eLife

Reviewed Preprint
v2 • March 21, 2025
Revised by authors

Reviewed Preprint
v1 • November 19, 2024

Medicine

Early menarche and childbirth accelerate aging-related outcomes and age-related diseases: Evidence for antagonistic pleiotropy in humans

Yifan Xiang, Vineeta Tanwar, Parminder Singh, Lizellen La Follette, Vikram Narayan, Pankaj Kapahi

The Buck Institute for Research on Aging, Novato, United States • Department of Urology, University of California, San Francisco, San Francisco, United States

eLife Assessment

This **important** study uses Mendelian Randomisation to show that early life phenotypes (i.e. onset of age at menarche and age at first birth) have an influence on a multitude of health outcomes later in life. The provided empirical evidence supporting the antagonistic pleiotropy theory is **solid**. However, some results seem improbable and need to be checked to make sure they are correct.

https://doi.org/10.7554/eLife.102447.2.sa3

Abstract

Aging can be understood as a consequence of the declining force of natural selection with age. Consistent with this, the antagonistic pleiotropy theory of aging proposes that aging arises from trade-offs that favor early growth and reproduction. However, evidence supporting antagonistic pleiotropy in humans remains limited. Using Mendelian Randomization (MR), we demonstrated that later ages of menarche or first childbirth were genetically associated with longer parental lifespan, decreased frailty index, slower epigenetic aging, later menopause, and reduced facial aging. Moreover, later menarche or first childbirth were also genetically associated with a lower risk of several age-related diseases, including late-onset Alzheimer's disease (LOAD), type 2 diabetes, heart disease, essential hypertension, and chronic obstructive pulmonary disease (COPD). We validated the associations between the age of menarche, childbirth, and the number of childbirths with several age-related outcomes in the UK Biobank by conducting regression analysis of nearly 200,000 subjects. Our results demonstrated that menarche before the age 11 and childbirth before 21 significantly accelerated the risk of several diseases, and almost doubled the risk for diabetes, heart failure, and quadrupled the risk of obesity, supporting the antagonistic pleiotropy theory. We identified 126 significant single nucleotide polymorphisms (SNPs) that influenced age-related outcomes, some of which were involved in known longevity pathways, including IGF1, growth hormone, AMPK, and mTOR signaling. Our study also identified higher BMI as a mediating factor in causing the increased risk of certain diseases, such as type 2 diabetes and heart failure, in women with early menarche or early pregnancy,

emphasizing the importance of the thrifty gene hypothesis in explaining in part the mechanisms behind antagonistic pleiotropy. Our study highlights the complex relationship between genetic legacies and modern diseases, emphasizing the need for gender-sensitive healthcare strategies that consider the unique connections between female reproductive health and aging.

Introduction

Given the decline in the force of natural selection with age, genes that influence aging and age-related diseases are likely to be selected for their influence on early life events [1]. The theory of antagonistic pleiotropy suggests that natural selection involves inherent trade-offs, where genes that are advantageous for early survival and reproduction may have detrimental effects later in life, contributing to the aging process and the development of age-related diseases. [1–3]. Though evidence for antagonistic pleiotropy has been observed in invertebrate models, evidence for causal relationships in mammals, especially humans, is largely lacking [4]. The timing of reproductive events, such as menarche and childbirth, has long been recognized as a crucial aspect of human life history and evolution. However, accumulating evidence suggests that these reproductive milestones may have far-reaching implications [5,6].

Considering reproductive events are partly regulated by genetic factors that can manifest the physiological outcome later in life, we hypothesized that earlier or later onset of menarche and first childbirth could reflect broader genetic influences on longevity and disease susceptibility, serving as proxies for biological aging processes (fig. S1 and fig. S2). Mendelian randomization (MR) is a term that applies to the use of genetic variation to address causal questions about how modifiable exposures influence different outcomes [7]. Two-sample and two-step MR analyses were adopted in our research to explore the genetic causal associations using the inverse variance weighted (IVW) model. We leveraged recent studies that have identified SNPs associated with early menarche or birth in humans, to examine their role in antagonistic pleiotropy. After preprocessing, there were 209 and 33 single nucleotide polymorphisms (SNPs) included in MR analyses for the two exposures, age at menarche [8] and age at first birth [9], respectively.

Later menarche was genetically associated with later aging outcomes

Since parental ages at death may reflect the genetic heritability of lifespan, parental ages at death were used as general aging outcomes [10]. Compared to early menarche, later age at menarche was significantly associated with later parental ages at death for mothers or fathers (Beta=2.27E-02, P=1.84×10^{-4}; Beta=1.97E-02, P=7.88×10^{-4} for father and mother's ages at death respectively) (**Fig. 1A** and **Table S1-S3**). Aging is the biggest risk factor for frailty and several age-related diseases. Later age at menarche was associated with a lower frailty index, which was calculated using a set of 49 self-reported questionnaire items on traits covering health, presence of diseases and disabilities, and mental well-being [11] (Beta=-2.36E-02, P=5.75×10^{-3}). After excluding the instrumental variables (IVs) associated with body mass index (BMI), the significant association between later menarche and outcomes remained based on the weighted median model (WMM) for frailty index (Beta=-2.88E-02, P=0.0202) and based on IVW for parental ages at death (Beta=1.47E-02, P=2.46×10^{-2}; Beta=1.24E-02, P=3.87×10^{-2}). Next, we examined the associations between the age of menarche and organ aging or age-related diseases (**Fig. 1B**). Later age of menarche was associated with later menopause onset (Beta=1.16E-01, P=2.11×10^{-2} with WMM), lower risks of late-onset Alzheimer's disease (LOAD) (OR=0.897 (0.810–0.994), P=3.88×10^{-2}), chronic heart failure (CHF) (OR=0.953 (0.913–0.995), P=$2.\ 80\times10^{-2}$), essential hypertension (OR=0.990 (0.986–0.993), P=1.99×10^{-8}), facial aging (Beta=-1.76E-02, P=2.16×10^{-9}), and early onset chronic obstructive pulmonary disease (COPD) (OR=0.838 (0.755–0.931), P=9.50×10^{-4}) and mildly higher risk of osteoporosis (OR=1.009 (1.006–1.013), P=1.95×10^{-7}). These associations were still significant

after excluding SNPs associated with BMI, except for CHF. One potential mechanism of antagonistic pleiotropy is accelerated cell growth and division; therefore, we examined the relationship between menarche and certain cancers. Later age at menarche was associated with lower risks of breast cancer (OR=0.859 (0.764–0.966), P=1.53×10^{-2}) and endometrial cancer (OR=0.896 (0.840–0.956), P=9.02×10^{-4}) compared with early menarche (**Fig. 1B**).

Later age at first birth was genetically associated with later aging outcomes

Next, we examined the associations between the age of childbirth and age-related outcomes. Compared to early first birth, later age at first birth was associated with lower frailty index (Beta=-6.48E-02, P=5.93×10^{-11}) and GrimAge, a measure of epigenetic aging (Beta=2.67E-01, P=4.37×10^{-2}), and older parental ages at death (Beta=3.43E-02, P=6.88×10^{-5} for father's age at death; Beta=3.37E-02, P=1.78×10^{-4} for mother's age at death) (**Fig. 1A**). Similar to results with the age of menarche, later age at first birth was associated with later menopause onset (Beta=2.67E-01, P=3.08×10^{-4}), lower risks of type 2 diabetes (OR=0.890 (0.847– 0.935), P=3.65×10^{-6}), CHF (OR=0.926 (0.866–0.991), P=2.56×10^{-2}), essential hypertension (OR=0.989 (0.984–0.995), P=1.54×10^{-4}), and gastrointestinal or abdominal disease (GAD) (OR=0.991 (0.985–0.996), P=1.29×10^{-3}). Most significant associations remained after the BMI-related SNPs were excluded, except for CHF (**Fig. 1C** and **Table S3**). Furthermore, later age at first birth was also significantly associated with a lower risk of cervical cancer (OR=0.999 (0.998–1.000), P=1.52×10^{-2}) but not breast and endometrial cancers (**Fig. 1C**).

BMI is an important mediator in significant associations

As BMI is an important modulator of aging [12], we examined its role in explaining these associations. Based on the two-sample MR analyses, significant associations between early menarche and CHF, and between early first birth and CHF and cervical cancer, were not observed after excluding SNPs related to BMI (**Fig. 1**). To estimate the effect of BMI as the mediator, we further conducted two-step MR. Exposures of age of menarche and age at first birth, and outcomes of frailty index, type 2 diabetes, and CHF, were included in the two-step MR analysis. Later ages of menarche and first birth were associated with lower BMI (Beta=-5.36E-02, P=2.73×10^{-13}; Beta=-4.49E-02, P=6.85×10^{-5}). Higher BMI was associated with higher risks of type 2 diabetes (OR=2.391 (2.243–2.550), P=4.65×10^{-156}), CHF (OR=1.690 (1.570–1.819), P=1.67×10^{-44}), and cervical cancer (OR=1.001 (1.000–1.002), P=3.16×10^{-2}). Significant pleiotropy was detected in the association between BMI and frailty index. Two-step MR showed that BMI significantly mediated the associations between menarche and, type 2 diabetes or CHF (**Table S4**). BMI also significantly mediated the associations between first birth and, type 2 diabetes, CHF, or cervical cancer (**Table S4**). As both early menarche and first birth were associated with higher BMI, the BMI had a significant effect to explain the associations partially. Results of colocalization analysis showed 5 SNPs associated with both the exposures and outcomes (**Table S5**). The genetic correlation results analyzed with LDSC method are listed in **Table S6**.

Mechanisms explaining antagonistic pleiotropy

We identified 126 significant SNPs in the genes from the MR analysis for their association between age of menarche or first birth with 15 aging and age-related disease outcomes. These SNPs were then listed according to (i) the number of study outcomes they were associated with and (ii) their frequency of occurrence in the European and global population. We chose to depict only the SNPs that were associated with greater than three outcomes as shown in **Fig. 2**. The SNP rs2003476 in the *CRTC1* gene was found to be associated with seven aging and age-related disease outcomes (frailty index, LOAD, CHF, father's age at death, mother's age at death, breast, and endometrial cancer). Previous studies have indicated that CRTC1 transcription domain is linked with extending the lifespan in C. elegans [13] and a decrease in CRTC1 levels was found to be associated with

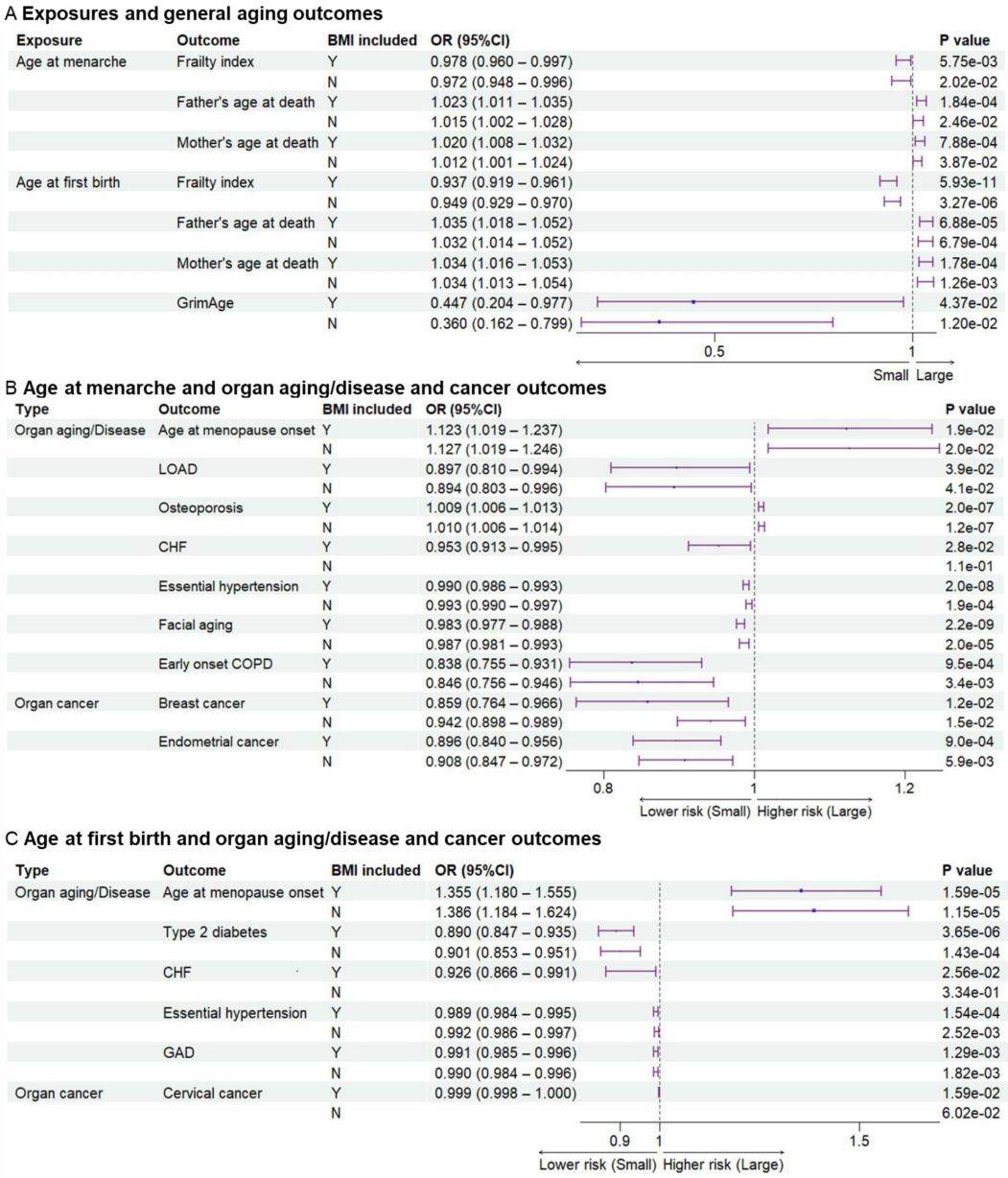

Figure 1.

