

OPEN

Age of puberty and Sleep duration: Observational and Mendelian randomization study

Jiao Wang¹, Man Ki Kwok¹, Shiu Lun Au Yeung¹, Jie Zhao¹, Albert Martin Li², Hugh Simon Lam², Gabriel Matthew Leung¹ & Catherine Mary Schooling^{1,3*}

Earlier age of puberty has detrimental consequences for many aspects of health. Here, for the first time, we assessed the association of earlier puberty with sleep duration observationally and with validation using Mendelian Randomization. In the “Children of 1997” birth cohort ($n = 8,327$), we used adjusted multivariable logistic regression to assess the associations of each clinically assessed marker of earlier puberty with self-report sleep duration in adolescence. Using two-sample MR, we assessed the effect of earlier puberty timing based on 203 single nucleotide polymorphisms applied to genome wide association studies of sleep duration in adults ($n = 335,410$). In “Children of 1997”, cross-sectionally, older age of menarche was associated with longer (9+ hours) sleep duration [odds ratio (OR) 1.11, 95% confidence interval (CI) 1.01 to 1.21] at 13.5 years. The other earlier puberty markers were unrelated to sleep duration. Using inverse variance weighting, later of age at menarche increased adult sleep duration [0.020 per category, 95% CI 0.006 to 0.034]. This study demonstrated a causal effect of age at menarche on adult sleep duration, since age of menarche also affects obesity, our novel finding may be relevant to the observed relation of sleep duration with obesity and poor health.

Sleep serves an important restorative function in humans. However, sleep duration has decreased for children and adolescents in the last 20 years^{1,2}. Short sleep duration is associated with a wide range of poor health outcomes including stress³, headache⁴, cardiovascular disease⁵, diabetes mellitus⁶ and even cancer⁷. Short sleep is also associated with screen time⁸, low socioeconomic position (SEP)⁹ and sedentary time¹⁰, making it difficult to ascertain whether these associations are causal or confounded. Sleep patterns change significantly at puberty with marked differences by sex^{11,12}, suggesting puberty may be a critical stage for the development of sleep patterns. Earlier puberty is associated with many non-communicable diseases¹³. However a Mendelian Randomization study was not completely supportive of sleep as causing all these outcomes¹⁴. As such sleep may also be a consequence of other attributes, such as pubertal timing, which have causal effects on some non-communicable diseases. Notably, puberty affecting sleep duration falls within the emerging paradigm of evolutionary public health, i.e., considering health within the well-established paradigm of natural selection favoring reproductive success rather than wellbeing¹⁵. Laboratory studies suggested sex hormones affect sleep, potentially with sex-specific consequences. For instance, a randomized controlled trial (RCT) found testosterone shortened sleep duration in older men¹⁶, while estrogen plus progestin has been associated with better sleep in postmenopausal women¹⁷.

Observationally earlier menarche is associated with sleep disorders in adolescents from Western and Chinese settings^{18,19}. However, observational studies are hard to interpret because they could be confounded by factors such as socio-economic position (SEP) and lifestyle. In this situation where experimental evidence is lacking Mendelian Randomization (MR), i.e., instrumental variable analysis with genetic instruments, may provide a way forward. Given genetic variants are allocated randomly at conception, MR, as a quasi-experimental study design, is less susceptible to confounding and so provides an alternative means of assessing the causal effect of puberty on sleep duration. Several studies using this approach have recently clarified the associations of age of menarche with adolescent depression²⁰, time spent in education²¹ and adult body mass index¹³, but no MR study has assessed the causal effects of earlier puberty on adult sleep duration¹⁴.

To clarify the effect of pubertal timing on sleep duration, we first conducted an observational study to assess associations of clinically assessed age of puberty with self-reported longer (9+ hours) compared to shorter

¹School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China.

