 2 diabetes, breast cancer, and prostate cancer). Filled and open diamonds represent effects from univariable and multivariable mendelian randomisation (MR) analyses, respectively. Estimates are displayed as odds ratios with 95% confidence intervals

premenopausal and postmenopausal women with breast cancer from large scale genome-wide association studies. This is particularly crucial given observational evidence suggesting that higher BMI increases the risk of invasive breast cancer in postmenopausal women.[56]

As an additional test of our multivariable framework, we undertook a negative control analysis to investigate the effect of early life body size on age at menarche in both a univariable and a multivariable framework accounting for adult body size. We found that estimates were consistent from both analyses for early life body size, whereas the effect estimates for adult body size in the multivariable analyses had a much smaller magnitude of effect in comparison. This provides a powerful proof of concept for our analytical framework, given that body size can only influence timing of puberty in early life whereas adult body size cannot as it occurs earlier in the life course. Moreover, our results support findings from the literature suggesting that higher BMI in childhood can lead to earlier timing of puberty.[57] Evidence is particularly strong in women, where an evolutionary mechanism rendering adequate fat storage to sustain both mother and growing fetus has been postulated.[58]

We undertook an additional analysis to investigate timing of puberty with respect to the putative protective effect of early life body size on risk of breast cancer. To do this we incorporated age at menarche as an additional exposure in our multivariable mendelian randomisation analysis. Our findings corroborate similar analyses undertaken using a multivariable framework, which suggest that later age at menarche has a protective effect on breast cancer risk but only when accounting for early life body size.[59] Evidence from the literature of a relation between timing of puberty and later life BMI is, however, strong.[60 61] Notably, our results build on previous findings by modelling early life and adult body size as two separate exposures, with the direct effect of early life body size providing strong evidence of a protective effect when accounting for age at menarche. Similar observational findings have been reported recently in premenopausal women after adjusting for age at menarche.[62] Developing insight into the underlying biological

BMJ: first published as 10.1136/bmj.m1203 on 6 May 2020. Downloaded from http://www.bmj.com/ on 25 June 2020 at BVA. Protected by copyright.

Table 1 | Estimates from univariable and multivariable mendelian randomisation analysis assessing effect of predicted early life and adult body size on age at menarche, and sensitivity analysis to investigate how both these exposures influence breast cancer risk when modelled with age at menarche as an additional third exposure

Outcomes and exposures	No of SNPs*	Univariable analysis		Multivariable analysis	
		β† (SE)	P value	β† (SE)	P value
Age at menarche[26]:					
Early life body size	102	−0.93 (0.14)	1.42×10^{-10}	−0.94 (0.15)	3.87×10^{-10}
Adult body size	142	-	-	−0.03 (0.14)	0.82
Breast cancer[24]:					
Early life body size	124	−0.46 (0.07)	3.84×10^{-11}	−0.64 (0.10)	1.57×10^{-10}
Adult body size	191	-	-	0.06 (0.09)	0.52
Age at menarche	60	−0.02 (0.04)	0.58	−0.10 (0.03)	4.84×10^{-03}

SNPs=single nucleotide polymorphisms.
*Number used as instrumental variables.
†Effect estimate coefficient for each standardised unit change in exposure.

mechanisms explaining this effect could highlight potentially modifiable pathways. In terms of proposed explanations, the earlier pubertal onset attributed to higher childhood adiposity has been postulated to result in slower pubertal growth.[63] This might therefore protect against rapid pubertal growth during adolescence, which has been linked to an increased risk of breast cancer.[64] Hormonal mechanisms might also play a role, such as higher oestrogen levels in early life produced by an increase in adipose tissue.[62] Oestrogenic effects have been reported to induce breast differentiation in early life, as well as to increase the expression of tumour suppressor genes.[65]