Genetic associations between exposures and outcomes.

(**A**), genetic associations between exposures and general aging. Later age at menarche was associated with a lower frailty index and higher parental ages at death. Later age at first birth was associated with lower frailty index and GrimAge and higher parental ages at death. (**B**), genetic associations between age at menarche and outcomes of organ aging and diseases as well as organ cancer. Later age at menarche was associated with later age at menopause onset and lower risks of LOAD, CHF, essential hypertension, facial aging, early onset COPD, breast cancer, and endometrial cancer. (**C**), genetic associations between age at first birth and outcomes of organ aging and diseases as well as organ cancer. Later age at first birth was associated with later age at menopause onset and lower risks of type 2 diabetes, CHF, essential hypertension, gastrointestinal or abdominal disease, and cervical cancer. The significant associations were not detected after BMI-related SNPs excluded for outcomes of CHF and cervical cancer. BMI included, BMI-related SNPs included; LOAD, late-onset Alzheimer's disease; CHF, chronic heart failure; COPD, chronic obstructive pulmonary; GAD, gastrointestinal or abdominal; BMI, body mass index.

human AD [14]. Two other genes involved in the glutathione metabolism pathway, *CHAC1* and *GGT7*, were also found to be associated with four different aging outcomes (*CHAC1* was associated with frailty index, diabetes, COPD, and BMI; *GGT7* was associated with father's age at death, hypertension, breast cancer and BMI) (**Fig. 2**). Importantly, the SNP in *CHAC1* is a missense variant, suggesting its impact on protein translation and potential contribution to disease biology. CHAC1 has recently been implicated in age-related macular degeneration [15] and muscle wasting [16]. Posttranslational protein-lipid modification processes are known to contribute to aging and age-related diseases [17]. N-myristoylation is one such process that is catalyzed by NMT (N-myristoyltransferase) [18], another intriguing gene candidate identified in our analysis. We found an association of *NMT1* with three aging outcomes: T2D, CHF, and endometrial cancer. Various myristoylated proteins have been implicated in diverse intracellular signaling pathways [19] including AMP-activated protein kinase (AMPK) [20,21]. It has been shown that NMT1 exerts synovial tissue-protective functions by promoting the recruitment of AMPK to lysosomes and inhibiting mTORC1 signaling [20]. In addition, drugs targeting NMTs (NMT1 and NMT2) have been suggested to be potent senolytics [22] and a target for treating or preventing various diseases such as cancer [23,24] and heart failure [25]. Furthermore, modulating the evolutionarily conserved N-myristoyl transferase NMT1 was shown to extend lifespan in yeast [26]. The full table of SNPs and their population frequency and association with aging outcomes is provided in Table S7.

Canonical signaling pathways analysis

A total of 498 canonical signaling pathways related to 126 SNPs/gene outcomes were identified based on the Ingenuity Pathway Knowledge Base (IPKB). After ranking the identified canonical signaling pathways according to their adjusted P-values, the top 25 canonical signaling pathways with a p-value < 10^{-2} enriched by SNPs/genes involved in the age of menarche and first birth are represented in **Fig. 3A**. The top enriched canonical signaling pathways fell into these broader categories: (1) Developmental and cellular signaling pathways such as BMP, IGF-1, and growth hormone signaling; (2) Neuronal signaling and neurological disorders that include amyloid processing, activation of NMDA receptors and postsynaptic events, and glioma signaling; (3) Metabolic signaling pathways such as insulin receptor, mTOR, and leptin signaling; (4) Immunological and inflammatory signaling pathways, which include lymphotoxin β receptor and TREM1 signaling; (5) Cardiovascular and muscular signaling pathways including the role of NFAT in cardiac hypertrophy. In addition to canonical pathways, SNPs/genes were shown to be further associated with diseases and functions. IPA showed a total of 80 diseases and functions associated with age at menarche and first birth. These associated diseases and functions were rated according to their adjusted P-values, and the top 20 enriched categories with a P-value < 10^{-5} were selected and depicted in **Fig. 3B**.

Gene network analysis and drug interactions

Besides pathway and disease associations, the 126 SNPs/genes were further connected to identify 11 functional networks using IPA. These IPA networks were ranked based on their consistency scores. The top 5 networks included a range of 13-17 genes with scores above 25, indicating robust regulatory analysis [27] (Table S8). The top networks were mainly connected to the following functions: Cellular interaction/signaling/development (**Fig. 3C**), cell death and survival, cancer, gastrointestinal, reproductive (**Fig. 3C**), neurological (**Fig. 3D**), and cardiovascular system, development, function, diseases. These results support the notion that dynamic cellular interactions and the development of various vital organ systems influenced by age at menarche and first birth also influence age-related diseases. The two top IPA networks were further outlined in **Fig. 3C** and **3D**. To further explore the drugs targeting genes associated with age at menarche and first birth, we conducted an analysis using ChEMBL[28] and DrugBank databases[29]. This examination revealed that a total of 11 FDA-approved drugs target the identified genes (*PRKAG2, SCN2A, DPYD, MC3R, AKT, MAPK, KCNK9, TACR3, HTR4, KCNB2, RXRG*) (Table S9).

Exposure	Gene	SNP	Eur SNP Frequency	Global SNP Frequency
Menarche	*CRTC1*	rs2003476[#]	C=0.40991	C=0.46903
Menarche	*DZANK1*	rs6045400[#]	G=0.78072	G=0.74726
Menarche	*TRPM3*	rs1360201[#]	T=0.48208	T=0.45743
Menarche	*NIM1K*	rs35375873[$]	C=0.12425	C=0.10196
Menarche	*USP35*	rs2297605[#]	A=0.42044	A=0.40012
Menarche	*CHAC1*	rs112036939[*]	T=0.02178	T=0.019887
Menarche	*HTR4*	rs9325098[†]	T=0.24744	T=0.24372
Menarche	*INO80*	Rs561821[&]	T=0.37147	T=0.37553
Menarche	*GGT7*	rs7273470[#]	C=0.22064	C=0.21895
Menarche	*HSPE1P19*	rs9681954[&]	C=0.60064	C=0.54979
Menarche	*MAPK3*	rs57149692[@]	C=0.40410	C=0.43346
Menarche	*C2orf74*	rs17006627[#]	C=0.1929	C=0.19119
Menarche	*KCTD9*	rs17053711[#]	A=0.26043	A=0.23871
Menarche	*C11orf58*	rs117947332[#]	C=0.04893	C=0.04341
Menarche	*DLGAP1*	rs12607903[#]	T=0.722209	T=0.709937
Menarche	*TKT*	rs141829833[&]	T=0.02352	T=0.01949
Menarche	*AKT3*	rs1973284[#]	G=0.67100	G=0.57113
Menarche	*EEFSEC*	rs3849531[#]	T=0.28151	T=0.35538
Menarche	*KCNK9*	rs62520182[#]	T=0.15894	T=0.12578
Menarche	*TRIM41*	rs7727787[#]	A=0.128488	A=0.121687
Menarche	*NR5A1*	rs2297605[#]	A=0.42044	A=0.40012
Menarche	*JADE2*	rs13179411[#]	T=0.10225	T=0.08157
Menarche	*BRWD1*	rs4818008[#]	A=0.38583	A=0.36226
Menarche	*BANF2*	rs852033[&]	C=0.7050	C=0.6298
Menarche	*TTC26*	rs113013060[#]	A=0.02158	A=0.01676
Menarche	*SULF2*	rs6066454[#]	T=0.35251	T=0.31344
Menarche	*NMT1*	rs9913787[#]	A=0.28846	A=0.26612
Menarche	*BMPR2*	rs190286519[#]	T=0.06894	T=0.05851
AFB	*TRAF3*	rs72704712[#]	G=0.16533	G=0.14575
Menarche	*PAPPA2*	rs10913213[#]	G=0.33651	G=0.32479
Menarche	*GREB1*	rs113467400[#]	T=0.08428	T=0.09952
Menarche	*ZC3HC1*	rs11556924[*]	T=0.384007	T=0.350133
Menarche	*ASCC3*	rs240766[&]	A=0.551850	A=0.571029
Menarche	*RORA*	rs3743266[#]	C=0.344786	C=0.333481
Menarche	*CCDC190*	rs2343506[&]	C=0.55437	C=0.59071
Menarche	*UNC79*	rs10150677[#]	G=0.6142	G=0.5898
Menarche	*ZBTB46*	rs6010651[#]	C=0.36875	C=0.39884
AFB	*NOL4L*	rs293566[#]	C=0.333621	C=0.370718
AFB	*SORCS3*	rs12250380[#]	G=0.51992	G=0.51992
AFB	*TEX14*	rs7359501[#]	T=0.3046	T=0.26923

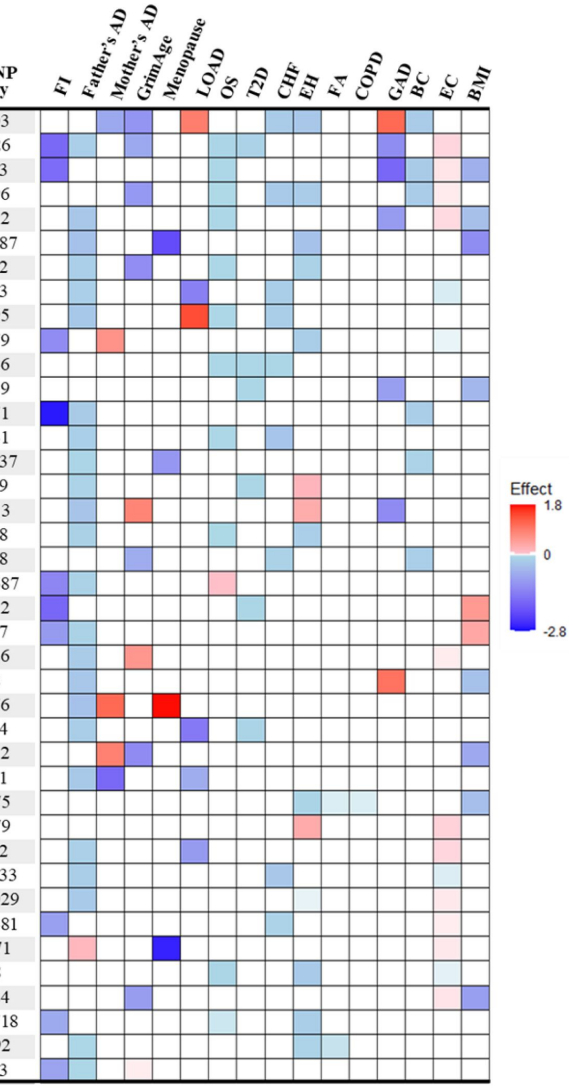

Figure 2.

List and heatmap of SNPs/genes from post-MR analysis showing an association between age at menarche and first birth with three or more aging outcomes.

Genomic region location for each variant is depicted by different characters: intron variant[#]; 2KB upstream variant[$]; Missense variant*; Intergenic variant[&]; Regulatory region variant[@]. 3′ UTR variant[†]. Heatmap is representative of an association of each gene/SNP with different aging outcomes. Red bar represents a harmful association (-beta for Father's DA, Mother's DA, and menopause; +beta for other outcomes); Blue bar represents a beneficial association (+beta for Father's DA, Mother's DA, and menopause; -beta for other outcomes); White bar represents no association. AFB; age at first birth; FI, frailty index; Father's DA, father's age at death; Mother's DA, mother's age at death; Menopause, age at menopause; LOAD, late-onset Alzheimer's disease; OS, osteoporosis; T2D, type II diabetes; CHF, chronic heart failure; EH, essential hypertension; FA, facial aging; COPD, chronic obstructive pulmonary disease; GAD, gastrointestinal or abdominal disease; BC, breast cancer; EC, endometrial cancer; BMI, body mass index.

