# Assessing the causal impact of adiposity variation on rates of hospital admission: Application of Mendelian randomization

Audinga-Dea Hazewinkel[1,2]*, Rebecca C. Richmond[1,2], Kaitlin H. Wade[1,2], Padraig Dixon[1,2]
1: Population Health Sciences, Bristol Medical School, University of Bristol
2: MRC Integrative Epidemiology Unit, Bristol Medical School, University of Bristol

## Abstract

*Background*
Body mass index (BMI) and waist-hip-ratio (WHR) are measures of adiposity, the former being a good marker for overall total body fat, the latter describing regional adiposity. Higher adiposity has been associated with the increased prevalence of many chronic diseases and a positive association between BMI and increased hospital admissions has previously been established. The aim of this study was to estimate the causal relationship between BMI, WHR and WHR adjusted for BMI (WHRadjBMI) and yearly hospital admission rates.

*Methods and Findings*
Mendelian randomization (MR) approaches were used to test the causal effect of BMI, WHR and WHRadjBMI on yearly hospital admission rates. Using data on 310,471 participants of White-British ancestry from the UK Biobank, we performed one-sample and two-sample MR analyses on the exposures individually and in a multivariable setting. MR analyses supported a causal role of adiposity on hospital admissions, with consistency across one- and two-sample MR methods. Primarily, one-sample MR analyses estimated fold-increases in yearly hospital admission rates of 1.13 (95% CI: 1.02, 1.27), 1.26 (95% CI: 1.00, 1.58) and 1.22 (95% CI: 1.01, 1.47) per SD for BMI, WHR and WHRadjBMI, respectively. A multivariable approach yielded estimates of 1.04 (95% CI: 0.99, 1.03) for BMI and 1.31 (95% CI: 1.04, 1.67) for WHR, while adjusting for WHR and BMI, respectively.

*Conclusions*
The results support a causal role of higher BMI and WHR in increasing the yearly hospital admission rate. The attenuation of the BMI effect, when adjusting for WHR in the multivariable MR analyses, suggested that an adverse fat distribution, rather than a higher BMI itself, may drive the relationship between adiposity and increased risk of hospital admission.

**Keywords:** Body mass index (BMI), waist-hip-ratio (WHR), hospital admission, Mendelian randomization

---

*Corresponding author. Audinga-Dea Hazewinkel, Oakfield House, Oakfield Grove, University of Bristol, Bristol, England, BS8 2BN, a.d.hazewinkel@bristol.ac.uk

## Introduction

The body stores fat in the form of energy in adipose tissue. Adiposity can be measured by body mass index (BMI, in kg/m$^2$) and by waist hip ratio (WHR), among others, the former being a good marker for overall total body fat, the latter describing regional adiposity. Both measures have been shown to be associated with increased incidence of various diseases [1-8] and all-cause and cause-specific mortality [8-10] Over the last 40 years, global increases in the mean and variance of BMI have occurred [11-13], with the 2016 World Health Organization statistics identifying 39% of men and 40% of women as overweight (BMI>25kg/m$^2$) and 11% of men and 15% of women as obese (BMI>30kg/m$^2$) worldwide [14].

The increased incidence and prevalence of many adiposity-related chronic diseases has implications for the sustainability of healthcare systems. A positive observational association between BMI and rates of hospital admission [15] and related outcomes such as healthcare costs has been shown [16-22]. However, causality in these associations is unclear. Establishing the causal impact of BMI and related measures on hospital admissions and related outcomes is an important step in understanding the impacts of adverse weight profiles on the health system.

Traditional observational analyses may be subject to unobserved confounding and reverse causation, precluding accurate causal inference. Mendelian randomization (MR) is an instrumental variable (IV) approach that allows for the robust estimation of the causal effect of a given exposure (e.g., BMI) on a given outcome (e.g., hospital admission rates) by using quasi-random allocation of genetic variants at conception [23,24]. MR reduces problems of confounding and various types of biases frequently encountered in traditional epidemiological studies. For example, MR can avoid biases due to measurement error in BMI and WHR (e.g., when individuals knowingly or unknowingly misreport their own weight). MR can rule out reverse causation, since germline variants are established at conception, and, in principle, should not be affected by confounding from unmeasured or incorrectly modelled variables, given the random allocation of variants at conception.

This paper models the causal effect of adiposity on inpatient hospital admission rates in a MR approach, using data from over 300,000 adults aged between 39 and 72 and with over 550,000 in-patient hospital admissions in the UK Biobank. We considered three related exposures: BMI, WHR and WHR adjusted for BMI (WHRadjBMI). The effect of the exposures on yearly hospital admission rate was investigated using one-sample MR and two-sample summary MR methods, in both a univariable and multivariable setting. A variety of estimators was employed with the purpose of testing the robustness of our conclusions, including pleiotropy-robust methods and principled identification of outlying variants.

## Data

*UK Biobank data*

The UK Biobank study is a unique resource of phenotypic, genetic, electronic health record and death registry data, collected from over 500,000 individuals, collected from 2006 to 2010 [25-27]. Participants were aged 39-72 years at recruitment and were predominantly of White British ancestry [27]. The UK Biobank received ethical approval from the North West-Haydock Research Ethics Committee (reference 11/NW/0382).

For 465,373 participants, information on hospital inpatient admissions was available through linked Hospital Episode Statistics (HES) data. From this set, 34 participants were removed on grounds of incorrect admission information (hospital admissions prior to study start date, after study censoring date, after recorded death date or recorded death data prior to study start). A further number of participants were removed for factors related to the genetic data according to a documented in-house quality control procedure (Figure 1, Table S1) [28]. Specifically, 1,675 participants failed the standard inclusion criteria (mismatch between genetically inferred and reported gender, sex chromosome types putatively different from XX/XY, outliers in heterozygosity and missing rate), 73,278 had non-white British ancestry, 73,893 were ineligible on grounds of relatedness, 13,797 participants had no genetic data available for the relevant single nucleotide polymorphisms (SNPs) and 12,899 had no genetic principal component information.

For 2,454 individuals, no BMI and/or WHR measurements were reported. Participants matching one or more of the exclusion criteria were removed (N=154,902). An overview of all relevant exclusion criteria is provided in the participant exclusion diagram in Figure 1, with Table S1 providing a cross-table overview of participant overlap for any two exclusion criteria. A total of 310,471 participants were considered eligible for analysis, with the same set of participants used for all analyses for a relevant comparison between estimates from the observational analyses and the MR analyses.

Patient demographics and missingness per covariate are summarized in Table S2.