We were unable to support previous evidence of an effect between genetically predicted body size and prostate cancer risk.[66 67] This relation is complex owing to the effect of obesity on various hormones in men, such as an inverse relation with prostate specific antigen[68] and a positive relation with oestrogen concentrations.[69] This outcome is therefore worth revisiting using this study's analytical framework when findings from a forthcoming large scale genome-wide association study of prostate cancer become available. A greater number of genetic instruments in men only is also likely to improve power for future endeavours. Evaluating the influence of pubertal development on this outcome would also be a worthwhile undertaking.[70]

We used univariable mendelian randomisation to estimate total effects of early body size, and multivariable mendelian randomisation to estimate direct effects of early and adult body size. In theory, the univariable estimate of the effect of adult body size (the total effect) and the multivariable estimate (the direct effect) should be equal. However, if some genetic variants influence both early and adult body size (albeit with different effect sizes), and early life body size has a direct effect on the outcome, then this will potentially generate bias in the univariable estimate of adult body size. This can be seen in the univariable estimate of the effect of adult body size on breast cancer risk, which is quite different to the multivariable estimate, probably related to the direct effect of early body size on breast cancer risk. For coronary artery disease, where there is no direct effect of early body size on the outcome, the univariable and multivariable effects of adult body size are the same. This highlights one of the problems with mendelian randomisation of time varying exposures—the univariable analyses cannot identify critical or sensitive periods of exposure but only an effect of a difference in the cumulative lifetime exposure.[9 71-73] Moreover, we assume that childhood body size has an effect on adulthood body size in all scenarios (see fig 1). Therefore, those who, for example, reduce their body size between childhood and adulthood might effectively break this mediated causal chain and reduce their increased risk for diseases such as coronary artery disease or type 2 diabetes.

Strengths and limitations of this study

The key strengths of our investigation include the large number of participants from the UK Biobank study with measures of both early life and later life adiposity (n=453 169). Although this sample size is large, a caveat is reliance on retrospective questionnaire based data for the early life variable. Our early life variable is therefore based on perceived early life body size, which also meant that for harmonisation purposes we had to generate a similarly categorised adult variable. As such there will likely be additional measurement error in the early life variable, which could be differential with respect to adult body size (ie, larger adults might misremember their size at age 10 differently from those who are thinner). The statistical power in our genome-wide association study therefore comes at a cost, although s we were able to recapitulate evidence of association between genetic variants and early life body size measurements in the literature.[74 75] Furthermore, we were able to validate our scores using individual level data from the ALSPAC cohort, showing that the score in early life was a superior predictor of childhood adiposity, whereas the score in adulthood performed better in adults. Along with the large sample size in UK Biobank, other benefits of this approach included being able to harmonise our early life and adult measures within the same sample of participants. This was undertaken to reduce the likelihood of bias from differing samples—for example, between two separate genome-wide association study consortiums involving differing populations. Moreover, measurements of BMI in childhood taken from current generations might not reflect those of older populations that contributed to the outcome estimates in our analysis, whereas the UK Biobank participants and the outcome genome-wide association studies are from similar birth cohorts.

Another strength of this study was that we used two sample mendelian randomisation to harness large scale summary statistics from genome-wide association studies. This circumvents the necessity of having both exposures and outcomes measured in the same sample. Moreover, along with using genetic variants to mitigate the influence of confounding and reverse causation, a multivariable mendelian randomisation framework enabled us to investigate the independent effect of early life body size, which is extremely challenging in an observational setting.

BMJ: first published as 10.1136/bmj.m1203 on 6 May 2020. Downloaded from http://www.bmj.com/ on 25 June 2020 at BVA. Protected by copyright.

This framework therefore presents a powerful means to investigate many other questions in epidemiology. These could be related to the effect of the same exposure at different time points in the life course, as investigated in this study of body size, or examining whether different risk factors influence disease risk independently or along the same causal pathway. Particularly given the large scale summary level data, which current health data researchers are fortunate to have access to, this presents an attractive opportunity to analyse the life course structure of the causes of disease using mendelian randomisation.