A

Canonical Signaling Pathways	-log$_{10}$(*p*-value)	Fraction of genes that are involved in the pathway	Genes belonging to the pathway
Transcriptional Regulatory Network in Embryonic Stem Cells	2.31	0.0427	AKT3,BMPR2,IGF1R,INHBA,MAPK1,MAPK3,SMARCAD1
Leptin Signaling in Obesity	2.31	0.0658	AKT3,LEPR,MAPK1,MAPK3,PRKAG2
Activation of NMDA receptors and postsynaptic events	2.31	0.0595	CAMK4,DLG2,MAPK1,MAPK3,PRKAG2
Cachexia Signaling Pathway	2.31	0.0245	AKT3,BMPR2,DDIT4,IGF1R,INHBA,MAPK1,MAPK3,PRKAG2,PTH2R
BMP signaling pathway	2.31	0.0549	BMPR2,CAMK4,MAPK1,MAPK3,PRKAG2
Signaling by Activin	2.31	0.176	INHBA,MAPK1,MAPK3
PFKFB4 Signaling Pathway	2.29	0.0816	MAPK1,MAPK3,PRKAG2,TKT
Amyloid Processing	2.28	0.0784	AKT3,MAPK1,MAPK3,PRKAG2
Intrinsic Pathway for Apoptosis	2.28	0.0727	AKT3,MAPK1,MAPK3,NMT1
Lymphotoxin β Receptor Signaling	2.28	0.0727	AKT3,MAPK1,MAPK3,TRAF3
IGF-1 Signaling	2.28	0.0476	AKT3,IGF1R,MAPK1,MAPK3,PRKAG2
Docosahexaenoic Acid (DHA) Signaling	2.23	0.028	AKT3,MAPK1,MAPK3,PRKAG2,RXRG,STXBP4,SYT14
IL-22 Signaling	2.18	0.125	AKT3,MAPK1,MAPK3
Glioma Signaling	2.03	0.04	AKT3,CAMK4,IGF1R,MAPK1,MAPK3
Growth Hormone Signaling	2.02	0.0563	IGF1R,MAPK1,MAPK3,RPS6KC1
CLEAR Signaling Pathway	2.02	0.0246	AKT3,BMPR2,DDIT4,IGF1R,MAPK1,MAPK3,PRKAG2
Melatonin Signaling	2.02	0.0556	MAPK1,MAPK3,PRKAG2,RORA
Extra-nuclear estrogen signaling	2.02	0.0548	AKT3,IGF1R,MAPK1,MAPK3
TREM1 Signaling	1.98	0.0519	AKT3,MAPK1,MAPK3,NLRC3
mTOR Signaling	1.98	0.028	AKT3,DDIT4,MAPK1,MAPK3,PRKAG2,RPS6KC1
Insulin Receptor Signaling	1.98	0.0357	AKT3,MAPK1,MAPK3,PRKAG2,STXBP4
Estrogen-Dependent Breast Cancer Signaling	1.94	0.0488	AKT3,IGF1R,MAPK1,MAPK3
Transcriptional Regulation by NPAS4	1.94	0.0857	BMAL1,MAPK1,MAPK3
Role of NFAT in Cardiac Hypertrophy	1.94	0.0268	AKT3,CAMK4,IGF1R,MAPK1,MAPK3,PRKAG2
RHO GTPases Activate WASPs and WAVEs	1.94	0.0833	CYFIP2,MAPK1,MAPK3
Dilated Cardiomyopathy Signaling Pathway	1.94	0.0333	AKT3,CAMK4,MAPK1,MAPK3,PRKAG2

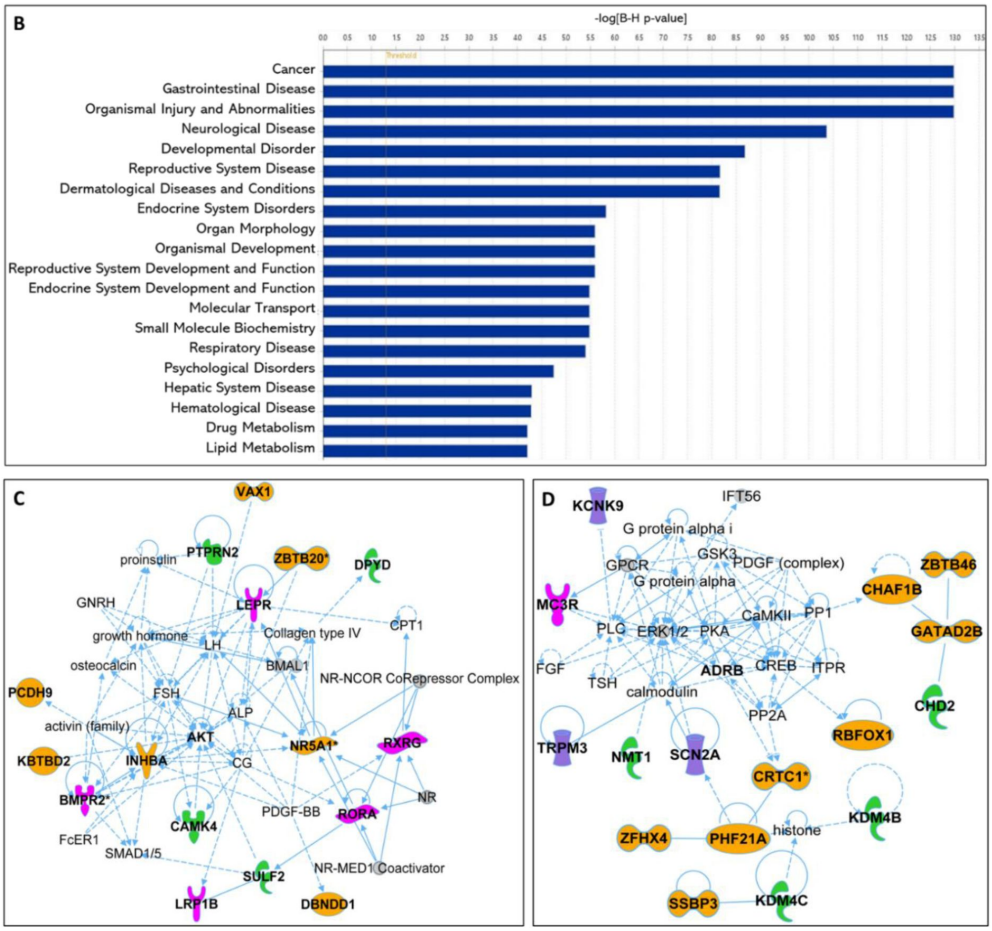

Figure 3.

Ingenuity Pathway Analysis (IPA). SNPs/gene outcomes from MR analysis were subjected to IPA.

(**A**) Canonical signaling pathways in age at menarche and first birth. The adjusted P-values of the top 25 signaling pathways are listed. (**B**) Disease and functions in age at first birth and menarche. The adjusted P-value of the top 20 significantly involved diseases and functions are listed. (**C**) IPA network 1 and (**D**) IPA network 2. The genes from our post-MR analysis are in bold. Solid lines indicate direct connections, while dotted lines indicate indirect connections (circular arrows mean influence itself). Color coding: pink-receptors, green-enzymes, purple-channels, orange-other proteins.

Population validation for associations between age at menarche, or first live birth, or number of births with age-related outcomes

Next, we validated the genetic associations in the population cohort of the UK Biobank. 264,335, 184,481, and 259,758 participants were included for independent variables of age at menarche, age at first live birth, and number of births, respectively (Table S10). Ages at menarche were divided into 5 age groups, <11y, 11-12y, 13-14y, 15-16y, and >16y. The probability of living to 80 years old, having menopause ≥50 years old, risk of diabetes, high blood pressure (HBP), heart failure, COPD, breast cancer, endometrial cancer, and obesity were compared among the 5 age groups of menarche. Each of the outcomes was compared between the highest and lowest education levels and 3 BMI categories (18.5-24.9, 25-29.9, and ≥30) among the 5 age groups (**Fig. 4A-I**). The ORs for each age group were calculated based on the coefficients (β) in logistic regression after including confounders of smoking, alcohol, education, and BMI. Compared to the group with menarche under 11 years old, which generally had the highest risk, females with menarche at 13-14y had the highest probability of living to 80 years old (OR=1.412, P<0.001) and having menopause ≥50 years old (OR=1.146, P<0.001), and had the lowest risk of diabetes (OR=0.773, P<0.001), heart failure (OR=0.835, P<0.001), and COPD (OR=0.728, P<0.001). The group with menarche at 15-16y had the lowest risks of high blood pressure (HBP) (OR=0.793, P<0.001) and obesity (OR=0.337, P<0.001). Females with menarche >16y had the lowest risks of breast cancer (OR=0.832, P<0.05) and endometrial cancer (OR=0.641, P<0.05) compared to females with menarche <11y (**Table S11**).

Ages at first birth were divided into 5 age groups, <21y, 21-25y, 26-30y, 31-35y, and >35y. The probability of living to 80 years old, having menopause ≥50 years old, risk of diabetes, AD, HBP, heart failure, cerebral infarction, digestive system disease, cervical cancer, and obesity were compared among the 5 age groups of first live birth. Compared to the group with first live birth under 21 years old, which generally had the highest risks, females with first live birth at 21-25y had the highest probability to live to 80 years old (OR=1.667, P<0.001), and females with first live birth at 26-30y had the highest probability to have menopause ≥50 years old (OR=1.184, P<0.001) and lowest risk of cervical cancer (OR=0.587, P<0.01).

Females with first live birth at 31-35y had the lowest risk of diabetes (OR=0.612, P<0.001). Females with first live birth >35y had the lowest risk of having AD (OR=0.474, P<0.01), HBP (OR=0.587, P<0.001), heart failure (OR=0.446, P<0.001), cerebral infarction (OR=0.485, P<0.001), digestive system disease (OR=0.569, P<0.001), and obesity (OR=0.416, P<0.001) (**Fig. 4J-S**).

Females with the number of births from 0 to 4 were included. Compared to no birth, females with 3 births had the highest probability of living to 80 years old (OR=1.818, P<0.001) and having menopause ≥50 years old (OR=1.562, P<0.001). Females with 4 births had the highest risk of diabetes (OR=1.399, P<0.001), AD (OR=1.676, P<0.001), HBP (OR=1.274, P<0.001), heart failure (OR=1.439, P<0.001), cerebral infarction (OR=1.470, P<0.001), digestive system disease (OR=1.370, P<0.001), and obesity (OR=1.327, P<0.001) but lowest risk of breast cancer (OR=0.869, P<0.001) (**Fig. 4T-AC**). The significant estimated effect (β) of each group compared to the youngest age groups of menarche and first live birth and 0 birth group and the ability to explain the variability of outcomes in regression analysis are exhibited in **Table S11**.

To estimate the combined effect of age at menarche and first live birth on diabetes, HBP, heart failure, and BMI≥30, participants were divided into 25 groups according to the menarche and first live birth ages.

Females with menarche<11y and first live birth<21y had higher risks of diabetes (OR=1.881, P<0.001), HBP (OR= 1.436, P<0.05), heart failure (OR= 1.776, P<0.01), and BMI≥30 (OR= 4.417, P<0.001) compared to females with menarche of 13-14y and first live birth of 26-30y (**Table S12**

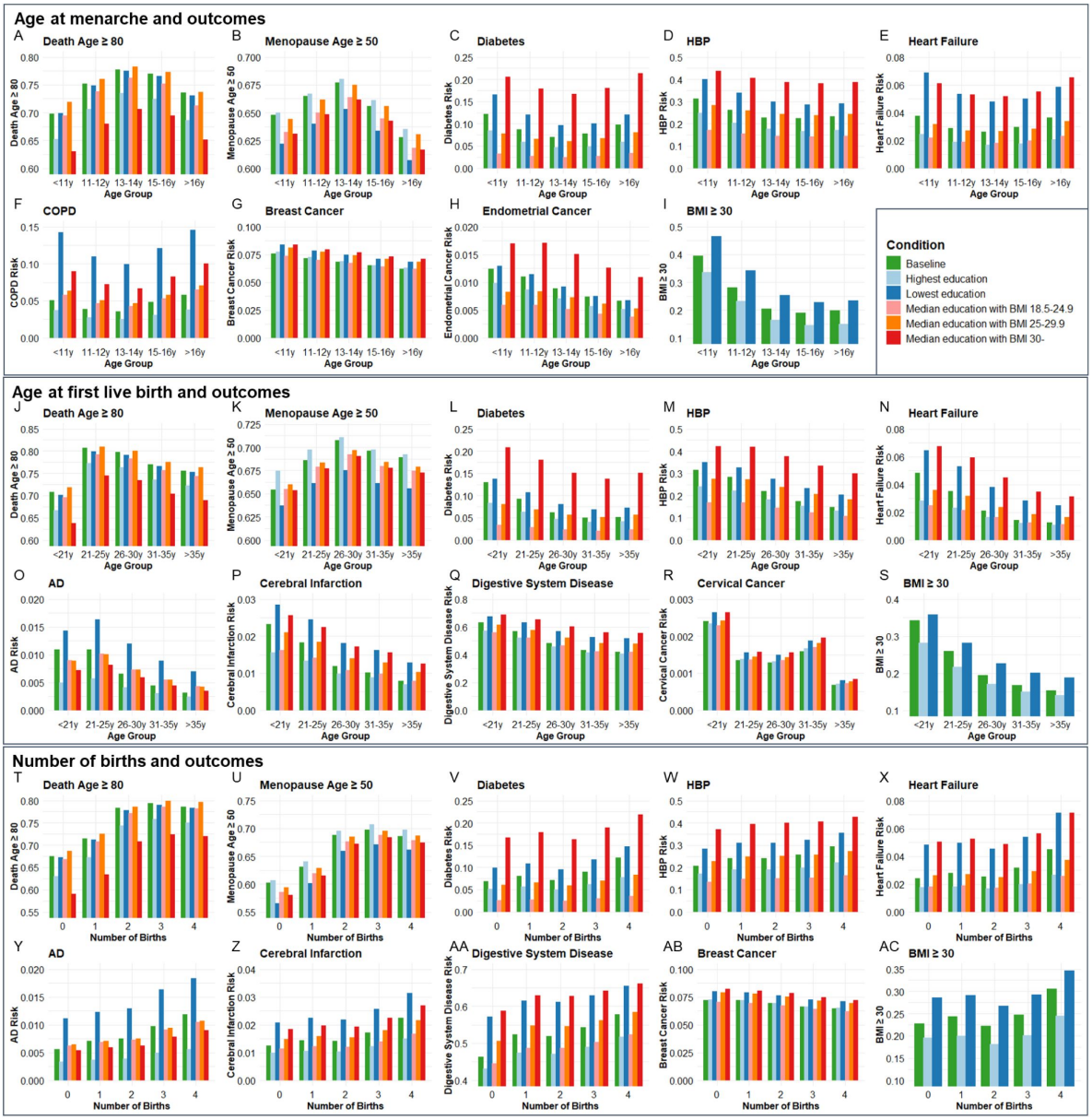

Figure 4.

The distributions of outcome age and risk according to age group of menarche and fist live birth as well as number of births before and after correcting confounders.