²Department of Pediatrics, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong SAR, China. ³City University of New York, School of Public Health and Health Policy, New York, USA. *email: cms1@hku.hk

Characteristic	Classification	n	Onset of Breast/Genitalia Mean (SD), y	P value	n	Onset of Pubic Hair Development Mean (SD), y	P value	n	Age of Menarche/Voice breaking, Mean (SD), y	P value
Gender				<0.01			<0.01			<0.01
Boys		1921	10.88 (1.08)		1428	11.48 (1.09)		1781	13.08 (1.18)	
Girls		2961	9.60 (1.28)		1736	10.56 (1.03)		3177	11.94 (1.08)	
Parents' birthplace				<0.01			<0.01			0.109
Both parents migrant		1156	10.00 (1.33)		716	10.86 (1.07)		1244	12.37 (1.26)	
One parent migrant		980	10.03 (1.42)		638	10.90 (1.16)		1055	12.30 (1.23)	
Both parents Hong Kong		2608	10.18 (1.34)		1717	11.06 (1.17)		2555	12.39 (1.25)	
Highest parental occupation				<0.01			0.02			<0.01
I (professional)		1046	10.2 (1.38)		691	11.06 (1.20)		1089	12.45 (1.26)	
II (managerial)		625	10.05 (1.30)		403	11.00 (1.14)		639	12.33 (1.24)	
IIINM (nonmanual skilled)		1298	10.14 (1.39)		855	10.97 (1.16)		1262	12.34 (1.21)	
IIIM (manual skilled)		700	10.03 (1.36)		454	10.94 (1.11)		748	12.35 (1.26)	
IV (semi-skilled)		404	9.96 (1.28)		261	10.87 (1.17)		420	12.26 (1.22)	
V (unskilled)		155	9.95 (1.30)		99	10.98 (1.08)		141	12.29 (1.30)	
Household income per head in quintiles				0.04			0.01			0.03
1st quintile (HK\$ 1751 ± 413)		806	10.03 (1.38)		519	10.87 (1.11)		835	12.26 (1.23)	
2nd quintile (HK\$ 2856 ± 325)		899	10.09 (1.40)		570	10.92 (1.21)		868	12.34 (1.27)	
3rd quintile (HK\$ 4362 ± 556)		845	10.06 (1.33)		555	10.99 (1.12)		897	12.38 (1.22)	
4th quintile (HK\$ 6822 ± 886)		901	10.16 (1.35)		616	11.06 (1.14)		892	12.36 (1.24)	
5th quintile (HK\$ 14850 ± 16050)		900	10.21 (1.34)		564	11.07 (1.22)		908	12.45 (1.23)	
Highest parental education level				<0.01			<0.01			<0.01
Grade 9 or below		1415	10.02 (1.37)		894	10.90 (1.12)		1457	12.25 (1.25)	
Grade 10–11		2115	10.12 (1.35)		1397	10.96 (1.13)		2121	12.40 (1.25)	
Grade 12 or above		1286	10.17 (1.37)		831	11.08 (1.21)		1333	12.41 (1.23)	

Table 1. Baseline Characteristics according to Age of Puberty from Hong Kong's "Children of 1997" Birth Cohort. (Available case Analysis).

(<9 hours) sleep duration in the large population-representative Hong Kong Chinese "Children of 1997" birth cohort. Second, we used two-sample MR to validate the causal effect of later puberty on adult sleep duration.

Results

Observational study in the Chinese "Children of 1997" birth cohort. Among the original 8,327 "Children of 1997" participants²², at the time of survey I (2008–09), 26 participants had permanently withdrawn, and 365 were not contactable. Of the 7,936 potential respondents, 3,603 provided sleep duration in Survey I (37.8% <9 hours and 62.2% 9+ hours), 3,933 provided sleep duration in Survey II (65.0% <9 hours and 35% 9+ hours) and 3,142 in the "Children of 1997" Biobank Clinical follow-up (79.0% <9 hours and 21% 9+ hours). In total 4,958 had age of menarche or voice breaking. The participants with and without puberty status differed in sex, parents' birthplace, highest parental occupation, household income per head in quintiles and highest parental education levels but the Cohen effect sizes indicated these differences were small (Appendix Table 1). Multicollinearity is also not a major issue in our study (Appendix Table 2). Mean age of onset of breast development was 9.6 years and 10.8 years for genitalia development. Girls (10.56 ± 1.03) had earlier age of onset of pubic hair than boys (11.48 ± 1.09). Mean age of menarche was 11.9 years and mean age of voice breaking was 13.1 years. Later pubertal development was associated with higher family SEP, such as parents' birthplace, highest parental occupation, household income per head in quintiles and highest parental education level (Table 1).