Our univariable analysis used all the single nucleotide polymorphisms associated with each exposure. Although there are variations in the single nucleotide polymorphisms associated with each of early life and adult BMI, considerable overlap also exists in the single nucleotide polymorphisms associated with each time point. Our univariable mendelian randomisation analyses therefore capture pleiotropic effects of the single nucleotide polymorphisms through the other time point as well as, for childhood BMI, the effect mediated by adult body size. This is a limitation of our univariable mendelian randomisation analyses but highlights the importance of the multivariable mendelian randomisation to disentangle the direct effects of child and adult body size. It should also be emphasised that we have used genetically determined body size as exposure in this work, which might not directly equate to weight loss or gain from lifestyle modifications. Moreover, we acknowledge that survival bias can distort findings from mendelian randomisation analyses, as discussed in a recent study.[76] Although this is more likely to be a problem for outcomes that typically have a later onset of disease than those studied in this work (eg, Alzheimer's disease), replication of our results is necessary to rule out the possibility of survival bias.

Another limitation of our study is that participants from the UK Biobank could have also contributed to the large scale genome-wide association study, the results of which were used in this study. Overlapping exposure and outcome samples in a two sample mendelian randomisation analysis might result in overfitting, although currently there is no way to discern whether this is the case unless anonymous identifiers between the UK Biobank and genome-wide association studies consortium can be linked together.[77] Moreover, estimates used in this study are based solely on body size and BMI data from the UK Biobank. These findings should therefore be evaluated in future cohorts when sample sizes make this possible. This is particularly important as it has been shown that UK Biobank participants are highly selected, which can be problematic for instrumental variables analyses.[78 79] However, severe selection bias that can result in reversed estimates is unlikely to be a problem in our study given that the analyses are consistent with previous mendelian randomisation studies.[55 80 81] Lastly, capturing non-linear effects between body size at different points in the life courses and disease outcomes is challenging in a two sample setting.[82] Development of methods is therefore warranted.

Conclusions

Using multivariable mendelian randomisation, we have provided strong evidence suggesting that early life adiposity has a causal influence on risk of breast cancer that acts independently of later life BMI. Conversely, our results suggest that the association between early life adiposity and risk of coronary artery disease and type 2 diabetes is likely due to those with a high BMI remaining overweight in later life. Our approach therefore yields insight into the pathway between early life risk factors such as BMI and disease outcomes. Furthermore, our approach provides scope to differentiate between effects when the conferred risk is or is not reversible by achieving and maintaining a healthy BMI in adulthood.

We thank the participants of the UK Biobank study and the genome-wide association study consortiums (CARDIoGRAMplusC4D, DIAGRAM, BCAC, and PRACTICAL) who made their summary statistics publicly available for this study. GDS conducts research at the National Institute for Health Research (NIHR) Biomedical Research Centre at the University Hospitals Bristol NHS Foundation Trust and the University of Bristol. The views expressed in this publication are those of the authors and not necessarily those of the national health service, the NIHR, or the Department of Health. UK Biobank data were analysed under application 15825. We thank the families who took part in this study, the midwives for their help in recruiting the families, and the ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses. The UK Medical Research Council and Wellcome Trust (grant ref 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. Data from genome-wide association studies were generated by Sample Logistics and Genotyping Facilities at the Wellcome Trust Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe.

Contributors: GDS conceived the study. TGR, ES, KT, and GDS contributed to the study design. TGR conducted the main analysis and wrote the original draft of the manuscript. ES undertook simulation and sensitivity analyses. BE applied genome-wide association studies to derive genetic instruments. All authors were involved in the interpretation of results, helped refine the manuscript, and approved its final version. TGR is the guarantor and attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: This work was supported by the Integrative Epidemiology Unit, which receives funding from the UK Medical Research Council and the University of Bristol (MC_UU_00011/1, MC_UU_00011/2 and MC_UU_00011/3). TGR is a United Kingdom Research and Innovation research fellow (MR/S003886/1). The research was designed, conducted, analysed, and interpreted by the authors independently of the funding sources.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: funding by grants from the Medical Research Council and Health Data Research UK for the submitted work; no support from any other organisation for the submitted work; no financial relationships with any other organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: UK Biobank has received ethical approval from the UK national health service's National Research Ethics Service (ref 11/NW/0382).

