PROTOCOL - AGE AT MENARCHE AND ADVERSE PREGNANCY OUTCOMES

CHANGE LOG

Date	Author	Notes
24/04/2024	Lizzy Aiton	Updated outcomes: • Three added due to availability in outcome datasets (very pre-term birth, post-term birth, continuous birthweight – all related to previously selected plausible outcomes) • One removed due to biological implausibility after discussions with coauthors (perinatal anaemia)
15/05/2023	Lizzy Aiton	Updated and created pdf for Github upload. Changed to include use of sex-combined BMI exposure GWAS.
01/03/2023	Lizzy Aiton	Initial version finalised.

1. BACKGROUND

Secular trends towards an earlier age at menarche have been observed globally since the nineteenth century (Harris, Prior and Koehoorn, 2008; Lehmann, Scheffler and Hermanussen, 2010; Song *et al.*, 2014). Similarly to increasing life expectancy, this demographic shift has been attributed to improved population health and nutrition (Tanner, 1968; Bosch *et al.*, 2008; Cheng *et al.*, 2012), though some authors suggest possible roles for endocrine-disrupting synthetic chemicals (Parent *et al.*, 2003) and psychosocial stress (Coall and Chisholm, 2003). Trends towards an earlier onset of female adolescence may have negative health implications (Ibitoye *et al.*, 2017), such as increased risk of cardiovascular diseases (Day *et al.*, 2015; Lee *et al.*, 2019) and endometrial cancer (Day *et al.*, 2017).

There is observational evidence that an earlier age at menarche is associated with increased risk of adverse pregnancy outcomes (APOs), including gestational diabetes mellitus (GDM) (Sun *et al.*, 2018; Wang *et al.*, 2019), gestational hypertension (Petry *et al.*, 2019), pre-term birth (Li *et al.*, 2017), miscarriage (Martin, Brinton and Hoover, 1983), ectopic pregnancy (Sandler, Wilcox and Horney, 1984), and low birth weight due to small-for-gestational age (Scholl *et al.*, 1989, p. 1).

However, whether these associations reflect a causal effect is unclear due to potential confounding by BMI, which is a risk factor for both earlier menarche (Zhou *et al.*, 2022) and APOs (Nelson, Matthews and Poston, 2010). For instance, adjustment for body mass index (BMI) attenuated the association of age at menarche with GDM (Sun *et al.*, 2018; Wang *et al.*, 2019) and with blood pressure (Petry *et al.*, 2019), but not with pre-term birth (Li *et al.*, 2017). Mendelian randomization (MR) studies have suggested that an earlier age at menarche can both cause a higher BMI in adulthood (Gill *et al.*, 2018), and be a consequence of higher childhood BMI (Mumby *et al.*, 2011). Yet there are other plausible mechanisms by which age at menarche may cause APOs. First, it may be mediated directly through increased lifetime exposure to biological processes initiated at menarche, such as estradiol fluctuations, and pro-inflammatory markers associated with severity of menstrual systems (Bertone-Johnson *et al.*, 2014). Second, it may be that both menarche and APOs are affected by an upstream factor such as insulin resistance (Li *et al.*, 2017), childhood illness (Downing and Bellis, 2009), or childhood psychosocial stress which has been associated with lower birth weight offspring (Coall and Chisholm, 2003).

MR circumvents the confounding in observational studies by exploiting genetic variants associated with the exposure of interest as a proxy for the exposure (here, age at menarche). Analogous to randomisation in a randomised clinical (Davies, Holmes and Smith, 2018), Mendel's law of independent assortment governs the random, independent inheritance of such variants during gamete formation. Multivariable MR (MVMR) extends this method to account for the effects of multiple genetically-proxied exposures on an outcome (Burgess and Thompson, 2015; Sanderson, Spiller and Bowden, 2021). The aim of this study is to investigate causality in the associations between age at menarche and APOs by triangulating evidence obtained through observational associations and MR analyses.

