



Wade, K., Carslake, D., Sattar, N., Davey Smith, G., & Timpson, N. (2018). BMI and Mortality in UK Biobank: Revised Estimates Using Mendelian Randomization. *Obesity*, 26(11), 1796-1806. <https://doi.org/10.1002/oby.22313>

Peer reviewed version

Link to published version (if available):  
[10.1002/oby.22313](https://doi.org/10.1002/oby.22313)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Wiley at <https://onlinelibrary.wiley.com/doi/10.1002/oby.22313> . 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/pure/user-guides/explore-bristol-research/ebr-terms/>

Body mass index and mortality in UK Biobank: revised estimates using Mendelian randomization

Kaitlin H Wade<sup>1,2</sup>, David Carslake<sup>1,2</sup>, Naveed Sattar<sup>3</sup>, George Davey Smith<sup>1,2</sup>, Nicholas J Timpson<sup>1,2</sup>

<sup>1</sup>MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN

<sup>2</sup>Population Health Sciences, Bristol Medical School, Faculty of Health Sciences, University of Bristol, BS8 2BN

<sup>3</sup>Institute of Cardiovascular and Medical Sciences, British Heart Foundation Glasgow Cardiovascular Research Centre, University of Glasgow, Glasgow, G12 8TA

**Corresponding author:**

Dr Kaitlin H Wade

Oakfield House

Oakfield Road

Clifton

Bristol

BS8 2BN

Telephone: +44(0)1173314092

Email: [kaitlin.wade@bristol.ac.uk](mailto:kaitlin.wade@bristol.ac.uk)

**Running title:** Mendelian randomization of BMI and mortality

**Word count:** 3999 (incl. introduction, methods, results, discussion and acknowledgments)

**Funding:** NJT is a Wellcome Trust Investigator (202802/Z/16/Z), a work-package lead in the Integrative Cancer Epidemiology Programme (ICEP) that is supported by a Cancer Research UK programme grant (C18281/A19169) and works within the University of Bristol NIHR Biomedical Research Centre (BRC). GDS is the director and a programme lead in the MRC-IEU (grant codes for 2013-2018: MC\_UU\_12013/1-6 and for 2018-2023: MC\_UU\_00011/1-7). At the start of this project, KHW was funded equally by two programs of the MRC-IEU (grant codes for 2013-2018: MC\_UU\_12013/3 and MC\_UU\_12013/4) and is now funded by the Wellcome Trust Investigator award (202802/Z/16/Z, PI: NJT). At the start of this project, DC was funded by the statistics theme of the MRC-IEU (grant code for 2013-2018: MC\_UU\_12013/9) and is now affiliated to program 1 of the MRC-IEU (grant code for 2013-2018: MC\_UU\_12013/1 and for 2018-

2023: MC\_UU\_00011/1). The funders of the study had no role in the study design, data collection, data analysis, data interpretation or writing of the report.

**Disclosures:** NS is a member of the UK Biobank international scientific advisory board and enhancements committee. This had no bearing on the study. Otherwise, the authors declared no conflicts of interest.

**Author Contributions:** All authors conceived and designed the study, interpreted results and approved the final manuscript. KHW had full access to all the data and performed all analyses with DC and takes responsibility for the integrity of the data and accuracy of the analysis. KHW, DC and NJT drafted the initial manuscript, the final version of which was critically appraised, revised and approved by all authors.

### **What is already known about this subject?**

- Whilst severe obesity clearly increases the risk of death, there have been inconsistencies within the current literature relating to the "obesity paradox", whereby being overweight can appear seemingly protective.
- Many studies also report a characteristic J-shaped curve in the association between body mass index and the risk of mortality from varying causes; however, there are discrepancies with the reporting of this pattern.
- Mendelian randomization is a well-documented approach that uses genetic variation to provide a relatively unbiased causal estimate of the effect of an exposure on an outcome, overcoming limitations inherent in these observational studies. Yet, no study has explicitly used this technique to explore the causal role of body mass index in all-cause and cause-specific mortality.

### **What does this study add?**

- We used Mendelian randomization to generate estimates of the causal role of body mass index in both all-cause and cause-specific mortality within the UK Biobank study, a powerful and large resource of comprehensive phenotypic, genetic and death registry data from the UK.
- Results supported the causal role of higher body mass index in increasing the risk of all-cause mortality and mortality specifically from cardiovascular diseases, various cancers and other causes. Whilst the characteristic J-shaped observational association between body mass index and mortality was visible with Mendelian randomization analyses, the apparent optimum body mass index for survival was lower and the association remained flatter over a larger range of body mass index.
- Our results further highlight the need for a global effort to reduce the rising population trends for excess weight and suggest that in most cases, any reduction in BMI is likely beneficial.

## ***Abstract***

**Objective:** Obtain estimates of the causal relationship between body mass index (BMI) and mortality.

**Methods:** Mendelian randomization (MR) using BMI-associated genotypic variation to test the causal effect of BMI on all-cause and cause-specific mortality in UK Biobank participants of White British ancestry.

**Results:** MR analyses supported a causal association between higher BMI and greater risk of all-cause mortality (hazard ratio (HR) per 1kg/m<sup>2</sup>: 1.02; 95% CI: 0.97,1.06) and mortality from cardiovascular diseases (HR: 1.12; 95% CI: 1.02,1.23), specifically coronary heart disease (CHD; HR: 1.19; 95% CI: 1.05,1.35) and those excluding CHD/stroke/aortic aneurysm, stomach cancer (HR: 1.30; 95% CI: 0.91,1.86) and oesophageal cancer (HR: 1.08; 95% CI: 0.84,1.38), and decreased risk of lung cancer mortality (HR: 0.97; 95% CI: 0.84,1.11). Sex-stratification supported the causal role of higher BMI increasing bladder cancer mortality risk (males), but decreasing respiratory disease mortality risk (males). The J-shaped observational association between BMI and mortality was visible with MR analyses but the BMI at which mortality was minimised was lower and the association was flatter over a larger BMI range.

**Conclusions:** Results support a causal role of higher BMI in increasing the risk of all-cause mortality and mortality from several specific causes.

**Key words:** Body mass index, BMI, mortality, genetic epidemiology

## **INTRODUCTION**

Whilst severe obesity (body mass index [BMI]  $\geq 35 \text{ kg/m}^2$ ) increases the risk of death, having a BMI  $> 25 \text{ kg/m}^2$  also increases the risk of all-cause mortality and mortality from vascular diseases, diabetes, respiratory diseases and cancer in a dose-response manner<sup>1-4</sup>. For example, each  $5 \text{ kg/m}^2$  higher BMI (a transition between BMI categories) increased the risk of mortality by  $> 30\%$ , vascular mortality by  $40\%$  and diabetic, renal and hepatic mortality by  $60-120\%$ <sup>1,5</sup>. Additionally,  $\sim 3.6\%$  of new adult cancer cases in 2012 (N  $\sim 481,000$ ; aged  $> 30$  after 10-years) were attributable to high BMI, a quarter of which could be attributed to rising BMI since 1982<sup>6</sup>.

However, there are inconsistencies within the literature relating to the “obesity paradox”, whereby being overweight can appear protective<sup>7,8</sup>. Most prominently, in a systematic review and meta-analysis ( $> 2.88$  million individuals), Flegal *et al.* showed  $\sim 6\%$  lower risk of all-cause mortality in overweight (i.e., BMI  $25.0-29.9 \text{ kg/m}^2$ ) compared to normal weight individuals (i.e., BMI  $18.5-24.9 \text{ kg/m}^2$ )<sup>7</sup>. Such controversial findings are not without limitation, as confounding by age, ill-health and lifestyle plus bias are likely<sup>9</sup>. Further, many studies report a characteristic J-shaped curve in the association between BMI and mortality<sup>1,2,5,8,10</sup>, where individuals at the lower tail of the BMI distribution (i.e., underweight [ $< 18.5 \text{ kg/m}^2$ ] or below  $22.5-24.9 \text{ kg/m}^2$ ) have an increased risk of mortality along with those above the ‘normal weight’ threshold<sup>1,2,5</sup>. However, there are discrepancies in the reporting of this pattern, specifically between condition-specific mortality and in populations of varying ancestries<sup>3,11-13</sup>.