Data sharing: All genetic instruments derived in this study are in the supplementary tables. Individual level data used to derive these results can be obtained with an approved application to the UK Biobank study. Validation data can be obtained with an approved application to the Avon Longitudinal Study of Parents and Children study. Outcome estimates were obtained using publicly accessible data made accessible by genome-wide association study consortiums. Code used to undertake mendelian randomisation analyses can be found as part of the "TwoSampleMR" R package.

BMJ: first published as 10.1136/bmj.m1203 on 6 May 2020. Downloaded from http://www.bmj.com/ on 25 June 2020 at BVA. Protected by copyright.

The manuscript's guarantor (TGR) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Dissemination to participants and related patient and public communities: We plan on disseminating the results of this study to any of the study participants or wider relevant communities on request.

This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/.

1 Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of childhood and adult obesity in the United States, 2011-2012. *JAMA* 2014;311:806-14. doi:10.1001/jama.2014.732.
2 Han JC, Lawlor DA, Kimm SY. Childhood obesity. *Lancet* 2010;375:1737-48. doi:10.1016/S0140-6736(10)60171-7.
3 Bibbins-Domingo K, Coxson P, Pletcher MJ, Lightwood J, Goldman L. Adolescent overweight and future adult coronary heart disease. *N Engl J Med* 2007;357:2371-9. doi:10.1056/NEJMsa073166.
4 Hannon TS, Rao G, Arslanian SA. Childhood obesity and type 2 diabetes mellitus. *Pediatrics* 2005;116:473-80. doi:10.1542/peds.2004-2536.
5 Baer HJ, Colditz GA, Rosner B, et al. Body fatness during childhood and adolescence and incidence of breast cancer in premenopausal women: a prospective cohort study. *Breast Cancer Res* 2005;7:R314-25. doi:10.1186/bcr998.
6 Swerdlow AJ, De Stavola BL, Floderus B, et al. Risk factors for breast cancer at young ages in twins: an international population-based study. *J Natl Cancer Inst* 2002;94:1238-46. doi:10.1093/jnci/94.16.1238.
7 Biro FM, Wien M. Childhood obesity and adult morbidities. *Am J Clin Nutr* 2010;91:1499S-505S. doi:10.3945/ajcn.2010.28701B.
8 Ogden CL, Carroll MD, Curtin LR, Lamb MM, Flegal KM. Prevalence of high body mass index in US children and adolescents, 2007-2008. *JAMA* 2010;303:242-9. doi:10.1001/jama.2009.2012.
9 Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?*Int J Epidemiol* 2003;32:1-22. doi:10.1093/ije/dyg070
10 Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* 2014;23(R1):R89-98. doi:10.1093/hmg/ddu328.
11 Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* 2018;362:k601. doi:10.1136/bmj.k601.
12 Davey Smith G, Lawlor DA, Harbord R, Timpson N, Day I, Ebrahim S. Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. *PLoS Med* 2007;4:e352. doi:10.1371/journal.pmed.0040352.
13 Burgess S, Thompson SG. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Am J Epidemiol* 2015;181:251-60. doi:10.1093/aje/kwu283.
14 Relton CL, Davey Smith G. Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. *Int J Epidemiol* 2012;41:161-76. doi:10.1093/ije/dyr233.
15 Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. *Int J Epidemiol* 2019;48:713-27. doi:10.1093/ije/dyy262.
16 Sutcliffe S, Colditz GA. Prostate cancer: is it time to expand the research focus to early-life exposures? [correction in: *Nat Rev Cancer* 2013;13:376.] *Nat Rev Cancer* 2013;13:208-518. doi:10.1038/nrc3434.
17 Couto Alves A, De Silva NMG, Karhunen V, et al. GWAS on longitudinal growth traits reveals different genetic factors influencing infant, child, and adult BMI. *Sci Adv* 2019;5:eaaw3095. doi:10.1126/sciadv.aaw3095
18 Sudlow C, Gallacher J, Allen N, 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* 2015;12:e1001779. doi:10.1371/journal.pmed.1001779.
19 Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;562:203-9. doi:10.1038/s41586-018-0579-z.