*Adiposity measures*

Measures on weight, bio-impedance, height, waist circumference and hip circumference were collected at the initial assessment at baseline. Weight and bio-impedance were measured using the Tanita BC-418MA body composition analyzer, which accurately measures body weight to within 0.1 kg. Standing height was measured using a Seca 202 height measure. BMI was calculated as weight divided by height squared (kg/m$^2$), with a second measure of BMI obtained using electrical impedance. When the first measure was unavailable, values were supplemented with the latter. WHR was calculated by dividing waist circumference by hip circumference, both of which were measured with a Wessex non-stretchable spurring tape measure [29].

*Hospital admission counts*
Study participants were considered from their time of assessment to ensure that hospital admissions could be appropriately related to baseline covariate measurements. Hospital admission count was derived using electronic health records linked to the UK Biobank study. In the UK healthcare system, hospital admissions are registered as episodes. For a given hospital admission, a patient may have multiple episodes. Admissions were defined where an individual had episodes starting on separate dates, excluding incomplete episodes, episodes with inconsistent or overlapping start and end dates and accounting for patient transfers. Code to define the admissions variable is available from https://github.com/pdixon-econ/admissions-biobank. Approximately half of the participants (47%) experienced no hospital admissions, while 81% of the remaining patients had four admissions or less during the time they were observed on study.

*Genetic variants*
Estimates for 77 genetic variants associated with BMI at a genome-wide significance level ($p<5x10^{-8}$) in the largest genome-wide association study meta-analysis of a combined number of up to 322,154 individuals of European descent (not including UK Biobank) were obtained from the Genetic Investigation of Anthropometric Traits (GIANT) consortium [30]. Individual-level genetic data of sufficient quality [28] was available from UK Biobank for 76 of these 77 SNPs, and were used to construct a genetic risk score (GRS), weighted with the exposure coefficients from the GIANT consortium. Genetic variants associated with WHR were also obtained from the GIANT consortium, with 39 SNPs identified in relation to unadjusted WHR at $p<5x10^{-8}$ in a meta-analysis of up to 210,088 individuals [31]. A second set of 48 SNPS was identified as being associated with WHRadjBMI at $p<5x10^{-8}$.[30]

# Methods

*Multivariable observational analyses*
Observational estimates were obtained by regressing each exposure directly on the outcome in a Poisson regression with time in UK Biobank as the offset. Unadjusted estimates were obtained alongside estimates adjusted for age at study entry, sex and an additional set of five categorical variables: alcohol frequency (ranging from never to daily or almost daily), days of exercise per week (ranging from 0 days to 7 days), educational/professional qualifications, employment, and material deprivation as measured by the Townsend deprivation index (divided into quintiles, with higher values indicating a greater degree of deprivation). In order to make full use of the available data, we imputed missing values in a 10-fold imputation approach and pooled the resulting coefficients and standard errors pooled using Rubin's rules [32].

*Mendelian randomization*

We conducted both one- and two-sample MR analyses. In a one-sample framework, individual-level genetic data is used to obtain estimates of both the gene-exposure association ($\beta_{exp}$) and the gene-outcome association ($\beta_{out}$), in a single sample using a linear regression for the former, and typically a linear or logistic regression for the latter for continuous and binary outcomes, respectively. Taking the ratio of the coefficients will give the causal IV estimate $\beta_{IV} = \frac{\beta_{out}}{\beta_{exp}}$. Equivalently, the fitted values obtained from the regression of the exposure on genetic variant can be regressed on the outcome. The former is referred to as the Wald ratio approach, the latter as the two-stage least squares (2SLS) approach. In a two-sample MR framework, Wald ratio-based methods are used to obtain the IV estimate with the exposure and outcome coefficients obtained from similar but separate populations [25].

In order to obtain a valid causal IV estimate in either framework, three main assumptions need to be satisfied: the genetic variant or variants must be correlated with the exposure, be independent of any unobserved confounders, and only affect the outcome through the exposure. Violation of this last assumption is known as horizontal pleiotropy – where a variant influences the outcome through a pathway other than via the exposure of interest. When multiple genetic variants are considered, heterogeneity amongst the Wald ratios may indicate that one or more SNPs are exhibiting horizontal pleiotropy [33,34].

*Poisson regression*

Both one and two-sample MR methods are traditionally applied to continuous or binary outcome data. While the use of MR in the context of different outcomes is less well-established, extensions to the traditional framework need to be considered for outcomes that cannot be accurately represented through linear or logistic models. Count data is frequently modeled using Poisson regression, which applies a generalized linear model with a logarithmic link function and can be used to obtain estimates of covariate effects on the absolute count or the count adjusted for a given measurement unit. Here, the outcome was given by hospital admission counts for subjects that were observed for varying lengths of time.

We employed a model that corrected for time at an individual level, with the yearly hospital admission rate as the outcome, and time for each individual defined as the observation time in years from study start to study censoring/death. The Poisson model assumes an outcome distribution such that the outcome mean is equal to the outcome variance. This stringent assumption may be relaxed by employing a quasi-Poisson model, which allows the variance to vary as a linear function of the mean. As the hospital admission count variance (55.3) is greater than the mean hospital admission count (1.89), we elected to use a quasi-Poisson model [35].

As the Poisson model is linear on the logarithmic scale, the second stage of the MR regression that estimates the gene-outcome association can be replaced by a Poisson regression. To obtain the causal IV estimate, the ratio of the $\beta_{out}$ and $\beta_{exp}$ coefficients is taken, with the $\beta_{out}$

coefficient on the log scale, and subsequently exponentiated to yield an estimate of the rate coefficient: $\beta_{IV} = \exp\left(\frac{\beta_{out}}{\beta_{\exp}}\right)$. Then, for a given exposure of interest, the coefficient $\beta_{IV}$ is the factor by which the yearly hospital admission rate increases for each unit of exposure, with values >1 indicating an increase in admission rate and values <1 indicating a decrease. Standard errors for the causal IV estimates were estimated using Taylor series expansions [36]. A more extensive description of the Poisson regression in general and in MR context is given in Appendix A.

*One-sample Mendelian randomization*

Weighted GRSs were constructed for the exposures of interest - BMI (76 SNPs) [30], WHR (39 SNPs) [31] and WHRadjBMI (48 SNPs) [31]. To ensure a meaningful interpretation of the score, the exposure-increasing allele for each genetic variant was chosen as the effect allele [9]. For each study participant, the dosage for each relevant genetic variant was extracted from the UK Biobank genetic data and weighted with the effect size reported by the GIANT consortium. Following this, the weighted dosages were summed and divided by the sum of all effect sizes, giving a GRS representing an estimate for the average number of exposure-increasing alleles. Estimates were adjusted for age, sex and the first 40 genetic principal components (PCAs).

