



Magnus, M., Guyatt, A., Lawn, R., Wyss, A. B., Trajanoska, K., Kupers, L., Rivadeneira, F., Tobin, M., London, S., Lawlor, D., Millard, L. A. C., & Fraser, A. (2020). Identifying potential causal effects of age at menarche: a Mendelian randomization phenome-wide association study. *BMC Medicine*, 18, [71 (2020)]. <https://doi.org/10.1186/s12916-020-01515-y>

Publisher's PDF, also known as Version of record

License (if available):  
CC BY

Link to published version (if available):  
[10.1186/s12916-020-01515-y](https://doi.org/10.1186/s12916-020-01515-y)

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

This is the final published version of the article (version of record). It first appeared online via BMC at <https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-020-01515-y>. Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

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

RESEARCH ARTICLE

Open Access



# Identifying potential causal effects of age at menarche: a Mendelian randomization phenome-wide association study

Maria C. Magnus<sup>1,2,3\*</sup>, Anna L. Guyatt<sup>4</sup>, Rebecca B. Lawn<sup>1,2,5</sup>, Annah B. Wyss<sup>6</sup>, Katerina Trajanoska<sup>7</sup>, Leanne K. Küpers<sup>1,2,8</sup>, Fernando Rivadeneira<sup>7</sup>, Martin D. Tobin<sup>4,9</sup>, Stephanie J. London<sup>6</sup>, Debbie A. Lawlor<sup>1,2,10†</sup>, Louise A. C. Millard<sup>1,2,11†</sup> and Abigail Fraser<sup>1,2,10†</sup>

## Abstract

**Background:** Age at menarche has been associated with various health outcomes. We aimed to identify potential causal effects of age at menarche on health-related traits in a hypothesis-free manner.

**Methods:** We conducted a Mendelian randomization phenome-wide association study (MR-pheWAS) of age at menarche with 17,893 health-related traits in UK Biobank ( $n = 181,318$ ) using PHEASANT. The exposure of interest was the genetic risk score for age at menarche. We conducted a second MR-pheWAS after excluding SNPs associated with BMI from the genetic risk score, to examine whether results might be due to the genetic overlap between age at menarche and BMI. We followed up a subset of health-related traits to investigate MR assumptions and seek replication in independent study populations.

**Results:** Of the 17,893 tests performed in our MR-pheWAS, we identified 619 associations with the genetic risk score for age at menarche at a 5% false discovery rate threshold, of which 295 were below a Bonferroni-corrected  $P$  value threshold. These included potential effects of younger age at menarche on lower lung function, higher heel bone-mineral density, greater burden of psychosocial/mental health problems, younger age at first birth, higher risk of childhood sexual abuse, poorer cardiometabolic health, and lower physical activity. After exclusion of variants associated with BMI, the genetic risk score for age at menarche was related to 37 traits at a 5% false discovery rate, of which 29 were below a Bonferroni-corrected  $P$  value threshold. We attempted to replicate findings for bone-mineral density, lung function, neuroticism, and childhood sexual abuse using 5 independent cohorts/consortia. While estimates for lung function, higher bone-mineral density, neuroticism, and childhood sexual abuse in replication cohorts were consistent with UK Biobank estimates, confidence intervals were wide and often included the null.

**Conclusions:** The genetic risk score for age at menarche was related to a broad range of health-related traits. Follow-up analyses indicated imprecise evidence of an effect of younger age at menarche on greater bone-mineral density, lower lung function, higher neuroticism score, and greater risk of childhood sexual abuse in the smaller replication samples available; hence, these findings need further exploration when larger independent samples become available.

**Keywords:** Menarche, Mendelian randomization, MR-pheWAS

\* Correspondence: [Maria.Christine.Magnus@fhi.no](mailto:Maria.Christine.Magnus@fhi.no)

†Debbie A. Lawlor, Louise A. C. Millard and Abigail Fraser contributed equally to this work.

<sup>1</sup>MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK

<sup>2</sup>Population Health Sciences, Bristol Medical School, Bristol, UK

Full list of author information is available at the end of the article



## Background

Menarche (onset of menses) is a hallmark event in a woman's life. Using conventional multivariable regression, a younger age at menarche has been associated with higher risk of death from all causes [1], death from cardiovascular disease [1, 2], reproductive cancers [3, 4], and depression [5], but lower risk of osteoporosis/fractures [6, 7]. It has been proposed that the well-known inverse association between childhood body mass index (BMI) and age at menarche explains some of the observed associations [8, 9]. However, due to the strong tracking of BMI across the life-course, it is challenging to disentangle the role of BMI as a potential confounder, as opposed to a mediator of associations between age at menarche and health outcomes [9]. It therefore remains unclear whether the previously reported associations between age at menarche and health outcomes reflect a causal effect.

One way of evaluating causality is to use single nucleotide polymorphisms (SNPs) as instrumental variables for the exposure of interest (here age at menarche), under the assumption that the allocation of SNPs at conception is random and unrelated to potential confounding factors [10]. Mendelian randomization has previously been used to explore effects of age at menarche on cardiometabolic traits [9, 11, 12], depression [13, 14], breast cancer [15, 16], educational level [17], lung function [18], osteoporosis [19], fracture risk [20], and reproductive/behavioral outcomes [21]. However, these have focused on hypothesized effects by exploring whether associations that have been widely examined in the literature have causal evidence from Mendelian randomization analyses. This approach can miss novel potentially important (unknown and unthought of effects). A hypothesis-free approach specifically aims to go beyond that narrow focus to gain new knowledge. For example, hypothesis-free genome-wide association studies, in comparison to candidate/hypothesized genetic association studies, have efficiently identified novel biological understanding. It is possible that a similar approach to a large group of non-genetic outcomes could yield novel causal understanding. Previous phenome-wide Mendelian randomization studies of BMI and smoking have provided novel evidence of effects on outcomes not previously identified as being associated with these exposures [22, 23]. Furthermore, as with previous candidate gene-association studies, previous Mendelian randomization studies of hypothesized associations had small sample sizes. The large sample size used here supports more precise estimates than these previous studies as well as the potential for novel etiological understanding.

The objective of this study was to systematically investigate causal effects of age at menarche on health-related traits, by conducting a Mendelian randomization phenome-wide association study (MR-pheWAS).

## Methods

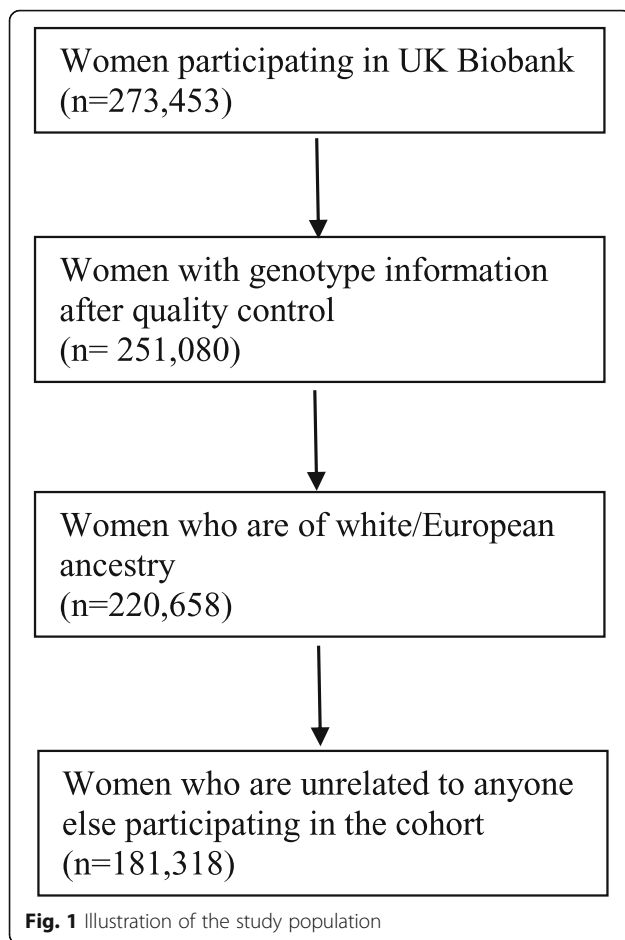
### UK Biobank

The MR-pheWAS was undertaken in the UK Biobank cohort. The UK Biobank cohort includes 503,325 people (273,453 women) between 40 and 69 years of age, who were recruited between 2006 and 2010, from 22 assessment centers across England, Scotland, and Wales [24, 25]. The response rate was 5%, and all participants gave written informed consent. Participants were followed prospectively after enrolment using Hospital Episode Statistics data, as well as data from cancer registries and the Office of National Statistics. Age at menarche in whole years was self-reported at the time of enrolment (mean 44 years after the event). Genotyping was performed using the Affymetrix UK BiLEVE Axiom array on an initial 50,000 participants; the remaining 450,000 participants were genotyped using the Affymetrix UK Biobank Axiom® array [26]. Quality control and imputation (to over 90 million SNPs, indels, and large structural variants) was performed by the Wellcome Trust Centre for Human Genetics [26, 27]. The data collection in UK Biobank was approved by the NHS National Research Ethics Service (ref 11/NW/0382). The current analysis included 181,318 unrelated genotyped women of European ancestry (Fig. 1). Relatedness was defined as third degree relatives or closer [27].

### Identifying genetic instruments for age at menarche

We used the findings from the most recent genome-wide association study (GWAS) of age at menarche to identify genetic instruments for our analysis [16]. This study included 40 studies from the ReproGen consortium ( $N=179,117$  women), in addition to 23andMe ( $N=76,831$  women) and UK Biobank ( $N=73,397$  women). They identified 37,925 variants associated with age at menarche ( $P$  value  $< 5 \times 10^{-8}$ ) that replicated across independent datasets, which constituted 389 independent signals. We generated an externally weighted genetic risk score ( $GRS_{all}$ ) as a weighted sum of the number of age at menarche decreasing alleles across 360 SNPs, weighted by the published GWAS effect estimates. The weights that we used included a subpopulation of women participating in UK Biobank included in our MR-PheWAS. Specifically, up to 73,397 (40%) of the 181,318 women included in our analyses contributed to the age at menarche GWAS effect estimates. Thus, a higher  $GRS_{all}$  reflects a younger age at menarche.

We hypothesized a priori that there could be horizontal pleiotropic effects via childhood BMI, and that there could plausibly be horizontal or vertical pleiotropic effects via adult BMI, which is itself influenced by childhood BMI (Fig. 2). Horizontal pleiotropy could bias our findings. By contrast, vertical pleiotropy is part of the



potential causal path between age at menarche and health outcomes that we are aiming to estimate. We used Steiger filtering to identify SNPs that explained more of the variation in adult BMI than the variation in age at menarche [28]. These SNPs are more likely to bias findings due to horizontal pleiotropy via BMI than SNPs that explain more of the variation of age at menarche. We generated a GRS ( $GRS_{Steiger}$ ) excluding the SNPs (27 SNPs) that explained more of the variation (quantified by the  $R^2$ ) in adult BMI than age at menarche. We only had a measure of adult BMI available in UK Biobank. Furthermore, 2 recent GWAS studies identified 15 and 941 SNPs associated with childhood and adult BMI, respectively [29, 30]. Of the 360 age at menarche SNPs included in our  $GRS_{all}$ , 7 and 199 SNPs were identified in, or situated near to (defined as being within 500,000 bp), SNPs identified in these childhood and adult BMI GWAS, respectively. We therefore also generated 2 alternative age at menarche GRS, excluding the 7 childhood BMI ( $GRS_{child\ BMI}$ ) and all 206 BMI ( $GRS_{BMI}$ ) associated SNPs, respectively.