Condition Baseline, no confounder was corrected. Conditions Highest education and Lowest education, confounders of education, smoking, and drinking were corrected, and bars were painted based on most smoking and drinking situations with the highest and lowest education levels. Conditions Median education with BMI 18.5-24.9, Median education with BMI 25-29.9, and Median education with BMI 30-, confounders of education, smoking, drinking, and BMI were corrected, and bars were painted based on median education and most smoking and drinking situation with 3 BMI categories. BMI, body mass index; HBP, high blood pressure; COPD, chronic obstructive pulmonary disease; AD, Alzheimer's disease.

Females with menarche<11y and first live birth<21y had higher risks of diabetes (OR=1.881, P<0.001), HBP (OR= 1.436, P<0.05), heart failure (OR= 1.776, P<0.01), and BMI≥30 (OR= 4.417, P<0.001) compared to females with menarche of 13-14y and first live birth of 26-30y (**Table S12** and **fig. S3**).

Discussion

The timing of reproductive milestones such as the age of menarche and age at first birth is partly governed by genetic factors, which may exert lasting impacts on an individual's health trajectory. Our study finds support for antagonistic pleiotropy by demonstrating that genetic variants that drive early menarche or early pregnancy accelerate several aging outcomes, including parental age, frailty index, and several age-related diseases such as diabetes and dementia. These results were validated in a cohort from the UK biobank, showing that women with early menarche, or early pregnancy, were associated with accelerated aging outcomes and increased risk of several age-related diseases.

Consistent with a recent study suggesting the relation of early first birth with a higher likelihood of frailty [30], our MR analysis provided a strong association between early age of first birth and GrimAge, a DNA methylation based marker to predict epigenetic age. Our findings align with the observed negative genetic correlation between reproductive traits and lifespan that individuals with higher polygenetic scores for reproduction have lower survivorships to age 76 [31]. Prior studies linking the female reproductive factors with aging are either limited to results obtained from observational studies [32,33] or are limited to few outcomes such as brain disorders [34,35]. Besides, emerging evidence suggests that early puberty in males is linked to adverse health outcomes, such as an increased risk of cardiovascular disease, type 2 diabetes, and hypertension in later life [36]. A Mendelian randomization study also reported a genetic association between the timing of male puberty and reduced lifespan [37]. These findings support the hypothesis that genetic variants associated with delayed reproductive timing in males might similarly confer health benefits or improved longevity, akin to the patterns observed in females. This would suggest that similar mechanisms of antagonistic pleiotropy could operate in males as well.

Using MR, we identified 126 SNPs associated with early menarche or first birth that significantly influence age-related outcomes. Among the top 25 enriched canonical signaling pathways, those involved in developmental and cellular processes are particularly significant in determining the timing of menarche and first childbirth, with nearly half of these pathways (12 out of 25) falling into this category. These developmental and cellular signaling pathways work together to regulate continuous growth and involution from puberty through menopause [38]. Interestingly, known longevity pathways such as insulin-like growth factor 1 (IGF1), growth hormone (GH) signaling, melatonin signaling, and bone morphogenetic protein (BMP) signaling, seem to play a crucial role in regulating the timing of these reproductive events. IGF-1 signaling promotes the growth and development of reproductive organs and tissues [39–41], but is also a conserved modulator of longevity [42–45]. Melatonin, known for regulating circadian rhythms, influences the timing of menarche through its effects on the hypothalamic-pituitary-gonadal axis [46] and is postulated to modulate oxidative, inflammatory, and autophagy states [47]. Furthermore, different processes in early life such as cell proliferation or differentiation [48], ovarian function, and follicular development [49–51] are dependent on BMP signaling. However, increased BMP signaling also contributes significantly to AD pathology [52] and impairments in neurogenesis and cognitive decline associated with aging [53]. Aberrant BMP signaling is also shown to be associated with age-related metabolic and cardiovascular diseases [54,55]. Additionally, growth hormone signaling is not only vital for development during childhood and puberty [56] but also impacts longevity [57]. Besides signaling pathways, genes in age at menarche and first childbirth were connected in gene-gene interaction networks. In IPA network 1, most genes were connected with

either *AKT* or genes encoding for follicle-stimulating hormone (*FSH*), luteinizing hormone (*LH*), leptin receptor (*LEPR*), and inhibin subunit beta A (*INHBA*). AKT signaling is implicated in oocyte maturation and embryonic development [58] and mediates pregnancy-induced cardiac adaptive responses [59]. AKT signaling also mediates age-related disease pathologies, such as AD [60]. Major ovarian functions are controlled by FSH (follicular growth, cellular proliferation, and estrogen production) and LH (oocyte maturation, ovulation, and terminal differentiation of follicles) which in turn are modulated by other ovarian factors such as INHBA (a subunit of activin and inhibin) [61,62]. Endocrine alterations at advanced aging or reproductive aging [63] are marked by changes in FSH and LH due to altered feedback resulting from the ovarian decline in sex steroids, inhibin A, and inhibin B production [64–66]. In IPA network 2, the hub genes are connected mostly with extracellular signal-regulated kinases (ERK), Similar to AKT, ERK is also a member of the mitogen-activated kinase family (MAPK), a highly conserved signaling pathway that plays a vital role in transducing extracellular signals into a wide range of cellular responses [67,68], maintaining tissue integrity during aging [69], and mediating age-related disease pathology [60,70]. These studies underscore the importance of considering that dysregulation or abnormal activation of any of these pathways could contribute to the onset of late-life diseases consistent with antagonistic pleiotropy in humans [2,71].

The thrifty gene hypothesis provides a compelling framework to explain antagonistic pleiotropy, particularly in the context of modern health challenges. The thrifty gene hypothesis suggests that genes favoring efficient energy storage were advantageous in historical periods of food scarcity, helping individuals survive through famines. However, in today's environment of abundant food, the same genes contribute to obesity and metabolic diseases like diabetes. Similarly, in the context of antagonistic pleiotropy, we hypothesize that genes that enhance early-life reproductive success favor efficient energy storage to support reproductive health. This is supported by our observation that early pregnancy is associated with increasing BMI, a recognized key factor in systemic aging [72,73]. In support of this, we found a lack of significant association between early menarche or age at first birth for CHF and cervical cancer after removal of BMI as a confounder in MR analysis. Based on regression analysis at the population level (**Fig. 4**), higher BMI was associated with higher risks of cardiovascular diseases and diabetes. Thus, an increase in BMI is one of the factors that explains the increased risk of age-related disease seen due to the exposure of early pregnancy or the prevalence of genes that enhance early reproductive success. These results are also consistent with calorie restriction being a robust way to extend healthspan and lifespan in many species [74].

Several limitations of our study need to be considered. Firstly, all the associations are explored at the genetic level with MR and population level with the UK Biobank. However, to confirm causal relationships, further validation through *in vitro* and *in vivo* research is essential. Secondly, though we excluded confounder-related SNPs and addressed potential pleiotropic effects, there may also be potential bias which could be addressed by extending these findings to other populations. Thirdly, significant heterogeneity and pleiotropy remained in some association analyses, which need further exploration. Fourthly, research regarding male-specific reproductive traits and their relationship to aging and health outcomes is needed to compare the difference between male and female aging. Given the increasing age of first childbirth in modern times, the ideal period of first childbirth for both slowing down female aging and benefiting fetal development needs more research [75]. In summary, our study underscores the complex relationship between genetic legacies and modern diseases. Understanding these genetic predispositions can inform public health strategies and their relevance to age-related disease outcomes in women of various ethnic groups. For example, interventions could be tailored not only to mitigate the risks associated with early reproductive timing but also to address lifestyle factors exacerbating conditions linked to thrifty genes, like diet and physical activity.

Methods

All the MR research data is from the public genome-wide association studies (GWAS) databases of the IEU Open GWAS project (*https://gwas.mrcieu.ac.uk/*) and PubMed. The target analysis is based on public datasets. The population data is from the UK Biobank. No definite personal information was included. No additional ethical approval was required for our study.

Exposure data

2 traits were included as female reproductive activity exposures, involving age at menarche [8] (GWAS ID: ebi-a-GCST90029036) and age at first birth [9] (GWAS ID: ebi-a-GCST90000048).

Outcome data

The aging outcomes involved general aging, organ aging/disease, and organ cancers (Table S1). Frailty index [76] (GWAS ID: ebi-a-GCST90020053), father's age at death [77] (GWAS ID: ebi-a-GCST006700), mother's age at death [77] (GWAS ID: ebi-a-GCST006699), and DNA methylation GrimAge acceleration [78] (GWAS ID: ebi-a-GCST90014294) were included as general aging outcomes. Specific aging diseases and aging levels included age at menopause onset [8] (GWAS ID: ebi-a-GCST90029037), late-onset Alzheimer's disease [79] (LOAD) (PMID: 30820047), osteoporosis [80] (GWAS ID: ebi-a-GCST90038656), and type 2 diabetes [81] (GWAS ID: ebi-a-GCST90018926), chronic heart failure [81] (CHF) (GWAS ID: ebi-a-GCST90018806), essential hypertension (GWAS ID: ukb-b-12493), facial aging (GWAS ID: ukb-b-2148), eye aging [82] (PMID: 36975205), cirrhosis (GWAS ID: finn-b-CIRRHOSIS_BROAD), chronic kidney disease [83] (CKD) (GWAS ID: ebi-a-GCST008026), early onset chronic obstructive pulmonary disease (COPD) (GWAS ID: finn-b-COPD_EARLY), and gastrointestinal or abdominal disease (GAD) [80] (GWAS ID: ebi-a-GCST90038597). Organ cancers included breast cancer [81] (GWAS ID: ebi-a-GCST90018799), ovarian cancer [81] (GWAS ID: ebi-a-GCST90018888), endometrial cancer [84] (GWAS ID: ebi-a-GCST006464), and cervical cancer (GWAS ID: ukb-b-8777).

SNPs selection

We identified single nucleotide polymorphisms (SNPs) associated with exposure datasets with $p < 5 \times 10^{-8}$ [85,86]. In this case, 249 SNPs and 67 SNPs were selected as eligible instrumental variables (IVs) for exposures of age at menarche and age at first birth, respectively. All selected SNPs for every exposure would be clumped to avoid the linkage disequilibrium (r^2 = 0.001 and kb = 10,000). Then we identified whether there were potential confounders of IVs associated with the outcomes based on a database of human genotype-phenotype associations, PhenoScanner V2 [87,88] (*http://www.phenoscanner.medschl.cam.ac.uk/*), with a threshold of $p < 1 \times 10^{-5}$. IVs associated with education, smoking, alcohol, activity, and other confounders related to outcomes would be excluded. During the harmonization process, we aligned the alleles to the human genome reference sequence and removed incompatible SNPs. Subsequent analyses were based on the merged exposure-outcome dataset. We calculated the F statistics to quantify the strength of IVs for each exposure with a threshold of F>10 [89]. If the effect allele frequency (EAF) was missing in the primary dataset, EAF would be collected from dsSNP (*https://www.ncbi.nlm.nih.gov/snp/*) based on the population to calculate the F value. As osteoporosis GWAS was processed by BOLT-LMM, we recalculated the beta value for MR analysis. To adjust the beta values (β_{BOLT}) estimated by BOLT-LMM for dichotomous, we applied the following formula to each SNP:

$$Adjusted\ beta = \frac{\beta_{BOLT}}{\sqrt{P(1-P)}}$$

The prevalence (P) of the binary outcome in the population is the proportion of individuals with the outcomes. The P value for osteoporosis was calculated based on the GWAS sample size (484,598) and disease cases (7,751).

MR analysis

All analysis was performed using R software (version 4.3.1, R Foundation for Statistical Computing, Vienna, Austria) and RStudio software (version 2023.09.0 Posit, PBC, Boston, USA) with TwoSampleMR (version 0.5.7) and MRPRESSO (version 1.0) packages.

Five two-sample MR methods were applied for analysis, including inverse variance weighted model (IVW), MR Egger regression model (MER), weighted median model (WMM), simple mode, and weighted mode. A pleiotropy test was used to check if the IVs influence the outcome through pathways other than the exposure of interest. A heterogeneity test was applied to ensure whether there is a variation in the causal effect estimates across different IVs. Significant heterogeneity test results indicate that some instruments are invalid or that the causal effect varies depending on the IVs used. MRPRESSO was applied to detect and correct potential outliers of IVs with NbDistribution = 10,000 and threshold p = 0.05. Outliers would be excluded for repeated analysis. The causal estimates were given as odds ratios (ORs) and 95% confidence intervals (CI). A leave-one-out analysis was conducted to ensure the robustness of the results by sequentially excluding each IV and confirming the direction and statistical significance of the remaining SNPs.

In two-step MR, we mainly use IVW to access the causal associations between exposure and outcome, exposure and mediator, and mediator and outcome. Body mass index (BMI) (GWAS ID: ukb-b-19953) was applied as a mediator. Similar steps of two-sample MR were repeated in the analysis after excluding confounders. The mediator effect (ME) and direct effect (DE) were calculated. The MEs were calculated with the formula: beta1 × beta2; DEs were calculated according to the formula: beta – (beta1 × beta2), beta stands for the total effect obtained from the primary analysis, beta1 stands for the effect of exposure on the mediator, and beta2 stands for the effect of the mediator on the outcome.

Colocalization analysis

To improve the robustness of our research, colocalization analysis was conducted with packages gwasglue (version 0.0.0.9000), coloc (version 5.2.3), and gassocplot (version 1.0) between the exposures and outcomes revealing significant associations after BMI-related SNPs were excluded. SNPs with significant P values in single SNP OR analysis were set as target SNPs. SNP-level colocalization was conducted with 50kb [90] windows around each target SNP. EAF was set to 0.5 if it was missed in the primary datasets. The Bayesian algorithm in the coloc package generates posterior probabilities (PP) for the hypothesis that both traits are associated and share the same single causal variant at a specific lo[91,92]cus (H_4) [93]. SNP with significant results in the colocalization analysis would be removed and two-sample MR analyses would be conducted again.