Mendelian randomization (MR) is a well-documented application of instrumental variable (IV) methodology using genetic variants (most commonly, single nucleotide polymorphisms [SNPs]) as IVs to provide relatively unbiased causal estimates of the effect of an exposure (i.e., BMI) on an outcome (i.e., mortality)<sup>14,15</sup>. MR has provided evidence to support a causal effect of higher BMI increasing the risk of cardiovascular diseases (CVDs), diabetes, cardiometabolic traits and various cancers<sup>16-27</sup>. However, no study has explicitly used MR to explore the causal role of BMI in all-cause and cause-specific mortality. Here, data from the UK Biobank study, a powerful and large resource of comprehensive phenotypic, genetic and death registry data from the UK, were used to generate overall and sex-stratified estimates of the causal role of BMI in all-cause and cause-specific mortality. This approach was chosen to reduce problems of confounding and bias (e.g., reporting and recall bias) seen in traditional epidemiological studies .

## ***METHODS***

### *The UK Biobank study*

UK Biobank recruited over 500,000 people aged 37-73 years (99.5% were 40-69 years) from the UK in 2006-2010. The study, participants and quality control (QC) have been described previously<sup>28-30</sup>. UK Biobank received ethical approval from the Research Ethics Committee (reference: 11/NW/0382). Details of BMI, mortality, covariables and genotyping are presented in the Supporting Information. At the time of this study (and after exclusions based on QC parameters for phenotypic and genetic data – Supporting Information, Figure S1), 335,308 participants of White British ancestry had valid BMI, genetic and survival data and 9,750 of these had died (Figure 1, Table S1 and Supporting Information).

### ***Statistical analysis***

Cox proportional hazards regression models were used to estimate hazard ratios (HRs) for all-cause and cause-specific mortality per unit increase ( $\text{kg/m}^2$ ) in BMI. The participant's age was used as a measure of time; thus, models were adjusted for age. Analyses were conducted with two models: i) adjusted for secular trends (date of birth, DOB) and ii) additionally adjusted for current occupation, qualifications, smoking status, alcohol intake and physical activity. Analyses were restricted to the conditions responsible for a minimum number of deaths ( $>40$ )<sup>31</sup> and performed in whole and sex-stratified samples; therefore, results for all-cause mortality include all individuals who had died by the 16<sup>th</sup> of February 2016 (N=9,750) but individual mortality causes presented may not equate to this number (Table S1 and Supporting Information).

To generate the weighted genetic risk score (GRS) for MR analyses, the dosage of each genetic variant was weighted by its relative effect size on BMI reported by the Genetic Investigation of Anthropometric Traits (GIANT) consortium<sup>32</sup> and summed across all variants (Table S2). The resulting total was rescaled by dividing by the sum of all effect sizes on BMI reported by the GIANT consortium<sup>32</sup> and multiplied by the number of genetic variants used; therefore, providing a variable reflecting the number of average BMI-increasing alleles each participant possessed<sup>33</sup>. The associations of the weighted GRS with BMI and of each covariable with BMI and the GRS were tested using linear regression and associations of each covariable with all-cause mortality was assessed using Cox proportional hazards regression models. Associations with the GRS were adjusted for the first ten genetic principal components (PCs).

For MR analyses, the instrumental variable ratio method was conducted. Firstly, BMI was regressed on a GRS comprising 77 SNPs (the denominator of the ratio method estimator) adjusted for the first ten genetic PCs. Secondly, Cox proportional hazards models were used to estimate the log(HR) of each mortality outcome per unit increase in the GRS (the numerator of the ratio method estimator) adjusted for secular trends (DOB) and the first ten genetic PCs. Exponentiating the resulting ratio of the numerator and denominator yielded a MR estimate of the HR of each mortality outcome per unit increase (kg/m<sup>2</sup>) in BMI (Box 1). Confidence intervals (CIs) were obtained using Taylor series expansions<sup>34</sup>. A simplification of the matrix-method for the Durbin-Wu-Hausman (DWH) test for endogeneity was used to compare the HR estimated from conventional Cox regression and MR (Supporting Methods). A priori, conclusions were based on effect estimates and their CIs, rather than using an arbitrary p-value threshold<sup>35</sup>. For example, given two effects with the same HR - one with narrow CIs, the other with wider CIs that included the null - both are described as showing the same effect, but one is more imprecisely estimated and should be treated with caution until replicated. All analyses were conducted using Stata v15.

#### *Linearity and proportional hazards assumption*

Cubic spline models for both BMI (adjusted for variables in model (ii), described above) and the GRS (adjusted for secular trends (DOB) and the first ten genetic PCs) were plotted to test their pattern of association with mortality. Linearity tests were conducted after removing data below/above the 1<sup>st</sup>/99<sup>th</sup> percentile, respectively, due to the scarcity of data towards the tails of the BMI distribution. In addition, an approximate MR analogue to the non-linear plot of mortality against BMI was obtained by estimating localized average causal effects (i.e., MR estimates of the log-linear effect of BMI on mortality, adjusted for secular trends (DOB) and the first ten genetic PCs) within percentiles (the 5<sup>th</sup>, 10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup> and 85<sup>th</sup> percentile) of the instrument-free exposure (i.e., BMI that is orthogonal to the GRS)<sup>36</sup>. These localised average causal effects were joined and plotted against corresponding quantiles of the original exposure<sup>37</sup>. HRs were calculated relative to the mean BMI (27kg/m<sup>2</sup>) and CIs were obtained using bootstrapping (N=1000). Meta-regression was used to test for a linear trend in the GRS-BMI association (i.e., denominator of the ratio method) over quantiles of the instrument-free exposure.

To check the proportional hazards assumption, Schoenfeld residuals for BMI from the cubic spline models of each mortality outcome were tested for association with rank-normalised natural log of

the follow-up time (age) using both Cox regression and MR (132 tests in the whole sample and sex-stratified analyses for both methodologies) using Pearson's correlation. If there was evidence for an association (using a Bonferroni-corrected alpha-level of  $0.05/132=0.0004$ ), an interaction term was fitted to the cubic spline model using the "tvc()" option in Stata.

### *Sensitivity analysis*

Sensitivity analyses were conducted to (i) investigate the validity of the GRS as an IV using the MR-Egger<sup>38</sup>, weighted median and mode-based estimators<sup>39</sup>, compared to the inverse-variance weighted (IVW) method for two-sample MR<sup>38,40</sup>; (ii) evaluate the impact of covariables associated with the GRS; and (iii) explore the sensitivity of the GRS by excluding genetic variants implicated as pleiotropic (N=7; leaving 70 SNPs, Table S3)<sup>17,33</sup>. Details are presented in the Supporting Information.

## **RESULTS**

Included participants had an average age (at initial assessment) of 56.9 years (SD=8.0) and BMI of 27.4kg/m<sup>2</sup> (SD=4.7) (Table 1). Of the 335,308 participants with required information for mortality analyses, 9,570 participants (N=5,882/3,688 males/females, respectively) had died by the 16<sup>th</sup> of February 2016 at an average age of 65.7 years old (SD=6.9) from various CVDs and cancers (Table S1a/1b).

### *Observational analyses*

Cox regression models provided evidence that BMI was associated with a higher risk of all-cause mortality (HR per 1kg/m<sup>2</sup> higher BMI: 1.02; 95% CI: 1.02, 1.03) and mortality from CVD (HR: 1.07; 95% CI: 1.06, 1.08), specifically CHD (HR: 1.08; 95% CI: 1.06, 1.09) and those excluding CHD/stroke/aortic aneurysm (HR: 1.10; 95% CI: 1.08, 1.12), alongside mortality from overall cancer (HR: 1.01; 95% CI: 1.00, 1.02) and cancers of the stomach (HR: 1.05; 95% CI: 1.01, 1.09), oesophagus (HR: 1.03; 95% CI: 1.00, 1.06), kidney (HR: 1.07; 95% CI: 1.04, 1.11) and liver (HR: 1.05; 95% CI: 1.02, 1.09)(Table 2). There was evidence of an inverse association between BMI and lung cancer mortality (HR: 0.97; 95% CI: 0.95, 0.99). There was also weak evidence to suggest that higher BMI marginally increased mortality from stroke, aortic aneurysm and cancers of the colorectum, pancreas and brain whilst decreasing mortality from

respiratory diseases, bladder cancer, malignant melanoma and external causes (but estimates had wide CIs).