To estimate the effect of WHRadjBMI on hospital admissions, we generated the residuals of a regression of WHR on BMI as an exposure, which gives an estimate for the predictive performance of the WHR component that cannot be linearly predicted by BMI.

We also considered the effects of BMI and WHR when estimated in a joint model, using a multivariable MR approach. Multivariable MR (MVMR) aims to estimate the causal effect of multiple exposures simultaneously. In contrast to univariable MR, which estimates the total effect of an exposure on the outcome, MVMR estimates the direct effect of each exposure conditioning on the causal effects of the SNPs on the other exposure [37]. Here, BMI and WHR were separately regressed on the full combined set of SNPs and the fitted values of both were regressed on hospital admission count in a Poisson regression with time as the offset. Standard errors were obtained through a 10,000-fold full-sample bootstrap [38].

For BMI, estimates of hospital admission rates per year were obtained per BMI unit (1 kg/m$^2$) and per BMI standard deviation (SD). For WHR, WHRadjBMI, and the WHR residuals, estimates were obtained per 0.1 WHR unit and per WHR SD. The relevant SDs were calculated directly from the UK Biobank data, yielding SDs of 4.74 kg/m$^2$ and 0.090 for BMI and WHR, respectively.

*Two-sample summary Mendelian randomization*

Gene-exposure association coefficients for the two-sample MR analyses were obtained from the GIANT consortium [30,31] and gene-outcome association coefficients from the UK Biobank data. A range of two-sample MR methods was applied for each exposure: 1) the regression-dilution bias robust random effects exact weights inverse-variance weighted (IVW) estimator

[39]; 2) the pleiotropy robust random effects MR-Egger estimator [40]; 3) the pleiotropy-robust penalized median estimator [41]; and 4) the pleiotropy-robust weighted mode estimator [42]. The consistency of the obtained estimates was evaluated across these four approaches. Each estimator is subject to different assumptions and methods of interpretation. Heterogeneity between the IV estimates of each genetic variant was investigated using the IVW-based Cochran's-Q and the MR-Egger based Rucker's Q [43]. Further detail on the methods and on assessing heterogeneity is provided in Appendix B.

Similar to one-sample MR analyses, two-sample MR estimates were obtained for BMI, WHR and WHRadjBMI individually and for BMI and WHR jointly in a multivariable two-sample MR analysis [37]. For BMI, estimates of hospital admission rates per year were obtained per BMI unit (1 $kg/m^2$) and SD. For WHR and WHRadjBMI estimates were obtained per 0.1 WHR unit and SD. SDs were calculated by taking the median SD across all studies used to obtain the summary measures, giving SDs of 4.6 $kg/m^2$ and 0.070 for BMI and WHR, respectively.

To satisfy the assumption of independent instruments, a threshold of $R^2<0.001$ reflecting linkage disequilibrium (LD) was employed for the two-sample MR analyses. Post LD-correction, 64, 34 and 45 SNPs were retained for BMI, WHR and WHRadjBMI, respectively. For the multivariable two-sample MR analysis, 70 SNPS were retained after a joint LD adjustment for BMI and WHR.

All analyses were performed in R version 3.6.1 [44]. R packages *TwoSampleMR* [45], *RadialMR* [43], and *MVMR* [37] were used for the two-sample summary MR analyses and *Amelia* [46] was used for the multiple imputation performed in the conventional multivariable analyses. An *R* code appendix is provided in the supplementary materials.

# Results

*Descriptives*

The 310,471 participants included in the analysis sample had an average age of 57.4 years (SD: 7.99) with a BMI of 27.4$kg/m^2$ (SD: 4.74) and WHR of 0.87 (SD: 0.09). Of these participants, 53.7% were female. Average follow-up time was 6.05 years (SD: 0.91). BMI, WHR, sex and age distributions were comparable across the UK Biobank and the GIANT consortium populations [27,30,31]. Demographics for the UK Biobank participants are given in Table S2. For the 310,471 participants, a total of 588,147 in-patient hospital admissions was recorded, with 79% of participants admitted twice or less and 47% without a single admission. Table S3 gives an overview of hospital admission counts across BMI categories, WHR quantiles, age quantiles and gender.

In order to facilitate comparisons of magnitude across the exposures, estimates of associations were obtained per exposure SD in addition to the standard per unit measures (with 0.1 units for WHR). Estimates per SD are reported in-text and in the summary figures (Figures 2-3), while estimates per SD and unit both are given in the supplementary tables (Tables S4, S5, S7 and S8).

*Observational multivariable analyses*

Poisson multivariable regression models provided evidence that BMI and WHR were associated with an increase in yearly hospital admission rate (Table S4, Figure 2). The adjusted regressions for the multivariable observational models yielded a 1.08-fold increase per BMI SD (95% CI: 1.07, 1.09) and a 1.16-fold increase per WHR SD (95% CI: 1.14, 1.18). Adjusting for BMI in the WHR regression resulted in a slightly attenuated association, with a 1.14-fold increase per WHR SD (95% CI: 1.12, 1.16).

Observational analyses regressing the outcome on both exposures simultaneously resulted in attenuated associations compared to the single exposure models, with a 1.03-fold increase per BMI SD (95% CI: 1.02, 1.04) and a 1.14-fold increase (95% CI: 1.12, 1.16) per WHR SD, respectively (Table S5, Figure 3). For all exposures, the estimates derived from adjusted models were lower than those derived from the unadjusted models, but with overlapping confidence intervals.

*Association between GRS and exposures*

Table S6 shows the associations between BMI and WHR and the three GRSs - the BMI GRS (comprising 76 SNPs), the WHR GRS (comprising 39 SNPs) and the WHRadjBMI GRS (comprising 48 SNPs). In UK Biobank participants of White British ancestry, each unit increase in BMI GRS was associated with a 0.11kg/m$^2$ higher BMI (95% CI: 0.11, 0.12), with the GRS explaining 1.69% of the variance. Each unit increase in WHR GRS was associated with a 0.01 higher WHR on the 0.1 unit scale (95% CI: 0.01, 0.01), explaining 0.41% of the variance, while each unit increase in WHRadjBMI GRS was associated with a 0.01 increase in WHR on the 0.1 unit scale (95% CI: 0.01, 0.02), explaining slightly more of the variance at 0.47%. $F$-statistics indicated that each GRS was a strong instrument for MR analyses (Table S6), with all $F$-statistics > 1267.