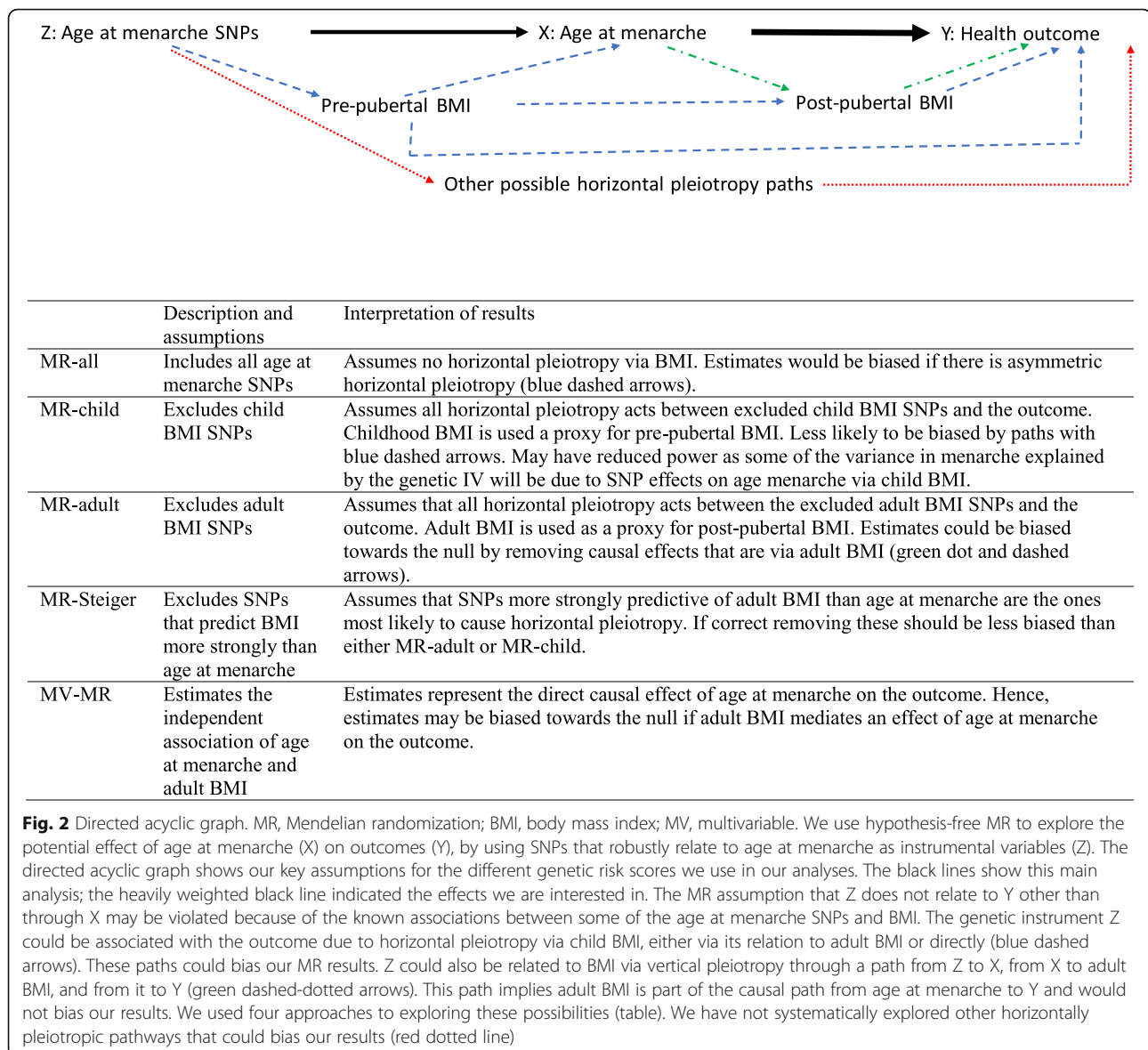
### Phenome-wide Mendelian randomization analysis

We conducted a MR-pheWAS using the publicly available PHENome Scan ANalysis Tool (PHESANT) version 0.17 which uses an automated rule-based method [31]. The decision rules start with identifying continuous, ordered categorical, unordered categorical, or binary variable fields. After outcome pre-processing (continuous traits were inverse normal rank transformed to ensure they were normally distributed), PHESANT runs linear (continuous outcomes), logistic (binary outcomes), ordered logistic (ordered categorical outcomes), and multinomial logistic (unordered categorical outcomes) regression, with the weighted allele score for age at menarche as the exposure. For unordered categorical outcomes analyzed using multinomial logistic regression, no beta coefficient or standard error is reported, but only the overall  $P$  value for the association based on a likelihood ratio test. All analyses are adjusted for age and the first 10 genetic principal components. The latter are used to account for population structure, which could produce confounded estimates.

We used PHESANT to examine the association of the age at menarche GRS with 17,893 health-related traits. To identify potential causal effects of age at menarche, we used 2 approaches that account for the number of tests performed, to help us evaluate the strength of the evidence from our MR-pheWAS. First, we derived a  $P$  value threshold setting the false discovery rate at 5%. After ranking the results by  $P$  value, this threshold is calculated as  $P_t(\text{rank}) = 0.05 \times \text{rank}/n$ , where  $n$  is the total number of tests in the phenome scan and rank is the largest rank position with a  $P$  value less than  $P_t$ . Second, we calculated a Bonferroni-corrected  $P$  value threshold, by dividing 0.05 by the number of tests performed. The Bonferroni-corrected  $P$  values assume that the tests conducted are independent. It is therefore likely to be an overly conservative correction for multiple testing, as several of the health-related traits evaluated are known to be associated. We reran our MR-pheWAS using our three alternative  $GRS_{Steiger}$ ,  $GRS_{child\ BMI}$ , and  $GRS_{BMI}$  scores as instruments. We used quantile-quantile (QQ) plots to illustrate how the distribution of the  $P$  values followed the distribution expected under the null hypothesis.

### Follow-up and replication analyses

We conducted follow-up and replication analyses of bone-mineral density (BMD), lung function, neuroticism, and childhood sexual abuse. The estimated relationships between the  $GRS_{all}$  and these outcomes were all below the Bonferroni-corrected  $P$  value threshold in our main MR-pheWAS including all SNPs in the genetic risk score. They were further selected because of the novelty (not previously studied in detail using Mendelian randomization) and pragmatically—due to the



availability of data to explore replication and potential assumption violations. Furthermore, we specifically chose not to follow-up any of the observed relationships with cardiometabolic health outcomes due to the close relationship with BMI.

We used information from 5 different cohorts/consortia to replicate our findings. These included 2 published GWAS studies of femoral neck ( $n = 22,990$ ), lumbar spine ( $n = 22,177$ ), and heel ( $n = 4566$ ) BMD from the Genetic Factors for Osteoporosis (GEFOS) consortium [32, 33]; 2 GWAS studies of lung function (1 from the Cohorts for Heart and Aging Research in Genetic Epidemiology (CHARGE) consortium ( $n = 60,552$ ) [34] and 1 from the SpiroMeta consortium ( $n = 79,055$ ) [35]); and 1 GWAS study of neuroticism from the Genetics of

Personality (GPC) consortium ( $n = 63,661$ ) [36]. Replication analysis of childhood sexual abuse was done in the Avon Longitudinal Study of Parents and Children (ALSPAC,  $n = 5953$ ) [37]. Detailed information of the replication cohorts and the definition of the outcomes used in the one- and two-sample Mendelian randomization analysis in both UK Biobank and the replications cohorts is provided in Additional file 1.

The first step in our follow-up of the findings in the MR-pheWAS was to estimate the magnitude of the causal effects. While estimates from PHESANT measure the association of the GRS with the outcomes (estimates reflect the mean difference or log-odds in outcomes per unit increase in the GRS), one- and two-sample Mendelian randomization analyses are required to estimate the



magnitude of the causal effect of the exposure (per year decrease in age at menarche). We estimated the magnitude of potential causal effects of age at menarche on BMD, lung function, and neuroticism in both UK Biobank and the replication cohorts using one- and two-sample Mendelian randomization. We estimated the effect of each age at menarche SNP on these outcomes, then generated Wald ratios (the effect estimate of each SNP on the outcome divided by the effect estimate of each SNP on age at menarche), and subsequently pooled these SNP-specific estimates using random-effects inverse-variance weighting (equivalent to random-effects meta-analysis). We used one-sample Mendelian randomization analysis to estimate the magnitude of the causal effect of age at menarche on childhood sexual abuse in both UK Biobank and ALSPAC. This approach entailed first estimating the genetically predicted age at menarche from a linear regression model including the weighted GRS as a predictor (independent variable). We then used this genetically predicted age at menarche as the exposure in a logistic regression model of childhood sexual abuse. We estimated the standard errors of the second step using bootstrapping. The potential causal effects were estimated using the  $GRS_{all}$ ,  $GRS_{Steiger}$ , and  $GRS_{BMI}$  as instruments.

Secondly, we conducted a number of sensitivity analyses in the two-sample Mendelian randomization analysis to evaluate the assumption of no unbalanced horizontal pleiotropic effects that underlies the approach (Fig. 2). We estimated the effect using Mendelian randomization-Egger (MR-Egger) regression. The estimate of the causal effect from the MR-Egger regression is unbiased if the strength of the gene-exposure association does not correlate with the strength of the bias due to horizontal pleiotropy (known as the Instrument Strength Independent of Direct Effect, or InSIDE assumption) [38]. A non-zero intercept from this regression model is an indicator of unbalanced horizontal pleiotropy [38]. Additional sensitivity analyses included simple and weighted mode-based and weighted median regression [39]. We also used the Mendelian randomization pleiotropy residual sum and outlier (MR-PRESSO) test to identify possible bias from horizontal pleiotropy [40]. This includes a global test which evaluates the overall evidence of horizontal pleiotropy, an outlier-corrected causal estimate which corrects for the detected horizontal pleiotropy, and a distortion test which estimates if the causal effect estimate is different before and after adjustment for outliers (with a  $P$  value of  $< 0.05$ ). We present the outlier-adjusted causal estimates for relationships where both the global and distortion tests provide evidence of horizontal pleiotropy (with both  $P$  values  $< 0.05$ ).

To further examine the role of adult BMI on the observed relationships with age at menarche, we also conducted multivariable Mendelian randomization in UK Biobank, which is conceptually equivalent to using traditional multivariable regression techniques to identify the independent associations between multiple exposures and an outcome of interest (see Additional file 1) to estimate effects of age at menarche independent of adult BMI (i.e., the effect of age at menarche on the health-related traits not via BMI). We also adjusted for height using multivariable Mendelian randomization (the genetic risk score for height included 3285 variants identified in the most recent GWAS) in the analyses of lung function [30].