20 Elsworth BL, Mitchell RE, Raistrick CA, MRC IEU UK Biobank GWAS pipeline version 2. 2019. doi:10.5523/bris.pnoat8cxo0u52p6ynfaekeigi
21 Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. *Nat Protoc* 2010;5:1564-73. doi:10.1038/nprot.2010.116.
22 Nikpay M, Goel A, Won HH, et al. A comprehensive 1,000 genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* 2015;47:1121-30. doi:10.1038/ng.3396.
23 Morris AP, Voight BF, Teslovich TM, et al, Wellcome Trust Case Control Consortium, Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) Investigators, Genetic Investigation of ANthropometric Traits (GIANT) Consortium, Asian Genetic Epidemiology Network–Type 2 Diabetes (AGEN-T2D) Consortium, South Asian Type 2 Diabetes (SAT2D) Consortium, DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet* 2012;44:981-90. doi:10.1038/ng.2383.
24 Michailidou K, Lindström S, Dennis J, et al, NBCS Collaborators, ABCTB Investigators, ConFab/AOCS Investigators. Association analysis identifies 65 new breast cancer risk loci. *Nature* 2017;551:92-4. doi:10.1038/nature24284.
25 Schumacher FR, Al Olama AA, Berndt SI, et al, Profile Study, Australian Prostate Cancer BioResource (APCB), IMPACT Study, Canary PASS Investigators, Breast and Prostate Cancer Cohort Consortium (BPC3), PRACTICAL (Prostate Cancer Association Group to Investigate Cancer-Associated Alterations in the Genome) Consortium, Cancer of the Prostate in Sweden (CAPS), Prostate Cancer Genome-wide Association Study of Uncommon Susceptibility Loci (PEGASUS), Genetic Associations and Mechanisms in Oncology (GAME-ON)/Elucidating Loci Involved in Prostate Cancer Susceptibility (ELLIPSE) Consortium. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci [correction in: *Nat Genet* 2019;51:363]. *Nat Genet* 2018;50:928-36. doi:10.1038/s41588-018-0142-8.
26 Perry JR, Day F, Elks CE, et al, Australian Ovarian Cancer Study, GENICA Network, kConFab, LifeLines Cohort Study, InterAct Consortium, Early Growth Genetics (EGG) Consortium. Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. *Nature* 2014;514:92-7. doi:10.1038/nature13545.
27 Boyd A, Golding J, Macleod J, et al. Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* 2013;42:111-27. doi:10.1093/ije/dys064.
28 Fraser A, Macdonald-Wallis C, Tilling K, et al. Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol* 2013;42:97-110. doi:10.1093/ije/dys066.
29 Loh PR, Tucker G, Bulik-Sullivan BK, et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat Genet* 2015;47:284-90. doi:10.1038/ng.3190.
30 Lee SH, Yang J, Goddard ME, Visscher PM, Wray NR. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* 2012;28:2540-2. doi:10.1093/bioinformatics/bts474.
31 Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 2015;4:7. doi:10.1186/s13742-015-0047-8.
32 Abecasis GR, Auton A, Brooks LD, et al, 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012;491:56-65. doi:10.1038/nature11632.
33 Bulik-Sullivan BK, Loh PR, Finucane HK, et al, Schizophrenia Working Group of the Psychiatric Genomics Consortium. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015;47:291-5. doi:10.1038/ng.3211.
34 Speliotes EK, Willer CJ, Berndt SI, et al, MAGIC, Procardis Consortium. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* 2010;42:937-48. doi:10.1038/ng.686.
35 Bradfield JP, Taal HR, Timpson NJ, et al, Early Growth Genetics Consortium. A genome-wide association meta-analysis identifies new childhood obesity loci. *Nat Genet* 2012;44:526-31. doi:10.1038/ng.2247.
36 Zheng J, Erzurumluoglu AM, Elsworth BL, et al, Early Genetics and Lifecourse Epidemiology (EAGLE) Eczema Consortium. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* 2017;33:272-9. doi:10.1093/bioinformatics/btw613.
37 Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife* 2018;7:e34408. doi:10.7554/eLife.34408.
38 Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* 2013;37:658-65. doi:10.1002/gepi.21758.