*One-sample MR analyses*

One-sample MR estimates of the effect of BMI, WHR and WHRadjBMI on hospital admission rates were obtained per exposure SD and unit (Table S4, Figure 2). The IV regressions, adjusted for age, sex and the first 40 genetic PCAs, yielded a 1.13-fold increase per BMI SD (95% CI: 1.02, 1.27) and a 1.26-fold increase per WHR SD (95% CI: 1.00, 1.58). Using the WHRadjBMI SNPs, a fold increase of 1.22 was observed WHR SD (95% CI: 1.01, 1.47). Adjusting for BMI in the WHR regression, by using the residuals from a linear regression of WHR on BMI as an exposure, resulted in a reduced effect of 1.16 (95% CI: 0.97, 1.39) per SD.

The multivariable one-sample MR analysis (Table S5, Figure 3), adjusted for age, sex and the first 40 genetic PCs, showed no strong evidence for an independent effect of BMI on hospital

admissions, with a fold increase of 1.04 (95% CI: 0.93, 1.15) per BMI SD, when controlling for WHR. Conversely, there was evidence for an independent effect of WHR on hospital admissions, with a fold increase of 1.31 (95% CI: 1.04, 1.67) per WHR SD, yielding a considerably higher estimate than observed for both the WHR and WHRadjBMI univariable analyses (Table S4); however, confidence intervals of these results all overlapped. Instrument strength, as assessed using the conditional Sanderson-Windmeijer *F*-statistic [37], was sufficient for the multivariable MR analysis, with *F*-statistics of 38.50 and 22.89 for BMI and WHR, respectively.

*Two-sample MR analyses*

The IVW estimator showed evidence for a causal effect of all three exposures on hospital admissions (Table S7, Figure 2), at magnitudes consistent with one-sample MR results. For BMI, we observed a fold increase of 1.10 (95% CI: 1.01, 1.19) per SD, for WHR a fold increase of 1.20 (95% CI: 1.05, 1.36) per SD and for WHRadjBMI a fold increase of 1.15 (95% CI: 1.03, 1.29) per SD. The penalized weighted median and weighted mode yielded near identical point estimates, paired with wider confidence intervals, now including the null. The widest confidence intervals were observed for the MR-Egger estimates, as is usually the case given the lower power of this estimator. While, for WHR, the MR-Egger point estimate was comparable to those of the previous estimators, it was markedly lower for BMI and WHRadjBMI. The MR-Egger intercept indicated no directional pleiotropy.

Limited heterogeneity as measured by Rücker's Q ($Q_R$) and Cochran's Q ($Q_C$) was observed. For all three exposures, the Q-statistics smaller than the number of SNPs used for estimation ($Q_{R,BMI}$=51.29, $Q_{C,BMI}$=52.06, $Q_{R,WHR}$=28.64, $Q_{C,WHR}$=28.92, $Q_{R,WHRadj}$=43.80, $Q_{C,WHRadj}$=44.01). Despite the lack of substantial heterogeneity, we performed a sensitivity analysis to investigate the potential presence of pleiotropy driven by outlying SNPs. Individual SNPs were identified in a visual inspection of leave-one-out plots (Figures S1-S3) and removed, with 3, 4 and 4 SNPs excluded for BMI, WHR, and WHRadjBMI, respectively. The two-sample MR analyses were repeated using the reduced set of SNPs, yielding comparable point estimates and confidence intervals overlapping with those from the original analyses (Table S8).

A multivariable two-sample analysis of the BMI and WHR simultaneously (Table S9, Figure 3) yielded no strong evidence for an association of BMI and the yearly hospital admission rate, with a fold increase of 0.99 per BMI SD (95% CI: 0.85, 1.14). Mirroring the pattern observed in the multivariable one-sample analysis, the multivariable two-sample analysis provided evidence for an association between WHR and hospital admission rate and had higher point estimates than the WHR and WHRadjBMI univariable two-sample analyses, with a fold increase of 1.30 per WHR SD (95% CI: 1.02, 1.65) (Table S9).

The gene-exposure associations and gene-outcome associations for the univariable two-sample MR analyses are given in Supplementary Tables S10, S11 and S12 for BMI, WHR, and WHRadjBMI, respectively. The same quantities are provided for the multivariable two-sample analysis in Supplementary Table S13.

*Comparison across methods*
Figure 2 summarizes the effect of each exposure on yearly hospital admission rate, as estimated from traditional multivariable analyses, and both one- and two-sample MR analyses. For all exposures, we observed a higher point estimate for the one-sample MR analyses when compared to the traditional multivariable analyses, more so for WHR and WHRadjBMI than for BMI. On the whole, the estimates of the effect of all three exposures on yearly hospital admission rate were consistent across estimators, with overlapping confidence intervals, and comparatively wider confidence intervals for WHR and WHRadjBMI. A similar pattern is observed in Figure 3, which provides a comparable summary for the multivariable estimators. Once more, the confidence intervals for both exposures overlapped, but with the BMI point estimates grouped more closely. In contrast to the univariable analyses summarized in Figure 2, the one- and two-sample BMI estimate confidence intervals now included the null, while for WHR the null was excluded for both the multivariable one-sample and two-sample analyses.

## Discussion

Our results, based on the first MR analysis of more than 550,000 hospital admissions measured in over 300,000 UK Biobank participants, supported the presence of causal effect of a higher BMI, WHR and WHRadjBMI on an increased hospital admission risk using one- and two-sample MR methods, in both a univariable and multivariable MR framework.

When evaluating the point estimates per exposure SD unit, WHR and WHRadjBMI had comparatively larger effect sizes than BMI for all estimators. In the multivariable one-sample and two-sample MR analyses, no strong evidence for an association was observed for BMI, while a strong positive association was measured between WHR and the yearly hospital admission rate. This suggests that the relationship between adiposity and hospital admissions may be driven by a detrimental distribution of fat and adipose tissue rather than by BMI itself.

In previous research, an observational study of approximately 1.3 million UK women found a yearly hospital admission rate increase of 1.12 (95% CI: 1.12, 1.13) for every 5kg/m$^2$ increase in BMI [15]. In just over 300,000 individuals, we observed estimates amounting to rate coefficients of $1.02^5 = 1.08$ (95% CI: 1.07, 1.10), 1.14 (95% CI: 1.02, 1.28) and 1.10 (95% CI: 1.01, 1.21) from the multivariable observational, one-sample MR and two-sample MR analyses, respectively, giving effect estimates of a magnitude comparable to those observed previously in UK women. Another multivariable study, examining the association between BMI and hospital admissions in 451,320 UK Biobank participants, found increases in yearly hospital admission rates, measured per 2kg/m$^2$ BMI, of 1.06 (95% CI: 1.05, 1.07) and 1.06 (95% CI: 1.05, 1.07) for male and female never-smokers, respectively [47]. We observed comparable estimates, with increases across both genders, per 2kg/m$^2$ BMI, of $1.02^2$=1.03 (95% CI: 1.03, 1.04), 1.05 (95% CI: 1.01, 1.11), and 1.04 (95% CI: 1.00, 1.08), for the multivariable observational, one-sample MR and two-sample MR analyses, respectively.