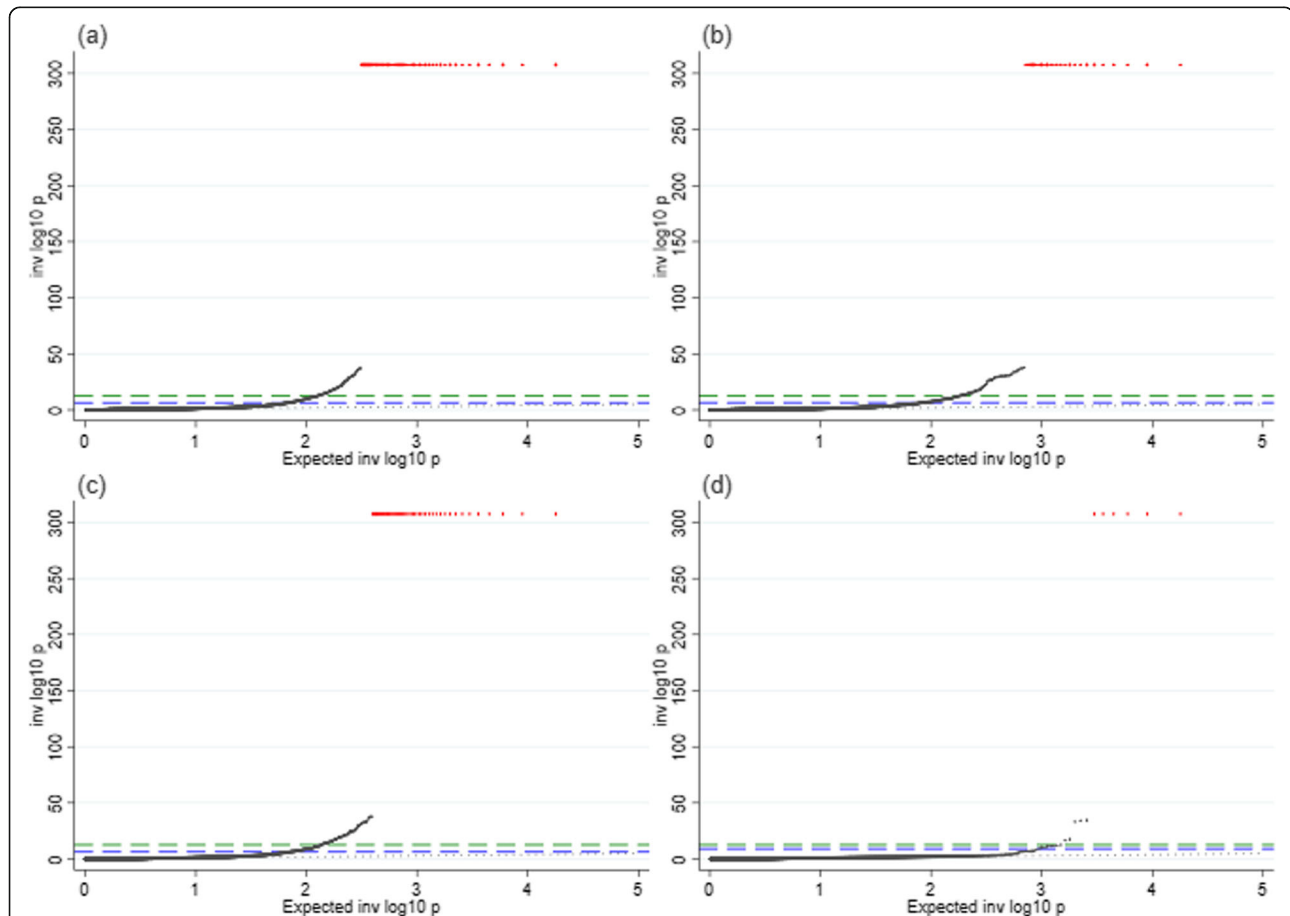
While one-sample and two-sample Mendelian randomization have the same three key assumptions, two-sample Mendelian randomization approaches vary in the extent and ways that the second assumption can be violated, which means that comparing results from both approaches is useful. In one-sample Mendelian randomization, causal estimates are robust to misspecification of the SNP-exposure association model (i.e., presence of interactions or nonlinear effects) [41]. In comparison, in two-sample Mendelian randomization, there is a risk of bias if the exposure/outcome relationship varies between the two samples used to obtain the necessary effect estimates. Furthermore, in one-sample Mendelian randomization, compared with two-sample MR, it is possible to avoid bias resulting from having to use summary data (in two-sample MR) that has been conditioned on other variables that can result in collider bias [42]. An important strength of the two-sample Mendelian randomization approach is the number of sensitivity analyses developed to explore potential presence of bias due to horizontal pleiotropy [41]. In the presence of weak instrument bias, estimates from one-sample Mendelian randomization will be biased towards the multivariable regression (observational) estimate and in the presence of residual confounding may be biased (commonly away from the null) [41]. By contrast, the estimate from the two-sample Mendelian randomization will be biased towards the null [41]. The statistical analyses were conducted using Stata version 15 (StataCorp, Texas) and R version 3.5.1 (R Foundation, [www.R-project.org](http://www.R-project.org)). The analysis code is provided in Additional file 2.

## Results

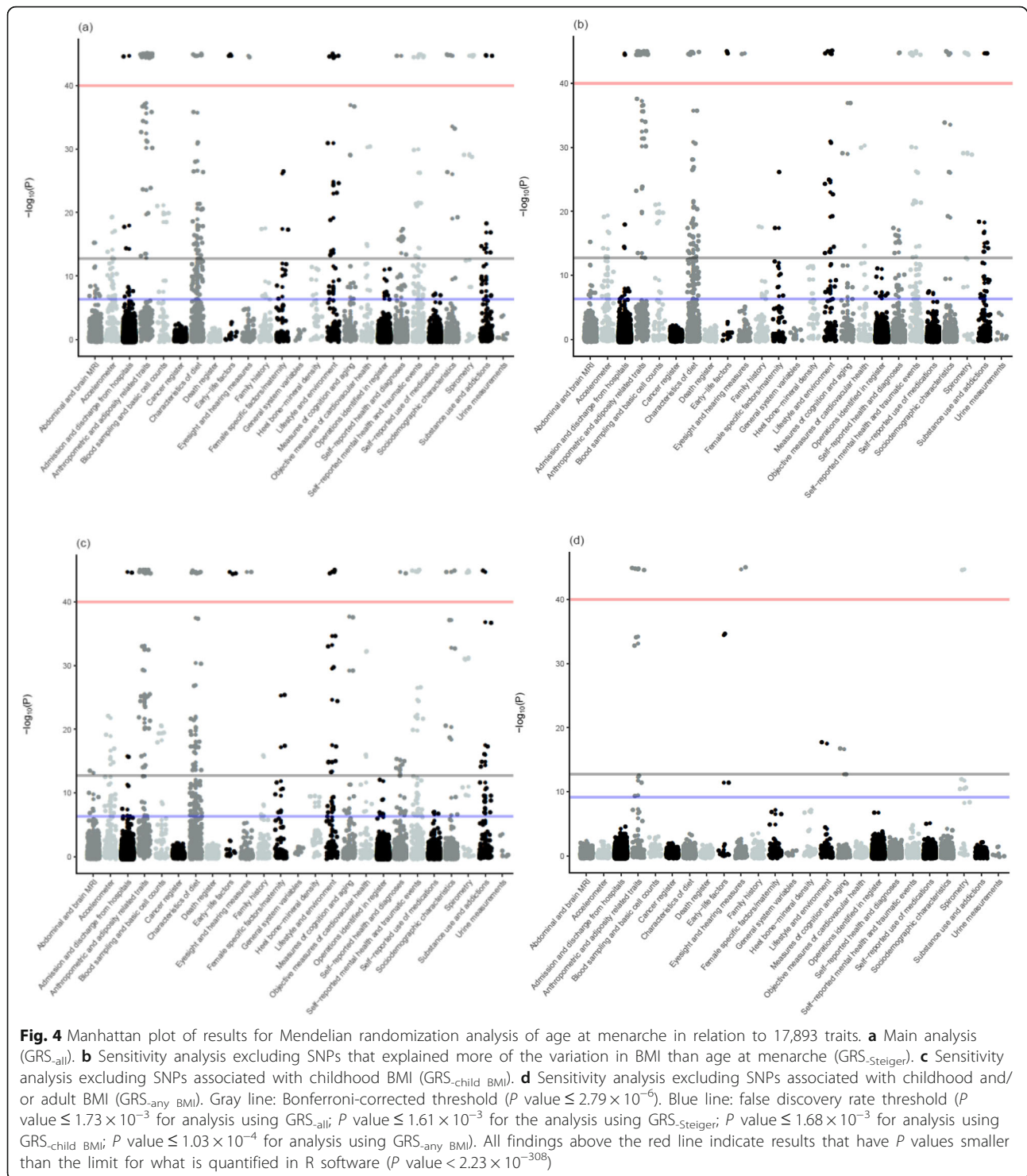
The mean age at menarche in UK Biobank was 12.9 years (standard deviation 1.6 years). The main GRS ( $GRS_{all}$ ), including 360 SNPs, explained 6.1% of the variation in age at menarche ( $F$ -statistic = 1043) (Additional file 3: Table S1). Steiger filtering found that 27 of the 360 SNPs included in the GRS for age at menarche

explained more of the variation in adult BMI than age at menarche. The genetic risk score including 333 SNPs after Steiger filtering ( $GRS_{Steiger}$ ) explained 5.8% of the variation in age at menarche ( $F$ -statistic = 979). After excluding SNPs associated with childhood BMI, the GRS including the remaining 353 SNPs ( $GRS_{child\ BMI}$ ) explained 6.0% of the variance in age at menarche ( $F$ -statistic = 1012), and after excluding childhood and/or adult BMI associated SNPs, the GRS including the remaining 154 SNPs ( $GRS_{any\ BMI}$ ) explained 3.0% of the variance ( $F$ -statistic = 489). The main GRS ( $GRS_{all}$ ) was not associated with study center or genotyping chip after adjusting for the first 10 genetic principal components ( $P$  value 0.2 for chip and  $\geq 0.6$  for study center). There was no strong evidence of a relationship between the  $GRS_{all}$  and age at recruitment ( $P$  value 0.4). The linear correlation coefficient was  $-0.002$ .

Of the 17,893 tests performed, our MR-pheWAS analysis (using  $GRS_{all}$ ) identified potential effects of age at menarche on 619 traits (3.5% of all traits) at a false discovery rate of 5% ( $P$  value  $\leq 1.73 \times 10^{-3}$ ), and 295 (1.6% of all traits) potential effects when using the more stringent Bonferroni-corrected threshold ( $P$  value  $\leq 2.79 \times 10^{-6}$ ). A quantile-quantile plot of  $P$  values is shown in Fig. 3a. The distribution of findings across categories of traits is shown in Fig. 4a, while a detailed list describing the findings and the direction of the effects is provided in Additional file 3: Table S2. Of the 619 potential effects of age at menarche, 88 were on BMI and other anthropometric traits, 111 were potential effects on diet, and 29 captured measures of physical activity. We also noted potential effects of younger age at menarche on lower lung function, greater heel BMD, increased risk of psychosocial/mental health issues, poorer cognition, increased report of physical health problems, and measures of reproductive health.