In males, results were similar to the whole sample but with additional evidence for an association between higher BMI and decreased respiratory disease mortality (HR: 0.91; 95% CI: 0.88, 0.95), which was weaker in the overall sample; an increased prostate cancer mortality (HR: 1.05; 95% CI: 1.02, 1.08), alongside greater magnitudes of association of higher BMI with a decreased risk of mortality from lung cancer (HR: 0.94; 95% CI: 0.91, 0.97) and bladder cancer (HR: 0.93; 95% CI: 0.86, 1.00) and increased risk of mortality from oesophageal cancer (HR: 1.07; 95% CI: 1.03, 1.11) and liver cancer (HR: 1.08; 95% CI: 1.03, 1.13) (Table 3). The estimate of association between BMI and brain cancer mortality was in the reverse direction to that obtained in the whole sample, but with wide CIs.

In females, results were similar to those in the whole sample but with additional evidence for an association between higher BMI and an increased respiratory disease mortality (HR: 1.06; 95% CI: 1.02, 1.09), the estimate of which was in the opposite direction in both the whole sample and males (Table 4). There was also evidence for an association between higher BMI and an increased risk of mortality from endometrial cancer (HR: 1.12; 95% CI: 1.07, 1.18) and both overall and post-menopausal breast cancer (HR: 1.02; 95% CI: 1.00, 1.04). There was no strong evidence of an association of BMI with lung cancer mortality and the estimate of association between higher BMI and oesophageal cancer mortality was in the opposite direction to that observed in the whole sample (HR: 0.87; 95% CI: 0.80, 0.95); however, all CIs overlapped.

#### *Association between the GRS and BMI*

Each unit increase in the GRS (comprising 77 SNPs) in the UK Biobank participants of White British ancestry was associated with 0.111kg/m<sup>2</sup> higher BMI (95% CI: 0.109, 0.114), explaining 1.8% of the variance and was slightly greater in females compared to males (Table 5).

#### *Covariable analysis*

Both BMI and mortality were associated with all covariables, including initial assessment age, sex, smoking status, alcohol consumption, qualifications, employment status and physical activity (Table S4 and Table S5 for BMI and all-cause mortality, respectively). Unlike the direct measurement of BMI, the GRS was associated with covariables to a much lesser extent, with all estimates near zero (Table S6).

### *MR analyses*

Within the whole UK Biobank sample, MR analyses provided estimates of a similar or greater magnitude to observational analyses (with wider CIs), supporting the causal role of higher BMI in increasing the risk of all-cause mortality (HR: 1.03; 95% CI: 0.99, 1.07) and mortality from CVD (HR: 1.10; 95% CI: 1.01, 1.19), specifically CHD (HR: 1.12; 95% CI: 1.00, 1.25) and those excluding CHD/stroke/aortic aneurysm (HR: 1.24; 95% CI: 1.03, 1.48), alongside mortality from stomach cancer (HR: 1.18; 95% CI: 0.87, 1.62) and oesophageal cancer (HR: 1.22; 95% CI: 0.98, 1.53) (Table 2). Although CIs were wide, the effect estimate for higher BMI on decreasing lung cancer mortality was consistent to that obtained in observational analyses (HR: 0.96; 95% CI: 0.85, 1.08). There was also evidence supporting the causal role of higher BMI in increasing mortality from external causes (HR: 1.30; 95% CI: 1.05, 1.61), unlike the inverse association obtained in observational analyses (DWH *P*-value for comparison=0.01). In contrast, the effect estimates for higher BMI on mortality from cancer, kidney cancer and liver cancer were attenuated or in the opposite direction, with CIs too wide for conclusive interpretation (Table 2).

Results for males were similar to those in the whole sample, as estimates of the causal role of higher BMI in increasing the risk of all-cause mortality and mortality from all CVDs, stomach cancer, oesophageal cancer and kidney cancer, alongside the decreased risk of mortality from lung cancer and bladder cancer, were consistent to or greater than the observational analyses (Table 3). The effect estimates for higher BMI on mortality from respiratory diseases, cancer, prostate cancer and liver cancer were attenuated or in the opposite direction, with CIs too wide for conclusive interpretation (Table 3).

In females, the effect estimates of higher BMI increasing the risk of all-cause mortality and mortality from all CVDs were consistent to the observational analyses (Table 4). The effect estimates for higher BMI on the risk of mortality from breast cancer (HR: 0.83; 95% CI: 0.70, 0.99), specifically post-menopausal breast cancer (HR: 0.84; 95% CI: 0.70, 1.00), endometrial cancer (HR: 0.63; 95% CI: 0.38, 1.07) and external causes (HR: 1.79; 95% CI: 1.23, 2.58) were in the opposite direction to those obtained in observational analyses (DWH *P*-values=0.02, 0.03, 0.04 and 0.002, respectively). Furthermore, the effect estimates for higher BMI on mortality from respiratory diseases, overall cancer, oesophageal cancer and kidney cancer were attenuated or in the opposite direction compared to observational analyses, but with CIs too wide for conclusive interpretation (Table 4).

Whilst there was some evidence for an observational relationship between higher BMI and mortality from other causes, CIs were too wide for conclusive interpretation in both adjusted observational and MR analyses, and with sex stratification (Table S7).

#### *Linearity and proportional hazards assumption*

The pattern of the GRS-mortality association appeared linear (Figure 2); however, the CIs were wide. The observational BMI-mortality relationship showed evidence of a J-shaped association (Figure 3A). The J-shaped BMI-mortality association remained in MR analyses (Figure 3B), but with a smaller value of BMI at which mortality risk was lowest (~23kg/m<sup>2</sup> vs. ~26kg/m<sup>2</sup> with observational analyses) and apparently flatter over a larger BMI range. Meta-regression provided some evidence that the GRS-BMI association was non-linear (*P*-value for linear trend=0.08 and *P*-value for heterogeneity <0.001). This was primarily driven by the extreme quantiles of BMI, as removal of these quantiles indicated a linear association (*P*-value for linear trend=0.999 and *P*-value for heterogeneity <0.001).

The proportional hazards assumption held for all mortality causes in both the conventional Cox regression and the MR analyses (Table S8a and S8b for observational and MR analyses, respectively).

#### *Sensitivity analyses*

Across all methods, which assume linearity (including the IVW, MR-Egger, weighted median- and mode-based estimators), MR-derived estimates were consistent (Table S9a, S9b and S9c for whole sample, males and females, respectively). The MR-Egger intercept estimate showed some evidence for pleiotropy in the association between BMI and mortality from other cancers in the whole sample (Figure S2a) and males (Figure S2b), suggesting an underestimated MR estimate with negative directional pleiotropy (which was likely driven by the rs17024393 SNP). There was no strong evidence of directional pleiotropy in female-specific analyses (Table S9c).

Additional adjustment for covariables made no substantive difference to the GRS-BMI association (Table S10a) and MR analyses (Table S10b). When excluding genetic variants implicated as pleiotropic (N=7; leaving 70 SNPs), there was no substantive difference in the GRS-BMI association (Table S11a) and MR analyses (Table S11b).

## **DISCUSSION**

Results supported the causal role of higher BMI in increasing the risk of all-cause mortality and mortality specifically from CVDs plus various cancers including oesophageal cancer and stomach cancer, as well as decreasing lung cancer mortality risk. Sex-stratified analyses were consistent with those in the whole sample and provided additional evidence for the causal role of higher BMI in increasing the risk of mortality from cancers of the kidney and liver in males and from external causes in females, whilst decreasing the risk of mortality from bladder cancer in males and breast cancer (specifically post-menopausal breast cancer) and endometrial cancer in females.

The current results for the common mortality causes are consistent with previous studies<sup>1-5,10</sup>. For example, the largest systematic review and meta-analysis of this relationship (including >30 million participants and ~3.7 million deaths) showed consistent evidence that each 5kg/m<sup>2</sup> increment in BMI was associated with a 5% increased risk (95% CI: 4-7%) of all-cause mortality<sup>10</sup>. Concordant with this, scaling the current results in UK Biobank suggested that each 5kg/m<sup>2</sup> increase in BMI was associated with a ~16% increased all-cause mortality risk (95% CI: -5%, 41%). Consistent with a collaborative analysis of >900,000 adults showing a ~40% increased risk of vascular mortality with each 5kg/m<sup>2</sup> higher BMI<sup>1</sup>, scaling the current results to reflect the same increase in BMI implied a ~61% increased risk of overall CVD (95% CI: 1.07, 2.43) and ~76% increased risk of CHD (95% CI: 1.00, 3.11).