While the effect of BMI/WHR on hospital admission rate has not previously been studied in an MR framework, the effect of BMI on hospital cost was examined in a two-sample MR analysis, using data from UK Biobank, which largely overlaps with our own study population [20]. A positive causal effect of BMI on hospital cost was found, in line with our own observation of increased hospital admissions for higher BMI.

We show the causal role of higher BMI and WHR in increasing the risk of hospital admission, using a large high-quality dataset. While a positive association of BMI and hospital admissions has previously been shown [15,47], this is the first time hospital admissions have been modeled using MR methods – the MR framework allowing us to obtain estimates of causal effects less sensitive to confounding and bias problems present in traditional epidemiological studies [23,24]. Our results further emphasize the necessity of increasing efforts towards the development of policies to combat the global rise in adiposity.

In addition to BMI, we also investigated the causal role of WHR on hospital admissions. WHR has not been investigated previously in the context of hospital admissions or health care costs. Importantly, we show here, through multivariable one-sample and two-sample MR analyses, that the observed effect of BMI on hospital admissions may actually be driven by abdominal adiposity, for which WHR is a more direct measure than BMI, the latter quantifying both overall adiposity and lean muscle mass.

Additionally, we demonstrate in this study that a non-standard outcome like hospital admission counts can be successfully modeled using MR methods, both in a one-sample and two-sample framework, by replacing the linear or logistic component (linear for continuous outcomes, logistic for binary ones) modeling the gene-outcome association with a Poisson regression.

### Limitations

MR methods make it possible to avoid certain biases common to traditional epidemiological studies, but also face several limitations, both in terms of interpretation and in terms of potential alternative sources of bias sources. When interpreting the results, it should be noted that MR does not estimate an average treatment effect, but rather a local average treatment effect (LATE) instead, under the assumption that the effect of IV on treatment for all IVs is in the same direction for all subjects – the condition of monotonicity [48]. In context of this particular study, this means that we estimate the effect of WHR and BMI in those subjects whose WHR/BMI exposure values differ on varying the levels of the respective IVs, under the condition that the change occurs in the same direction for all participants.

As the IVs are comprised of genetic markers, which are 'assigned' at conception, the estimated LATE is a measure of the effects of a lifelong exposure to BMI-increasing alleles/WHR-increasing alleles. Additionally, we should note that, for all three exposures, a relatively modest

percentage of variance is explained by the genetic variants, reducing statistical power for detecting the effect of a change in BMI/WHR and consequently less precise estimates [23].

A key interpretive assumption of MR is that of gene-environment equivalence: that genetically influenced BMI and WHR will have the same effect on hospital admission risk as, for example, adiposity modified by diet and/or exercise. The included SNPs, however, may not meet the stable unit treatment assumption (SUTVA) [49] and as such, our estimates of the effect of BMI and WHR will not necessarily be representative of the increase or reduction in hospital admission rate when adiposity is altered through interventions.

We were limited to inpatient hospital admissions as the source of hospital data linked to UK Biobank at the time of writing. It is possible that other forms of hospital and primary care may substitute for inpatient care, in which case our estimates may overstate the effect of adiposity on overall healthcare admissions and care episodes. On the other hand, if inpatient care is complementary to other forms of care, then we may have understated these episodes. In practice, both effects may be present, and we do not have data to definitively parameterize the relative balance of each effect. Our results are best interpreted in relation to our outcome of inpatient admissions.

While UK Biobank is a unique and high-quality source, the underlying demographic structure of the data imposes various restrictions in terms of generalizability. The UK Biobank sample is healthier and wealthier than the population from which it is drawn and consequently is likely to not be representative of the wider UK. Indeed, we observed lower rates of mortality than in the general population [25], better health-related behaviour and a higher level of education [50]. As less healthy individuals are less likely to participate in the study, the observed sample may be subject to selection bias [51-53], which will impact the estimates obtained from both conventional multivariable analyses and MR analyses – where estimates obtained may even be an underestimation of the effect of adiposity on hospital admissions in the general population. It should also be noted that current analyses have been limited to individuals of White British ancestry, and as such the results will not necessarily generalize to other ancestral groups.

A potential further limitation is the possibility of cohort effects. UK Biobank participants are aged 39-72 years, giving rise to range of birth cohorts in our data sample. There is some evidence from other sources to suggest that SNPs exert a greater influence on BMI for those born in more recent decades, possibly because of an increasingly obesogenic environment [55,56].

Additional sources of bias that should be considered in the context of MR and the UK Biobank data are assortative mating [56-59], confounding due to population stratification bias [59-62] and the presence of dynastic effects [59,61,63], which have all be implicated with respect to BMI previously. While we attempted to minimize the impact of population stratification with the inclusion of 40 genetic principle components, latent data structure may still exist for some exposures such as BMI [60]. Within-family analysis may address some of these issues [61,64],

although statistical power for this form of modelling is currently limited given relatively modest available sample sizes of related individuals in UK Biobank [65].

## Conclusion

This study describes the first MR analysis to estimate the causal effect of BMI, WHR and WHRadjBMI on yearly hospital admission rates. Results supported the causal role of greater adiposity in increasing the risk of hospital admissions. Multivariable MR analyses suggested that the effect of BMI on hospital admission rates may be mediated by WHR and that an unfavourable fat distribution may drive the relationship between increased adiposity and higher hospital admission rates.

In applying robust causal inference methods that are able to account for biases common to conventional epidemiological analyses, we have obtained results that highlight the negative effect of adverse weight profiles on the health system, further emphasizing the necessity of developing policies to combat the global rise in adiposity.

## Declarations

**Funding statement:** AH, RCR, KHW and PD are members of the MRC Integrative Epidemiology Unit at the University of Bristol which is supported by the Medical Research Council and the University of Bristol (MC_UU_12013/1, MC_UU_12013/9). PD acknowledges support from a Medical Research Council Skills Development Fellowship (MR/P014259/1). RCR is a de Pass Vice Chancellor's Research Fellow at the University of Bristol. KHW was supported by the Elizabeth Blackwell Institute for Health Research, University of Bristol and the Wellcome Trust Institutional Strategic Support Fund (204813/Z/16/Z) and works within a group funded by the Wellcome Trust Investigator Award (202802/Z/16/Z).

**Conflict of interest statement**: The authors declare no conflicts of interest.

**Acknowledgments**

This research has been conducted using the UK Biobank Resource as part of application numbers 16391 and 29294.