**Fig. 3** QQ-plots for the Mendelian randomization analysis of age at menarche in relation to 17,893 traits. **a** Main analysis ( $GRS_{all}$ ). **b** Sensitivity analysis excluding SNPs that explained more of the variation in BMI than age at menarche ( $GRS_{Steiger}$ ). **c** Sensitivity analysis excluding SNPs associated with childhood BMI ( $GRS_{child\ BMI}$ ). **d** Sensitivity analysis excluding SNPs associated with childhood and/or adult BMI ( $GRS_{any\ BMI}$ ). Green dashed line: Bonferroni-corrected threshold ( $P$  value  $\leq 2.79 \times 10^{-6}$ ). Blue dashed line: false discovery rate threshold ( $P$  value  $\leq 1.73 \times 10^{-3}$  for analysis using  $GRS_{all}$ ;  $P$  value  $\leq 1.61 \times 10^{-3}$  for the analysis using  $GRS_{Steiger}$ ;  $P$  value  $\leq 1.68 \times 10^{-3}$  for analysis using  $GRS_{child\ BMI}$ ;  $P$  value  $\leq 1.03 \times 10^{-4}$  for analysis using  $GRS_{any\ BMI}$ ). Black dotted line: actual = expected. Black points: results of tests performed in MR-pheWAS. Red stars: results with  $P$  values  $< 2.23 \times 10^{-308}$



For the reproductive health indicators, we observed potential effects of younger age at menarche on younger age at first birth, younger age at last birth, earlier use of exogenous hormones (both oral contraceptives and hormone replacement therapy), lower risk of bilateral

oophorectomy, higher likelihood of pregnancy terminations, lower risk of miscarriage, and higher likelihood of having already experienced menopause. Regarding specific health outcomes, we identified potential effects of younger age at menarche on greater risk of



cardiometabolic outcomes (type 2 diabetes, hypertension, ischemic heart disease, cerebral infarction, and angina). Other potential effects of younger age at menarche included increased risk of self-reported osteoporosis, arthrosis, osteoarthritis, and cholecystitis. More novel findings included potential effects of younger age at menarche on increased risk of neuroticism, sleep disorders/chronotype, and having been sexually abused during childhood. Results of our sensitivity MR-pheWAS analysis adjusting for genotyping chip were similar to the main analysis, with 617 findings below a 5% false discovery rate  $P$  value threshold and 295 below a Bonferroni-corrected  $P$  value threshold, with similar traits identified in these and the main analysis (Additional file 3: Table S3).

The sensitivity analysis using Steiger filtering to exclude SNPs that explained more of the variation in BMI from the genetic risk score yielded a slightly smaller number of findings, with 576 at the 5% false discovery rate threshold, out of which 245 were below the Bonferroni-corrected  $P$  value threshold (Figs. 3b, 4b, Additional file 3: Table S4). A total of 207 of the 245 findings that were below the Bonferroni-corrected  $P$  value threshold after Steiger filtering also passed this threshold in the main analysis.

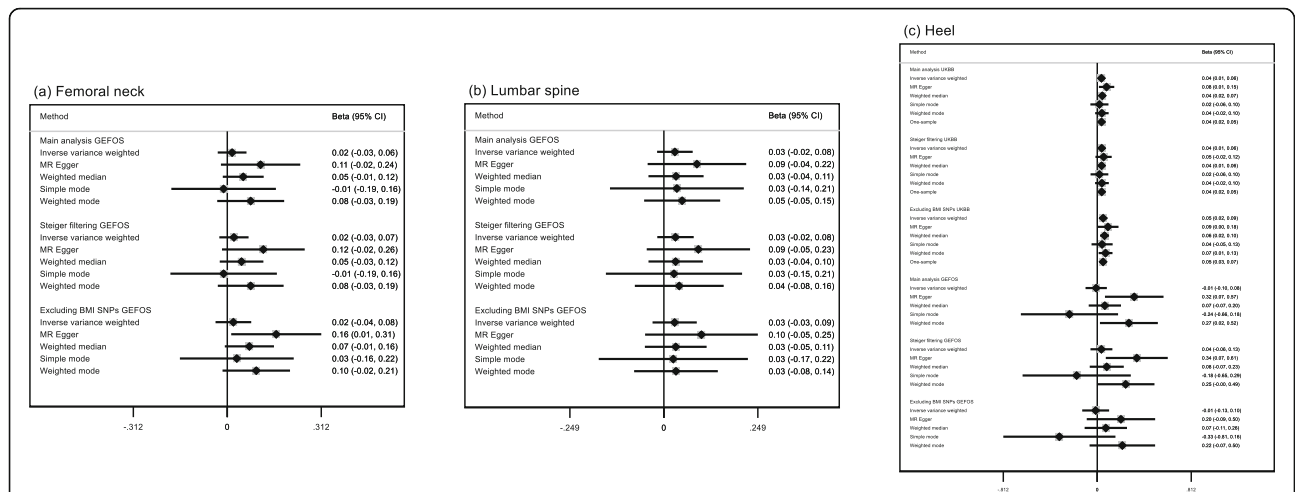
Overall, the results were also similar when we excluded SNPs associated with childhood BMI from the GRS for age at menarche ( $GRS_{child\ BMI}$ ), yielding 601 significant findings below the 5% false discovery rate ( $P$  value  $\leq 1.68 \times 10^{-3}$ ), out of which 290 were below the Bonferroni-corrected threshold, and the outcomes showing potential effects consistent with those in the main analysis (Figs. 3c, 4c, and Additional file 3: Table S5).

When we excluded all SNPs associated with childhood or adult BMI from the GRS for age at menarche ( $GRS_{BMI}$ ), we identified 37 potential effects of age at menarche below false discovery rate ( $P$  value  $\leq 1.03 \times 10^{-4}$ ), out of which 29 were below the Bonferroni-corrected threshold (Figs. 3d, 4d, and Additional file 3: Table S6). These 37 identified relationships were a subset of those identified in our main analyses ( $GRS_{all}$ ). Despite the exclusion of all SNPs associated (at a genome-wide significant level) with BMI, we still observed a potential effect of younger age at menarche on higher adult BMI (at a mean age of 57), lower height, lower lung function, greater BMD, and higher risk of hypertension. We also found a potential effect of younger age at menarche on younger age at first sexual intercourse, younger age of starting to use oral contraceptives, and younger age at first delivery among others.

We attempted to replicate some of our findings in independent samples. For ease of comparability, we show the effect estimates from an instrumental variable analysis in

UK Biobank using the same standardization/adjustment strategy applied in the published GWAS studies (see Additional file 1 for details). We also show the results for UK Biobank only adjusting for age and the first 10 genetic principal components, for BMD and lung function (Additional file 4: Figure S1 and S2, respectively). For heel BMD, we found little evidence of a difference between the estimates in UK Biobank and GEFOS, although the confidence intervals in GEFOS were wide and included the null value (Fig. 5;  $I^2$  heterogeneity statistic  $P$  values  $\geq 0.08$ ). We observed little evidence for an increase in both femoral neck and lumbar spine BMD per year decrease in age at menarche in GEFOS as all confidence intervals included the null (Fig. 5). The intercept from the MR-Egger regression was consistent with the null (no directional horizontal pleiotropy) for all BMD replication analyses except the main analysis ( $GRS_{all}$ ) of heel BMD in GEFOS (Additional file 3: Table S7; MR-Egger intercept  $P$  value = 0.01).

For  $FEV_1$  and FVC, the potential effects of age at menarche in UK Biobank and the replication cohorts (SpiroMeta and CHARGE) were consistent, although the majority of confidence intervals in replication analyses included the null value (Figs. 6 and 7; all  $I^2$  heterogeneity statistic  $P$  values  $\geq 0.07$ ). We observed no strong evidence of an effect of age at menarche on the  $FEV_1/FVC$  ratio (Figs. 6 and 7). The intercepts from the MR-Egger regression for the lung function measures were consistent with the null (Additional file 3: Table S7;  $P$  values  $> 0.2$ ), except for  $FEV_1/FVC$  in SPIROMETA ( $P$  value = 0.04). There was no strong evidence to support an effect of age at menarche on neuroticism in the GPC consortium (Fig. 8), although estimates were largely consistent with those in UK Biobank (all  $I^2$  heterogeneity statistic  $P$  values  $\geq 0.14$  except for weighted median approach for which  $I^2$   $P$  value = 0.02). The MR-Egger intercept for the regression of age at menarche on neuroticism in the GPC consortium also indicated evidence of unbalanced directional pleiotropy (Additional file 3: Table S7;  $P$  value 0.02). In the Avon Longitudinal Study of Parents and Children, the GRS for age at menarche explained 7.6% of the variation in age at menarche (Additional file 3: Table S8). The results from our replication analysis of childhood sexual abuse were consistent with the results in UK Biobank ( $I^2$  heterogeneity statistic  $P$  value = 0.9), although the confidence interval was very wide and included the null value (Fig. 9). The MR-PRESSO results from the global test for horizontal pleiotropy and the distortion test for all of the relationships evaluated in the follow-up analyses (presented in Additional file 3: Table S9) found little evidence of horizontal pleiotropy, as none of the effect estimates examined

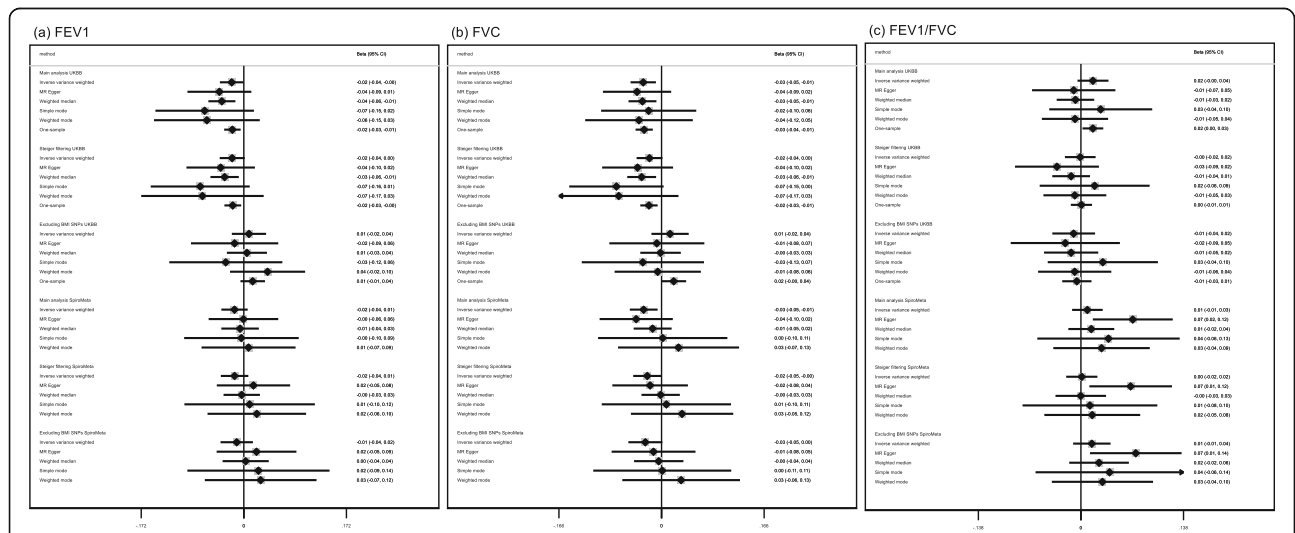


**Fig. 5** Estimates of the potential causal effect of age at menarche on bone-mineral density. BMD, bone-mineral density; BMI, body mass index; CI, confidence interval; UKBB, UK Biobank. The results reflect standard deviation difference in BMD measurements per year decrease in age at menarche. The BMD measurements were standardized by age, weight, height (heel BMD only), and genomic principal components. The measurement of femoral neck BMD was available for 22,990 women of European ancestry from the GEFOS consortium, lumbar spine BMD was available for 22,177 women of European ancestry from the GEFOS consortium, and heel BMD was available for 4566 individuals of European ancestry. For the GEFOS consortium, the main analysis of femoral and lumbar spine BMD included 263 autosomal SNPs in the genetic risk score for age at menarche, while the main analysis of heel BMD included 252 SNPs. The sensitivity analysis of femoral and lumbar spine BMD excluding BMI-related SNPs included 166 SNPs in the genetic risk score for age at menarche, while the sensitivity analysis of heel BMD included 158 SNPs