For cancer, many MR-derived effect estimates were in the same direction as those derived from previous large-scale meta-analyses and reviews. For example, the association of BMI on incidence of 22 cancer sites in 5.24 million individuals suggested linear positive relationships with cancers of the kidney, liver, colorectal and ovary and inverse associations with prostate, pre-menopausal breast cancer and lung cancer, the latter being strongly driven by smoking status<sup>11</sup>. Consistent with this, despite estimates from the Cox regression suggesting a positive association between BMI and prostate cancer mortality in UK Biobank, MR analyses provided evidence (with wide CIs) in the opposite direction (i.e., higher BMI reducing prostate risk). Additionally, in the Million Women Study, incrementally higher BMI was associated with an increased risk of mortality from cancers of the endometrium, oesophagus, kidney, pancreas, lymphatic system, ovary, breast cancer (in post-menopausal women) and colorectal cancer (in pre-menopausal women)<sup>3</sup>. Whilst there was observational evidence for a positive association on mortality from endometrial cancer and post-menopausal breast cancer in the current study, estimates were inverse in MR analyses. However, analyses of cancer-specific mortality in the current study were limited by the rarity of these deaths (i.e., many cancers had <300 cases), which was accentuated further in

sex-stratified analyses, where many estimates derived from MR analyses were opposite to those from observational analyses or had CIs too wide for interpretation.

The association between BMI and all-cause mortality in MR analyses showed a J-shaped pattern but appeared flatter over a larger range of BMI compared to the observational association, with a smaller value of BMI at which mortality risk was lowest. This difference may be suggestive of confounding in previous observational associations, which overestimate the harmful effects of being underweight whilst underestimating the harmful effects of being overweight/obese. For example, studies using populations comprising older individuals with likely existing illnesses can generate spurious associations between lower BMI and increased risk of mortality (i.e., those who lose weight due to disease)<sup>9,31,41</sup>. Indeed, in the largest study to date, overestimation of estimates and this characteristic J-shaped association were reported greatest in analyses with the most potential for bias (including all participants; current, former or never smokers; and studies with short follow-up of <5 years), highlighting the importance for unbiased modes of estimation (such as those used here)<sup>10</sup>. Those that attempt to appropriately control for such effects (i.e., adjusting for baseline traits, restricting analyses to individuals who never smoked or had a longer follow-up), observe an emerging linear association<sup>2,10,42,43</sup>. Whilst it is plausible that individuals considered severely and unhealthily underweight have a higher risk of mortality than those within the normal BMI range<sup>44</sup>, the current findings in this large population of healthy individuals support a more linear association, with lower BMI being protective over most of the observed range. Furthermore, the lowest risk of mortality occurred at approximately 23kg/m<sup>2</sup> with MR, as opposed to being overweight (i.e., a BMI of 25.0-29.9kg/m<sup>2</sup>), which was observed in the current observational analyses, and has been implied previously by some existing observational studies<sup>7</sup>. Therefore, a stable BMI within the 'normal' range (i.e., 18.5-24.9kg/m<sup>2</sup>) may be the most beneficially healthy in reducing mortality risk, with any reduction within that range likely to be favourable<sup>5,10</sup>.

The MR concept rests on several key assumptions<sup>14,15</sup>: (i) the GRS must be associated with BMI; (ii) the GRS must be independent of the confounding factors that of the association between BMI and mortality; and (iii) there must be no independent pathway between the GRS and mortality other than through BMI – horizontal pleiotropy<sup>15</sup>. These assumptions were tested where possible and sensitivity analyses conducted in the current study provided little evidence of confounding or pleiotropy and awarded greater confidence in the validity of the instrument used and, thus, MR-derived estimates. Notably, the GRS was associated (with very small effect sizes) with covariables. The sheer presence of an

association between traditionally considered confounders with the GRS is interesting and could be due to (i) vertical pleiotropy (i.e., the GRS being associated with smoking status, for example, because of the potentially causal relationship between BMI and smoking) or (ii) co-incident genetic and phenotypic variation due to population structure or selection/collider bias; both reasons of which are increasingly easier to detect with the advent of very large studies such as UK Biobank<sup>45</sup>. Nevertheless, the magnitude of these relationships was marginal and MR analyses adjusting for these covariables were consistent with main analyses, suggesting little impact. Reverse causality is an important source of bias in observational estimates of the association between BMI and mortality and may be the driver of the characteristic J-shaped association. Whilst it is possible that mortality may influence the relative distribution of genetic variants within a selected sample<sup>45</sup>, it is likely that this potential bias is less marked than that seen in observational studies. Whilst there are limitations to this current study, triangulation of different methodologies (each with orthogonal sources of bias) is important for drawing causal inference within this context and these findings add to the current body of evidence aiming to estimate the role played by BMI in mortality.

The UK Biobank study is a unique opportunity to undertake these analyses, however, there are important aspects to consider. Firstly, current analyses were restricted to those of White British ancestry, limiting the generalisability of results to other ancestral groups. Secondly, one cannot rule the coincident structure in both genotype and phenotype out of any potential biasing role in genetic analyses within a study of this scale. Lastly, the power to detect associations with MR analyses remains low for many mortality causes even in a study comprising ~500,000 participants. Despite these, and given the incidence of the outcomes tested (where incidence of mortality from many causes will approximately double by 2022<sup>28</sup>), UK Biobank provides a unique opportunity to analyse and revise these estimates further over the coming years.

### *Conclusions*

This study represents the application of MR to assess the causal effect of higher BMI on the risk of mortality. Results supported the causal role of higher BMI in increasing the risk of all-cause mortality and mortality from CVDs, various cancers and several specific causes. Alongside more large-scale, comprehensive studies and the application of robust causal inference methods that appropriately account for the heavy burden of confounding, reverse causation and bias within observational epidemiological

designs, our results further highlight the need for a global effort to reduce the rising population trends for excess weight.

### ***Acknowledgements***

The authors are grateful to the UK Biobank participants. This research was conducted using the UK Biobank Resource (Application Number: 16391).

## References

1. Prospective Studies Collaboration. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. *Lancet*. 2009;373:1083-1096.
2. Berrington de Gonzalez A, Hartge P, Cerhan JR, *et al*. Body-Mass Index and Mortality among 1.46 Million White Adults. *New England Journal of Medicine*. 2010;363:2211-2219.
3. Reeves GK, Pirie K, Beral V, Green J, Spencer E and Bull D. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. *BMJ*. 2007;335:1134.
4. Kyrgiou M, Kalliala I, Markozannes G, *et al*. Adiposity and cancer at major anatomical sites: umbrella review of the literature. *BMJ*. 2017;356.
5. The Global BMI Mortality Collaboration. Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. *Lancet (London, England)*. 2016;388:776-786.
6. Arnold M, Pandeya N, Byrnes G, *et al*. Global burden of cancer attributable to high body-mass index in 2012: a population-based study. *The Lancet Oncology*. 2015;16:36-46.
7. Flegal KM, Kit BK, Orpana H and Graubard BI. Association of all-cause mortality with overweight and obesity using standard body mass index categories: A systematic review and meta-analysis. *JAMA*. 2013;309:71-82.
8. Lee J, Meyerhardt JA, Giovannucci E and Jeon JY. Association between Body Mass Index and Prognosis of Colorectal Cancer: A Meta-Analysis of Prospective Cohort Studies. *PLoS ONE*. 2015;10:e0120706.
9. Willett WC, Hu FB and Thun M. Overweight, obesity, and all-cause mortality. *JAMA*. 2013;309:1681-1682.
10. Aune D, Sen A, Prasad M, *et al*. BMI and all cause mortality: systematic review and non-linear dose-response meta-analysis of 230 cohort studies with 3.74 million deaths among 30.3 million participants. *BMJ*. 2016;353.
11. Bhaskaran K, Douglas I, Forbes H, dos-Santos-Silva I, Leon DA and Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5.24 million UK adults. *The Lancet*. 2014;384:755-765.
12. Secord AA, Hasselblad V, Von Gruenigen VE, *et al*. Body mass index and mortality in endometrial cancer: A systematic review and meta-analysis. *Gynecologic Oncology*. 2016;140:184-190.