2. AIMS

This study aims to investigate the associations of age at menarche with adverse pregnancy outcomes and determine whether any observed associations are confounded by childhood BMI. The central research question is: is an earlier age at menarche associated with an increased risk of a range of adverse pregnancy outcomes, independently of its association with BMI?

The following specific hypotheses will be tested, in the following order.

a) Using observational multivariable regression, accounting for confounders:

1. Is an earlier age at menarche associated with adverse pregnancy outcomes?

2. Is an earlier age at menarche associated with adverse pregnancy outcomes after adjustment for adult BMI (as a proxy for pre-pubertal BMI)?

b) Using two-sample MR:

3. Does a (genetically instrumented) earlier age at menarche cause adverse pregnancy outcomes?
4. Accounting for (genetically instrumented) pre-pubertal BMI, does a (genetically instrumented) earlier age at menarche cause adverse pregnancy outcomes?
5. Accounting for (genetically instrumented) pre-pubertal BMI, is the causal effect of a (genetically instrumented) earlier age at menarche on adverse pregnancy outcomes mediated by (genetically instrumented) adult BMI?

3. DATA

3.1 SUMMARY DATA

Published summary GWAS data will be used to construct genetic instruments.

To instrument age at menarche we will use the latest GWAS, which identified 389 loci in 329,345 women of European ancestry (Day *et al.*, 2017).

Our genetic instrument for pre-pubertal BMI we will use a GWAS on the outcome of self-reported perceived body size at age 10 in the UK Biobank, validated in ALSPAC which used direct measures of pre-pubertal BMI (Richardson *et al.*, 2020). This GWAS identified 299 independent genetic variants associated with pre-pubertal BMI at genome-wide significance (n=453 169 participants of European ancestry; supplementary table 2). For adulthood BMI we will use results from the same study, which also identified 561 independent loci associated with adulthood BMI in the same sample (supplementary table 3). Though the discovery sample, UK Biobank, is one of the constituent MR-PREG datasets, these results have also recently been externally validated in HUNT where variance explained by the sex-combined polygenic risk scores was 6.7% for pre-pubertal BMI and 3.9% for adulthood BMI (n = 66,963) (Brandkvist *et al.*, 2021). Sex-combined BMI instruments were chosen over the available female-sex specific ones since sex-combined scores can explain more variance in BMI even in female-only samples (Waterfield *et al.*, 2023).

Summary data from a recent within-sibship GWAS which included adult BMI and age at menarche as outcomes will be used to identify genetic instruments for sensitivity analyses accounting for genetic confounding (Howe *et al.*, 2022).

3.1 INCLUSION CRITERIA

For the observational analysis, all datasets must include information for women on age at menarche and at least one of the following APOs of interest:

- Gestational hypertension (GH), pre-eclampsia (PE), hypertensive disorders (HDP, combined GH and PE), gestational diabetes mellitus (GDM), perinatal depression, preterm birth PTB, large- and small-for-gestational age LGA, SGA, low birthweight, high birthweight, very pre-term birth, post-term birth, and continuous birthweight.

3.2 DATASETS

The observational analysis will be conducted in ALSPAC. If possible within the project timeframe, observational analyses will also be conducted in Born in Bradford and MoBa.

For MR analysis outcome data, we will use meta-analysed associations with APOs across all MR-PREG cohorts.

Study	Observational analyst	MR analyst	Comments
ALSPAC	Lizzy	Lizzy – analysis will be conducted on MR-PREG GWAS meta-analysis results	
Born in Bradford			
MoBa			
UK Biobank			
FinnGenn			

3.3 DATA PREPARATION

Exposures and covariates

Genetic instruments will be constructed using SNPs identified as associated with age at menarche (in years), pre-pubertal BMI (in kg/m^2), and adult BMI (in kg/m^2) at the genome-wide significant level (see 3.1). Where SNPs are not available in the MR-PREG GWAS meta-analysis, we will select appropriate proxies (R^2 >0.9). We will perform clumping as appropriate to remove SNPs in linkage disequilibrium.

Outcomes

Where studies have included multiple pregnancies per woman, we will include only one pregnancy for each woman as per MR-PREG protocol. Multiple births will be excluded from the analysis.