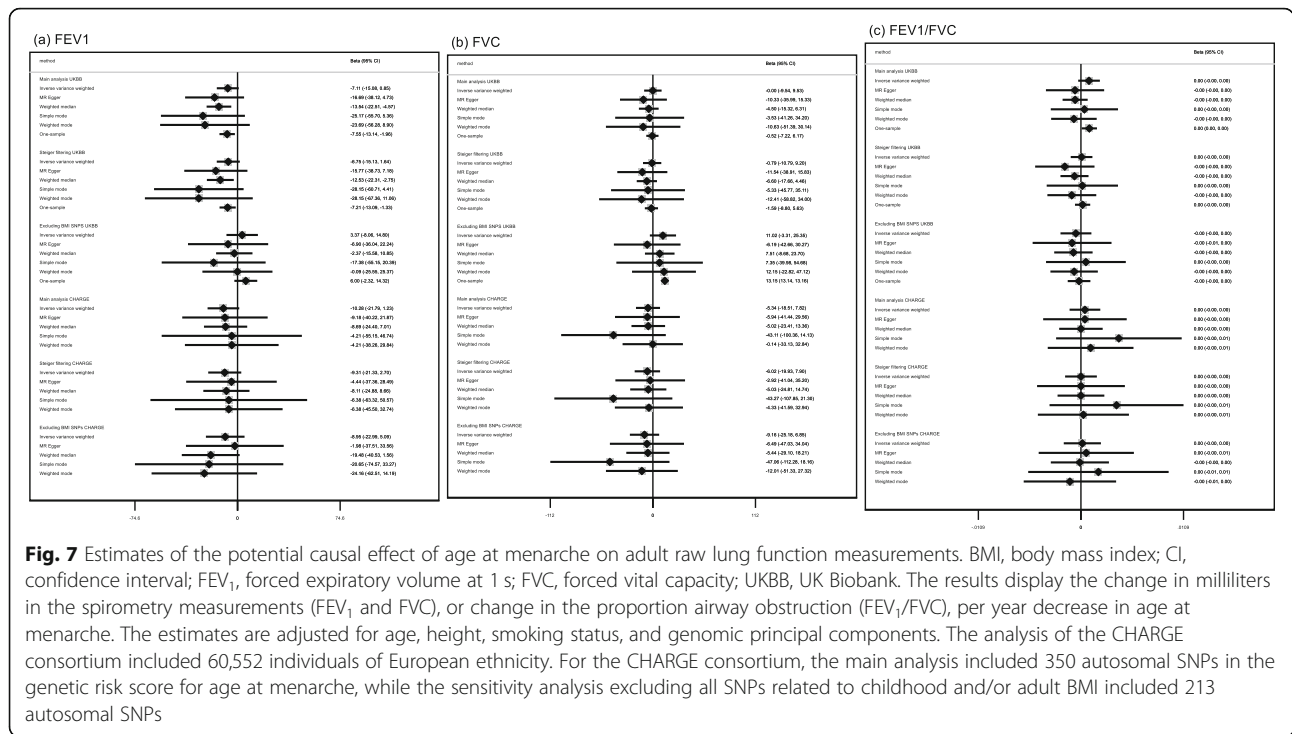
showed both global and distortion tests with  $P$  values  $< 0.05$ . We therefore do not present the causal effect estimates after adjustment for outliers.

The genetic risk scores for adult BMI and height explained 6% and 27% of the variation in these traits in UK Biobank, respectively (Additional file 3: Table S10).

Using multivariable Mendelian randomization, we found evidence of an independent effect of earlier menarche on heel BMD, FEV<sub>1</sub>, FVC, and neuroticism after accounting for adult BMI, but little evidence of an independent effect on risk of childhood sexual abuse and obstructive lung disease (Additional file 3: Table S11).



**Fig. 6** Estimates of the potential causal effect of age at menarche on adult standardized lung function measurements. BMI, body mass index; CI, confidence interval; FEV<sub>1</sub>, forced expiratory volume at 1 s; FVC, forced vital capacity; UKBB, UK Biobank. The results display the change in the ranked-based inverse normal transformed spirometry measurements per year decrease in age at menarche. The spirometry measurements were standardized by age, height, smoking status, and genomic principal components. The analysis in SpiroMeta included 79,055 individuals of European ethnicity. For the SpiroMeta consortium, the main analysis included 328 autosomal SNPs in the genetic risk score for age at menarche, while the sensitivity analysis excluding all SNPs related to childhood and/or adult BMI included 200 autosomal SNPs



**Fig. 7** Estimates of the potential causal effect of age at menarche on adult raw lung function measurements. BMI, body mass index; CI, confidence interval; FEV<sub>1</sub>, forced expiratory volume at 1 s; FVC, forced vital capacity; UKBB, UK Biobank. The results display the change in milliliters in the spirometry measurements (FEV<sub>1</sub> and FVC), or change in the proportion airway obstruction (FEV<sub>1</sub>/FVC), per year decrease in age at menarche. The estimates are adjusted for age, height, smoking status, and genomic principal components. The analysis of the CHARGE consortium included 60,552 individuals of European ethnicity. For the CHARGE consortium, the main analysis included 350 autosomal SNPs in the genetic risk score for age at menarche, while the sensitivity analysis excluding all SNPs related to childhood and/or adult BMI included 213 autosomal SNPs

There was little evidence to support a relationship between age at menarche and lung function independently of height (Additional file 3: Table S11), with effect estimates attenuating towards the null.

**Discussion**

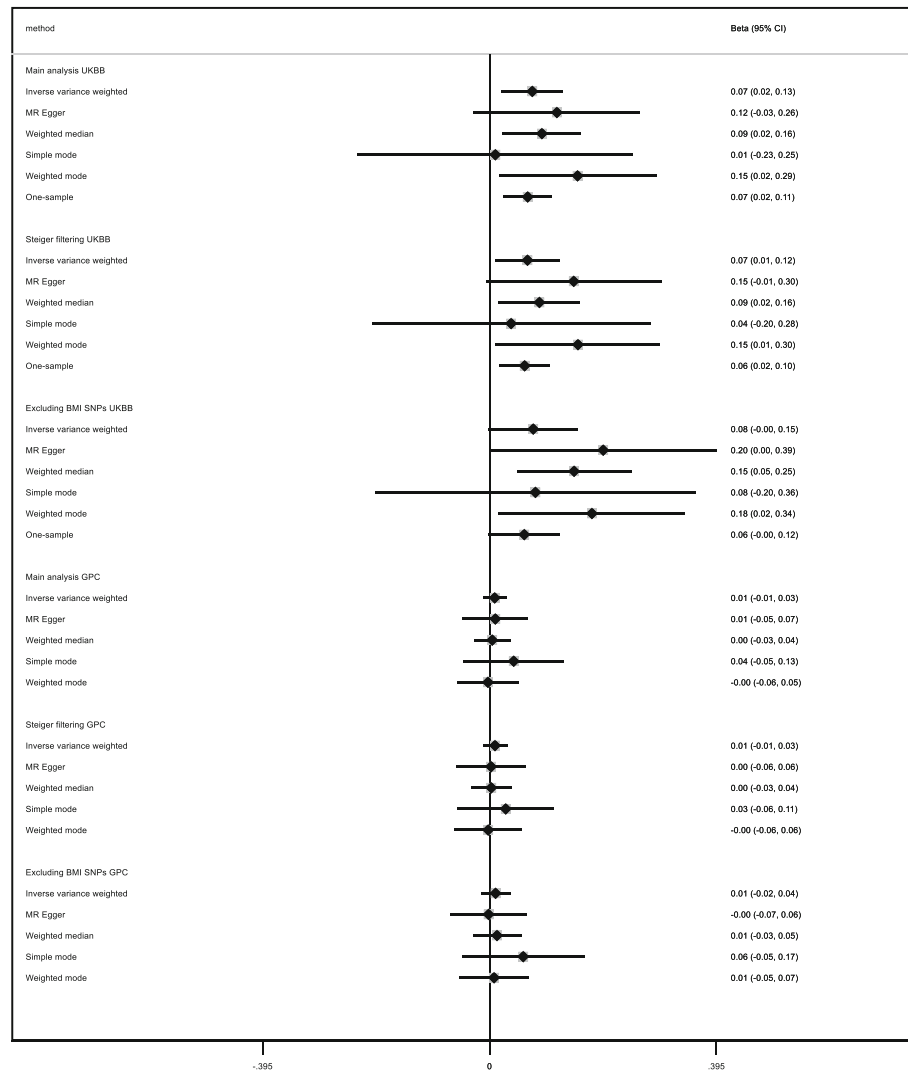
We found that younger age at menarche has a potential effect on a broad range of health-related traits. When we accounted for the genetic overlap between age at menarche and BMI, the number of potential effects identified in our MR-pheWAS at a 5% false discovery rate dropped from 619 to 37, though whether removing all BMI-related SNPs reduced bias is unclear as this could be removing a potential causal path between age at menarche and the outcome that is mediated by adult BMI. In relation to this, the majority of the potential causal effects (576/619) remained when we removed the genetic variants that explained more of the variation in BMI than age at menarche (the variants most likely to reflect horizontal pleiotropic pathways).

We conducted replication and follow-up analyses to test Mendelian randomization assumptions following our MR-pheWAS for four of the potential effects identified. While the estimated effects of younger age at menarche on lower lung function and higher BMD were consistent in UK Biobank and the replication cohorts, our replication analyses were undertaken in smaller samples than UK Biobank and had limited statistical power particularly for the binary childhood sexual abuse. This meant that with the exception of an effect of younger

age at menarche with higher BMD, the confidence intervals for the replication estimates included the null. Additional replication studies of these findings are therefore necessary when larger study populations (or published GWAS studies) of these outcomes become available. Future studies should also follow-up other potential effects identified in our MR-pheWAS which we did not follow-up to estimate the magnitude of the causal effects and investigate validity of Mendelian randomization assumptions.

In common with most existing Mendelian randomization studies, including previous studies of the effects of age at menarche on hypothesized outcomes [9, 11, 12, 14–20], we did not explore potential nonlinear effects of age at menarche on outcomes. The reason we have not done this and it is rarely done in other studies is because current methods are only feasible in one-sample MR and require very large sample sizes, and the choice of where to put thresholds (for examining MR effects in subsets across the distribution) is unclear [43]. As methods are further developed, including for potential use in two-sample MR, this could be further explored in future studies that follow up specific findings from our MR-PheWAS.