13. Jiang Wubin CB. Does body mass index correlate with mortality of prostate cancer? A dose-response meta-analysis of cohort studies. . *International Journal of Clinical and Experimental Medicine*. 2017;10:88-96.
14. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiology studies. *Hum Mol Genet*. 2014;23(R1):R89-R98.
15. Haycock PC, Burgess S, Wade KH, Bowden J, Relton C and Davey Smith G. Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. *The American Journal of Clinical Nutrition*. 2016;103(4):965-978.
16. Dale CE, Fatemifar G, Palmer TM, *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;135:2373.
17. Corbin LJ, Richmond RC, Wade KH, *et al*. BMI as a Modifiable Risk Factor for Type 2 Diabetes: Refining and Understanding Causal Estimates Using Mendelian Randomization. *Diabetes*. 2016;65:3002-3007.
18. Fall T, Hägg S, Ploner A, *et al*. Age- and Sex-Specific Causal Effects of Adiposity on Cardiovascular Risk Factors. *Diabetes*. 2015;64:1841-1852.
19. Timpson NJ, Harbord R, Davey Smith G, Zacho J, Tybjaerg-Hansen A and Nordestgaard BG. Does Greater Adiposity Increase Blood Pressure and Hypertension Risk? *Hypertension*. 2009;54:84.
20. Holmes MV, Lange LA, Palmer T, L *et al*. Causal Effects of Body Mass Index on Cardiometabolic Traits and Events: A Mendelian Randomization Analysis. *American Journal of Human Genetics*. 2014;94:198-208.
21. Gao C, Patel CJ, Michailidou K, *et al* P. Mendelian randomization study of adiposity-related traits and risk of breast, ovarian, prostate, lung and colorectal cancer. *International Journal of Epidemiology*. 2016;45:896-908.
22. Brennan P, McKay J, Moore L, *et al*. Obesity and cancer: Mendelian randomization approach utilizing the FTO genotype. *International Journal of Epidemiology*. 2009;38:971-975.
23. Painter JN, O'Mara TA, Marquart L, *et al*. Genetic risk score Mendelian randomization shows obesity measured as body mass index, but not waist:hip ratio, is causal for endometrial cancer. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2016;25:1503-1510.

24. Mao Y, Yan C, Lu Q, *et al.* Genetically predicted high body mass index is associated with increased gastric cancer risk. *Eur J Hum Genet.* 2017;25:1061-1066.
25. Thrift AP, Shaheen NJ, Gammon MD, *et al.* Obesity and Risk of Esophageal Adenocarcinoma and Barrett's Esophagus: A Mendelian Randomization Study. *JNCI Journal of the National Cancer Institute.* 2014;106:dju252.
26. Dixon SC, Nagle CM, Thrift AP, *et al.* Adult body mass index and risk of ovarian cancer by subtype: a Mendelian randomization study. *International Journal of Epidemiology.* 2016;45:884-895.
27. Carreras-Torres R, Johansson M, Haycock PC, *et al.* Obesity, metabolic factors and risk of different histological types of lung cancer: A Mendelian randomization study. *PLOS ONE.* 2017;12:e0177875.
28. Collins R. What makes UK Biobank special? *The Lancet.* 2012;379:1173-1174.
29. Bycroft C, Freeman C, Petkova D, *et al.* Genome-wide genetic data on ~500,000 UK Biobank participants. *bioRxiv.* 2017; doi:10.1101/166298.
30. Allen NE, Sudlow C, Peakman T and Collins R. UK Biobank Data: Come and Get It. *Science Translational Medicine.* 2014;6:224ed4.
31. Wade KH, Carslake D, Ivar Nilsen T, Timpson NJ, Davey Smith G and Romundstad P. Blood pressure and mortality: using offspring blood pressure as an instrument for own blood pressure in the HUNT study. *Scientific Reports.* 2015;5:12399.
32. Locke AE, Kahali B, Berndt SI, *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature.* 2015;518:197-206.
33. Tyrrell J, Jones SE, Beaumont R, *et al.* Height, body mass index, and socioeconomic status: mendelian randomisation study in UK Biobank. *BMJ.* 2016;352.
34. Thomas DC, Lawlor DA and Thompson JR. Re: Estimation of Bias in Nongenetic Observational Studies Using "Mendelian Triangulation" by Bautista *et al.* *Annals of Epidemiology.* 2007;17:511-513.
35. Sterne JAC, Cox DR and Davey Smith G. Sifting the evidence—what's wrong with significance tests? Another comment on the role of statistical methods. *BMJ.* 2001;322:226.
36. Burgess S, Davies NM, Thompson SG and EPIC-InterAct Consortium. Instrumental variable analysis with a nonlinear exposure-outcome relationship. *Epidemiology.* 2014;25:877-885.

37. Staley JR and Burgess S. Semiparametric methods for estimation of a nonlinear exposure-outcome relationship using instrumental variables with application to Mendelian randomization. *Genet Epidemiol.* 2017;41:341-352.
38. Bowden J, Davey Smith G and Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *International Journal of Epidemiology.* 2015;44(2):512-525.
39. Bowden J, Davey Smith G, Haycock PC and Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genetic Epidemiology.* 2016;40:304-314.
40. Hartwig FP, Davey Smith G and Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *International Journal of Epidemiology.* 2017;46:1985-1998.
41. Davey Smith G, Sterne JA, Fraser A, Tynelius P, Lawlor DA and Rasmussen F. The association between BMI and mortality using offspring BMI as an indicator of own BMI: large intergenerational mortality study. *BMJ.* 2009;339.
42. Manson JE, Stampfer MJ, Hennekens CH and Willett WC. Body weight and longevity: A reassessment. *JAMA.* 1987;257:353-358.
43. Twig G, Yaniv G, Levine H, *et al.* Body-Mass Index in 2.3 Million Adolescents and Cardiovascular Death in Adulthood. *New England Journal of Medicine.* 2016;374:2430-2440.
44. Roh L, Braun J, Chiolero A, Bopp M, Rohrmann S and Faeh D. Mortality risk associated with underweight: a census-linked cohort of 31,578 individuals with up to 32 years of follow-up. *BMC Public Health.* 2014;14:371.
45. Haworth S, Mitchell R, Corbin L, *et al.* Common genetic variants and health outcomes appear geographically structured in the UK Biobank sample: Old concerns returning and their implications. *bioRxiv* 294876. 2018; doi: <https://doi.org/10.1101/294876>.

Table 1. Descriptive statistics for UK Biobank participants of White British Ancestry included in the main analyses

<b>Variable</b>	<b>N</b>	<b>Mean (SD) or percentage</b>
Age (years) at initial assessment	335,308	56.87 (8.00)
Sex (% of males)	335,308	46.22
Body mass index (kg/m <sup>2</sup> )	335,308	27.38 (4.74)
<i>Smoking status</i>	334,142	
Never	183,170	54.82
Former	117,838	35.27
Current	33,134	9.92
<i>Alcohol drinker status</i>	335,074	
Never	10,311	3.08
Former	11,368	3.39
Current	313,395	93.53
<i>Highest qualifications</i>	275,544	
College or University degree	106,280	38.57
A-levels	38,271	13.89
O-levels	73,770	26.77
CSEs	18,016	6.54
NVQ/HND/HNC	22,012	7.99
Other professional qualifications	17,195	6.24
<i>Current employment status</i>	332,835	
In paid employment or self-employed	190,085	57.11
Retired	117,615	35.34
Looking after home/family	8,690	2.61
Unable to work due to sickness/disability	9,982	3.00
Unemployed	4,436	1.33
Doing unpaid or voluntary work	1,404	0.42
Full or part-time student	623	0.19
Days/week spent doing vigorous physical activity	319,813	1.82 (1.94)
Genotyping chip <sup>1</sup>	335,308	9.24
Age at death (years)	9,570	65.66 (6.88)
Date of death <sup>2</sup>	9,570	06/02/2013 (07/07/2007-16/02/2016)

*CSE = certificate of secondary education; HNC = higher national certificate; HND = higher national diploma; NVQ = national vocational qualification; SD = standard deviation*

<sup>1</sup>*UK BiLEVE participants genotyped on the Affymetrix Axiom Array*

<sup>2</sup>*Recorded as the mean (minimum and maximum) date of death*

Table 2. Observational and MR analyses of all-cause and cause-specific mortality by BMI in UK Biobank participants of White British ancestry (men and women)