BMJ: first published as 10.1136/bmj.m1203 on 6 May 2020. Downloaded from http://www.bmj.com/ on 25 June 2020 at BVA. Protected by copyright.

39 Burgess S, Thompson SG, Collaboration CCG, CRP CHD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol* 2011;40:755-64. doi:10.1093/ije/dyr036.
40 Bowden J, Hemani G, Davey Smith G. Invited commentary: detecting individual and global horizontal pleiotropy in mendelian randomization-a job for the humble heterogeneity statistic?*Am J Epidemiol* 2018;187:2681-5. doi:10.1093/aje/kwy185.
41 Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015;44:512-25. doi:10.1093/ije/dyv080.
42 Rees JMB, Wood AM, Burgess S. Extending the MR-Egger method for multivariable Mendelian randomization to correct for both measured and unmeasured pleiotropy. *Stat Med* 2017;36:4705-18. doi:10.1002/sim.7492.
43 Ginestet C. ggplot2: elegant graphics for data analysis. *J R Stat Soc Stat* 2011;174:245-45. doi:10.1111/j.1467-985X.2010.00676_9.x
44 Viechtbauer W. Conducting meta-analyses in R with the metafor Package. *J Stat Softw* 2010;36:1-48. doi:10.18637/jss.v036.i03.
45 Guo SS, Wu W, Chumlea WC, Roche AF. Predicting overweight and obesity in adulthood from body mass index values in childhood and adolescence. *Am J Clin Nutr* 2002;76:653-8. doi:10.1093/ajcn/76.3.653.
46 Goran MI, Ball GD, Cruz ML. Obesity and risk of type 2 diabetes and cardiovascular disease in children and adolescents. *J Clin Endocrinol Metab* 2003;88:1417-27. doi:10.1210/jc.2002-021442.
47 Inge TH, Miyano G, Bean J, et al. Reversal of type 2 diabetes mellitus and improvements in cardiovascular risk factors after surgical weight loss in adolescents. *Pediatrics* 2009;123:214-22. doi:10.1542/peds.2008-0522.
48 Franks PW, Hanson RL, Knowler WC, Sievers ML, Bennett PH, Looker HC. Childhood obesity, other cardiovascular risk factors, and premature death. *N Engl J Med* 2010;362:485-93. doi:10.1056/NEJMoa0904130.
49 Bjerregaard LG, Baker JL. Change in overweight from childhood to early adulthood and risk of type 2 diabetes. *N Engl J Med* 2018;378:2537-8. doi:10.1056/NEJMoa1713231.
50 Baker JL, Olsen LW, Sørensen TI. Childhood body-mass index and the risk of coronary heart disease in adulthood. *N Engl J Med* 2007;357:2329-37. doi:10.1056/NEJMoa072515.
51 Fang X, Zuo J, Zhou J, et al. Childhood obesity leads to adult type 2 diabetes and coronary artery diseases: A 2-sample mendelian randomization study. *Medicine (Baltimore)* 2019;98:e16825. doi:10.1097/MD.0000000000016825.
52 Tretli S. Height and weight in relation to breast cancer morbidity and mortality. A prospective study of 570,000 women in Norway. *Int J Cancer* 1989;44:23-30. doi:10.1002/ijc.2910440105
53 Vatten LJ, Kvinnsland S. Body mass index and risk of breast cancer. A prospective study of 23,826 Norwegian women. *Int J Cancer* 1990;45:440-4. doi:10.1002/ijc.2910450311
54 Ooi BNS, Loh H, Ho PJ, et al. The genetic interplay between body mass index, breast size and breast cancer risk: a Mendelian randomization analysis. *Int J Epidemiol* 2019;48:781-94. doi:10.1093/ije/dyz124.