Age of menarche was positively associated with sleep duration [odds ratio (OR) 1.11, 95% CI (1.01 to 1.21)] at age 13.5 years, but other measures of puberty timing were unrelated to sleep duration. (Table 2).

Mendelian randomization study. *Genetic predictors of pubertal timing (exposure).* In total 389 single nucleotide polymorphisms (SNPs) predicting age at menarche (per year) at genome-wide significance (p -value $< 5 \times 10^{-8}$) were obtained from summary genetic associations concerning 329,345 women of European ancestry. These summary genetic associations are based on the ReproGen consortium ($N = 179,117$ from 40 studies), in addition to 23andMe ($N = 76,831$) and the UK Biobank ($N = 73,397$)²³, with F statistics ranging from 29 to 953 (Appendix Table 3). The 389 loci explained 7.4% of the variance in age of menarche and the overall F -statistic was about 63. Of these 389 SNPs, 181 SNPs were excluded because of linkage disequilibrium ($R^2 < 0.001$). Of the remaining 208 SNPs, 203 SNPs were available for sleep duration and no proxy was found for the 5 missing SNPs, giving 203 SNPs. Figure 1 shows the selection of SNPs related to age of menarche used as instruments. For age of voice breaking 11 genome-wide significant (p -value $< 5 \times 10^{-8}$) signals for age of voice breaking were obtained from 55,871 European men from the 23andMe study²⁴ After excluding 7 correlated SNPs, 4 independent SNPs

	Age 11–12 OR [#] (95% CI)	Age 13–14 OR [#] (95% CI)	Age 17–18 OR [#] (95% CI)
Overall			
Age of menarche/voice breaking		1.05 (0.98, 1.11)	1.01 (0.94, 1.09)
Onset of Breast/Genitalia	1.02 (0.96, 1.08)	1.02 (0.96, 1.09)	1.05 (0.98, 1.12)
Onset of Pubic Hair Development	1.03 (0.97, 1.10)	1.04 (0.97, 1.12)	1.02 (0.93, 1.11)
Girls			
Age of menarche/voice breaking		1.11 (1.01, 1.21)	1.06 (0.94, 1.18)
Onset of Breast/Genitalia	1.04 (0.96, 1.13)	1.02 (0.93, 1.11)	1.08 (0.99, 1.19)
Onset of Pubic Hair Development	1.06 (0.96, 1.17)	1.09 (0.98, 1.22)	1.04 (0.92, 1.18)
Boys			
Age of menarche/voice breaking		1.00 (0.92, 1.10)	0.97 (0.88, 1.08)
Onset of Breast/Genitalia	0.995 (0.91, 1.09)	1.03 (0.95, 1.12)	0.99 (0.89, 1.11)
Onset of Pubic Hair Development	0.996 (0.90, 1.10)	1.01 (0.92, 1.10)	0.98 (0.87, 1.11)

Table 2. Adjusted* Associations of Age of pubertal status (year) with sleep duration (9+ hours versus <9 hours) in Hong Kong's "Children of 1997" Birth Cohort. *Adjusted for parents' place of birth, highest parental occupation, household income per head and highest parental education levels. #Odds ratio (OR) per 1-year older age of puberty; thus, a significant OR >1 indicates that older age of puberty is associated with higher odds of longer sleep duration (9+ hours). Bold font: Statistical significance.

