Additional file 1: Table S1. Design table with a summary of questions, hypotheses, power analyses, analysis plan, and interpretation of different potential outcomes.

Question	Hypothesis	Power analysis	Analysis plan	Interpretation of different outcomes
To what extent is age at menarche associated with adolescent depression?	H1a) Earlier age at menarche will be associated with elevated depressive symptoms at age 14.	Power was calculated in R based on simulated data (1000 replicates). For this analysis, we simulated a depressive symptoms variable (M = 5.71, SD = 4.93; values taken from Sequeira et al. Results indicated 95% power to detect an effect of Cohen's $D \geq 0.08$ at N = 12,000 and ≥ 0.06 at N = 13,000 for the association between age at menarche and depressive symptoms.	Linear regression models with age of menarche as the independent variable and depressive symptoms at age 14 as the dependent variable. For this analysis, we will also add depressive symptoms at age 8 ("pre-pubertal symptoms") as a covariate. Assess whether the 90% CIs for the association between age at menarche and depressive symptoms overlaps with D = 0.23 (the smallest effect size of interest; SESOI).	A negative and statistically significant association between age at menarche and depressive symptoms at age 14 will suggest that adolescent girls with earlier age at menarche show elevated depressive symptoms. A non-significant association will suggest an absence of evidence for this. If the 90% CIs for the association is lower than the SESOI, this suggests that earlier age at menarche is not associated with a meaningful increase in depressive symptoms at age 14. If the 90% CIs overlap with the SESOI, this suggests that the association is of a meaningful magnitude. Hypothesis 1a will be supported if 1) the coefficient for the effect of age at menarche on 14-year depressive symptoms is significantly less than zero (one-tailed test; alpha 5%) in the pre-pubertal symptoms-adjusted model; and 2) we fail to reject the null hypothesis that this effect in the population is at least as large as the SESOI (one-tailed test; alpha 5%). If we can reject an effect size at least as large as the SESOI, then the effect will be declared practically equivalent to zero, irrespective of the result of the NHST. We would explicitly remain undecided on the hypothesis if we cannot reject an effect of exactly zero (in the NHST) *or* an effect at least as large as the SESOI. Both these conditions apply to all remaining hypotheses.

Question	Hypothesis	Power analysis	Analysis plan	Interpretation of different outcomes
To what extent is age at menarche associated with adolescent depression?	H1b) Earlier age at menarche will be associated with higher rates of depression diagnoses during adolescence.	For this power analysis, we simulated a binary depression diagnosis variable, based on a pre-specified prevalence and association with age at menarche (varied across scenarios at 2%, 6%, 10%, 14% and $D = 0$ to -0.26 in increments of 0.02). Results indicated 95% power to detect an effect of Cohen's $D \geq 0.14$ at both $N = 12,000$ and $N = 13,000$ for depression prevalence of 6% or higher.	Logistic regression models with covariates and age of menarche as independent variables and depression diagnoses during adolescence (age 10-17) as the dependent variable. For this analysis, we will also add depression diagnoses during childhood (age 0-8; "pre-pubertal depression status") as a covariate. Assess whether the 90% CIs for the association between age at menarche and depression diagnoses overlaps with $D = 0.23$ (the SESOI).	A negative and statistically significant association between age at menarche and depression diagnoses during adolescence will suggest that adolescent girls with earlier age at menarche show higher rates of depression. A non-significant association will suggest an absence of evidence for this. If the 90% CIs for the association is lower than the SESOI, this suggests that earlier age at menarche is not associated with a meaningful increase in depression diagnoses during adolescence. If the 90% CIs overlap with the SESOI, this suggests that the association is of a meaningful magnitude. Hypothesis 1b will be supported if 1) the coefficient for the effect of age at menarche on odds of depression diagnoses is significantly less than zero (one-tailed test; alpha 5%) in the pre-pubertal depression status-adjusted model; and 2) we fail to reject the null hypothesis that this effect in the population is at least as large as the SESOI (one-tailed test; alpha 5%).