Cause of death	N <sup>1</sup>	Observational				MR-analyses		DWH <sup>5</sup>
		Unadjusted		Adjusted		HR (95% CI) <sup>4</sup>	P-value	
		HR (95% CI) <sup>2</sup>	P-value	HR (95% CI) <sup>3</sup>	P-value			
All-cause <sup>6</sup>	9,570	1.03 (1.02, 1.03)	1.16x10 <sup>-35</sup>	1.02 (1.02, 1.03)	1.20x10 <sup>-14</sup>	1.03 (0.99, 1.07)	0.17	0.96
Cardiovascular disease <sup>6</sup>	1,967	1.07 (1.06, 1.08)	1.67x10 <sup>-65</sup>	1.07 (1.06, 1.08)	3.15x10 <sup>-38</sup>	1.10 (1.01, 1.19)	0.04	0.62
Coronary heart disease	1,087	1.07 (1.06, 1.09)	3.16x10 <sup>-40</sup>	1.08 (1.06, 1.09)	1.35x10 <sup>-25</sup>	1.12 (1.00, 1.25)	0.06	0.51
Stroke	346	1.02 (1.00, 1.04)	0.12	1.01 (0.98, 1.04)	0.53	0.98 (0.80, 1.20)	0.84	0.70
Aortic aneurysm	109	1.03 (0.99, 1.07)	0.10	1.03 (0.98, 1.08)	0.32	0.80 (0.56, 1.15)	0.23	0.17
Other cardiovascular diseases	425	1.11 (1.09, 1.13)	1.19x10 <sup>-40</sup>	1.10 (1.08, 1.12)	5.74x10 <sup>-22</sup>	1.24 (1.03, 1.48)	0.02	0.23
Respiratory diseases	532	1.00 (0.98, 1.01)	0.65	0.98 (0.96, 1.01)	0.19	1.03 (0.88, 1.22)	0.68	0.64
Cancer <sup>6</sup>	5,613	1.01 (1.01, 1.02)	1.53x10 <sup>-06</sup>	1.01 (1.00, 1.02)	0.01	0.99 (0.94, 1.04)	0.68	0.34
Lung cancer	993	0.99 (0.97, 1.00)	0.10	0.97 (0.95, 0.99)	0.01	0.96 (0.85, 1.08)	0.49	0.62
Colorectal cancer	552	1.01 (1.00, 1.03)	0.14	1.01 (0.99, 1.04)	0.18	1.06 (0.90, 1.25)	0.46	0.56
Pancreatic cancer	388	1.01 (0.99, 1.03)	0.45	1.01 (0.99, 1.04)	0.34	1.10 (0.91, 1.33)	0.34	0.38
Stomach cancer	144	1.06 (1.03, 1.09)	0.0003	1.05 (1.01, 1.09)	0.03	1.18 (0.87, 1.62)	0.29	0.48
Oesophageal cancer	283	1.04 (1.01, 1.06)	0.002	1.03 (1.00, 1.06)	0.05	1.22 (0.98, 1.53)	0.08	0.15
Malignant melanoma	119	1.00 (0.97, 1.04)	0.86	0.98 (0.93, 1.03)	0.38	1.18 (0.83, 1.66)	0.36	0.36
Kidney cancer	181	1.08 (1.05, 1.11)	1.94x10 <sup>-09</sup>	1.07 (1.04, 1.11)	3.41x10 <sup>-05</sup>	0.93 (0.71, 1.23)	0.62	0.30
Bladder cancer	101	1.02 (0.98, 1.06)	0.40	0.97 (0.92, 1.03)	0.33	0.79 (0.54, 1.15)	0.21	0.18
Brain cancer	280	1.01 (0.99, 1.04)	0.37	1.01 (0.98, 1.04)	0.46	1.02 (0.81, 1.27)	0.89	0.97
Liver cancer	169	1.07 (1.04, 1.10)	1.04x10 <sup>-06</sup>	1.05 (1.02, 1.09)	0.005	0.99 (0.74, 1.32)	0.95	0.60
Lymphatic cancer	528	1.00 (0.98, 1.02)	0.88	1.00 (0.98, 1.02)	0.87	1.04 (0.88, 1.22)	0.67	0.68
Other cancers	755	1.00 (0.99, 1.02)	0.92	1.00 (0.98, 1.02)	0.87	0.95 (0.83, 1.09)	0.46	0.45
External causes	306	0.99 (0.97, 1.01)	0.44	0.97 (0.95, 1.00)	0.07	1.30 (1.05, 1.61)	0.02	0.01

BMI = body mass index; CI = confidence interval; DWH = Durbin-Wu-Hausman; HR = hazard ratio; MR = Mendelian randomization

<sup>1</sup>Number of deaths from all causes or cause-specific mortality

<sup>2</sup>Adjusted for secular trends (date of birth), estimates represent HR with each unit increase in BMI (kg/m<sup>2</sup>)

<sup>3</sup>Adjusted for secular trends (date of birth), highest household occupation, education, smoking status, alcohol intake and physical activity

<sup>4</sup>Adjusted for secular trends (date of birth) and the first ten genetic principal components

<sup>5</sup>P-value for comparing estimates derived from observational and MR analyses using a simplification of the matrix method for DWH test statistic (see Supporting Methods)

<sup>6</sup>Total number of UK Biobank participants who had died by 16<sup>th</sup> February 2016 from any cause (or those specifically defined as cardiovascular disease or cancer), which was stratified further into primary diseases of focus (excluding the mortality causes with fewer than 40 deaths and all other causes; see Table S1a and Supporting Information)

Table 3. Observational and MR analyses of all-cause and cause-specific mortality by BMI in male UK Biobank participants of White British ancestry

Cause of death	N <sup>1</sup>	Observational				MR-analyses		DWH <sup>5</sup>
		Unadjusted		Adjusted		HR (95% CI) <sup>4</sup>	P-value	
		HR (95% CI) <sup>2</sup>	P-value	HR (95% CI) <sup>3</sup>	P-value			
All-cause <sup>6</sup>	5,882	1.03 (1.02, 1.03)	4.00x10 <sup>-18</sup>	1.02 (1.01, 1.03)	1.59x10 <sup>-07</sup>	1.03 (0.98, 1.08)	0.26	0.93
Cardiovascular disease <sup>6</sup>	1,467	1.08 (1.07, 1.09)	1.23x10 <sup>-48</sup>	1.08 (1.06, 1.09)	7.39x10 <sup>-28</sup>	1.09 (0.98, 1.20)	0.10	0.88
Coronary heart disease	906	1.08 (1.07, 1.09)	8.84x10 <sup>-32</sup>	1.08 (1.06, 1.10)	4.80x10 <sup>-18</sup>	1.12 (0.98, 1.27)	0.09	0.62
Stroke	194	1.03 (0.99, 1.06)	0.10	1.02 (0.98, 1.07)	0.29	1.01 (0.76, 1.33)	0.96	0.88
Aortic aneurysm	83	1.04 (0.99, 1.09)	0.14	1.03 (0.97, 1.09)	0.40	0.80 (0.52, 1.21)	0.29	0.22
Other cardiovascular diseases	284	1.11 (1.09, 1.14)	9.96x10 <sup>-25</sup>	1.11 (1.08, 1.14)	9.99x10 <sup>-15</sup>	1.16 (0.92, 1.45)	0.21	0.75
Respiratory diseases	361	0.94 (0.91, 0.97)	1.08x10 <sup>-05</sup>	0.91 (0.88, 0.95)	2.22x10 <sup>-06</sup>	1.04 (0.85, 1.27)	0.71	0.32
Cancer <sup>6</sup>	3,113	1.01 (1.00, 1.02)	0.002	1.01 (1.00, 1.02)	0.06	1.00 (0.93, 1.07)	0.98	0.72
Lung cancer	571	0.96 (0.94, 0.98)	0.0002	0.94 (0.91, 0.97)	4.27x10 <sup>-05</sup>	0.92 (0.78, 1.08)	0.29	0.57
Prostate cancer	308	1.03 (1.01, 1.06)	0.01	1.05 (1.02, 1.08)	0.004	0.87 (0.70, 1.08)	0.21	0.12
Colorectal cancer	329	1.03 (1.00, 1.05)	0.04	1.02 (0.99, 1.05)	0.23	1.09 (0.88, 1.34)	0.43	0.59
Pancreatic cancer	201	1.01 (0.97, 1.04)	0.76	1.00 (0.96, 1.04)	0.97	1.18 (0.90, 1.54)	0.24	0.25
Stomach cancer	105	1.07 (1.03, 1.11)	0.001	1.06 (1.01, 1.12)	0.02	1.15 (0.79, 1.68)	0.45	0.70
Oesophageal cancer	226	1.06 (1.03, 1.09)	1.06x10 <sup>-04</sup>	1.07 (1.03, 1.11)	1.17x10 <sup>-04</sup>	1.28 (0.99, 1.65)	0.06	0.14
Malignant melanoma	78	0.99 (0.94, 1.05)	0.85	0.97 (0.91, 1.04)	0.42	0.99 (0.64, 1.53)	0.96	0.98
Kidney cancer	137	1.09 (1.05, 1.12)	8.27x10 <sup>-07</sup>	1.08 (1.03, 1.13)	5.72x10 <sup>-04</sup>	1.04 (0.75, 1.44)	0.82	0.79
Bladder cancer	78	0.98 (0.93, 1.04)	0.58	0.93 (0.86, 1.00)	0.05	0.73 (0.47, 1.13)	0.16	0.18
Brain cancer	169	1.01 (0.97, 1.05)	0.59	0.98 (0.94, 1.03)	0.47	1.15 (0.85, 1.54)	0.36	0.39
Liver cancer	100	1.11 (1.07, 1.15)	3.18x10 <sup>-08</sup>	1.08 (1.03, 1.13)	0.003	1.03 (0.70, 1.52)	0.86	0.73
Lymphatic cancer	329	1.00 (0.97, 1.03)	0.91	1.01 (0.98, 1.04)	0.58	1.03 (0.83, 1.27)	0.81	0.79
Other cancers	460	0.99 (0.96, 1.01)	0.22	1.00 (0.98, 1.03)	0.88	0.89 (0.74, 1.06)	0.20	0.26
External causes	206	0.97 (0.94, 1.01)	0.12	0.97 (0.93, 1.01)	0.11	1.11 (0.85, 1.45)	0.44	0.32