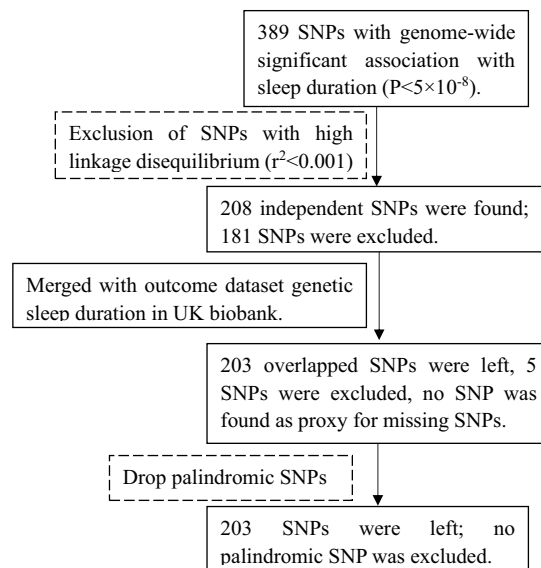


Figure 1. Selection of SNPs for age of menarche related to sleep duration used in Mendelian Randomization.

were left²⁴. Two independent SNPs (p -value $< 5 \times 10^{-6}$) predicting age at Tanner stage age in girls and one in boys were obtained from 11,000 Europeans from the Early Growth Genetics (EGG) Consortium²⁵. Appendix Table 3 summarizes the information extracted for each SNP.

Genetic associations with sleep duration (outcome). Genetic associations with sleep duration were obtained from the UK Biobank²⁶, which recruited more than 500 000 people of white British ancestry (intended age range 40–69 years) in Great Britain from 2006 to 2010.

Later age at menarche was associated with longer adult sleep duration using all methods except MR-Egger where the confidence interval included the null value. Heterogeneity was high ($p < 0.0001$), but the MR-Egger intercept did not indicate horizontal pleiotropy ($p = 0.95$). MR-PRESSO removed some SNPs as outliers ($P < 0.0001$), but the corrected estimates also showed a positive causal effect of age of menarche on sleep duration. The results were similar after removing 10 SNPs relevant to potential confounders and 7 potentially pleiotropic SNPs (Appendix Table 4). In boys, age of voice breaking was positively associated with sleep duration but the confidence interval included the null value. Tanner stage age in girls and boys was unrelated to sleep duration (Table 3).

	Puberty timing	SNPs	Average F statistic	Mendelian Randomization Method	β	95% Confidence Interval	I ² (p-value for heterogeneity)	MR-Egger intercept (p-value)	Outliers from MR PRESSO
Women	Age of menarche	203 (P < 5 × 10 ⁻⁸)	73.2	IVW with random-effects	0.020	0.006, 0.034	46.0% (<0.0001)	0.0001 (0.95)	rs9972653 rs35935052
				WM	0.021	0.003, 0.040			
				MR Egger	0.020	-0.018, 0.058			
				Corrected MR PRESSO	0.017	0.003, 0.030			
	Tanner stage	2 (P < 5 × 10 ⁻⁶)	26	IVW with random-effects	0.008	-0.035, 0.051	82.9% (0.02)	NA	NA
Men	Age of voice breaking	4 (P < 5 × 10 ⁻⁸)	84.5	IVW with random-effects	0.015	-0.035, 0.064	57.2% (0.06)	0.10 (0.425)	NA
				WM	0.003	-0.034, 0.04			
				MR Egger	-0.086	-0.338, 0.166			
				Corrected MR PRESSO	0.015	-0.035, 0.064			
	Tanner stage	1 (P < 5 × 10 ⁻⁶)	25	Wald ratio method	0.005	-0.037, 0.048	NA	NA	NA

Table 3. Mendelian Randomization Estimates of the Effect of puberty timing (year) on adult sleep duration (category). IVW: Inverse Variance Weighting WM: Weighted median method MR-PRESSO: Mendelian randomization pleiotropy residual sum and outlier Bold font: Statistical significance.