Due to the number of potential effects on health-related traits by age at menarche identified in the MR-PheWAS, we could not follow-up all the findings with sensitivity and replication analyses. Several other health-related traits were related to the genetic risk score for age at menarche in our MR-pheWAS, and their causal

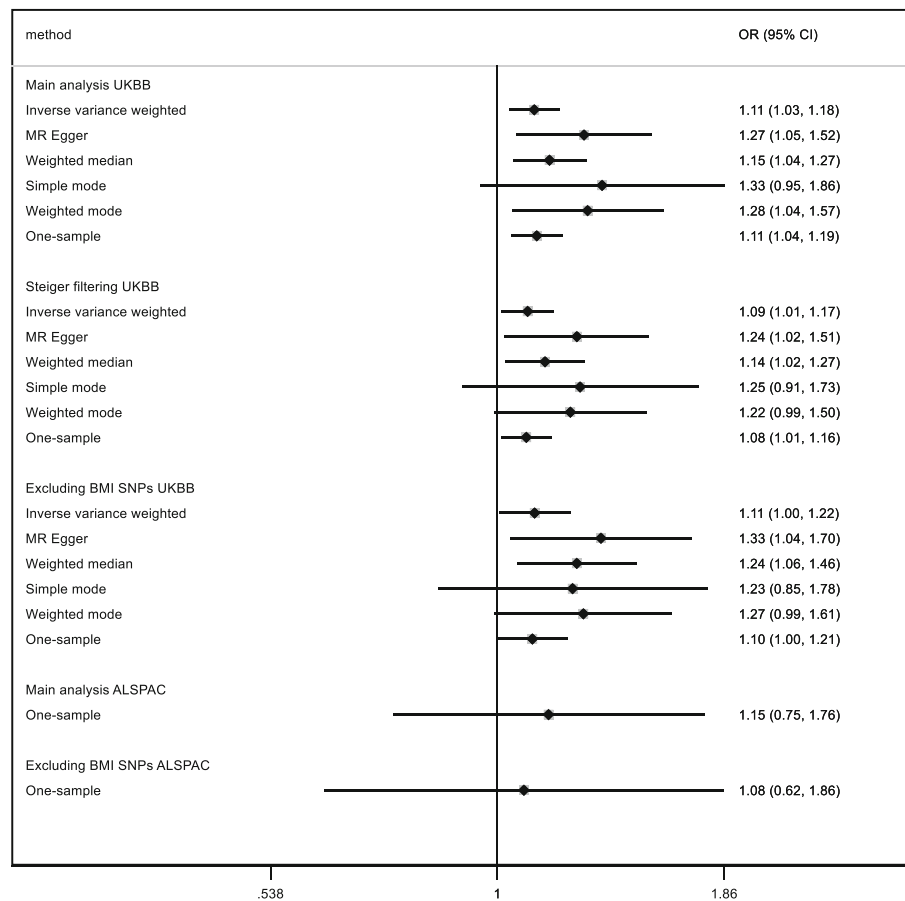


**Fig. 8** Estimates of the potential causal effect of age at menarche on neuroticism. BMI, body mass index; CI, confidence interval; GPC, Genetics of Personality Consortium; UKBB, UK Biobank. The estimates reflect the change in the harmonized neuroticism score per year decrease in age at menarche adjusted for age and principal components. The Genetics of Personality Consortium (GPC) analysis included 63,661 individuals. For the GPC consortium, the main analysis included 344 SNPs in the genetic risk score for age at menarche, while the sensitivity analysis excluding SNPs associated with BMI included 208 SNPs

relationship with age at menarche should be further explored in future studies. An underlying genetic predisposition to younger age at menarche was for example linked to a younger age at first birth, higher number of offspring, higher risk of miscarriages/stillbirths, and higher risk of sleeplessness/insomnia, depression, osteoarthritis, arthrosis, and cholecystitis among others. Our MR-PheWAS included some early-life health-related outcomes that predate menarche (maternal smoking around the time of delivery, comparative body size at age 10, and whether the woman was breastfed). We note that the MR-PheWAS estimates the association of a genetic predisposition to younger age at menarche with these outcomes, and this does not necessarily reflect an

effect of age at menarche on these outcomes since there may have been violations in the instrumental variable assumptions. We chose to follow-up childhood sexual abuse, but we were not able to distinguish whether the abuse occurred before or after menarche, as detailed information on the particular ages during which the abuse occurred was not available. Considered in the context of a genetic predisposition to earlier maturation (age at menarche occurs at a specific time but pubertal changes will have started earlier), the potential effect we see here may also reflect a tendency for girls who undergo puberty earlier to be more prone to abuse.

One limitation of our analysis is the low response in UK Biobank (5%), which could have resulted in selection



**Fig. 9** Estimates of the potential causal effect of age at menarche on risk of sexual abuse. ALSPAC, Avon Longitudinal Study of Parents and Children; BMI, body mass index; CI, confidence interval; UKBB, UK Biobank. The ordinal response scale to sexual abuse in UK Biobank was converted to a binary variable denoting whether the participant reported any history of childhood sexual abuse, to be comparable to the replication cohort. The estimates reflect the change in risk of sexual abuse per year decrease in age at menarche after adjusting for 5 principal components. The analysis of the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort included 5953 women. For ALSPAC, the main analysis included 342 SNPs in the genetic risk score for age at menarche, while the sensitivity analysis excluding SNPs associated with BMI included 208 SNPs

bias [44, 45]. Participants in UK Biobank have been shown to be healthier and of a higher socioeconomic position compared to other estimates for the British population [46]. This could have resulted in a lower burden of some of the health-related outcomes evaluated, such as a lower proportion of smokers, lower mean BMI, lower overall CVD risk (less diabetes and less hypertension), and less psychological problems, among others. The effect of potential selection bias likely varies across the large number of health-related outcomes evaluated. Reassuringly, the mean age at menarche was as expected based on the estimated age at menarche in the general population [47]. Recruitment into the cohort is also restricted to individuals who had survived until the age at which time they were recruited, meaning that if age at menarche is causally related to (premature) mortality, survival bias may have influenced our findings, as it would have in any previous prospective studies.

A second limitation is the statistical power of our replication analyses. We conducted post hoc power calculations to evaluate the power of our replication analyses. Assuming a type 1 error rate of 5%, and that the instrument explains 5% of the variation in age at menarche, we calculated the minimal effect detectable with 80% power (Supplement Table S12). The effect estimates reflect the change in the standard deviation of the continuous outcomes per standard deviation decrease in age at menarche. These results indicated that we were adequately powered (80%) to detect a standard deviation in the continuous outcomes ranging from 0.04 for the lung function measures in the CHARGE consortium to 0.17 for heel bone-mineral density in the GEFOS consortium. We were further powered to detect a twofold increase in the odds of sexual abuse per standard deviation increase in age at menarche in Avon Longitudinal Study of Parents and Children. The replication analyses were



therefore powered to detect modest effect estimates similar to that observed in UK Biobank, with the exception of the analysis in Avon Longitudinal Study of Parents and Children. Thus, for this outcome in particular, further replication in larger studies is necessary. To our knowledge, we have used the largest sample sizes currently available for replication (i.e., the largest samples with the relevant outcome and genetic data).

A strength of our study is the use of genetic variants as instrumental variables to reduce the risk of confounding. However, the weights we used to generate our GRS included a subpopulation of women participating in UK Biobank. This overfitting of the estimated coefficients for the relationship between the SNPs and age at menarche to UK Biobank could therefore have contributed to an overestimation of the amount of variation in age at menarche explained by the GRS. On the other hand, the known “winner’s curse” in GWAS studies might also lead to an underestimation of causal effect estimates using Mendelian randomization [48]. We attempted to minimize the risk of bias due to population stratification by restricting to individuals of European ethnicity and adjusting for genetic principal components. However, this limits the generalizability of our results to people from other backgrounds. A core Mendelian randomization assumption states that the genetic instruments should only affect the outcome through pathways that are via the exposure of interest (which may be violated by horizontal pleiotropy). The GRS for age at menarche includes a large number of genetic variants. This increases the likelihood that horizontal pleiotropy may be biasing our results. However, we estimated causal effects on our follow-up analyses using two-sample Mendelian randomization methods that require different assumptions about the extent that the exclusion restriction can be violated (e.g., the extent that horizontal pleiotropy can occur) and tested for unbalanced horizontal pleiotropy using MR-Egger regression and found little evidence of this for these outcomes. The self-reported information used to create a score for neuroticism and to define a history of childhood sexual abuse did not directly compare across UK Biobank and the replication cohorts. However, the point estimates particularly for sexual abuse were remarkably similar.

Due to the genetic correlation between age at menarche and BMI, we repeated the analyses excluding SNPs associated with BMI at a Bonferroni-corrected level from the age at menarche GRS. Notably, the  $GRS_{BMI}$  for age at menarche excluding SNPs associated with childhood and adult BMI was still associated with adult BMI (the PHESANT estimated change in the inverse ranked transformed BMI per allele increase was 0.006 (95% CI, 0.005, 0.006) for  $GRS_{all}$ , while it was 0.002 (95% CI, 0.002, 0.003) for  $GRS_{BMI}$ ). This is likely due to the

presence of SNPs with modest effects on BMI that did not reach the Bonferroni-corrected  $P$  value threshold in the published GWAS of childhood and adult BMI. Steiger filtering indicated that only 27 of the 360 SNPs included in the GRS for age at menarche explained more of the variation in BMI than age at menarche. The low proportion of the GRS SNPs that explained a greater proportion of variation in BMI than age at menarche suggests that most of the SNPs are valid instruments for age at menarche and their relationship with BMI is downstream of age at menarche (reflecting that BMI is a potential mediator of the effects of interest). This is why we present the findings from the genetic risk score including all genetic variants as the main analysis. We conducted multivariable Mendelian randomization in UK Biobank, which also suggested effects of age at menarche on heel BMD, FEV1, FVC, neuroticism, and childhood sexual abuse independent of BMI. In our replication analysis of lung function and neuroticism, we had to use estimates of the effect of the SNPs on the outcomes including both sexes, as GWAS results of these traits by sex were not available. However, there is evidence that a large proportion of the SNPs included in the GRS for age at menarche also predict puberty development in a similar way in males (i.e., those predicting early pubertal development in females also reflect an earlier puberty development in males) [16]. We did not conduct multivariable Mendelian randomization for the entire MR-PheWAS, as MR-pheWAS is hypothesis-generating, identifying potential effects which subsequently need to be followed up with further analyses. Multivariable Mendelian randomization also has some limitations, as it requires the variables used in the model to have similar instrument strength [49]. The stronger instrument strength for BMI compared with age at menarche could therefore have attenuated our findings and caused us to miss potential causal effects of age at menarche. PHESANT performs the test of the association of a genetic instrumental variable with a wide range of outcomes (here ~18,000 outcomes) to identify outcomes for which there is evidence of a potential causal effect, providing the opportunity to identify novel causal effects. It was not possible for us to complete replication and assumption testing for all of the potential effects identified (over 600 at a 5% false discovery rate), and we therefore followed up 4 selected examples identified in our MR-pheWAS to estimate the magnitude of the causal effect and explore validity of instrumental variable assumptions. Thus, a useful further advance may be automation of follow-up analyses and replications, for example, using a system like MR-Base. The lack of available data to undertake replication and assumption testing for all outcomes will remain an issue, but with increasing

amounts of GWAS studies, this will improve over time.