We will use MR-PREG's definitions for all APOs and MR-PREG Stata/R scripts to construct the variables within observational datasets.

4. STATISTICAL METHODS

All analyses will be conducted in R 4.2.3.

Separate analysis will be conducted for each APO, using all participants with data available for each outcome to maximise power.

We will provide summary statistics for each cohort study describing means, medians and prevalence of each APO.

Descriptive statistics for cohorts will be provided covering age, sex, ethnicity, smoking, alcohol, parity, offspring sex, BMI, socioeconomic position (proxied by occupation class, highest level of educational attainment and/or household income) (dependent on variables available in each cohort).

4.1 OBSERVATIONAL MULTIVARIABLE REGRESSION

Observational effect estimates will be obtained for each study using multivariable logistic and linear regression as appropriate, accounting for confounders.

Models including pre-pubertal BMI as a covariate will be compared to those not accounting for it. Where pre-pubertal BMI is not available, pre-pregnancy BMI will be used as the best available proxy; there is an observed correlation between these measures: r = 0.51 for measures obtained at age 3-7 to those from age 18-37 years (Freedman *et al.*, 2004), r = 0.37 for measures obtained at age 7 and again at 33 years (Power, Lake and Cole, 1997).

The models tested will be:

1) APO ~ age at menarche + confounders

2) APO ~ age at menarche + pre-pubertal BMI/proxy + confounders

Where available, the following confounders will be adjusted for:

- Maternal grandfather's socioeconomic position (SEP) as a proxy for mother's pre-pubertal SEP
- Mother's SEP
- Ethnicity

Parity, age at delivery and offspring sex will also be adjusted for to reduce noise since these are strongly correlated to APOs.

Non-linear exposure-outcome relationships will be considered for each APO. Age at menarche will be coded as three categories (early, normative, late) based on standard deviations from the mean, as per (Roberts *et al.*, 2019), and a likelihood ratio test of model fit conducted comparing a categoric and linear coding of these categories.

Results across studies will be meta-analysed using fixed-effects or random-effects meta-analysis (if heterogeneity between studies is high (Cochrane's Q P value < 0.1)).

4.2 UNIVARIATE MENDELIAN RANDOMISATION

We will then conduct two-sample MR using the inverse variance weighted method to generate causal effect estimates, using the TwoSampleMR R package (Hemani *et al.*, 2018).

The model tested will be:

1) APO ~ (genetically instrumented) age at menarche

Sample overlap between GWAS used to construct instruments and MR-PREG may generate bias due to overfitting, biasing causal estimates towards the observational effect estimate of exposure on outcome (Burgess, Davies and Thompson, 2016). Overlap is as follows:

- Day et al.'s age at menarche GWAS included UK Biobank and ALSPAC, alongside other studies.
- Richardson et al.'s pre-pubertal and adulthood BMI GWAS used only UK Biobank, but has been externally validated in HUNT.

Sample overlap will be quantified as an estimated percentage, and we will rerun analyses using the MRlap method to adjust for sample overlap as a sensitivity analysis (Mounier and Kutalik, 2022).

MR assumptions will be tested as described in Table 1, below.

Table 1. MR assumptions tested in sensitivity analyses.

Assumption	**Approach**
IV1: Genetic instrument is strongly associated with exposure	Estimate r^2 and F from summary data (Bowden *et al.*, 2016; Burgess, Davies and Thompson, 2016).
IV2: Genetic instrument is not confounded by population stratification, dynastic effects, assortative mating, or transmission ratio distortion.	To account for population stratification, MR-PREG GWAS have already been performed stratified by ethnicity and adjusting for PC's. Analysis will be restricted to white European ancestry. (With potential for future work to repeat this analysis stratified by ancestry.)

	Summary data from a recent within-sibship GWAS including BMI and age at menarche will be used as alternative genetic instruments unbiased by demographic and indirect genetic effects (Howe *et al.*, 2022).
IV3: Genetic instrument is independent of outcome, given exposure and confounders (no horizontal pleiotropy).	Examine heterogeneity between SNP causal estimates with Cochrane's Q statistic (Greco M *et al.*, 2015) and leave out one analysis. We will examine asymmetry with funnel plots. Include sensitivity analyses using methods such as MR-Egger (Burgess and Thompson, 2017), weighted-median and weighted-mode analysis. Analyses adjusting for offspring and father genotype will also be conducted.