Genetic correlation analysis

Genetic correlation analysis helps clarify whether there is a shared genetic architecture between two traits, which can support the interpretation of our MR results. LDSC software, a robust framework to estimate genetic correlation using GWAS summary statistics, was applied to calculate the genetic correlation [91,92]. GWAS summary statistics were filtered with HapMap3 ref. Variants, which were repeated or had a minor allele frequency (MAF) ≤0.01 were excluded. Genetic Correlation (r_g) was calculated for all pairwise comparisons among age at menarche, age at first birth, BMI, and age at menopause onset using LDSC method.**Target Analysis**

We employed the Ingenuity Pathway Analysis (IPA) software (version 01-22-01; Ingenuity Systems; QIAGEN) to investigate various gene-related aspects associated with age at first birth and menarche. IPA is a widely used bioinformatics tool for interpreting high-throughput data [94]. Briefly, the SNPs/genes from MR analysis were uploaded into Qiagen's IPA system for core analysis and then the outcome was overlaid with the global molecular network in the Ingenuity pathway knowledge base (IPKB). IPA was performed to identify canonical pathways, diseases, and functions, and to investigate gene networks. Additionally, the Chemical Biology Database (ChEMBL) (*https://www.ebi.ac.uk/chembl/*), a large-scale database of bioactive molecules for drug discovery [28], and the DrugBank database [29], a comprehensive database on FDA-approved drugs, drug targets, mechanisms of action, and interactions (*https://go.drugbank.com/*), were used to identify candidate genes targeted by approved drugs in clinical trials or under current development phases.

Population validation

The MR results were further validated based on the UK Biobank with package stats (version 4.3.1). In addition to age at menarche and first live birth, the number of births is listed as an independent variable for the validation analysis. Based on the significant associations, regression analysis was adopted to explore the effect of independent variables and related confounders on outcomes. Ages at menarche were divided into 5 age groups, <11y, 11-12y, 13-14y, 15-16y, and >16y. Ages at first live birth were divided into 5 age groups, <21y, 21-25y, 26-30y, 31-35y, and >35y. Females with the number of births from 0 to 4 were included. Education was divided into 7 levels based on the degrees. Smoking was divided into 2 categories, ever smoking or not.

Drinking was divided into 3 categories, never, previous, and current drinking. BMI was divided into four categories, <18.5, 18.5-24.9, 25-29.9, and ≥30 based on the average value of 4 records. The outcome results were preprocessed based on the baseline information collection, follow-up outcome, surgical history, and ICD-10 diagnosis summary. Logistic regression was used for the analysis of disease risks (categorical variables). The coefficients (β) in logistic regression were log-odds ratios. Although the continuous variables exhibited mild to moderate skewed distributions, considering the large sample size, linear regression was used for the analysis of age at death, menopause age, and BMI in the first step. Then logistic regression was applied to confirm the robustness of the results with the outcome of death age ≥ 80, menopause age ≥ 50, and BMI ≥ 30. Confounders for regression analysis I include education, smoking, and drinking. Confounders for regression analysis II include education, smoking, drinking, and BMI. Variance Inflation Factors (VIF) were calculated for each predictor in regression analysis to avoid multicollinearity. Values of VIF were less than 1.5 for each predictor. For outcomes of diabetes, HBP, heart failure, and BMI≥30, the combined effect of age at menarche and first birth was analyzed. 179,821 participants were divided into 25 groups according to the menarche and first live birth age groups. All other groups were compared to the group with menarche of "13-14y" and the first birth of "26-30y" in logistic regression (confounders including education, smoking, drinking, and BMI for diabetes, HBP, and heart failure; including education, smoking, and drinking for BMI≥30).

Data Availability

All data and code are available in the main text, supplementary materials, or public database. The population data is available through the UK Biobank.

Acknowledgements

We thank the Kapahi Lab at Buck Institute for helpful discussions. This research has been conducted using the UK Biobank Resource under Application Number 151101.

Additional information

Funding

Hevolution Foundation (PK). National Institute of Health grant R01AG068288 and R01AG045835 (PK). Larry L. Hillblom Foundation (PK), Larry L. Hillblom Foundation (PS).

Author contributions

Conceptualization: YX, VT, PS, LLF, and PK. MR analysis and UK biobank validation: YX. Gene target analysis: VT. Funding acquisition: PK and PS. Writing original draft: YX, PS, and VT. All authors reviewed and edited the manuscript.

Additional files

Tables S1 to S12

Figs S1 to S3

References

1 Medawer P. B. (1952) **An Unsolved Problem of Biology**

2 Byars S. G., Voskarides K (2020) **Antagonistic Pleiotropy in Human Disease** *J Mol Evol* **88**:12–25 https://doi.org/10.1007/s00239-019-09923-2

3 Wang Z., Cong H (2021) **Antagonistic pleiotropy can promote adaptation to patchy environments** *Evolution* **75**:197–199 https://doi.org/10.1111/evo.14133

4 Kumar J., et al. (2016) **Zinc Levels Modulate Lifespan through Multiple Longevity Pathways in Caenorhabditis elegans** *PLoS One* **11**:e0153513 https://doi.org/10.1371/journal.pone.0153513

5 Grundy E., Tomassini C (2005) **Fertility history and health in later life: a record linkage study in England and Wales** *Soc Sci Med* **61**:217–228 https://doi.org/10.1016/j.socscimed.2004.11.046

6 Day F. R., et al. (2016) **Physical and neurobehavioral determinants of reproductive onset and success** *Nat Genet* **48**:617–623 https://doi.org/10.1038/ng.3551

7 Sanderson E., et al. (2022) **Mendelian randomization** *Nat Rev Methods Primers* **2** https://doi.org/10.1038/s43586-021-00092-5

8 Loh P. R., Kichaev G., Gazal S., Schoech A. P., Price A. L (2018) **Mixed-model association for biobank-scale datasets** *Nat Genet* **50**:906–908 https://doi.org/10.1038/s41588-018-0144-6

9 Mills M. C., et al. (2021) **Identification of 371 genetic variants for age at first sex and birth linked to externalising behaviour** *Nat Hum Behav* **5**:1717–1730 https://doi.org/10.1038/s41562-021-01135-3

10 Ng J. C. M., Schooling C. M (2023) **Effect of basal metabolic rate on lifespan: a sex-specific Mendelian randomization study** *Sci Rep* **13**:7761 https://doi.org/10.1038/s41598-023-34410-6

11 Williams D. M., Jylhava J., Pedersen N. L., Hagg S (2019) **A Frailty Index for UK Biobank Participants** *J Gerontol A Biol Sci Med Sci* **74**:582–587 https://doi.org/10.1093/gerona/gly094

12 Locke A. E., et al. (2015) **Genetic studies of body mass index yield new insights for obesity biology** *Nature* **518**:197–206 https://doi.org/10.1038/nature14177

13 Silva-Garcia C. G., et al. (2023) **The CRTC-1 transcriptional domain is required for COMPASS complex-mediated longevity in C. elegans** *Nat Aging* **3**:1358–1371 https://doi.org/10.1038/s43587-023-00517-8

14 Mendioroz M., et al. (2016) **CRTC1 gene is differentially methylated in the human hippocampus in Alzheimer's disease** *Alzheimers Res Ther* **8**:15 https://doi.org/10.1186/s13195-016-0183-0

15 Liu Y., et al. (2023) **CHAC1 as a Novel Contributor of Ferroptosis in Retinal Pigment Epithelial Cells with Oxidative Damage** *Int J Mol Sci* **24** https://doi.org/10.3390/ijms24021582

16 Li J., et al. (2023) **CHAC1 inactivation is effective to preserve muscle glutathione but is insufficient to protect against muscle wasting in cachexia** *PLoS One* **18**:e0283806 https://doi.org/10.1371/journal.pone.0283806

17 Cloos P. A., Christgau S (2004) **Post-translational modifications of proteins: implications for aging, antigen recognition, and autoimmunity** *Biogerontology* **5**:139–158 https://doi.org/10.1023/B:BGEN.0000031152.31352.8b

18 Udenwobele D. I., et al. (2017) **Myristoylation: An Important Protein Modification in the Immune Response** *Front Immunol* **8**:751 https://doi.org/10.3389/fimmu.2017.00751

19 Hayashi N., Titani K (2010) **N-myristoylated proteins, key components in intracellular signal transduction systems enabling rapid and flexible cell responses** *Proc Jpn Acad Ser B Phys Biol Sci* **86**:494–508 https://doi.org/10.2183/pjab.86.494

20 Wen Z., et al. (2019) **N-myristoyltransferase deficiency impairs activation of kinase AMPK and promotes synovial tissue inflammation** *Nat Immunol* **20**:313–325 https://doi.org/10.1038/s41590-018-0296-7

21 Liang J., et al. (2015) **Myristoylation confers noncanonical AMPK functions in autophagy selectivity and mitochondrial surveillance** *Nat Commun* **6**:7926 https://doi.org/10.1038/ncomms8926

22 McHugh D., et al. (2023) **COPI vesicle formation and N-myristoylation are targetable vulnerabilities of senescent cells** *Nat Cell Biol* **25**:1804–1820 https://doi.org/10.1038/s41556-023-01287-6

23 Thinon E., Morales-Sanfrutos J., Mann D. J., Tate E. W (2016) **N-Myristoyltransferase Inhibition Induces ER-Stress, Cell Cycle Arrest, and Apoptosis in Cancer Cells** *ACS Chem Biol* **11**:2165–2176 https://doi.org/10.1021/acschembio.6b00371

24 Zhu G., et al. (2021) **N-Myristoylation by NMT1 Is POTEE-Dependent to Stimulate Liver Tumorigenesis via Differentially Regulating Ubiquitination of Targets** *Front Oncol* **11**:681366 https://doi.org/10.3389/fonc.2021.681366

25 Tomita Y., et al. (2023) **Targeting N-Myristoylation Through NMT2 Prevents Cardiac Hypertrophy and Heart Failure** *JACC Basic Transl Sci* **8**:1263–1282 https://doi.org/10.1016/j.jacbts.2023.06.006

26 Ashrafi K., Lin S. S., Manchester J. K., Gordon J. I (2000) **Sip2p and its partner snf1p kinase affect aging in S. cerevisiae** *Genes Dev* **14**:1872–1885

27 Kramer A., Green J., Pollard J., Tugendreich S (2014) **Causal analysis approaches in Ingenuity Pathway Analysis** *Bioinformatics* **30**:523–530 https://doi.org/10.1093/bioinformatics/btt703

28 Gaulton A., et al. (2012) **ChEMBL: a large-scale bioactivity database for drug discovery** *Nucleic Acids Res* **40**:D1100–1107 https://doi.org/10.1093/nar/gkr777

29 Wishart D. S., et al. (2018) **DrugBank 5.0: a major update to the DrugBank database for 2018** *Nucleic Acids Res* **46**:D1074–D1082 https://doi.org/10.1093/nar/gkx1037

30 Guo H. J., Ye Y. L., Gao Y. F., Liu Z. H (2023) **Age at first birth is associated with the likelihood of frailty in middle-aged and older women: A population-based analysis from NHANES 1999-2018** *Maturitas* **181**:107904 https://doi.org/10.1016/j.maturitas.2023.107904

31 Long E., Zhang J (2023) **Evidence for the role of selection for reproductively advantageous alleles in human aging** *Sci Adv* **9**:eadh4990 https://doi.org/10.1126/sciadv.adh4990

32 Liu D., et al. (2022) **Vitamin D and Multiple Health Outcomes: An Umbrella Review of Observational Studies, Randomized Controlled Trials, and Mendelian Randomization Studies** *Adv Nutr* **13**:1044–1062 https://doi.org/10.1093/advances/nmab142

33 Chen X., et al. (2021) **Kidney damage causally affects the brain cortical structure: A Mendelian randomization study** *EBioMedicine* **72**:103592 https://doi.org/10.1016/j.ebiom.2021.103592

34 Barth C., de Lange A. G (2020) **Towards an understanding of women's brain aging: the immunology of pregnancy and menopause** *Front Neuroendocrinol* **58**:100850 https://doi.org/10.1016/j.yfrne.2020.100850

35 Sherwin B. B., Henry J. F (2008) **Brain aging modulates the neuroprotective effects of estrogen on selective aspects of cognition in women: a critical review** *Front Neuroendocrinol* **29**:88–113 https://doi.org/10.1016/j.yfrne.2007.08.002

36 Day F. R., Elks C. E., Murray A., Ong K. K., Perry J. R (2015) **Puberty timing associated with diabetes, cardiovascular disease and also diverse health outcomes in men and women: the UK Biobank study** *Sci Rep* **5** https://doi.org/10.1038/srep11208

37 Hollis B., et al. (2020) **Genomic analysis of male puberty timing highlights shared genetic basis with hair colour and lifespan** *Nat Commun* **11**:1536 https://doi.org/10.1038/s41467-020-14451-5

38 Jorge S., Chang S., Barzilai J. J., Leppert P., Segars J. H (2014) **Mechanical signaling in reproductive tissues: mechanisms and importance** *Reprod Sci* **21**:1093–1107 https://doi.org/10.1177/1933719114542023

39 Baumgarten S. C., Armouti M., Ko C., Stocco C (2017) **IGF1R Expression in Ovarian Granulosa Cells Is Essential for Steroidogenesis, Follicle Survival, and Fertility in Female Mice** *Endocrinology* **158**:2309–2318 https://doi.org/10.1210/en.2017-00146