# References


1. Dalton M, Cameron AJ, Zimmet PZ, Shaw JE, Jolley D, Dunstan DW, et al. Waist circumference, waist-hip ratio and body mass index and their correlation with cardiovascular disease risk factors in Australian adults. J Intern Med. 2003;254(6): 555-563.
2. Hu G, Tuomilehto J, Silventoinen K, Sarti C, Männistö S, Jousilahti P. Body mass index, waist circumference, and waist-hip ratio on the risk of total and type-specific stroke. Arch Intern Med. 2007;167(13): 1420-1427.
3. Folsom AR, Kushi LH, Anderson KE, Mink PJ, Olson JE, Hong CP, et al. Associations of general and abdominal obesity with multiple health outcomes in older women: the Iowa Women's Health Study. *Arch Intern Med*. 2000;160(14): 2117-2128.
4. Lyall DM, Celis-Morales C, Ward J, Iliodromiti S, Anderson JJ, Gill JMR, et al. Association of Body Mass Index With Cardiometabolic Disease in the UK Biobank: A Mendelian Randomization Study. JAMA Cardiol. 2017;2(8): 882-889.
5. Corbin LJ, Richmond RC, Wade KH, Burgess S, Bowden J, Davey Smith G, et al. Body mass index as a modifiable risk factor for type 2 diabetes: Refining and understanding causal estimates using Mendelian randomisation. Diabetes. 2016;65(10): 3002–3007.
6. Timpson NJ, Harbord R, Davey Smith G, Zacho J, Tybjærg-Hansen A, Nordestgaard BG. Does greater adiposity increase blood pressure and hypertension risk? Mendelian randomization using the FTO/MC4R genotype. Hypertension. 2009;54(1): 84–90
7. Dale CE, Fatemifar G, Palmer TM, White J, Prieto-Merino D, Zabaneh D, et al. Causal Associations of Adiposity and Body Fat Distribution With Coronary Heart Disease, Stroke Subtypes, and Type 2 Diabetes Mellitus: A Mendelian Randomization Analysis. Circulation. 2017; 135(24): 2373-2388.
8. Staiano AE, Reeder BA, Elliott S, Joffres MR, Pahwa P, Kirkland SA, et al. Body mass index versus waist circumference as predictors of mortality in Canadian adults. *Int J Obes. 2005; 36*(11): 1450–1454.
9. Wade KH, Carslake D, Sattar N, Davey Smith G, Timpson NJ. BMI and Mortality in UK Biobank: Revised Estimates Using Mendelian Randomization [published correction appears in Obesity (Silver Spring). 2019;27(2): 349]. *Obesity (Silver Spring)*. 2018;26(11): 1796-1806.
10. Srikanthan P, Seeman TE, Karlamangla AS. Waist-hip-ratio as a predictor of all-cause mortality in high-functioning older adults. *An Epidemiol.* 2009;*19*(10): 724–731.
11. Finucane MM, Stevens GA, Cowan MJ, Danaei G, Lin JK, Paciorek CJ, et al. National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9·1 million participants. The Lancet. 2011;377(9765): 557-567.
12. N. C. D. Risk Factor Collaboration. Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19·2 million participants. The Lancet. 2016;387(10026): 1377-1396.
13. Davey Smith G. A fatter, healthier but more unequal world. The Lancet. 2016;387(10026): 1349-1350.
14. World Health Organization. Obesity and overweight. *WHO. 2016. [Cited 2020 June 24]* https://www.who.int/mediacentre/factsheets/fs311/en/
15. Reeves GK, Balkwill A, Cairns BJ, Green J, Beral V, the Million Women Study C. Hospital admissions in relation to body mass index in UK women: a prospective cohort study. BMC Med. 2014;12(1):45. doi: 10.1186/1741-7015-12-45
16. Cawley J. An economy of scales: A selective review of obesity's economic causes, consequences, and solutions. J Health Econ. 2015;43: 244-268.
17. Finkelstein E, Yang H. Obesity and medical costs. In: Cawley J, editor. The Oxford HAndbook of the Social Science of Obesity. New York: Oxford University Press; 2011.
18. Withrow D, Alter DA. The economic burden of obesity worldwide: a systematic review of the direct costs of obesity. Obes Rev. 2011;12(2): 131-141.

19. Dixon P, Davey Smith G, Hollingworth W. The Association Between Adiposity and Inpatient Hospital Costs in the UK Biobank Cohort. *Appl Health Econ Health Policy*. 2019;17(3):359-370. doi:10.1007/s40258-018-0450-2
20. Dixon P, Hollingworth W, Harrison S, Davies NM, Davey Smith G. Mendelian Randomization analysis of the causal effect of adiposity on hospital costs. *J Health Econ*. 2020;70: 102300. doi: 10.1016/j.jhealeco.2020.102300
21. Kent S, Fusco F, Gray A, Jebb SA, Cairns BJ, Mihaylova B. Body mass index and healthcare costs: a systematic literature review of individual participant data studies. Obes Rev. 2017;18(8): 869-879.
22. Kent S, Green J, Reeves G, Beral V, Gray A, Jebb SA, et al. Hospital costs in relation to body-mass index in 1·1 million women in England: a prospective cohort study. Lancet Public Health. 2017;2(5):e214-e222.
23. Haycock PC, Burgess S, Wade KH, Bowden J, Relton C, Davey Smith G. Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. *Am J Clin Nutr*. 2016;103(4): 965-978.
24. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiology studies. *Hum Mol Genet* 2014;23(R1): R89-R98. doi: 10.1093/hmg/ddu328
25. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018;562(7726): 203-209.
26. Collins R. What makes UK Biobank special?. *Lancet*. 2012;379(9822): 1173-1174.
27. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, 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(3): e1001779. doi: 10.1371/journal.pmed.1001779
28. In-house proving document. https://data.bris.ac.uk/data/dataset/41a12b3c-7717-4c91-a838-63557ae78992
29. UK Biobank. Protocol for a large-scale prospective epidemiological resource. 2007. http://www.ukbiobank.ac.uk/key-documents/
30. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015;518(7538): 197-206.
31. Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira T, Locke AE, Mägi R, *et al*. New genetic loci link adipose and insulin biology to body fat distribution. *Nature*. 2015;518(7538):187-196.
32. Rubin, D. Multiple imputation after 18+ years. J Am Stat Assoc. 1996;91(434): 473-489.
33. 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
34. Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum Mol Genet*. 2018;27(R2): R195-R208. doi: 10.1093/hmg/ddy163
35. Ver Hoef JM, Boveng PL. Quasi-Poisson vs. negative binomial regression: how should we model over dispersed count data? Ecology. 2007;88(11): 2766–2772.
36. Thomas DC, Lawlor DA, Thompson JR. Re: Estimation of bias in nongenetic observational studies using "Mendelian triangulation" by Bautista et al. *Ann Epidemiol*. 2007;17(7): 511-513.
37. 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(3): 713-727.
38. Efron B, Tibshirani R. *An Introduction to the Bootstrap*. Chapman & Hall, New York; 1993.
39. Bowden J, Del Greco FM, Minelli C, Zhao Q, Lawlor D, Sheehan NA *et al.* Improving the accuracy of two-sample summary-data Mendelian randomization: moving beyond the NOME assumption, *Int J Epidemiol. 2019;* 48(3): 728–742
40. 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(2): 512–525.
41. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol. 2016;40(4): 304-314.

42. Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum Mol Genet*. 2018;27(R2): R195-R208. doi: 10.1093/hmg/ddy163
43. Bowden J, Spiller W, Del Greco MF, Sheehan N, Thompson J, Minelli C, et al. Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the Radial plot and Radial regression. Int J Epidemiol. 2018;47(4): 1264-1278.
44. R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/
45. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. *eLife*. 2018;7:e34408. doi: 10.7554/eLife.34408
46. Honaker J, King G, Blackwell M. Amelia II: A Program for Missing Data. Unpubl. users guide. Cambridge, MA: Harvard Univ; 2007. http://gking.harvard.edu/amelia/
47. O'Halloran R, Mihaylova B, Cairns BJ, Kent S. BMI and Cause-Specific Hospital Admissions and Costs: The UK Biobank Cohort Study. Obesity. 2020;28: 1332-1341.
48. von Hinke S, Davey Smith G, Lawlor DA, Propper C, Windmeijer F. Genetic markers as instrumental variables. J Health Econ. 2016;45: 131-148.
49. Burgess S, Thompson S. Mendelian Randomization: Methods for Using Genetic Variants in Causal Estimation. Boca Raton, Florida: CRC Press; 2015.
50. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants with the General Population. Am J Epidemiol. 2017;186: 1026-1034.
51. Glymour MM. Using causal diagrams to understand common problems in social epidemiology. In: Oakes JM, Kaufman JS, editors. Methods in Social Epidemiology. San Francisco: Jossey-Bass-John Wiley & Sons 2006. p. 393-428.
52. Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. Int J Epidemiol. 2017;47(1): 226-235.
53. Spirtes P, Glymour CN, Scheines R, Heckerman D, Meek C, Cooper G, et al. Causation, prediction, and search. 2nd edition, MIT press; 2000.
54. Rosenquist JN, Lehrer SF, O'Malley AJ, Zaslavsky AM, Smoller JW, Christakis NA. Cohort of birth modifies the association between FTO genotype and BMI. Proceedings of the National Academy of Sciences. 2015;112(2): 354.
55. Walter S, Mejía-Guevara I, Estrada K, Liu SY, Glymour MM. Association of a genetic risk score with body mass index across different birth cohorts. JAMA. 2016;316(1): 63-69.
56. Hartwig FP, Davies NM, Davey Smith G. Bias in Mendelian randomization due to assortative mating. Genet Epidemiol. 2018;42(7): 608-620.
57. Tenesa A, Rawlik K, Navarro P, Canela-Xandri O. Genetic determination of height-mediated mate choice [published correction appears in Genome Biol. 2016;17:26]. *Genome Biol*. 2016;16: 269. doi:10.1186/s13059-015-0833-8
58. Jacobson P, Torgerson JS, Sjöström L, Bouchard C. Spouse resemblance in body mass index: effects on adult obesity prevalence in the offspring generation. Am J Epidemiol. 2007;165(1): 101-108.
59. Morris TT, Davies NM, Hemani G, Davey Smith G. Why are education, socioeconomic position and intelligence genetically correlated? bioRxiv. 2019. doi: 10.1101/630426
60. Haworth S, Mitchell R, Corbin L, Wade KH, Dudding T, Budu-Aggrey A, et al. Apparent latent structure within the UK Biobank sample has implications for epidemiological analysis. Nature Communications. 2019;10(1): 333.
61. Brumpton B, Sanderson E, Hartwig FP, Harrison S, Vie GÅ, Cho Y, et al. Within-family studies for Mendelian randomization: avoiding dynastic, assortative mating, and population stratification biases. bioRxiv. 2019. doi: 10.1101/602516.
62. Koellinger PD, de Vlaming R. Mendelian randomization: the challenge of unobserved environmental confounds. Int J Epidemiol. 2019;48(3): 665-671.
63. Fletcher JM. The promise and pitfalls of combining genetic and economic research. Health Econ. 2011;20(8): 889-892.

64. Davies NM, Howe LJ, Brumpton B, Havdahl A, Evans DM, Smith GD. Within family Mendelian randomization studies, *Human Molecular Genetics*. 2019;28(R2): R170–R179. doi: 10.1093/hmg/ddz204
65. Howe LD, Kanayalal R, Harrison S, Beaumont RN, Davies AR, Frayling TM, *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. 2019. doi: 10.1093/ije/dyz240

# Figures

**Figure 1.** Participant inclusion diagram. Eight exclusion categories are shown with the corresponding numbers of participants per category. On removal of duplicates a total of 154902 unique participants are considered ineligible for analysis. a) admissions prior to study start date, post death/censoring date or registered death prior to study start; b) individuals that have a mismatch between genetically inferred and reported gender, individuals with sex chromosome types putatively different from XX or XY and individuals that are outliers in heterozygosity and missing rate; c) individuals related to more than 200 other participants; d) on exclusion a maximal set of unrelated individuals is retained; e) not genotyped for the exposures of interest (BMI, WHR, WHRadjBMI)