Meta-analysed MR-PREG outcome results will be used in the first instance, but individual study cohorts will be considered in sensitivity analyses e.g. leave-one-out analysis.

4.3 MULTIVARIABLE MENDELIAN RANDOMISATION

MVMR will be conducted on the following regression models, restricted to APOs where a causal effect was observed in the univariable model.

First, to account for pre-pubertal BMI as a confounder:

1) APOs ~ (genetically instrumented) pre-pubertal BMI + (genetically instrumented) age at menarche

For APOs where a causal effect is still seen in the first model above, we plan to test the following model to ascertain whether adulthood BMI acts as a mediator:

2) APOs ~ (genetically instrumented) pre-pubertal BMI + (genetically instrumented) adulthood BMI + (genetically instrumented) age at menarche

The MVMR R package will be used, following the workflow outlined in Sanderson, Spiller and Bowden (2021) (also described: https://wspiller.github.io/MVMR/articles/MVMR.html) . First, a covariance matrix of the effects of all instrumented variants on both exposures of interest will be generated or approximated within ALSPAC.

Sensitivity analyses will be conducted as described for univariate MR (table 1). Additionally, since in MVMR the estimated effect of each exposure may be biased either towards or away from the null due to weak instrument bias, we will test for this with the F_{SW} statistic (Sanderson, Spiller and Bowden, 2021).

5. BIBLIOGRAPHY

Bertone-Johnson, E.R. *et al.* (2014) 'Association of inflammation markers with menstrual symptom severity and premenstrual syndrome in young women', *Human Reproduction*, 29(9), pp. 1987–1994. Available at: https://doi.org/10.1093/humrep/deu170.

Bosch, A.M. *et al.* (2008) 'Association between Age at Menarche and Early-life Nutritional Status in Rural Bangladesh', *Journal of Biosocial Science*, 40(2), pp. 223–237. Available at: https://doi.org/10.1017/S0021932007002490.

Bowden, J. *et al.* (2016) 'Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic', *International Journal of Epidemiology*, 45(6), pp. 1961–1974. Available at: https://doi.org/10.1093/ije/dyw220.

Brandkvist, M. *et al.* (2021) 'Separating the genetics of childhood and adult obesity: a validation study of genetic scores for body mass index in adolescence and adulthood in the HUNT Study', *Human Molecular Genetics*, 29(24), pp. 3966–3973. Available at: https://doi.org/10.1093/hmg/ddaa256.

Burgess, S., Davies, N.M. and Thompson, S.G. (2016) 'Bias due to participant overlap in two-sample Mendelian randomization', *Genetic Epidemiology*, 40(7), pp. 597–608. Available at: https://doi.org/10.1002/gepi.21998.

Burgess, S. and Thompson, S.G. (2015) 'Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects', *American Journal of Epidemiology*, 181(4), pp. 251–260. Available at: https://doi.org/10.1093/aje/kwu283.

Burgess, S. and Thompson, S.G. (2017) 'Interpreting findings from Mendelian randomization using the MR-Egger method', *European Journal of Epidemiology*, 32(5), pp. 377–389. Available at: https://doi.org/10.1007/s10654-017-0255-x.

Cheng, G. *et al.* (2012) 'Beyond overweight: nutrition as an important lifestyle factor influencing timing of puberty', *Nutrition Reviews*, 70(3), pp. 133–152. Available at: https://doi.org/10.1111/j.1753-4887.2011.00461.x.

Coall, D.A. and Chisholm, J.S. (2003) 'Evolutionary perspectives on pregnancy: maternal age at menarche and infant birth weight', *Social Science & Medicine (1982)*, 57(10), pp. 1771–1781. Available at: https://doi.org/10.1016/s0277-9536(03)00022-4.