40 Botkjaer J. A., et al. (2024) **Dynamics of IGF signalling during the ovulatory peak in women undergoing ovarian stimulation** *J Clin Endocrinol Metab* https://doi.org/10.1210/clinem/dgae132

41 Stubbs S. A., et al. (2013) **Role of Insulin-like growth factors in initiation of follicle growth in normal and polycystic human ovaries** *J Clin Endocrinol Metab* **98**:3298–3305 https://doi.org/10.1210/jc.2013-1378

42 Papaconstantinou J (2009) **Insulin/IGF-1 and ROS signaling pathway cross-talk in aging and longevity determination** *Mol Cell Endocrinol* **299**:89–100 https://doi.org/10.1016/j.mce.2008.11.025

43 Kalman S. M (1969) **Effects of oral contraceptives** *Annu Rev Pharmacol* **9**:363–378 https://doi.org/10.1146/annurev.pa.09.040169.002051

44 van Heemst D. (2010) **Insulin, IGF-1 and longevity** *Aging Dis* **1**:147–157

45 Junnila R. K., List E. O., Berryman D. E., Murrey J. W., Kopchick J. J (2013) **The GH/IGF-1 axis in ageing and longevity** *Nat Rev Endocrinol* **9**:366–376 https://doi.org/10.1038/nrendo.2013.67

46 Aleandri V., Spina V., Ciardo A (1997) **[The role of the pineal body in the endocrine control of puberty]** *Minerva Ginecol* **49**:43–48

47 Olcese J. M (2020) **Melatonin and Female Reproduction: An Expanding Universe** *Front Endocrinol (Lausanne)* **11**:85 https://doi.org/10.3389/fendo.2020.00085

48 Chen D., Zhao M., Mundy G. R. (2004) **Bone morphogenetic proteins** *Growth Factors* **22**:233–241 https://doi.org/10.1080/08977190412331279890

49 Magro-Lopez E., Munoz-Fernandez M. A (2021) **The Role of BMP Signaling in Female Reproductive System Development and Function** *Int J Mol Sci* **22** https://doi.org/10.3390/ijms222111927

50 Childs A. J., et al. (2010) **BMP signaling in the human fetal ovary is developmentally regulated and promotes primordial germ cell apoptosis** *Stem Cells* **28**:1368–1378 https://doi.org/10.1002/stem.440

51 da Cunha E. V., et al. (2017) **Effects of bone morphogenetic protein 4 (BMP4) on in vitro development and survival of bovine preantral follicles enclosed in fragments ovarian tissue** *Zygote* **25**:256–264 https://doi.org/10.1017/S0967199417000089

52 Zhang X., et al. (2021) **BMP4 overexpression induces the upregulation of APP/Tau and memory deficits in Alzheimer's disease** *Cell Death Discov* **7**:51 https://doi.org/10.1038/s41420-021-00435-x

53 Meyers E. A., et al. (2016) **Increased bone morphogenetic protein signaling contributes to age-related declines in neurogenesis and cognition** *Neurobiol Aging* **38**:164–175 https://doi.org/10.1016/j.neurobiolaging.2015.10.035

54 Ye D., et al. (2023) **Insights into bone morphogenetic proteins in cardiovascular diseases** *Front Pharmacol* **14**:1125642 https://doi.org/10.3389/fphar.2023.1125642

55 Baboota R. K., Bluher M., Smith U (2021) **Emerging Role of Bone Morphogenetic Protein 4 in Metabolic Disorders** *Diabetes* **70**:303–312 https://doi.org/10.2337/db20-0884

56 Saenger P (2003) **Dose effects of growth hormone during puberty** *Horm Res* **60**:52–57 https://doi.org/10.1159/000071226

57 Bartke A. (2021) **Growth hormone and aging** *Rev Endocr Metab Disord* **22**:71–80 https://doi.org/10.1007/s11154-020-09593-2

58 Kalous J., Aleshkina D., Anger M (2023) **A Role of PI3K/Akt Signaling in Oocyte Maturation and Early Embryo Development** *Cells* **12** https://doi.org/10.3390/cells12141830

59 Chung E., Yeung F., Leinwand L. A (2012) **Akt and MAPK signaling mediate pregnancy-induced cardiac adaptation** *J Appl Physiol (1985)* **112**:1564–1575 https://doi.org/10.1152/japplphysiol.00027.2012

60 Chen Y. R., et al. (2019) **Aging-induced Akt activation involves in aging-related pathologies and Abeta-induced toxicity** *Aging Cell* **18**:e12989 https://doi.org/10.1111/acel.12989

61 Howard S. R (2021) **Interpretation of reproductive hormones before, during and after the pubertal transition-Identifying health and disordered puberty** *Clin Endocrinol (Oxf)* **95**:702–715 https://doi.org/10.1111/cen.14578

62 Bhartiya D., Sriraman K., Gunjal P., Modak H (2012) **Gonadotropin treatment augments postnatal oogenesis and primordial follicle assembly in adult mouse ovaries?** *J Ovarian Res* **5**:32 https://doi.org/10.1186/1757-2215-5-32

63 Danilovich N., Javeshghani D., Xing W., Sairam M. R (2002) **Endocrine alterations and signaling changes associated with declining ovarian function and advanced biological aging in follicle-stimulating hormone receptor haploinsufficient mice** *Biol Reprod* **67**:370–378 https://doi.org/10.1095/biolreprod67.2.370

64 Hale G. E., et al. (2007) **Endocrine features of menstrual cycles in middle and late reproductive age and the menopausal transition classified according to the Staging of Reproductive Aging Workshop (STRAW) staging system** *J Clin Endocrinol Metab* **92**:3060–3067 https://doi.org/10.1210/jc.2007-0066

65 Santoro N., Randolph J. F. (2011) **Reproductive hormones and the menopause transition** *Obstet Gynecol Clin North Am* **38**:455–466 https://doi.org/10.1016/j.ogc.2011.05.004

66 Brink Vanden, et al. (2015) **Associations Between Antral Ovarian Follicle Dynamics and Hormone Production Throughout the Menstrual Cycle as Women Age** *J Clin Endocrinol Metab* **100**:4553–4562 https://doi.org/10.1210/jc.2015-2643

67 Plotnikov A., Zehorai E., Procaccia S., Seger R (2011) **The MAPK cascades: signaling components, nuclear roles and mechanisms of nuclear translocation** *Biochim Biophys Acta* **1813**:1619–1633 https://doi.org/10.1016/j.bbamcr.2010.12.012

68 Zhang W., Liu H. T (2002) **MAPK signal pathways in the regulation of cell proliferation in mammalian cells** *Cell Res* **12**:9–18 https://doi.org/10.1038/sj.cr.7290105

69 Yuan W., et al. (2023) **Modulating p38 MAPK signaling by proteostasis mechanisms supports tissue integrity during growth and aging** *Nat Commun* **14**:4543 https://doi.org/10.1038/s41467-023-40317-7

70 An Y., et al. (2022) **Oridonin Delays Aging Through the AKT Signaling Pathway** *Front Pharmacol* **13**:888247 https://doi.org/10.3389/fphar.2022.888247

71 Abdellatif M., et al. (2022) **Fine-Tuning Cardiac Insulin-Like Growth Factor 1 Receptor Signaling to Promote Health and Longevity** *Circulation* **145**:1853–1866 https://doi.org/10.1161/CIRCULATIONAHA.122.059863

72 Lundgren S., et al. (2022) **BMI is positively associated with accelerated epigenetic aging in twin pairs discordant for body mass index** *J Intern Med* **292**:627–640 https://doi.org/10.1111/joim.13528

73 Quach A., et al. (2017) **Epigenetic clock analysis of diet, exercise, education, and lifestyle factors** *Aging (Albany NY)* **9**:419–446 https://doi.org/10.18632/aging.101168

74 Wilson K. A., et al. (2021) **Evaluating the beneficial effects of dietary restrictions: A framework for precision nutrigeroscience** *Cell Metab* **33**:2142–2173 https://doi.org/10.1016/j.cmet.2021.08.018

75 Balasch J. (2010) **Ageing and infertility: an overview** *Gynecol Endocrinol* **26**:855–860 https://doi.org/10.3109/09513590.2010.501889

76 Atkins J. L., et al. (2021) **A genome-wide association study of the frailty index highlights brain pathways in ageing** *Aging Cell* **20**:e13459 https://doi.org/10.1111/acel.13459

77 Pilling L. C., et al. (2017) **Human longevity: 25 genetic loci associated in 389,166 UK biobank participants** *Aging (Albany NY)* **9**:2504–2520 https://doi.org/10.18632/aging.101334

78 McCartney D. L., et al. (2021) **Genome-wide association studies identify 137 genetic loci for DNA methylation biomarkers of aging** *Genome Biol* **22**:194 https://doi.org/10.1186/s13059-021-02398-9

79 Kunkle B. W., et al. (2019) **Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing** *Nat Genet* **51**:414–430 https://doi.org/10.1038/s41588-019-0358-2

80 Donertas H. M., Fabian D. K., Valenzuela M. F., Partridge L., Thornton J. M (2021) **Common genetic associations between age-related diseases** *Nat Aging* **1**:400–412 https://doi.org/10.1038/s43587-021-00051-5

81 Sakaue S., et al. (2021) **A cross-population atlas of genetic associations for 220 human phenotypes** *Nat Genet* **53**:1415–1424 https://doi.org/10.1038/s41588-021-00931-x

82 Ahadi S., et al. (2023) **Longitudinal fundus imaging and its genome-wide association analysis provide evidence for a human retinal aging clock** *eLife* **12** https://doi.org/10.7554/eLife.82364

83 Wojcik G. L., et al. (2019) **Genetic analyses of diverse populations improves discovery for complex traits** *Nature* **570**:514–518 https://doi.org/10.1038/s41586-019-1310-4

84 O'Mara T. A., et al. (2018) **Identification of nine new susceptibility loci for endometrial cancer** *Nat Commun* **9**:3166 https://doi.org/10.1038/s41467-018-05427-7

85 Savage J. E., et al. (2018) **Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence** *Nat Genet* **50**:912–919 https://doi.org/10.1038/s41588-018-0152-6

86 Gao X., et al. (2019) **The bidirectional causal relationships of insomnia with five major psychiatric disorders: A Mendelian randomization study** *Eur Psychiatry* **60**:79–85 https://doi.org/10.1016/j.eurpsy.2019.05.004

87 Staley J. R., et al. (2016) **PhenoScanner: a database of human genotype-phenotype associations** *Bioinformatics* **32**:3207–3209 https://doi.org/10.1093/bioinformatics/btw373

88 Kamat M. A., et al. (2019) **PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations** *Bioinformatics* **35**:4851–4853 https://doi.org/10.1093/bioinformatics/btz469

89 Burgess S., Small D. S., Thompson S. G (2017) **A review of instrumental variable estimators for Mendelian randomization** *Stat Methods Med Res* **26**:2333–2355 https://doi.org/10.1177/0962280215597579

90 Venkateswaran S., et al. (2018) **Enhanced Contribution of HLA in Pediatric Onset Ulcerative Colitis** *Inflamm Bowel Dis* **24**:829–838 https://doi.org/10.1093/ibd/izx084

91 Bulik-Sullivan B., et al. (2015) **An atlas of genetic correlations across human diseases and traits** *Nat Genet* **47**:1236–1241 https://doi.org/10.1038/ng.3406

92 Bulik-Sullivan B. K., et al. (2015) **LD Score regression distinguishes confounding from polygenicity in genome-wide association studies** *Nat Genet* **47**:291–295 https://doi.org/10.1038/ng.3211

93 Wallace C (2021) **A more accurate method for colocalisation analysis allowing for multiple causal variants** *PLoS Genet* **17**:e1009440 https://doi.org/10.1371/journal.pgen.1009440

94 Ghosh S., et al. (2017) **Core Canonical Pathways Involved in Developing Human Glioblastoma Multiforme (GBM)** *Int J Sci Res Sci Eng Technol* **3**:458–465

Author information

Yifan Xiang
The Buck Institute for Research on Aging, Novato, United States

Vineeta Tanwar
The Buck Institute for Research on Aging, Novato, United States

Parminder Singh
The Buck Institute for Research on Aging, Novato, United States
ORCID iD: 0000-0002-2851-6805

Lizellen La Follette
The Buck Institute for Research on Aging, Novato, United States

Vikram Narayan
The Buck Institute for Research on Aging, Novato, United States

Pankaj Kapahi
The Buck Institute for Research on Aging, Novato, United States, Department of Urology, University of California, San Francisco, San Francisco, United States
ORCID iD: 0000-0002-5629-4947

For correspondence: pkapahi@buckinstitute.org

Editors

Reviewing Editor
Joris Deelen
Leiden University Medical Center, Leiden, Netherlands

Senior Editor
Matthias Barton
University of Zurich, Zurich, Switzerland

Reviewer #1 (Public review):

Summary:

The present study aims to associate reproduction with age-related disease as support of the antagonistic pleiotropy hypothesis of ageing predominantly using Mendelian Randomization. The authors found evidence that early-life reproductive succes is associated with advanced ageing.

Strengths:

Large sample size. Many analyses

Weaknesses:

Still a number of doubts with regard to some of the results and their interpretation.

https://doi.org/10.7554/eLife.102447.2.sa2

Reviewer #2 (Public review):

Summary:

The authors present an interesting paper where they test the antagonistic pleiotropy theory. Based on this theory they hypothesize that genetic variants associated with later onset of age at menarche and age at first birth may have a positive effect on a multitude of health outcomes later in life, such as epigenetic aging and prevalence of chronic diseases. Using a mendelian randomization and colocalization approach, the authors show that SNPs associated with later age at menarche are associated with delayed aging measurements, such as slower epigenetic aging and reduced facial aging and a lower risk of chronic diseases, such as type 2 diabetes and hypertension. Moreover, they identify 128 fertility-related SNPs that associate with age-related outcomes and they identified BMI as a mediating factor for disease risk, discussing this finding in the context of evolutionary theory.