BMI = body mass index; CI = confidence interval; DWH = Durbin-Wu-Hausman; HR = hazard ratio; MR = Mendelian randomization

<sup>1</sup>Number of deaths from all causes or cause-specific mortality

<sup>2</sup>Adjusted for secular trends (date of birth), estimates represent HR with each unit increase in BMI (kg/m<sup>2</sup>)

<sup>3</sup>Adjusted for secular trends (date of birth), highest household occupation, education, smoking status, alcohol intake and physical activity

<sup>4</sup>Adjusted for secular trends (date of birth) and the first ten genetic principal components

<sup>5</sup>P-value for comparing estimates derived from observational and MR analyses using a simplification of the matrix method for DWH test statistic (see Supporting Methods)

<sup>6</sup>Total number of male UK Biobank participants who had died by 16<sup>th</sup> February 2016 from any cause (or those specifically defined as cardiovascular disease or cancer), which was stratified further into primary diseases of focus (excluding the mortality causes with fewer than 40 deaths and all other causes; see Table S1b and Supporting Information)

Table 4. Observational and MR analyses of all-cause and cause-specific mortality by BMI in female UK Biobank participants of White British ancestry

Cause of death	N <sup>1</sup>	Observational				MR-analyses		DWH <sup>5</sup>
		Unadjusted		Adjusted		HR (95% CI) <sup>4</sup>	P-value	
		HR (95% CI) <sup>2</sup>	P-value	HR (95% CI) <sup>3</sup>	P-value			
All-cause <sup>6</sup>	3,688	1.02 (1.01, 1.03)	1.84x10 <sup>-11</sup>	1.02 (1.01, 1.02)	3.10x10 <sup>-05</sup>	1.03 (0.96, 1.09)	0.42	0.90
Cardiovascular disease <sup>6</sup>	500	1.06 (1.04, 1.08)	6.64x10 <sup>-14</sup>	1.06 (1.04, 1.08)	1.55x10 <sup>-08</sup>	1.12 (0.95, 1.32)	0.19	0.53
Coronary heart disease	181	1.06 (1.03, 1.09)	5.05x10 <sup>-06</sup>	1.08 (1.04, 1.12)	6.87x10 <sup>-06</sup>	1.12 (0.85, 1.47)	0.43	0.71
Stroke	152	1.01 (0.97, 1.04)	0.75	1.00 (0.95, 1.04)	0.84	0.95 (0.70, 1.28)	0.72	0.70
Other cardiovascular diseases	141	1.11 (1.08, 1.14)	2.01x10 <sup>-17</sup>	1.10 (1.06, 1.13)	8.41x10 <sup>-09</sup>	1.42 (1.04, 1.93)	0.03	0.12
Respiratory diseases	171	1.05 (1.02, 1.08)	0.0004	1.06 (1.02, 1.10)	0.002	1.02 (0.77, 1.36)	0.88	0.86
Cancer <sup>6</sup>	2,500	1.01 (1.00, 1.02)	0.01	1.01 (1.00, 1.02)	0.20	0.98 (0.91, 1.05)	0.52	0.35
Lung cancer	422	1.01 (0.99, 1.03)	0.53	1.00 (0.97, 1.03)	0.97	1.02 (0.85, 1.22)	0.86	0.91
Breast cancer	468	1.02 (1.00, 1.03)	0.05	1.02 (1.00, 1.04)	0.13	0.83 (0.70, 0.99)	0.03	0.02
Pre-menopausal	48	1.00 (0.95, 1.06)	0.94	1.00 (0.95, 1.06)	0.90	0.77 (0.45, 1.32)	0.35	0.34
Post-menopausal	420	1.02 (1.00, 1.04)	0.04	1.02 (1.00, 1.04)	0.13	0.84 (0.70, 1.00)	0.05	0.03
Colorectal cancer	223	0.99 (0.97, 1.02)	0.59	1.01 (0.98, 1.04)	0.73	1.02 (0.80, 1.31)	0.85	0.80
Pancreatic cancer	187	1.01 (0.98, 1.04)	0.57	1.02 (0.99, 1.05)	0.27	1.02 (0.78, 1.34)	0.88	0.93
Ovarian cancer	211	1.00 (0.97, 1.02)	0.76	1.00 (0.96, 1.03)	0.82	1.19 (0.92, 1.53)	0.19	0.17
Endometrial cancer	50	1.10 (1.06, 1.15)	3.29x10 <sup>-06</sup>	1.12 (1.07, 1.18)	1.23x10 <sup>-05</sup>	0.63 (0.38, 1.07)	0.09	0.04
Oesophageal cancer	57	0.95 (0.90, 1.01)	0.10	0.87 (0.80, 0.95)	0.001	1.04 (0.64, 1.70)	0.87	0.71
Malignant melanoma	41	1.00 (0.94, 1.06)	0.95	0.97 (0.90, 1.04)	0.40	1.61 (0.91, 2.87)	0.10	0.10
Kidney cancer	44	1.08 (1.03, 1.13)	0.002	1.06 (1.00, 1.13)	0.06	0.67 (0.38, 1.17)	0.16	0.09
Brain cancer	111	1.00 (0.97, 1.04)	0.80	1.03 (0.99, 1.07)	0.19	0.85 (0.60, 1.21)	0.36	0.34
Liver cancer	69	1.03 (0.99, 1.08)	0.18	1.03 (0.97, 1.09)	0.34	0.94 (0.60, 1.46)	0.77	0.67
Lymphatic cancer	199	1.00 (0.97, 1.02)	0.76	0.98 (0.94, 1.01)	0.21	1.05 (0.81, 1.37)	0.70	0.68
Other cancers	295	1.01 (0.98, 1.03)	0.52	0.99 (0.96, 1.02)	0.60	1.04 (0.84, 1.29)	0.71	0.76
External causes	100	0.99 (0.95, 1.03)	0.72	0.96 (0.91, 1.01)	0.09	1.79 (1.23, 2.58)	0.002	0.002

BMI = body mass index; CI = confidence interval; DWH = Durbin-Wu-Hausman; HR = hazard ratio; MR = Mendelian randomization

<sup>1</sup>Number of deaths from all causes or cause-specific mortality

<sup>2</sup>Adjusted for secular trends (date of birth), estimates represent HR with each unit increase in BMI (kg/m<sup>2</sup>)

<sup>3</sup>Adjusted for secular trends (date of birth), highest household occupation, education, smoking status, alcohol intake and physical activity