Davies, N.M., Holmes, M.V. and Smith, G.D. (2018) 'Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians', *BMJ*, 362, p. k601. Available at: https://doi.org/10.1136/bmj.k601.

Day, F.R. *et al.* (2015) 'Puberty timing associated with diabetes, cardiovascular disease and also diverse health outcomes in men and women: the UK Biobank study', *Scientific Reports*, 5(1), p. 11208. Available at: https://doi.org/10.1038/srep11208.

Day, F.R. *et al.* (2017) 'Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk', *Nature Genetics*, 49(6), pp. 834–841. Available at: https://doi.org/10.1038/ng.3841.

Downing, J. and Bellis, M.A. (2009) 'Early pubertal onset and its relationship with sexual risk taking, substance use and anti-social behaviour: a preliminary cross-sectional study', *BMC Public Health*, 9(1), p. 446. Available at: https://doi.org/10.1186/1471-2458-9-446.

Freedman, D.S. *et al.* (2004) 'Inter-relationships among childhood BMI, childhood height, and adult obesity: the Bogalusa Heart Study', *International Journal of Obesity*, 28(1), pp. 10–16. Available at: https://doi.org/10.1038/sj.ijo.0802544.

Gill, D. *et al.* (2018) 'Age at menarche and adult body mass index: a Mendelian randomization study', *International Journal of Obesity (2005)*, 42(9), pp. 1574–1581. Available at: https://doi.org/10.1038/s41366-018-0048-7.

Greco M, F.D. *et al.* (2015) 'Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome', *Statistics in Medicine*, 34(21), pp. 2926–2940. Available at: https://doi.org/10.1002/sim.6522.

Harris, M.A., Prior, J.C. and Koehoorn, M. (2008) 'Age at Menarche in the Canadian Population: Secular Trends and Relationship to Adulthood BMI', *Journal of Adolescent Health*, 43(6), pp. 548–554. Available at: https://doi.org/10.1016/j.jadohealth.2008.07.017.

Hemani, G. *et al.* (2018) 'The MR-Base platform supports systematic causal inference across the human phenome', *eLife*, 7, p. e34408. Available at: https://doi.org/10.7554/eLife.34408.

Howe, L.J. *et al.* (2022) 'Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects', *Nature Genetics*, 54(5), pp. 581–592. Available at: https://doi.org/10.1038/s41588-022-01062-7.

Ibitoye, M. *et al.* (2017) 'Early menarche: A systematic review of its effect on sexual and reproductive health in low- and middle-income countries', *PLOS ONE*, 12(6), p. e0178884. Available at: https://doi.org/10.1371/journal.pone.0178884.

Lee, J.J. *et al.* (2019) 'Age at Menarche and Risk of Cardiovascular Disease Outcomes: Findings From the National Heart Lung and Blood Institute-Sponsored Women's Ischemia Syndrome Evaluation', *Journal of the American Heart Association*, 8(12), p. e012406. Available at: https://doi.org/10.1161/JAHA.119.012406.

Lehmann, A., Scheffler, C. and Hermanussen, M. (2010) 'The variation in age at menarche: an indicator of historic developmental tempo', *Anthropologischer Anzeiger*, 68(1), pp. 85–99.

Li, H. *et al.* (2017) 'Age at menarche and prevalence of preterm birth: Results from the Healthy Baby Cohort study', *Scientific Reports*, 7(1), p. 12594. Available at: https://doi.org/10.1038/s41598-017-12817-2.

Martin, E., Brinton, L. and Hoover, R. (1983) 'Menarcheal Age and Miscarriage', *American Journal of Epidemiology*, 117(5), pp. 634–636. Available at: https://doi.org/10.1093/oxfordjournals.aje.a113587.

Mounier, N. and Kutalik, Z. (2022) 'Bias correction for inverse variance weighting Mendelian randomization'. bioRxiv, p. 2021.03.26.437168. Available at: https://doi.org/10.1101/2021.03.26.437168.

Mumby, H.S. *et al.* (2011) 'Mendelian Randomisation Study of Childhood BMI and Early Menarche', *Journal of Obesity*, 2011, p. 180729. Available at: https://doi.org/10.1155/2011/180729.