Strengths:

The major strength of this manuscript is that it addresses the antagonistic pleiotropy theory in aging. Aging theories are not frequently empirically tested although this is highly necessary. The work is therefore relevant for the aging field as well as beyond this field, as the antagonistic pleiotropy theory addresses the link between fitness (early life health and reproduction) and aging.

The authors addressed the remarks on the previous version very well. Addressing the two points below would further increase the quality of the manuscript.

(1) In the previous version the authors mentioned that their results are also consistent with the disposable soma theory: "These results are also consistent with the disposable soma theory that suggests aging as an outcome tradeoff between an organism's investment in reproduction and somatic maintenance and repair."

Although the antagonistic pleiotropy and disposable soma theories describe different mechanisms, both provide frameworks for understanding how genes linked to fertility influence health. The antagonistic pleiotropy theory posits that genes enhancing fertility early in life may have detrimental effects later. In contrast, the disposable soma theory suggests that energy allocation involves a trade-off, where investment in fertility comes at the expense of somatic maintenance, potentially leading to poorer health in later life.

To strengthen the manuscript, a discussion section should be added to clarify the overlap and distinctions between these two evolutionary theories and suggest directions for future research in disentangling their specific mechanisms.

(2) In response to the question why the authors did not include age at menopause in addition to the already included age at first child and age at menarche the following explanation was provided: "Our manuscript focuses on the antagonistic pleiotropy theory, which posits that inherent trade-off in natural selection, where genes beneficial for early survival and reproduction (like menarche and childbirth) may have costly consequences later. So, we only included age at menarche and age at first childbirth as exposures in our research."

It remains, however, unclear why genes beneficial for early survival and reproduction would be reflected only in age at menarche and age at first childbirth, but not in age at menopause. While age at menarche marks the onset of fertility, age at menopause signifies its end. Since evolutionary selection acts directly until reproduction is no longer possible (though indirect evolutionary pressures persist beyond this point), the inclusion of additional fertility-related measures could have strengthened the analysis. A more detailed justification for focusing exclusively on age at menarche and first childbirth would enhance the clarity and rigor of the manuscript.

https://doi.org/10.7554/eLife.102447.2.sa1

Author response:

The following is the authors' response to the original reviews.

> ***Reviewer #1 (Public review):***
>
> *Summary:*
>
> *The present study aims to associate reproduction with age-related disease as support of the antagonistic pleiotropy hypothesis of ageing, predominantly using Mendelian Randomization. The authors found evidence that early-life reproductive success is associated with advanced ageing.*
>
> *Strengths:*
>
> *Large sample size. Many analyses.*
>
> *Weaknesses:*
>
> *There are some errors in the methodology, that require revisions.*
>
> *In particular, the main conclusions drawn by the authors refer to the Mendelian Randomization analyses. However, the authors made a few errors here that need to be reconsidered:*
>
> *(1) Many of the outcomes investigated by the authors are continuous outcomes, while the authors report odds ratios. This is not correct and should be revised.*

Thank you for your observation. We have revised the manuscript to ensure that the results for continuous outcomes are appropriately reported using beta coefficients, which indicate the change in the outcome per unit increase in exposure. This will accurately reflect the nature of the analysis and provide a clearer interpretation of continuous outcomes (lines 56-109).

> *(2) Some of the odds ratios (for example the one for osteoporosis) are really small, while still reaching the level of statistical significance. After some checking, I found the GWAS data used to generate these MR estimates were processed by the program BOLT-LLM. This program is a linear mixed model program, which requires the transformation of the beta estimates to be useful for dichotomous outcomes. The authors should check the manual of BOLT-LLM and recalculate the beta estimates of the SNP-outcome associations prior to the Mendelian Randomization analyses. This should be checked for all outcomes as it doesn't apply to all.*

Thank you for your detailed feedback. We have reviewed all the GWAS data used in our MR analyses and confirmed that all GWAS of continuous traits have already been processed using the BOLT-LMM, including age at menarche, age at first birth, BMI, frailty index, father's age at death, mother's age at death, DNA methylation GrimAge acceleration, age at menopause, eye age, and facial aging. Most of the dichotomous outcomes have not been processed by BOLT-LMM, including late-onset Alzheimer's disease, type 2 diabetes, chronic heart failure, essential hypertension, cirrhosis, chronic kidney disease, early onset chronic obstructive pulmonary disease, breast cancer, ovarian cancer, endometrial cancer, and cervical cancer, except osteoporosis. We have reprocessed the GWAS beta values of osteoporosis and re-conducted the MR analysis (lines 74-75; lines 366-373).

> *(3) The authors should follow the MR-Strobe guidelines for presentation.*

Thank you for your suggestion to follow the MR-STROBE guidelines for the presentation of our study. We appreciate the importance of adhering to these standardized guidelines to ensure clarity and transparency in reporting Mendelian Randomization (MR) analyses. We confirm that the MR components of our research are structured and presented following the MR-STROBE checklist. In addition to the MR analyses, our study also integrates Colocalization analysis, Genetic correlation analysis, Ingenuity Pathway Analysis (IPA), and population validation to provide a more comprehensive understanding of the genetic and biological context. While these analyses are not strictly covered by MR-STROBE guidelines, they complement the MR results by offering additional validation and mechanistic insights.

We have structured our manuscript to separate these complementary analyses from the core MR results, maintaining alignment with MR-STROBE for the MR-specific components. The additional analyses are discussed in dedicated sections to highlight their unique contributions and avoid conflating them with the MR findings.

> *(4) The authors should report data in the text with a 95% confidence interval.*

Thank you for your feedback. We have added the 95% confidence intervals for the reported data within the main text to enhance clarity and provide comprehensive context (lines 56-109). Additionally, the complete analysis data, including all detailed results, can be found in Table S3.

> *(5) The authors should consider correction for multiple testing*

Thank you for your comment regarding the need to consider correction for multiple testing. We agree that correcting for multiple comparisons is an important step to control for the possibility of false-positive findings, particularly in studies involving large numbers of statistical tests. In our study, we carefully considered the issue of multiple testing and adopted the following approach:

Context of Multiple Testing: The tests we conducted were hypothesis-driven, focusing on specific relationships (e.g., genetic correlation, colocalization, and Mendelian

Randomization). These analyses are based on priori hypotheses supported by existing literature or biological relevance.

Statistical Methods: Where applicable, we applied appropriate measures to account for multiple tests. For instance, in Mendelian Randomization, sensitivity analyses serve to validate the robustness of the results.

We believe that the methodology and corrections applied in our study appropriately address concerns about multiple testing, given the hypothesis-driven nature of our analyses and the rigorous steps taken to validate our findings. If you feel that additional corrections are required for specific parts of the analysis, we would be happy to further clarify or revise as needed.

> ***Reviewer #2 (Public review):***
>
> *Summary:*
>
> *The authors present an interesting paper where they test the antagonistic pleiotropy theory. Based on this theory they hypothesize that genetic variants associated with later onset of age at menarche and age at first birth have a positive causal effect on a multitude of health outcomes later in life, such as epigenetic aging and prevalence of chronic diseases. Using a mendelian randomization and colocalization approach, the authors show that SNPs associated with later age at menarche are associated with delayed aging measurements, such as slower epigenetic aging and reduced facial aging, and a lower risk of chronic diseases, such as type 2 diabetes and hypertension. Moreover, they identified 128 fertility-related SNPs that are associated with age-related outcomes and they identified BMI as a mediating factor for disease risk, discussing this finding in the context of evolutionary theory.*
>
> *Strengths:*
>
> *The major strength of this manuscript is that it addresses the antagonistic pleiotropy theory in aging. Aging theories are not frequently empirically tested although this is highly necessary. The work is therefore relevant for the aging field as well as beyond this field, as the antagonistic pleiotropy theory addresses the link between fitness (early life health and reproduction) and aging.*
>
> *Points that have to be clarified/addressed:*
>
> *(1) The antagonistic pleiotropy is an evolutionary theory pointing to the possibility that mutations that are beneficial for fitness (early life health and reproduction) may be detrimental later in life. As it concerns an evolutionary process and the authors focus on contemporary data from a single generation, more context is necessary on how this theory is accurately testable. For example, why and how much natural variation is there for fitness outcomes in humans?*

Thank you for these insightful questions. We appreciate the opportunity to clarify how we approach the testing of AP theory within a contemporary human cohort and address the evolutionary context and comparative considerations with the disposable soma theory.

We recognize that modern human populations experience selection pressures that differ from those in the past, which may affect how well certain genetic variants reflect historical fitness benefits. Nonetheless, the genetic variation present today still offers valuable insights into potential AP mechanisms through statistical associations in contemporary cohorts. We believe that AP can indeed be explored in current populations by examining genetic links between reproductive traits and age-related health outcomes. In our study, we investigate whether certain genetic variants linked to reproductive timing—such as age at menarche and

age at first birth—also correlate with late-life health risks. By identifying SNPs associated with both early-life reproductive success and adverse aging outcomes, we aim to capture the evolutionary trade-offs that AP theory suggests.

Despite contemporary selection pressures that differ from historical conditions, there remains natural genetic variation in traits like reproductive timing and longevity in humans today. This diversity allows us to apply MR to test causal relationships between reproductive traits and aging outcomes, providing insights into potential AP mechanisms. Prior studies have demonstrated that reproductive behaviors exhibit significant heritability and have identified genetic loci associated with reproductive timing (1,2). This genetic variation facilitates causal inference in modern cohorts, despite environmental and healthcare advances that might modulate these associations (3). By leveraging genetic risk scores for reproductive timing, our study captures the necessary variability to assess potential AP effects, thus providing valuable insights into how evolutionary trade-offs may continue to influence human health outcomes.

> *How do genetic risk score distributions of the exposure data look like?*

Thank you for your question. Our study is focused on Mendelian Randomization (MR) analysis, which aims to infer causal relationships between exposures and outcomes. While genetic risk scores (GRS) provide valuable insights at an individual level, they do not directly align with our study's objective, which is centered on population-level causal inference rather than individual-level genetic risk assessment. In MR, we use genetic variants as instrumental variables to determine the causal effect of an exposure on an outcome. GRS analysis typically focuses on summarizing an individual's risk based on multiple genetic variants, which is outside the scope of our current research. Therefore, we did not perform or analyze the distribution of genetic risk scores, as our primary goal was to understand broader causal relationships using established genetic instruments.

> *Also, how can the authors distinguish in their data between the antagonistic pleiotropy theory and the disposable soma theory, which considers a trade-off between investment in reproduction and somatic maintenance and can be used to derive similar hypotheses? There is just a very brief mention of the disposable soma theory in lines 196-198.*

In our manuscript, we test AP theory specifically by examining genetic variants associated with reproductive timing and their association with age-related health risks in later life. MR and genetic risk scores allow us to assess these associations, directly testing the hypothesis that certain alleles enhancing reproductive success might have adverse effects on aging outcomes. This gene-centered approach aligns with AP's premise of genetic trade-offs, enabling us to observe whether alleles associated with early-life reproductive traits correlate with increased risks of age-related diseases. Distinguishing from disposable soma theory, which would predict a general trade-off in energy allocation affecting somatic maintenance and not specific genetic effects, our data focuses on how certain alleles have differential impacts across life stages. Our findings thus support AP theory over disposable soma by highlighting the effects of specific genetic loci on both reproductive and aging phenotypes. However, future research could indeed explore the intersection of these theories, for example, by examining how resource allocation and genetic predispositions interact to influence longevity in various environmental contexts.

> *(2) The antagonistic pleiotropy theory, used to derive the hypothesis, does not necessarily distinguish between male and female fitness. Would the authors expect that their results extrapolate to males as well? And can they test that?*

Emerging evidence suggests that early puberty in males is linked to adverse health outcomes, such as an increased risk of cardiovascular disease, type 2 diabetes, and hypertension in later life (4). A Mendelian randomization study also reported a genetic association between the timing of male puberty and reduced lifespan (5). These findings support the hypothesis that genetic variants associated with delayed reproductive timing in males might similarly confer health benefits or improved longevity, akin to the patterns observed in females. This would suggest that similar mechanisms of antagonistic pleiotropy could operate in males as well.

In our study, BMI was identified as a mediator between reproductive timing and disease risk. Given that BMI is a common risk factor for age-related diseases in both males and females (6-9), it is plausible that similar mechanisms involving BMI, reproductive timing, and disease risk could exist in males. This shared mediator points to the possibility that, while reproductive timelines may differ, the pathways through which these traits influence aging outcomes may be consistent across genders.

AP theory could potentially be tested in males, as the principles of the theory may extend to analogous reproductive traits in males, such as age at puberty and testosterone levels, which could similarly influence health outcomes later in life. However, as our current study focuses specifically on female reproductive traits, testing the AP theory in males is outside the scope of this work. We acknowledge the importance of exploring these mechanisms in males, and we hope that future research will address this by investigating male-specific reproductive traits and their relationship to aging and health outcomes.

> *(3) There is no statistical analyses section providing the exact equations that are tested. Hence it's not clear how many tests were performed and if correction for multiple testing is necessary. It is also not clear what type of analyses have been done and why they have been done. For example in the section starting at line 47, Odds Ratios are presented, indicating that logistic regression analyses have been performed. As it's not clear how the outcomes are defined (genotype or phenotype, cross-sectional or longitudinal, etc.) it's also not clear why logistic regression analysis was used for the analyses.*

Thank you for your thoughtful comments regarding the statistical analyses and the clarification of methods and variables used in the study.