55 Guo Y, Warren Andersen S, Shu XO, et al. Genetically predicted body mass index and breast cancer risk: mendelian randomization analyses of data from 145,000 women of European descent. *PLoS Med* 2016;13:e1002105. doi:10.1371/journal.pmed.1002105.
56 Neuhouser ML, Aragaki AK, Prentice RL, et al. Overweight, obesity, and postmenopausal invasive breast cancer risk: a secondary analysis of the Women's Health Initiative randomized clinical trials. *JAMA Oncol* 2015;1:611-21. doi:10.1001/jamaoncol.2015.1546.
57 Power C, Lake JK, Cole TJ. Body mass index and height from childhood to adulthood in the 1958 British born cohort. *Am J Clin Nutr* 1997;66:1094-101. doi:10.1093/ajcn/66.5.1094.
58 Kaplowitz PB. Link between body fat and the timing of puberty. *Pediatrics* 2008;121(Suppl 3):S208-17. doi:10.1542/peds.2007-1813F.
59 Burgess S, Thompson DJ, Rees JMB, Day FR, Perry JR, Ong KK. Dissecting causal pathways using mendelian randomization with summarized genetic data: application to age at menarche and risk of breast cancer. *Genetics* 2017;207:481-7. doi:10.1534/genetics.117.300191.
60 Kaplowitz PB, Slora EJ, Wasserman RC, Pedlow SE, Herman-Giddens ME. Earlier onset of puberty in girls: relation to increased body mass index and race. *Pediatrics* 2001;108:347-53. doi:10.1542/peds.108.2.347.
61 Lee JM, Kaciroti N, Appugliese D, Corwyn RF, Bradley RH, Lumeng JC. Body mass index and timing of pubertal initiation in boys. *Arch Pediatr Adolesc Med* 2010;164:139-44. doi:10.1001/archpediatrics.2009.258.
62 Schoemaker MJ, Nichols HB, Wright LB, et al, Premenopausal Breast Cancer Collaborative Group. Association of body mass index and age with subsequent breast cancer risk in premenopausal women. *JAMA Oncol* 2018;4:e181771. doi:10.1001/jamaoncol.2018.1771.
63 Martí-Henneberg C, Vizmanos B. The duration of puberty in girls is related to the timing of its onset. *J Pediatr* 1997;131:618-21. doi:10.1016/S0022-3476(97)70073-8.
64 Berkey CS, Frazier AL, Gardner JD, Colditz GA. Adolescence and breast carcinoma risk. *Cancer* 1999;85:2400-9. doi:10.1002/(SICI)1097-0142(19990601)85:11<2400::AID-CNCR15>3.0.CO;2-O.
65 Hilakivi-Clarke L, Forsén T, Eriksson JG, et al. Tallness and overweight during childhood have opposing effects on breast cancer risk. *Br J Cancer* 2001;85:1680-4. doi:10.1054/bjoc.2001.2109.
66 Rodriguez C, Freedland SJ, Deka A, et al. Body mass index, weight change, and risk of prostate cancer in the Cancer Prevention Study II Nutrition Cohort. *Cancer Epidemiol Biomarkers Prev* 2007;16:63-9. doi:10.1158/1055-9965.EPI-06-0754.
67 Benn M, Tybjærg-Hansen A, Davey Smith G, Nordestgaard BG. High body mass index and cancer risk-a Mendelian randomisation study. *Eur J Epidemiol* 2016;31:879-92. doi:10.1007/s10654-016-0147-5.
68 Harrison S, Tilling K, Turner EL, et al. Investigating the prostate specific antigen, body mass index and age relationship: is an age-BMI-adjusted PSA model clinically useful?*Cancer Causes Control* 2016;27:1465-74. doi:10.1007/s10552-016-0827-1.