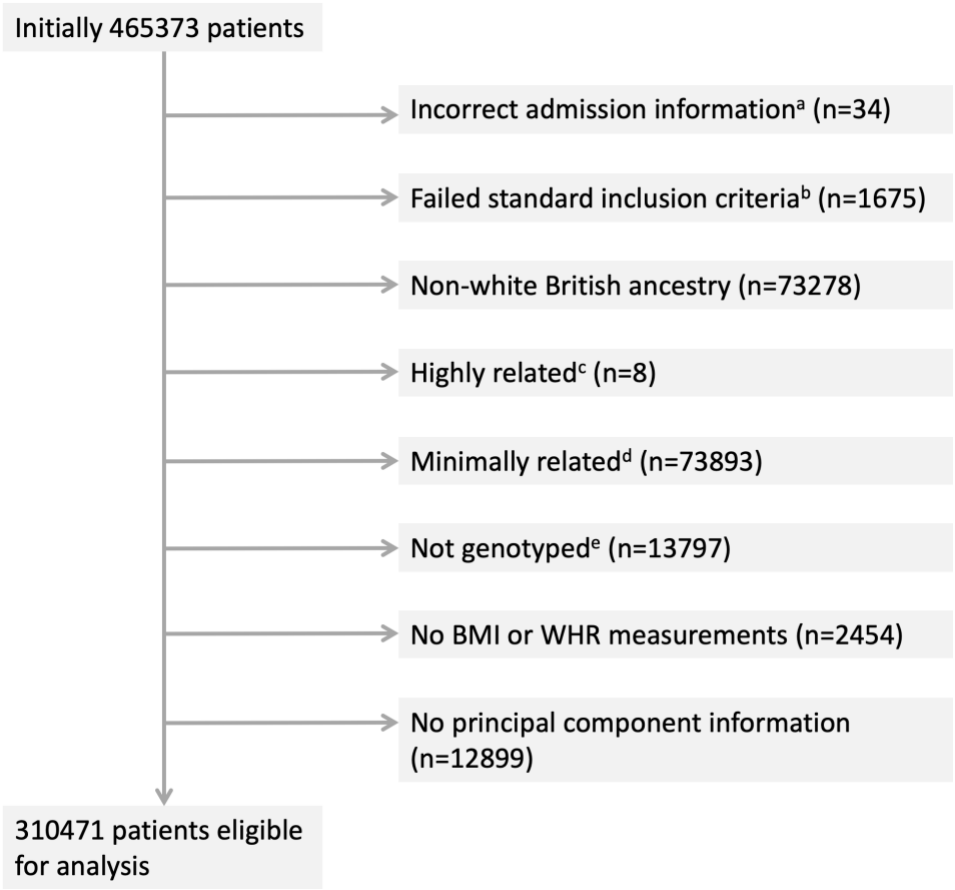

**Figure 2.** Estimates from multivariable observational analyses, one-sample MR analyses and two-sample MR IVW analyses for exposures BMI, WHR and WHRadjBMI per SD unit. Shown are point estimates alongside 95% CIs for the effect of exposure on yearly hospital admission rate, given on the log scale. Rate estimates and x-axis values are given on the exponential scale. All MR analyses were adjusted for age, sex and the first 40 genetic PCAs. The multivariable observational analyses were adjusted for a range of baseline patient characteristics (Table S2).

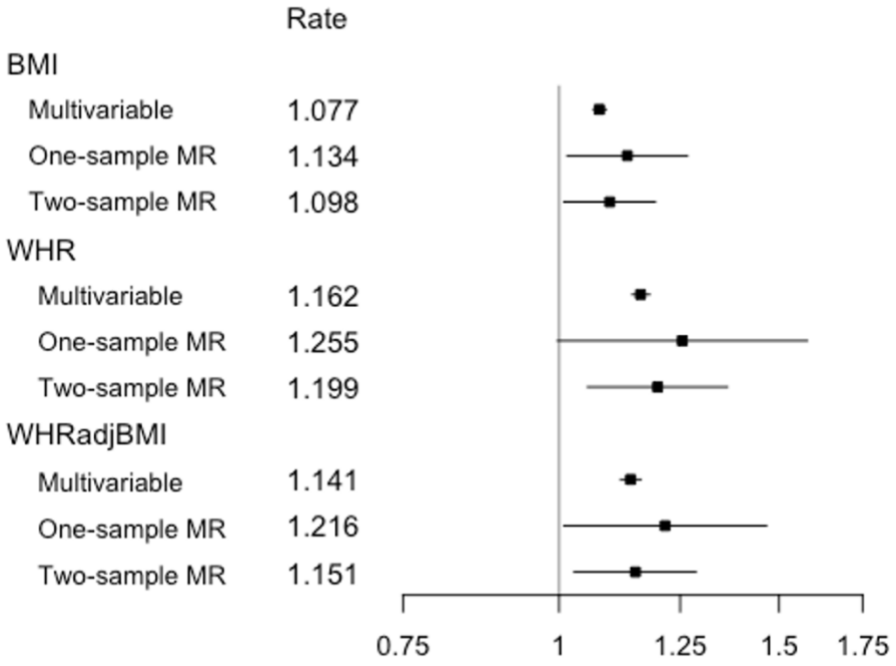
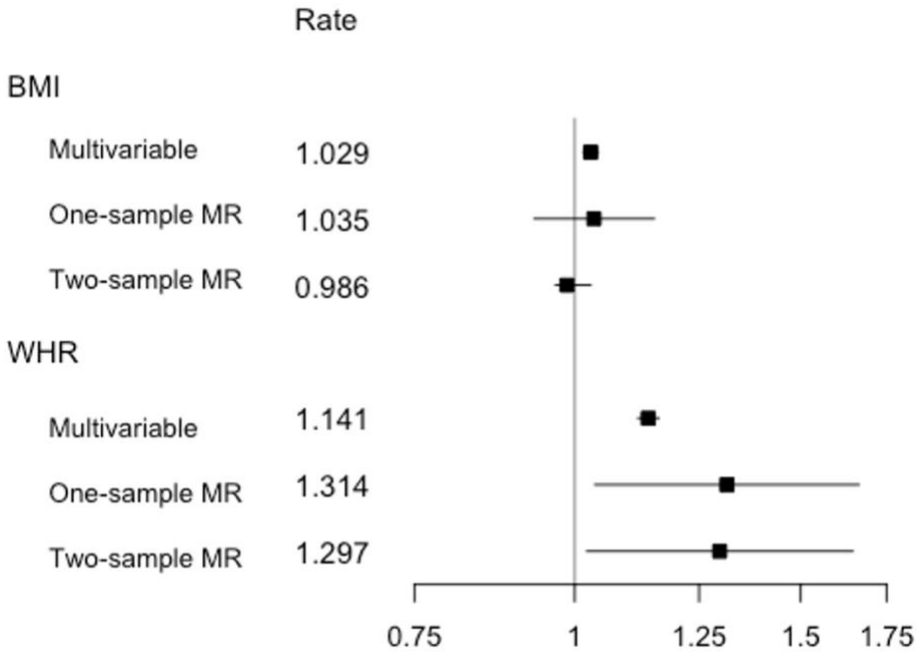
**Figure 3.** Estimates from multivariable observational analyses, multivariable one-sample MR analyses and multivariable two-sample MR analyses for exposures BMI and WHR per SD unit. Shown are point estimates alongside 95% CIs for the effect of exposure on yearly hospital admission rate, given on the log scale. Rate estimates and x-axis values are given on the exponential scale. MR analyses were adjusted for age, sex and the first 40 genetic PCAs. The multivariable observational analyses were adjusted for a range of baseline patient characteristics (Table S2), with the BMI estimate adjusted for WHR and vice versa.