Statistical Analyses Section: We have included a detailed explanation of all statistical analyses in the Methods section (lines 291–408), specifying the rationale for the choice of methods, the variables analyzed, and their relationships. Additionally, we have provided the relevant equations or statistical models used where appropriate to ensure transparency.

Beta Values and Odds Ratios: In the Results section (starting at line 56), both Beta values and Odds Ratios are presented: Beta values were used for analyses of continuous outcomes to quantify the linear relationship between predictors and outcomes. Odds Ratios (ORs) were calculated for binary or categorical disease outcomes to describe the relative odds of an outcome given specific exposures or independent variables.

Validation and Regression Analyses: For further validation of the MR results, we conducted analyses using the UK Biobank dataset (starting at line 162). Logistic regression analysis was then employed for disease risk assessments involving categorical outcomes (e.g., diseased or not).

We hope that this clarifies the methods and their applicability to our study, as well as the rationale for the presentation of Beta values and Odds Ratios. If further details or refinements are required, we are happy to incorporate them.

> *(4) Mendelian Randomization is an important part of the analyses done in the manuscript. It is not clear to what extent the MR assumptions are met, how the assumptions were tested, and if/what sensitivity analyses are performed; e.g. reverse MR, biological knowledge of the studied traits, etc. Can the authors explain to what extent the genetic instruments represent their targets (applicable expression/protein levels) well?*

Thank you for your insightful comments regarding the Mendelian Randomization (MR) analysis and the evaluation of its assumptions. Below, we provide additional clarification on how the MR assumptions were addressed, sensitivity analyses performed, and the representativeness of the genetic instruments (starting at line 314):

Relevance Assumption (Genetic instruments are associated with the exposure): "We identified single nucleotide polymorphisms (SNPs) associated with exposure datasets with $p < 5 \times 10^{-8}$ (10,11). In this case, 249 SNPs and 67 SNPs were selected as eligible instrumental variables (IVs) for exposures of age at menarche and age at first birth, respectively. All selected SNPs for every exposure would be clumped to avoid the linkage disequilibrium (r^2 = 0.001 and kb = 10,000)." "During the harmonization process, we aligned the alleles to the human genome reference sequence and removed incompatible SNPs. Subsequent analyses were based on the merged exposure-outcome dataset. We calculated the F statistics to quantify the strength of IVs for each exposure with a threshold of F>10 (12)."

Independence Assumption (Genetic instruments are not associated with confounders, Genetic instruments affect the outcome only through the exposure): Then we identified whether there were potential confounders of IVs associated with the outcomes based on a database of human genotype-phenotype associations, PhenoScanner V2 (13,14) (http://www.phenoscanner.medschl.cam.ac.uk/), with a threshold of $p < 1 \times 10^{-5}$. IVs associated with education, smoking, alcohol, activity, and other confounders related to outcomes would be excluded.

Sensitivity Analyses Performed: A pleiotropy test was used to check if the IVs influence the outcome through pathways other than the exposure of interest. A heterogeneity test was applied to ensure whether there is a variation in the causal effect estimates across different IVs. Significant heterogeneity test results indicate that some instruments are invalid or that the causal effect varies depending on the IVs used. MRPRESSO was applied to detect and correct potential outliers of IVs with NbDistribution = 10,000 and threshold p = 0.05. Outliers would be excluded for repeated analysis. The causal estimates were given as odds ratios (ORs) and 95% confidence intervals (CI). A leave-one-out analysis was conducted to ensure the robustness of the results by sequentially excluding each IV and confirming the direction and statistical significance of the remained remaining SNPs.

Supplemental post-GWAS analysis: Colocalization analysis (starting at line 356), Genetic correlation analysis (starting at line 366).

Our MR analysis adheres to the guidelines for causal inference in MR studies. By combining multiple sensitivity analyses and ensuring the quality of genetic instruments, we demonstrate that the results are robust and unlikely to be driven by confounding or pleiotropy.

> *(5) It is not clear what reference genome is used and if or what imputation panel is used. It is also not clear what QC steps are applied to the genotype data in order to construct the genetic instruments of MR.*

Starting in line 314, the steps of SNPs selection were included in the Methods part. "We identified single nucleotide polymorphisms (SNPs) associated with exposure datasets with $p < 5 \times 10^{-8}$ (10,11). In this case, 249 SNPs and 67 SNPs were selected as eligible instrumental variables (IVs) for exposures of age at menarche and age at first birth, respectively. All

selected SNPs for every exposure would be clumped to avoid the linkage disequilibrium (r^2 = 0.001 and kb = 10,000). Then we identified whether there were potential confounders of IVs associated with the outcomes based on a database of human genotype-phenotype associations, PhenoScanner V2 (13,14) (http://www.phenoscanner.medschl.cam.ac.uk/), with a threshold of $p < 1 \times 10^{-5}$. IVs associated with education, smoking, alcohol, activity, and other confounders related to outcomes would be excluded. During the harmonization process, we aligned the alleles to the human genome reference sequence and removed incompatible SNPs. Subsequent analyses were based on the merged exposure-outcome dataset. We calculated the F statistics to quantify the strength of IVs for each exposure with a threshold of F>10 (12). If the effect allele frequency (EAF) was missing in the primary dataset, EAF would be collected from dsSNP (https://www.ncbi.nlm.nih.gov/snp/) based on the population to calculate the F value." The SNP numbers of exposures for each outcome and F statistics results were listed in supplemental table S2.

> *(6) A code availability statement is missing. It is understandable that data cannot always be shared, but code should be openly accessible.*

We have added it to the manuscript (starting at line 410).

> ***Reviewer #2 (Recommendations for the authors):***
>
> *(1) The outcomes seem to be genotypes (lines 274-288). In MR, genotypes are used as an instrument, representing an exposure, which is then associated with an outcome that is typically observed and measured at a later moment in time than the predictors. If both exposure and outcome are genotypes it is not clear how this works in terms of causality; it would rather reflect a genetic correlation. One would expect the genotypes that function as instruments for the exposure to have a functional cascade of (age-related) effects, leading to an (age-related) outcome. From line 149 the outcomes seem to be phenotypes. Can the authors please clearly explain in each section what is analyzed, how the analyses were done, and why the analyses were done that way?*

Thank you for your insightful comment. We understand the concern regarding the use of genotypes as both exposures and outcomes and the implications this has for interpreting causality versus genetic correlation. To clarify, in our study, the outcomes analyzed in the MR framework are indeed genotypes, starting from line 47. We use genotypes as instrumental variables for exposures, which are then linked to phenotypic outcomes observed at a later stage, in line with standard MR principles.

To improve the robustness of the MR results, we validated the genetic associations in the population with phenotype data from UK Biobank (lines 162-203), and the detailed methods were listed in lines 385-408.

> *(2) Overall, the English writing is good. However, some small errors slipped in. Please check the manuscript for small grammar mistakes like in sentences 10 (punctuation) and 33 (grammar).*

Thank you for your feedback. We appreciate your careful review and attention to detail. We thoroughly rechecked the manuscript for any grammatical errors, including punctuation and sentence structure, especially in sentences 11 and 35 in revised manuscript, as suggested.

> *(3) There is currently no results and discussion section.*

The manuscript was submitted as Short Reports article type with a combined Results and Discussion section. We have added the section title of Discussion.

> *(4) Why did the authors not include SNPs associated with age at menopausal onset? See for example: https://www.nature.com/articles/s41586-021-03779-7https://urldefense.com/v3/__https://www.nature.com/articles/s41586-021-03779-7__;!!HYjtAOY1tjP_!Kl_ZKCmWOQEnvEbl46TG0TuhlsxapwvFdAFfZJkMvz8z7XhX5VEA1cT8CVvNu8xrv9k679Kl0XTrxwSajUeiXWm04X*

Thank you for your information. Our manuscript focuses on the antagonistic pleiotropy theory, which posits that inherent trade-off in natural selection, where genes beneficial for early survival and reproduction (like menarche and childbirth) may have costly consequences later. So, we only included age at menarche and age at first childbirth as exposures in our research.

> *(5) Can the authors include genetic correlations between menarche, age at first child, BMI, and preferably menopause?*

Thank you for your suggestion. We acknowledge that including genetic correlations between age at menarche, age at first childbirth, BMI, and menopause can provide valuable context to our analysis. While our current MR study sets age at menarche and age at first childbirth as exposures and menopause as the outcome, and we have already included results that account for BMI-related SNPs before and after correction, we recognize the importance of assessing genetic correlations.

To address this, we calculated the genetic correlations between these traits to provide insight into their shared genetic architecture. This analysis helps clarify whether there is a significant genetic overlap between the two exposures and between exposure and outcome, which can inform and support the interpretation of our MR results. We appreciate your suggestion and include these calculations to enhance the robustness and comprehensiveness of our study. In the genetic correlations analysis, LDSC software was applied and the genetic correlation values for all pairwise comparisons among age at menarche, age at first birth, BMI, and age at menopause onset were calculated(15,16). The results are listed in Table S6.

> *(6) Line 39-40: that is not entirely true. There is also amounting evidence that socioeconomic factors cause earlier onset of menarche through stress-related mechanisms: https://doi.org/10.1016/j.annepidem.2010.08.006https://urldefense.com/v3/__https://doi.org/10.1016/j.annepidem.2010.08.006__;!!HYjtAOY1tjP_!Kl_ZKCmWOQEnvEbl46TG0TuhlsxapwvFdAFfZJkMvz8z7XhX5VEA1cT8CVvNu8xrv9k679Kl0XTrxwSajUeiXZ4vbX0*

Thank you so much for your information. We changed it to “Considering reproductive events are partly regulated by genetic factors that can manifest the physiological outcome later in life”.

> *(7) Why did the authors choose to work with studies derived from IEU Open GWAS? as it is often does not contain the most recent and relevant GWAS for a specific trait.*

We chose to work with studies derived from the IEU Open GWAS database after careful consideration of several sources, including the GWAS Catalog database and recently published GWAS papers. Our selection criteria focused on publicly available GWAS with large sample sizes and a higher number of SNPs to ensure robust analysis. For specific traits such as late-onset Alzheimer's disease and eye aging, we used GWAS data published in scientific articles to ensure that our research reflects the latest findings in the field.

(1) Barban, N. et al. Genome-wide analysis identifies 12 loci influencing human reproductive behavior. Nat Genet 48, 1462-1472 (2016). https://doi.org/10.1038/ng.3698

(2) Tropf, F. C. et al. Hidden heritability due to heterogeneity across seven populations. Nat Hum Behav 1, 757-765 (2017). https://doi.org/10.1038/s41562-017-0195-1

(3) Stearns, S. C., Byars, S. G., Govindaraju, D. R. & Ewbank, D. Measuring selection in contemporary human populations. Nat Rev Genet 11, 611-622 (2010). https://doi.org/10.1038/nrg2831

(4) Day, F. R., Elks, C. E., Murray, A., Ong, K. K. & Perry, J. R. Puberty timing associated with diabetes, cardiovascular disease and also diverse health outcomes in men and women: the UK Biobank study. Sci Rep 5, 11208 (2015). https://doi.org/10.1038/srep11208

(5) Hollis, B. et al. Genomic analysis of male puberty timing highlights shared genetic basis with hair colour and lifespan. Nat Commun 11, 1536 (2020). https://doi.org/10.1038/s41467-020-14451-5

(6) Field, A. E. et al. Impact of overweight on the risk of developing common chronic diseases during a 10-year period. Arch Intern Med 161, 1581-1586 (2001). https://doi.org/10.1001/archinte.161.13.1581

(7) Singh, G. M. et al. The age-specific quantitative effects of metabolic risk factors on cardiovascular diseases and diabetes: a pooled analysis. PLoS One 8, e65174 (2013). https://doi.org/10.1371/journal.pone.0065174

(8) Kivimaki, M. et al. Obesity and risk of diseases associated with hallmarks of cellular ageing: a multicohort study. Lancet Healthy Longev 5, e454-e463 (2024). https://doi.org/10.1016/S2666-7568(24)00087-4

(9) Kivimaki, M. et al. Body-mass index and risk of obesity-related complex multimorbidity: an observational multicohort study. Lancet Diabetes Endocrinol 10, 253-263 (2022). https://doi.org/10.1016/S2213-8587(22)00033-X

(10) Savage, J. E. et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. Nat Genet 50, 912-919 (2018). https://doi.org/10.1038/s41588-018-0152-6

(11) Gao, X. et al. The bidirectional causal relationships of insomnia with five major psychiatric disorders: A Mendelian randomization study. Eur Psychiatry 60, 79-85 (2019). https://doi.org/10.1016/j.eurpsy.2019.05.004

(12) Burgess, S., Small, D. S. & Thompson, S. G. A review of instrumental variable estimators for Mendelian randomization. Stat Methods Med Res 26, 2333-2355 (2017). https://doi.org/10.1177/0962280215597579

(13) Staley, J. R. et al. PhenoScanner: a database of human genotype-phenotype associations. Bioinformatics 32, 3207-3209 (2016). https://doi.org/10.1093/bioinformatics/btw373

(14) Kamat, M. A. et al. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. Bioinformatics 35, 4851-4853 (2019). https://doi.org/10.1093/bioinformatics/btz469

(15) Bulik-Sullivan, B. et al. An atlas of genetic correlations across human diseases and traits. Nat Genet 47, 1236-1241 (2015). https://doi.org/10.1038/ng.3406

(16) Bulik-Sullivan, B. K. et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 47, 291-295 (2015). https://doi.org/10.1038/ng.3211

https://doi.org/10.7554/eLife.102447.2.sa0