69 Giovannucci E, Rimm EB, Liu Y, et al. Body mass index and risk of prostate cancer in U.S. health professionals. *J Natl Cancer Inst* 2003;95:1240-4. doi:10.1093/jnci/djg009.
70 Bonilla C, Lewis SJ, Martin RM, et al, PRACTICAL consortium. Pubertal development and prostate cancer risk: Mendelian randomization study in a population-based cohort. *BMC Med* 2016;14:66. doi:10.1186/s12916-016-0602-x.
71 Labrecque JA, Swanson SA. Interpretation and potential biases of mendelian randomization estimates with time-varying exposures. *Am J Epidemiol* 2019;188:231-8. doi:10.1093/aje/kwy204.
72 Holmes MV, Ala-Korpela M, Davey Smith G. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. *Nat Rev Cardiol* 2017;14:577-90. doi:10.1038/nrcardio.2017.78.
73 Labrecque JA, Swanson SA. Mendelian randomization with multiple exposures: the importance of thinking about time. *Int J Epidemiol* 2019;dyz234. doi:10.1093/ije/dyz234.
74 Warrington NM, Howe LD, Paternoster L, et al. A genome-wide association study of body mass index across early life and childhood. *Int J Epidemiol* 2015;44:700-12. doi:10.1093/ije/dyv077.
75 Felix JF, Bradfield JP, Monnereau C, et al, Bone Mineral Density in Childhood Study (BMDCS), Early Genetics and Lifecourse Epidemiology (EAGLE) consortium, Early Growth Genetics (EGG) Consortium, Bone Mineral Density in Childhood Study BMDCS. Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index. *Hum Mol Genet* 2016;25:389-403. doi:10.1093/hmg/ddv472.
76 Smit RAJ, Trompet S, Dekkers OM, Jukema JW, le Cessie S. Survival bias in mendelian randomization studies: a threat to causal inference. *Epidemiology* 2019;30:813-6. doi:10.1097/EDE.0000000000001072.
77 Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. *Genet Epidemiol* 2016;40:597-608. doi:10.1002/gepi.21998.
78 Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am J Epidemiol* 2017;186:1026-34. doi:10.1093/aje/kwx246.
79 Hughes RA, Davies NM, Davey Smith G, et al. Selection bias in instrumental variable analyses. bioRxiv [Preprint] 2018:192237. doi:10.1101/192237%J.
80 Nordestgaard BG, Palmer TM, Benn M, et al. The effect of elevated body mass index on ischemic heart disease risk: causal estimates from a Mendelian randomisation approach. *PLoS Med* 2012;9:e1001212. doi:10.1371/journal.pmed.1001212.
81 Censin JC, Peters SAE, Bovijn J, et al. Causal relationships between obesity and the leading causes of death in women and men. *PLoS Genet* 2019;15:e1008405. doi:10.1371/journal.pgen.1008405.
82 Staley JR, Burgess S. Semiparametric methods for estimation of a nonlinear exposure-outcome relationship using instrumental variables with application to Mendelian randomization. *Genet Epidemiol* 2017;41:341-52. doi:10.1002/gepi.22041.

Web appendix: Supplementary material
Web appendix: Supplementary tables

BMJ: first published as 10.1136/bmj.m1203 on 6 May 2020. Downloaded from http://www.bmj.com/ on 25 June 2020 at BVA. Protected by copyright.