We found some evidence of an effect of younger age at menarche on lower lung function in UK Biobank, but while the confidence intervals were consistent in both the SpiroMeta and CHARGE consortia replication samples, the majority of confidence intervals included the null. Previously, a Mendelian randomization analysis of age at menarche and lung function using data from the earlier release of the genetic data from UK Biobank, where UK Biobank was 1 of 3 cohorts used to estimate the effect on adult lung function (total  $n = 46,944$  women), also reported lower adult FEV<sub>1</sub> and FVC with younger age at menarche [18]. One potential explanation for this finding is that younger age at menarche/earlier exposure to female sex hormones terminates lung growth at a younger age, and subsequently leads to a lower maximally attained lung function and lower lung function in adulthood [50]. This potential explanatory mechanism was further supported by the fact that we observed little evidence to support a relationship between age at menarche and lung function after accounting for adult height in multivariable Mendelian randomization.

Our results in UK Biobank suggest that younger age at menarche may result in greater adult heel BMD, but while our replication estimates were largely consistent with our UK Biobank results, confidence intervals were wide and included the null. A beneficial effect of age at menarche on BMD, if one exists, would be consistent with results of a recent Mendelian randomization study indicating that a delayed onset of puberty may causally affect fracture risk [20]. The importance of female sex hormones for bone health is established and is clearly reflected in the reduction of bone mass and increased risk of osteoporosis among postmenopausal women [51].

Our results in UK Biobank also suggested that younger age at menarche might result in higher levels of neuroticism. Notably, a previously published BMI MR-pheWAS supported a causal effect of BMI on neuroticism [22]. Any potential explanatory mechanisms for a relationship between age at menarche and neuroticism are less evident than for lung function and BMD and need to be further investigated. We also observed evidence of a higher risk of sexual abuse among women with a younger age at menarche, although the replication analysis was underpowered. This association has been previously reported [52–54].

The main contribution of our study is to increase understanding of potential effects of age at menarche on health-related traits. While it is not possible to directly intervene on age at menarche, it is important to understand the causal pathways to disease and the role that pubertal timing might play. Our study therefore lays the

foundations for future research into the biosocial mechanisms that might explain how age at menarche might influence the health-related traits, some of which are likely to be modifiable. As several of the health-related traits are common and complex with multiple risk factors, if our findings are further replicated in larger cohorts providing more precise estimates, they can be used to give advice on risk. For example, if it is established that younger age at menarche is causally related to lower lung function, girls with younger age at menarche might be advised to remain physically fit to maintain their lung function. Furthermore, knowing that you have one (unmodifiable) risk factor for some health outcomes is useful for women, as it means they should attempt to avoid any additional adverse modifiable risk factors.

## Conclusions

Our results suggest that younger age at menarche has potential effects on a broad range of health-related traits. Follow-up analysis indicated imprecise evidence of an effect of younger age at menarche on higher bone-mineral density, lower lung function, greater score for neuroticism, and greater risk of childhood sexual abuse in the smaller replications samples available, and these relationships should therefore be re-examined when larger study populations become available. Future studies are needed to further investigate the potential effects which we did not follow-up here. Additional studies using other designs with different biases and sufficient statistical power to replicate our findings would also be useful. Where future studies provide strong evidence for causal effects of age at menarche on several outcomes, studies to explore potential modifiable mechanisms to mitigate any effect of age at menarche will be important.

## Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s12916-020-01515-y>.

**Additional file 1.** Online Methods.

**Additional file 2.** Analysis code.

**Additional file 3: Table S1.** Strength of the genetic instrument in predicting age at menarche in UK Biobank. **Table S2.** Findings from the main analysis of the genetic risk score for age at menarche. **Table S3.** Finding from the main analysis of the genetic risk score for age at menarche after additional adjustment for chip used for the genotyping. **Table S4.** Finding from the main analysis of the genetic risk score for age at menarche after Steiger filtering. **Table S5.** Findings from the sensitivity analysis of the genetic risk score for age at menarche excluding SNPs associated with childhood body-mass index. **Table S6.** Findings from the sensitivity analysis of the genetic risk score for age at menarche excluding SNPs associated with childhood and/or adult body-mass index. **Table S7.** MR-Egger intercept. **Table S8.** Results from MR-PRESSO global test for pleiotropy and distortion tests. **Table S9.** Strength of the genetic instrument in predicting age at menarche in ALSPAC. **Table S10.** Strength of the genetic instrument in predicting body-mass index and height in UK Biobank. **Table S11.** Findings from the multivariable mendelian randomization analysis further adjusting for

body-mass index and height. **Table S12.** Post hoc power calculation analyses for the replication analyses.

**Additional file 4: Figure S1.** Estimates of the potential causal effect of age at menarche on bone-mineral density in UK Biobank. **Figure S2.** Estimates of the potential causal effect of age at menarche on adult lung function in UK Biobank.

#### Acknowledgements

This research has been conducted using the UK Biobank Resource (application number 16729). We are extremely grateful to all participants in UK Biobank, ALSPAC, GOPEC, GEFOS, CHARGE, and SpiroMeta, and to the people taking part in the infrastructure surrounding these cohorts.

#### Authors' contributions

This study was initially conceived and preliminarily designed by MCM, DAL, LACM, and AF. The analysis was undertaken by MCM, ALG, LK, RBL, ABW, and KT. SJL, MDT, and FR all contributed to providing access to data for replication analyses. The manuscript was authored and approved by all named authors.

#### Funding

This work was partly supported by the Research Council of Norway through its Centres of Excellence funding scheme, project number 262700. MCM, AF, RL, LACM, and DAL work at the Medical Research Council Integrative Epidemiology Unit at the University of Bristol which receives infrastructure funding from the UK MRC (MC\_UU\_00011/6). The contributions of AF and DAL were supported by the National Institute for Health Research Biomedical Research Centre at University Hospitals Bristol National Health Service Foundation Trust and the University of Bristol. MCM and AF are supported by a UK MRC fellowship (MR/M009351/1) awarded to AF. LACM is funded by a University of Bristol Vice-Chancellor's Fellowship. DAL's contribution is supported by the US National Institutes of Health (R01 DK10324) and a European Research Council Advanced Grant (DevelopObese; 669545); DAL is an NIHR Senior Investigator (NF-SI-0616-10102). SJL and ABW are supported by the Intramural Research Program of the NIH (NIEHS Z01 ES49019 and for ABW contract no. HHSN273201800005I). UK Biobank has received funding from the UK Medical Research Council, Wellcome Trust, Department of Health, British Heart Foundation, Diabetes UK, Northwest Regional Development Agency, Scottish Government, and Welsh Assembly Government. The data collection in ALSPAC is supported by the UK MRC, the Wellcome Trust (grant ref: 102215/2/13/2), and the University of Bristol. Genetic analyses in ALSPAC were supported by the Wellcome Trust (WT088806). Genetic analyses of lung function in SpiroMeta were supported by the Wellcome Trust (WT202849) and UK MRC MR/N011317/1 and G0902313. Infrastructure for the CHARGE Consortium is supported by the NHLBI grant R01HL105756. K. T and F. R are supported by the Netherlands Scientific Organization (NWO) and ZonMW Project number: NW O/ZONMW-VIDI-0 16-136-367. The GEFOS Consortium (www.gefos.org) was funded by the European Commission (HEALTH-F2-2008-201865-GEFOS). No funding bodies had any role in the study design, data collection and analysis, decision to publish, or preparation of the paper.

#### Availability of data and materials

The data underlying the results presented in this study can be made available on request from the UK Biobank. Questions and applications for access to data can be submitted to [access@ukbiobank.ac.uk](mailto:access@ukbiobank.ac.uk).

#### Ethics approval and consent to participate

The data collection in UK Biobank was approved by the NHS National Research Ethics Service (Ref 11/NW/0382). Ethical approval for the study was granted by the ALSPAC Law and Ethics Committee and the Local Research Ethics Committees. Ethical approval for the different cohorts contributing to the CHARGE, SPIROMETA, and GPC consortia was obtained from local institution ethical committees. All participants provided written informed consent.

#### Consent for publication

Not applicable.

#### Competing interests

DAL receives (or has received in the last 10 years) research support from national and international government and charitable bodies, Roche Diagnostics, and Medtronic for research unrelated to the current work. The other authors declare that they have no competing interests.

#### Author details

<sup>1</sup>MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK. <sup>2</sup>Population Health Sciences, Bristol Medical School, Bristol, UK. <sup>3</sup>Centre for Fertility and Health, Norwegian Institute of Public Health, P.O. Box 222 Skøyen, 0213 Oslo, Norway. <sup>4</sup>Department of Health Sciences, University of Leicester, Leicester, UK. <sup>5</sup>School of Experimental Psychology, University of Bristol, Bristol, UK. <sup>6</sup>Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, NC, USA. <sup>7</sup>Department of Internal Medicine, Erasmus MC, University Medical Center, Rotterdam, The Netherlands. <sup>8</sup>Division of Human Nutrition and Health, Wageningen University & Research, Wageningen, The Netherlands. <sup>9</sup>National Institute for Health Research, Leicester Respiratory Biomedical Research Centre, Glenfield Hospital, Leicester, UK. <sup>10</sup>NIHR Bristol Biomedical Research Centre at the University Hospitals Bristol NHS Foundation Trust and the University of Bristol, Bristol, UK. <sup>11</sup>Intelligent Systems Laboratory, Department of Computer Science, University of Bristol, Bristol, UK.

Received: 12 August 2019 Accepted: 10 February 2020

Published online: 23 March 2020

#### References

- Charalamopoulos D, McLoughlin A, Elks CE, Ong KK. Age at menarche and risks of all-cause and cardiovascular death: a systematic review and meta-analysis. *Am J Epidemiol*. 2014;180:29–40.
- Luijken J, van der Schouw YT, Mensink D, Onland-Moret NC. Association between age at menarche and cardiovascular disease: a systematic review on risk and potential mechanisms. *Maturitas*. 2017;104:96–116.
- Gong TT, Wang YL, Ma XX. Age at menarche and endometrial cancer risk: a dose-response meta-analysis of prospective studies. *Sci Rep*. 2015;5:14051.
- Gong TT, Wu QJ, Vogtmann E, Lin B, Wang YL. Age at menarche and risk of ovarian cancer: a meta-analysis of epidemiological studies. *Int J Cancer*. 2013;132:2894–900.
- Miller LJ, Girgis C, Gupta R. Depression and related disorders during the female reproductive cycle. *Womens Health (Lond)*. 2009;5:577–87.
- Bonjour JP, Chevalley T. Pubertal timing, bone acquisition, and risk of fracture throughout life. *Endocr Rev*. 2014;35:820–47.
- Clarke BL, Khosla S. Female reproductive system and bone. *Arch Biochem Biophys*. 2010;503:118–28.
- Ahmed ML, Ong KK, Dunger DB. Childhood obesity and the timing of puberty. *Trends Endocrinol Metab*. 2009;20:237–42.
- Bell JA, Carlslake D, Wade KH, Richmond RC, Langdon RJ, Vincent EE, Holmes MV, Timpson NJ, Davey SG. Influence of puberty timing on adiposity and cardiometabolic traits: a Mendelian randomisation study. *PLoS Med*. 2018;15:e1002641.
- Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey SG. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med*. 2008;27:1133–63.
- Au Yeung SL, Jiang C, Cheng KK, Xu L, Zhang W, Lam TH, Leung GM, Schooling CM. Age at menarche and cardiovascular risk factors using Mendelian randomization in the Guangzhou Biobank Cohort Study. *Prev Med*. 2017;101:142–8.
- Gill D, Brewer CF, Del Greco MF, Sivakumaran P, Bowden J, Sheehan NA, Minelli C. Age at menarche and adult body mass index: a Mendelian randomization study. *Int J Obes*. 2018;42:1574–81.
- Sequeira ME, Lewis SJ, Bonilla C, Smith GD, Joinson C. Association of timing of menarche with depressive symptoms and depression in adolescence: Mendelian randomisation study. *Br J Psychiatry*. 2017;210:39–46.
- Au Yeung SL, Jiang C, Cheng KK, Xu L, Zhang W, Lam TH, Leung GM, Schooling CM. Age at menarche and depressive symptoms in older Southern Chinese women: a Mendelian randomization study in the Guangzhou Biobank Cohort Study. *Psychiatry Res*. 2017;259:32–5.
- Burgess S, Thompson DJ, Rees JMB, Day FR, Perry JR, Ong KK. Dissecting causal pathways using Mendelian randomization with summarized genetic