Question	Hypothesis	Power analysis	Analysis plan	Interpretation of different outcomes
Does age at menarche associate with symptoms or diagnoses in other domains of mental health, independent of depression?	H2.1-4a) Age at menarche will be associated with symptoms of anxiety (H2.1a), CD (H2.2a), ODD (H2.3a), and ADHD (H2.4a) at age 14, independent of depressive symptoms.	For this power analysis, we simulated anxiety, CD, ODD, and ADHD symptoms, with distributions based on the same variables in MoBa data at 8 years. Results indicated 95% power to detect an effect of Cohen's $D \geq$ 0.12 at both simulated sample sizes (N = 12,000 and N = 13,000).	Separate multiple linear regression models with covariates and self-reported age at menarche included as independent variables, and either symptoms of anxiety, CD, ODD, or ADHD as the dependent variable. Add depressive symptoms as a covariate in each model to examine whether any associations are independent of co-occurring depressive symptoms. Then add a measure of each domain at age 8 as a covariate in the age 14 model for that domain, to examine whether associations are independent of pre-pubertal symptoms ("fully adjusted" models). Assess whether the 90% CIs for the associations between age at menarche and any of the mental health domains lie between -0.22 - 0.22 (equivalence bounds).	Any association between age at menarche and anxiety, CD, ODD, or ADHD that significantly differs from zero at age 14 will suggest that there is an association with age at menarche in that domain. No significant associations will suggest an absence of evidence for this. If the 90% CIs for all associations fall entirely within the equivalence bounds in the fully adjusted models, this suggests that age at menarche is not associated with meaningful variation in other domains independent of depression. If the CIs for any association is not entirely within the equivalence bounds, this suggests that there is an association in that domain that is of a meaningful magnitude. Each of the hypotheses 2.1-4a will be supported if 1) the coefficient for the association between age at menarche and that domain of 14-year symptoms in the fully adjusted model is different from zero (two-tailed tests, 5% alpha); and 2) we fail to reject the null hypothesis that the association in the population is at least as extreme as the SESOI in either direction (two one-tailed tests, 5% alpha).

Question	Hypothesis	Power analysis	Analysis plan	Interpretation of different outcomes
Does age at menarche associate with symptoms or diagnoses in other domains of mental health, independent of depression?	H2.1-3b) Age at menarche will be associated with diagnoses in other domains: anxiety disorders (H2.1b); conduct disorders (H2.2b), including CD and ODD; and ADHD (H2.3b) during adolescence, independent of depressive disorders.	This power analysis was similar to H2.1-4a above, apart from different equivalence bounds and the use of two-tailed tests. Results indicated 95% power to detect an association of Cohen's $D \geq 0.20$ at $N = 12{,}000$ and ≥ 0.18 at $N = 13{,}000$ between age at menarche and diagnoses of anxiety, ADHD, and conduct disorders (CD and ODD), for diagnosis prevalence of 2% or higher.	Separate multivariate logistic regression models with covariates and self-reported age at menarche included as independent variables and either diagnoses of anxiety, ADHD, or conduct disorders as the dependent variable. Add depression diagnostic status (age 10-17) as a covariate in each model to examine whether any associations are independent of comorbid depression. Then add pre-pubertal diagnostic status (age 0-8) as a covariate in the age 14 model for that domain, to examine whether associations are independent of prior diagnoses ("fully adjusted" models). Assess whether the 90% CIs for the associations between age at menarche and any of the mental health domains lie between -0.22 - 0.22 (equivalence bounds).	Any association between age at menarche and anxiety disorders, ADHD or conduct disorders will suggest that there is an association with age at menarche in that domain. No significant associations will suggest an absence of evidence for this. If the 90% CIs for all associations fall entirely within the equivalence bounds in the fully adjusted models, this suggests that age at menarche is not associated with meaningful variation in other domains independent of depression. If the 90% CIs for any association is not entirely within the equivalence bounds, this suggests that there is an association in that domain that is of a meaningful magnitude. Each of the hypotheses 2.1-3b will be supported if 1) the coefficient for the association between age at menarche and diagnoses in that domain in the fully adjusted model is different from zero (two-tailed tests, 5% alpha); and 2) we fail to reject the null hypothesis that the association in the population is at least as extreme as the SESOI in either direction (two one-tailed tests, 5% alpha).