<sup>4</sup>Adjusted for secular trends (date of birth) and the first ten genetic principal components

<sup>5</sup>P-value for comparing estimates derived from observational and MR analyses using a simplification of the matrix method for DWH test statistic (see Supporting Methods)

<sup>6</sup>Total number of female UK Biobank participants who had died by 16<sup>th</sup> February 2016 from any cause (or those specifically defined as cardiovascular disease or cancer), which was stratified further into primary diseases of focus (excluding the mortality causes with fewer than 40 deaths and all other causes; see Table S1b and Supporting Information)

Table 5. Association between weighted GRS (comprising 77 SNPs) and BMI in UK Biobank participants of White British ancestry

<b>Sample</b>	<b>N</b>	<b>Effect estimate (95% CI)<sup>1</sup></b>	<b>P-value</b>	<b>R<sup>2</sup> (%)<sup>2</sup></b>
Whole sample	335,308	0.111 (0.109 0.114)	<1.20x10 <sup>-307</sup>	1.82
Males	154,967	0.105 (0.101, 0.109)	<1.20x10 <sup>-307</sup>	2.06
Females	180,341	0.117 (0.112, 0.121)	<1.20x10 <sup>-307</sup>	1.70

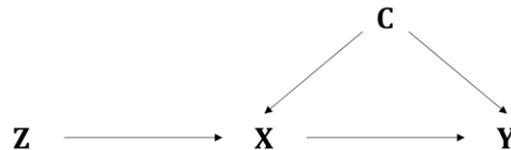
*BMI = body mass index; CI = confidence interval; GRS = genetic risk score; SNP = single nucleotide polymorphism*

<sup>1</sup>*Effect estimate (and corresponding P-value) represents the change in BMI (kg/m<sup>2</sup>) per BMI-increasing allele in individuals of White British ancestry adjusted for the first ten genetic principal components*

<sup>2</sup>*Variance in BMI explained by the GRS*

### Box: Mendelian randomization in the context of survival analyses

Mendelian randomization (MR) is a well-documented method that uses genetic variation (most commonly, single nucleotide polymorphisms [SNPs] or genetic risk score [GRS] comprising multiple SNPs) as a proxy for an exposure of interest in an instrumental variable analysis to provide an unbiased and unconfounded causal estimate of the effect of the exposure (here, body mass index [BMI]) on an outcome (here, mortality). MR relies on three key assumptions (i) the instrument (Z) is associated with the exposure (X); (ii) the instrument is independent of confounding factors (C) of the association between the exposure (X) and outcome (Y); and (iii) there must be no independent pathway between the instrument (Z) and outcome (Y) other than through exposure (X) – horizontal pleiotropy (see Figure below).



Initially, MR was applied within large-scale cohort studies and consortia that had available genetic, exposure and outcome data in one sample, where the causal estimate could be calculated in a variety of ways<sup>1</sup>. However, having all information available for MR analyses (genetic, exposure and outcome data) within one sample is difficult in large enough samples for adequate statistical power. More recently, and with the rise in genome-wide association studies (GWAS), two-sample MR methods have been developed to overcome the necessity of having all information within one sample and have proved useful in situations where both genetic and exposure data are present in one sample and both genetic and outcome data are present in a second sample. Here, the causal estimate can be calculated in a variety of ways, each of which has different assumptions and provides the ability to test the validity of the MR estimate:

- Inverse variance weighted<sup>2</sup>
- Weighted median- and mode-based estimators<sup>3,4</sup>
- MR-Egger regression<sup>5</sup>

Whilst MR is an established technique within population health sciences, the application in longitudinal studies and survival analyses is new; therefore, there is no 'gold standard'. For this manuscript, the instrumental variable ratio estimate was used in primary analyses, separating out the analyses that generated the numerator and denominator:

$$\beta_{IV} = \beta_{YZ} / \beta_{XZ}$$

where  $\beta_{IV}$  is the instrumental variable causal estimate of the association between BMI and mortality;  $\beta_{YZ}$  (numerator) is the log hazard ratio (HR) of each mortality outcome (Y) with each unit increase in a GRS (Z) derived from the Cox proportional hazards model and  $\beta_{XZ}$  (denominator) is the change in BMI (X) with each unit increase in the GRS (Z). Exponentiating the resulting ratio of the numerator and denominator yielded a MR estimate of the HR of each mortality outcome per unit increase (kg/m<sup>2</sup>) in BMI. For primary analyses in the current study, the instrument used was a GRS comprising 77 SNPs associated with BMI reported in the Genetic Investigation of ANthropometric Traits (GIANT) consortium. The GRS was generated in UK Biobank by weighting the genetic dosage of each of the 77 SNPs by its relative effect size reported in the GIANT consortium, then summed across all SNPs, divided by the combined effect size and multiplied by the number of SNPs available (N=77). The GRS therefore represented the number of average BMI-increasing variants that each individual possessed. In sensitivity analyses in this study, each of the 77 SNPs was used individually and combined using the various two-sample MR techniques (inverse variance weighted, weighted median, weighted mode and the MR-Egger estimators) to test the validity of MR assumptions.

For more details on each method discussed, see the following published articles:

- 1) **One-sample MR methods:** Haycock *et al.* Best (but oft-forgotten) practices: the design, analysis and interpretation of Mendelian randomization studies. *The American Journal of Clinical Nutrition* 2016; 103(4): 965-978.
- 2) **Inverse-variance weighted:** Burgess *et al.* Mendelian randomization analysis with multiple genetic variants using summarized data. *Genetic Epidemiology* 2013; 37(7): 658-665.
- 3) **Weighted median:** Bowden *et al.* Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genetic Epidemiology* 2016; 40(4): 304-314.
- 4) **Weighted mode:** Hartwig *et al.* Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *International Journal of Epidemiology* 2017; 46(6): 1985-1998.
- 5) **MR-Egger:** Bowden *et al.* Mendelian randomization with invalid instruments; effect estimation and bias detection through Egger regression. *International Journal of Epidemiology* 2015; 44(2): 512-525.

## FIGURE LEGENDS

### Figure 1. Flow-chart of those included in main analyses

*BMI = body mass index*

*<sup>1</sup>Of those with valid BMI, genetic and survival data, 335,308 were of White British ancestry*

*<sup>2</sup>Of those who had died by the 16<sup>th</sup> February 2016, 9,570 were of White British ancestry.*

### Figure 2. Assessment of linearity in associations of the GRS (comprising 77 SNPs) and all-cause mortality in the UK Biobank sample of White British ancestry

*CI = confidence intervals; GRS = genetic risk score; SNPs = single nucleotide polymorphisms*

*Association between the GRS (comprising 77 SNPs) and all-cause mortality, adjusted for secular trends (date of birth) and the first ten genetic principal components. Linearity tests were conducted after removing data below/above the 1<sup>st</sup>/99<sup>th</sup> percentile of BMI, respectively, due to the scarcity of data towards the tails of the BMI distribution. Hazard ratios (HRs) were calculated relative to the mean GRS value with 1000 bootstrap resamples to obtain 95% confidence intervals (CIs). The black lines represent the fitted HRs from cubic spline models (with the mean value of the GRS as the reference).*

### Figure 3. Assessment of linearity in associations of BMI and all-cause mortality in the UK Biobank sample of White British ancestry using BMI (A) and instrument-free BMI (B).

*BMI = body mass index; CI = confidence intervals*

*A: Observational associations between BMI and all-cause mortality obtained using conventional Cox regression (adjusted for secular trends (date of birth), current occupation, qualifications, smoking status, alcohol intake and physical activity).*

*B: An approximate analogue using MR stratified by categories of the instrument-free exposure (divided at the 5<sup>th</sup>, 10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup> and 85<sup>th</sup> percentile) adjusted for secular trends (date of birth) and first ten genetic principal components. These localised average causal effects were then joined together and plotted against the corresponding percentiles of the original exposure. Linearity tests were conducted after removing data below/above the 1<sup>st</sup>/99<sup>th</sup> percentile, respectively, due to the scarcity of data towards the tails of the BMI distribution. Hazard ratios (HRs) were calculated relative to the mean BMI (27kg/m<sup>2</sup>), with 1000 bootstrap resamples to obtain 95% confidence intervals (CIs). The darker lines represent the fitted HRs from cubic spline models (with mean BMI as the reference).*