Nelson, S.M., Matthews, P. and Poston, L. (2010) 'Maternal metabolism and obesity: modifiable determinants of pregnancy outcome', *Human Reproduction Update*, 16(3), pp. 255–275. Available at: https://doi.org/10.1093/humupd/dmp050.

Parent, A.-S. *et al.* (2003) 'The Timing of Normal Puberty and the Age Limits of Sexual Precocity: Variations around the World, Secular Trends, and Changes after Migration', *Endocrine Reviews*, 24(5), pp. 668–693. Available at: https://doi.org/10.1210/er.2002-0019.

Petry, C.J. *et al.* (2019) 'Age at Menarche and Blood Pressure in Pregnancy', *Pregnancy hypertension*, 15, pp. 134–140. Available at: https://doi.org/10.1016/j.preghy.2019.01.004.

Power, C., Lake, J.K. and Cole, T.J. (1997) 'Body mass index and height from childhood to adulthood in the 1958 British born cohort', *The American Journal of Clinical Nutrition*, 66(5), pp. 1094–1101. Available at: https://doi.org/10.1093/ajcn/66.5.1094.

Richardson, T.G. *et al.* (2020) 'Use of genetic variation to separate the effects of early and later life adiposity on disease risk: mendelian randomisation study', *BMJ*, 369, p. m1203. Available at: https://doi.org/10.1136/bmj.m1203.

Roberts, E. *et al.* (2019) 'Timing of menarche and self-harm in adolescence and adulthood: a population-based cohort study', *Psychological Medicine*, 50(12), pp. 2010–2018. Available at: https://doi.org/10.1017/S0033291719002095.

Sanderson, E., Spiller, W. and Bowden, J. (2021) 'Testing and correcting for weak and pleiotropic instruments in two-sample multivariable Mendelian randomization', *Statistics in Medicine*, 40(25), pp. 5434–5452. Available at: https://doi.org/10.1002/sim.9133.

Sandler, D.P., Wilcox, A.J. and Horney, L. (1984) 'Age at Menarche and Subsequent Reproductive Events', *American Journal of Epidemiology*, 119(5), pp. 765–774. Available at: https://doi.org/10.1093/oxfordjournals.aje.a113797.

Scholl, T.O. *et al.* (1989) 'Effects of early maturation on fetal growth', *Annals of Human Biology*, 16(4), pp. 335–345. Available at: https://doi.org/10.1080/03014468900000462.

Song, Y. *et al.* (2014) 'Trends of Age at Menarche and Association with Body Mass Index in Chinese School-Aged Girls, 1985-2010', *The Journal of Pediatrics*, 165(6), pp. 1172-1177.e1. Available at: https://doi.org/10.1016/j.jpeds.2014.08.013.

Sun, X. *et al.* (2018) 'Age at menarche and the risk of gestational diabetes mellitus: a systematic review and meta-analysis', *Endocrine*, 61(2), pp. 204–209. Available at: https://doi.org/10.1007/s12020-018-1581-9.

Tanner, J.M. (1968) 'Earlier Maturation in Man', *Scientific American*, 218(1), pp. 21–27.

Wang, H. *et al.* (2019) 'Genotype-by-environment interactions inferred from genetic effects on phenotypic variability in the UK Biobank', *Science Advances*, 5(8), p. eaaw3538. Available at: https://doi.org/10.1126/sciadv.aaw3538.

Waterfield, S. *et al.* (2023) 'Life course effects of genetic susceptibility to higher body size on body fat and lean mass: prospective cohort study', *International Journal of Epidemiology*, p. dyad029. Available at: https://doi.org/10.1093/ije/dyad029.

Zhou, X. *et al.* (2022) 'Overweight/Obesity in Childhood and the Risk of Early Puberty: A Systematic Review and Meta-Analysis', *Frontiers in Pediatrics*, 10. Available at: https://www.frontiersin.org/articles/10.3389/fped.2022.795596 (Accessed: 24 February 2023).