Question	Hypothesis	Power analysis	Analysis plan	Interpretation of different outcomes
What is the evidence for a causal link between age at menarche and depression?	H3a) Earlier age at menarche will result in elevated depressive symptoms at age 14.	Power was calculated for simulated genetic instruments with an R^2 for the instrument-exposure association of 0.05, 0.075 and 0.1. Results indicated 95% power to detect an average causal effect (using two-stage least squares regression; 2SLS) of Cohen's $D \geq 0.2$ when the R^2 of the instrument is 0.075 or above at either simulated sample size ($N = 9,500$ and $N = 10,500$).	2SLS regression with the genetic risk score for age at menarche as the instrument, reported age at menarche as the exposure, and depressive symptoms at age 14 as the outcome. 2SLS regression with depressive symptoms at age 8 as the outcome, as a negative control outcome. Assess whether the 90% CIs for the MR estimates of the association between age at menarche and depressive symptoms overlaps with $D = 0.25$ (the SESOI). Two-sample multivariable MR (MVMR) with genetic risk scores for age at menarche, childhood body size and adult body mass index (BMI) as instruments, to estimate the direct effect of age at menarche on depressive symptoms. Two-sample MR sensitivity analyses: MR-Egger, MR-PRESSO, weighted median and contamination mixture.	In the 2SLS regression, a negative and statistically significant estimate of the link between genetically predicted age at menarche and depressive symptoms at age 14 will suggest that earlier age at menarche is causally linked with adolescent depressive symptoms. A non-significant association will suggest an absence of evidence for this. If the 90% CIs for the causal estimate are lower than the SESOI, this suggests that earlier genetically predicted age at menarche is not associated with a meaningful increase in depressive symptoms at age 14. If the 90% CIs overlap with the SESOI, this suggests that the causal relationship is of a meaningful magnitude. Hypothesis 3a will be supported if 1) the coefficient for the causal effect of age at menarche on 14-year depressive symptoms is significantly less than zero (one-tailed test; alpha 5%); 2) we fail to reject the null hypothesis that this causal effect in the population is at least as large as the SESOI (one-tailed test; alpha 5%); and 3) we fail to reject the null hypothesis that this causal effect in the population is at least as large as the upper bound of the negative control (8-year) estimate (one-tailed test; alpha 5%). If the hypothesis is supported, consistent results across the two-sample MR sensitivity analyses will provide further evidence of causality.

Question	Hypothesis	Power analysis	Analysis plan	Interpretation of different outcomes
What is the evidence for a causal link between age at menarche and depression?	H3b) Earlier age at menarche will result in higher rates of depression diagnoses during adolescence.	Power was calculated for simulated genetic instruments with an R^2 for the instrument-exposure association of 0.05, 0.075 and 0.1. Results indicated 95% power to detect a causal effect (using logistic 2SLS MR with robust standard errors) of Cohen's $D \geq 0.24$ when the R^2 of the instrument is ≥ 0.075 and depression prevalence $\geq 10\%$ at either simulated sample size ($N = 9{,}500$ and $N = 10{,}500$).	2SLS regression with the genetic risk score for age at menarche as the instrument, self-reported age at menarche as the exposure, and depression diagnoses during adolescence (age 10-17) as the outcome. 2SLS regression with depression diagnoses during childhood (age 0-8) as the outcome, as a negative control outcome. Assess whether the 90% CIs for the MR estimates of the association between age at menarche and depression diagnoses overlaps with $D = 0.25$ (the SESOI) Two-sample MVMR with genetic risk scores for age at menarche, childhood body size and adult BMI as instruments, to estimate the direct effect of age at menarche on depression diagnoses. Two-sample MR sensitivity analyses: MR-Egger, MR-PRESSO, weighted median and contamination mixture.	In the 2SLS regression, a negative and statistically significant association between genetically predicted age at menarche and depression diagnoses during adolescence will suggest that earlier age at menarche causes higher rates of adolescent depression. A non-significant association will suggest an absence of evidence for this. If the 90% CI for the causal estimate is lower than the SESOI, this suggests that earlier genetically predicted age at menarche is not associated with a meaningful increase in depression diagnoses during adolescence. If the 90% CI overlaps with the SESOI, this suggests that the causal relationship is of a meaningful magnitude. Hypothesis 3b will be supported if 1) the coefficient for the causal effect of age at menarche on depressive disorders is significantly less than zero (one-tailed test; alpha 5%); and 2) we fail to reject the null hypothesis that this causal effect in the population is at least as large as the SESOI (one-tailed test; alpha 5%). If the hypothesis is supported, consistent results across the two-sample MR sensitivity analyses will provide further evidence of causality.