- data: application to age at menarche and risk of breast cancer. *Genetics*. 2017;207:481–7.
16. Day FR, Thompson DJ, Helgason H, Chasman DI, Finucane H, Sulem P, Ruth KS, Whalen S, Sarkar AK, Albrecht E, et al. Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk. *Nat Genet*. 2017;49:834–41.
  17. Gill D, Del Greco MF, Rawson TM, Sivakumaran P, Brown A, Sheehan NA, Minelli C. Age at menarche and time spent in education: a Mendelian randomization study. *Behav Genet*. 2017;47:480–5.
  18. Gill D, Sheehan NA, Wielscher M, Shrine N, Amaral AFS, Thompson JR, Granell R, Leynaert B, Real FG, Hall IP, et al. Age at menarche and lung function: a Mendelian randomization study. *Eur J Epidemiol*. 2017;32:701–10.
  19. Zhang Q, Greenbaum J, Zhang WD, Sun CQ, Deng HW. Age at menarche and osteoporosis: a Mendelian randomization study. *Bone*. 2018;117:91–7.
  20. Trajanoska K, Morris JA, Oei L, Zheng HF, Evans DM, Kiel DP, Ohlsson C, Richards JB, Rivadeneira F. Assessment of the genetic and clinical determinants of fracture risk: genome wide association and mendelian randomisation study. *BMJ*. 2018;362:k3225.
  21. Lawn RB, Sallis H, Wootton RE, Taylor AE, Demange P, Fraser A, Penton-Voak IS, Munafo MR. The effects of age at menarche and first sexual intercourse on reproductive and behavioural outcomes: a Mendelian randomization study. *bioRxiv*. 2018; <https://doi.org/10.1101/423251>.
  22. Millard LAC, Davies NM, Tilling K, Gaunt TR, Davey SG. Searching for the causal effects of body mass index in over 300 000 participants in UK Biobank, using Mendelian randomization. *PLoS Genet*. 2019;15:e1007951.
  23. Millard LAC, Munafo MR, Tilling K, Wootton RE, Davey SG. MR-pheWAS with stratification and interaction: searching for the causal effects of smoking heaviness identified an effect on facial aging. *PLoS Genet*. 2019;15:e1008353.
  24. Allen NE, Sudlow C, Peakman T, Collins R. UK biobank data: come and get it. *Sci Transl Med*. 2014;6:224ed4.
  25. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015;12:e1001779.
  26. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562:203–9.
  27. Mitchell R, Hermani G, Dudding T, Corbin L, Harrison S, Paternoster L. UK biobank genetic data: MRC-IEU quality control, version 2. 2019. <https://data.bris.ac.uk/data/dataset/1ovauu5sxunp2cv8rcy88688v>. Accessed 1 Feb 2019.
  28. Hermani G, Tilling K, Davey SG. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet*. 2017;13:e1007081.
  29. Felix JF, Bradfield JP, Monnereau C, van der Valk RJ, Stergiakouli E, Chesi A, Gaillard R, Feenstra B, Thiering E, Kreiner-Moller E, et al. Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index. *Hum Mol Genet*. 2016;25:389–403.
  30. Yengo L, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN, Frayling TM, Hirschhorn J, Yang J, Visscher PM. Meta-analysis of genome-wide association studies for height and body mass index in approximately 700000 individuals of European ancestry. *Hum Mol Genet*. 2018;27:3641–9.
  31. Millard LAC, Davies NM, Gaunt TR, Davey Smith G, Tilling K. Software application profile: PHESANT: a tool for performing automated phenome scans in UK Biobank. *Int J Epidemiol*. 2017. <https://doi.org/10.1093/ije/dyx204>.
  32. Estrada K, Stykarsdottir U, Evangelou E, Hsu YH, Duncan EL, Ntzani EE, Oei L, Albarga OM, Amin N, Kemp JP, et al. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat Genet*. 2012;44:491–501.
  33. Moayyeri A, Hsu YH, Karasik D, Estrada K, Xiao SM, Nielson C, Srikanth P, Giroux S, Wilson SG, Zheng HF, et al. Genetic determinants of heel bone properties: genome-wide association meta-analysis and replication in the GEFOS/GENOMOS consortium. *Hum Mol Genet*. 2014;23:3054–68.
  34. Wyss AB, Sofer T, Lee MK, Terzikhan N, Nguyen JN, Lahousse L, Latourelle JC, Smith AV, Bartz TM, Feitosa MF, et al. Multiethnic meta-analysis identifies ancestry-specific and cross-ancestry loci for pulmonary function. *Nat Commun*. 2018;9:2976.
  35. Shrine N, Guyatt AL, Erzurumluoglu AM, Jackson VE, Hobbs BD, Melbourne C, Batini C, Fawcett KA, Song K, Sakornsakolpat P, et al. New genetic signals for lung function highlight pathways and pleiotropy, and chronic obstructive pulmonary disease associations across multiple ancestries. *BioRxiv*. 2018. <https://doi.org/10.1101/343293>.
  36. de Moor MH, van den Berg SM, Verweij KJ, Krueger RF, Luciano M, Arias Vasquez A, Matteson LK, Derringer J, Esko T, Amin N, et al. Meta-analysis of genome-wide association studies for neuroticism, and the polygenic association with major depressive disorder. *JAMA Psychiatry*. 2015;72:642–50.
  37. Fraser A, Macdonald-Wallis C, Tilling K, Boyd A, Golding J, Davey Smith G, Henderson J, Macleod J, Molloy L, Ness A, et al. Cohort profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol*. 2013;42:97–110.
  38. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;44:512–25.
  39. Zheng J, Baird D, Borges MC, Bowden J, Hemani G, Haycock P, Evans DM, Smith GD. Recent developments in Mendelian randomization studies. *Curr Epidemiol Rep*. 2017;4:330–45.
  40. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet*. 2018;50:693–8.
  41. Hemani G, Bowden J, Davey SG. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum Mol Genet*. 2018;27:R195–r208.
  42. Hartwig FP, Tilling K, Davey Smith G, Lawlor D, Borges MC. Bias in two-sample Mendelian randomization by using covariable-adjusted summary associations. *bioRxiv*. 2019; doi: <https://doi.org/10.1101/816363>.
  43. Staley JR, Burgess S. Semiparametric methods for estimation of a nonlinear exposure-outcome relationship using instrumental variables with application to Mendelian randomization. *Genet Epidemiol*. 2017;41:341–52.
  44. Munafo MR, Tilling K, Taylor AE, Evans DM, Davey SG. Collider scope: when selection bias can substantially influence observed associations. *Int J Epidemiol*. 2018;47:226–35.
  45. Taylor AE, Jones HJ, Sallis H, Euesden J, Stergiakouli E, Davies NM, Zammit S, Lawlor DA, Munafo MR, Davey Smith G, et al. Exploring the association of genetic factors with participation in the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol*. 2018;47:1207–16.
  46. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, Collins R, Allen NE. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am J Epidemiol*. 2017;186:1026–34.
  47. Morris DH, Jones ME, Schoemaker MJ, Ashworth A, Swerdlow AJ. Secular trends in age at menarche in women in the UK born 1908–93: results from the Breakthrough Generations Study. *Paediatr Perinat Epidemiol*. 2011;25:394–400.
  48. Taylor AE, Davies NM, Ware JJ, VanderWeele T, Smith GD, Munafo MR. Mendelian randomization in health research: using appropriate genetic variants and avoiding biased estimates. *Econ Hum Biol*. 2014;13:99–106.
  49. Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. *Int J Epidemiol*. 2019;48:713–27.
  50. Twisk JW, Staal BJ, Brinkman MN, Kemper HC, van Mechelen W. Tracking of lung function parameters and the longitudinal relationship with lifestyle. *Eur Respir J*. 1998;12:627–34.
  51. Almeida M, Laurent MR, Dubois V, Claessens F, O'Brien CA, Bouillon R, Vanderschueren D, Manolagas SC. Estrogens and androgens in skeletal physiology and pathophysiology. *Physiol Rev*. 2017;97:135–87.
  52. Chen FR, Rothman EF, Jaffee SR. Early puberty, friendship group characteristics, and dating abuse in US girls. *Pediatrics*. 2017;139. <https://doi.org/10.1542/peds.2016-2847>.
  53. Magnus MC, Anderson EL, Howe LD, Joinson CJ, Penton-Voak IS, Fraser A. Childhood psychosocial adversity and female reproductive timing: a cohort study of the ALSPAC mothers. *J Epidemiol Community Health*. 2018;72:34–40.
  54. Zabin LS, Emerson MR, Rowland DL. Childhood sexual abuse and early menarche: the direction of their relationship and its implications. *J Adolesc Health*. 2005;36:393–400.
  55. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, et al. A second generation human haplotype map of over 3.1 million SNPs. *Nature*. 2007;449:851–61.
  56. Servin B, Stephens M. Imputation-based analysis of association studies: candidate regions and quantitative traits. *PLoS Genet*. 2007;3:e114.



57. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet.* 2007;39:906–13.
58. Li Y, Willer C, Sanna S, Abecasis G. Genotype imputation. *Annu Rev Genomics Hum Genet.* 2009;10:387–406.
59. Estrada K, Abuseiris A, Grosveld FG, Uitterlinden AG, Knoch TA, Rivadeneira F. GRIMP: a web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. *Bioinformatics.* 2009;25:2750–2.
60. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet.* 2002; 30:97–101.
61. Aulchenko YS, Struchalin MV, van Duijn CM. ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinformatics.* 2010;11:134.
62. Devlin B, Roeder K, Wasserman L. Genomic control, a new approach to genetic-based association studies. *Theor Popul Biol.* 2001;60:155–66.
63. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics.* 2010;26:2190–1.
64. Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA. A map of human genome variation from population-scale sequencing. *Nature.* 2010;467:1061–73.
65. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, Kang HM, Fuchsberger C, Danecek P, Sharp K, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet.* 2016;48:1279–83.
66. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med.* 1999;159:179–87.
67. van den Berg SM, de Moor MH, McGue M, Pettersson E, Terracciano A, Verweij KJ, Amin N, Derringer J, Esko T, van Grootheest G, et al. Harmonization of neuroticism and extraversion phenotypes across inventories and cohorts in the genetics of personality consortium: an application of item response theory. *Behav Genet.* 2014;44:295–313.
68. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007;81:559–75.

## Publisher's Note

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

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