Question	Hypothesis	Power analysis	Analysis plan	Interpretation of different outcomes
Is there evidence of causal links between age at menarche and other domains of mental health?	H4.1-4a) Age at menarche will show a causal relationship with symptoms in other domains: anxiety (H4.1a); CD (H4.2a); ODD (H4.3a); and ADHD (H4.4a) at age 14.	The simulations for this power analysis were essentially identical to H3a (but with different equivalence bounds and the use of two-tailed tests). Results indicated 95% power to detect an average causal effect (using 2SLS MR) of Cohen's $D \geq 0.2$ when the R^2 of the instrument is 0.10 at either simulated sample size and 80% power to detect Cohen's $D \geq 0.2$ with the R^2 of the instrument at ≥ 0.075 for $N = 9500$, as well as with the R^2 of the instrument at ≥ 0.05 for $N = 10500$.	2SLS regression model with covariates and genetically predicted age at menarche included as independent variables and either symptoms of anxiety, CD, ODD, or ADHD at age 14 as the dependent variable.

Run the same models with other symptom domains at age 8 as dependent variables, as negative control outcomes.

Assess whether the 90% CIs for the associations between genetically predicted age at menarche and symptoms in any of the mental health domains lie between -0.20 - 0.20 (equivalence bounds).

Two-sample MR sensitivity analyses: MR-Egger, MR-PRESSO, weighted median and contamination mixture. | Any significant links between age at menarche and symptoms of anxiety, CD, ODD, or ADHD at age 14 will suggest that causal effects of age at menarche extend to those other symptom domains. Non-significant estimates will suggest an absence of evidence for this.

If the 90% CIs for a relationship falls entirely within the equivalence bounds, this suggests that the effect of genetically predicted age at menarche does not extend to that domain. If the CIs for a relationship are not entirely within the equivalence bounds, this suggests the causal effect of age at menarche is of a meaningful magnitude in that domain.

Each of the hypotheses 4.1-4a will be supported if 1) the coefficient for the effect of age at menarche on a domain of 14-year symptoms is different from zero (two-tailed tests, 5% alpha); 2) we fail to reject the null hypothesis that the causal effect in the population is at least as extreme as the SESOI in either direction (two one-tailed tests, 5% alpha); and 3) we fail to reject the null hypothesis that this causal effect in the population is at least as large as the upper bound of the negative control (8-year) estimate (one-tailed test; alpha 5%).

If a hypothesis is supported, consistent results across the two-sample MR sensitivity analyses will provide further evidence of causality. |

Question	Hypothesis	Power analysis	Analysis plan	Interpretation of different outcomes
Is there evidence of causal links between age at menarche and other domains of mental health?	H4.1-3b) Age at menarche will show a causal relationship with rates of diagnoses in other domains: anxiety disorders (H.4.1b); conduct disorders (H4.2b), including CD and ODD; and ADHD (H4.3b) during adolescence.	The simulations for this power analysis were essentially identical to H3b (but with different equivalence bounds and the use of two-tailed tests). Results indicated 95% power to detect a causal effect (using logistic 2SLS MR with robust standard errors) of Cohen's $D \geq 0.26$ when the R^2 of the instrument is ≥ 0.075 and diagnosis prevalence is 14% in either simulated sample size ($N = 9,500$ and $N = 10500$), and 80% power for Cohen's $D \geq 0.20$ in the same scenarios.	Multivariate logistic regression model with covariates included as independent variables and either diagnoses of anxiety, conduct disorders, or ADHD during adolescence (age 10-17) as the dependent variable. Run the same models with diagnoses during childhood (age 0-8) as dependent variables, as negative control outcomes. Assess whether the 90% CIs for the associations between genetically predicted age at menarche and diagnoses in any of the mental health domains lie between -0.20 - 0.20 (equivalence bounds). Two-sample MR sensitivity analyses: MR-Egger, MR-PRESSO, weighted median and contamination mixture.	Any significant links between age at menarche and anxiety disorders, conduct disorders, or ADHD during adolescence will suggest that causal relationships with age at menarche extend to those other mental health domains. Non-significant estimates will suggest an absence of evidence for this. If the 90% CIs for a relationship fall entirely within the equivalence bounds, this suggests that the effect of genetically predicted age at menarche does not extend to that domain. If the CIs for a relationship are not entirely within the equivalence bounds, this suggests the causal effect of age at menarche is of a meaningful magnitude in that domain. Each of hypotheses 4.1-3b will be supported if 1) the coefficient for the effect of age at menarche on that diagnosis is different from zero (two-tailed tests, 5% alpha); and 2) we fail to reject the null hypothesis that the causal effect in the population is at least as extreme as the SESOI in either direction (two one-tailed tests, 5% alpha). If a hypothesis is supported, consistent results across the two-sample MR sensitivity analyses will provide further evidence of causality.