Discussion

In a population-representative Chinese birth cohort (“Children of 1997”)²² from an understudied non-Western population, age of menarche was positively associated with sleep duration at about 13.5 years. This finding was validated in a Mendelian Randomization study of the association of age of menarche with adult sleep duration. The evidence was not very supportive of pubertal markers causing sleep duration in boys or for other indicators of pubertal timing.

Our observational study extends previous studies by using separate indicators of the onset of puberty clinically assessed by a trained physician evaluation, rather than self-reports^{19,27}. A previous genetic study using linkage disequilibrium (LD) score regression found genetic correlations between sleep duration and age of menarche²⁸. Our two-sample MR study confirmed these findings with directional estimates using more genetic instruments and a larger sample. However, earlier age of voice breaking in boys was unrelated to sleep in both the observational and the MR study.

We used observational and two-sample MR designs to assess the association of earlier puberty with sleep duration, but limitations exist. First, child sleep duration was reported, which might be less accurate than actigraphy or polysomnography. However, self- and parentally-reported sleep duration has been widely used in previous studies²⁹. Second, observationally sleep duration in children could only be analyzed in two groups, reducing discrimination. However, the National Sleep Foundation has suggested that 9–11 hours sleep is appropriate for 6–13 year olds³⁰. Third, adjustment was not made for other factors such as diet and physical activity which may affect pubertal timing and sleep duration. However, we expected systematic rather than differential effects from residual or unmeasured confounding. Fourth, MR has three assumptions. First, strong associations of the genetic predictors with the exposure are required. In the current study, the SNPs for age of menarche and voice breaking all reached genome-wide significance with a high average F-statistic (73.2 for age of menarche and 84.5 for age of voice breaking). Some of these SNPs are in genes functionally relevant to puberty timing. For example, rare coding mutations in *MKRN3* cause central precocious puberty³¹. The *DLK1* locus confers a substantial decrease in the age of pubertal timing and *FSHB* confers earlier puberty timing through promoting higher levels of hypothalamic–pituitary–gonadal axis activity. Although we cannot rule out the possibility that some SNPs of uncertain biological function are included with corresponding risk of pleiotropy, we thoroughly investigated pleiotropic effects through multiple sensitivity analysis. Second, similar results after removing potentially confounding SNPs show the independence of the genetic variants from the confounders. Third, SNPs affecting sleep duration via mechanisms other than via puberty may generate a bias. In terms of such pleiotropy, we found no statistical evidence (from MR-Egger). However, the I² statistic suggested a certain level of heterogeneity, but the weighted median and MR-Egger provided consistent positive estimates similar to those from IVW. We found some outliers from MR-PRESSO, but the main findings were not changed after excluding these outliers. Estimates were similar after removing potentially pleiotropic SNPs, again suggesting that the results were less likely to result from pleiotropy. We have adequate power to detect the observed effect size for age of menarche. However, the limited number of SNPs for voice breaking (4) and Tanner stage (2 for girls, 1 for boys) means we do not have adequate power to detect the small observed effect sizes for the other earlier puberty markers (Appendix Table 5). Replication is required when a larger study becomes possible. Fifth, the MR estimates could be confounded by population stratification, but we predominantly used genetic studies in people of European ancestry with genomic control used for both exposure and outcome, which should minimize any such bias. Sixth, we performed the observational study in a Chinese population, but the MR study was restricted to people largely of European descent, which may reduce the comparability. However, we would normally expect causal factors to act consistently in different populations unless they act by a mechanism whose relevance varies across populations³². We know of no reason why age of puberty should have different effects on sleep in Chinese and people of European descent. Finally,

since only summary statistics, rather than individual level data, for the exposure and outcome from two different samples were used, we were unable to check for possible non-linear associations of earlier puberty with sleep duration. However, the risk of chance associations resulting from the underlying data structure in single sample is reduced when using separate sample instrumental variable analysis³³.

Although the mechanisms underlying the causal effect of earlier puberty on sleep duration are unclear, several potential explanations exist for our findings. First, sex hormones may play a role, since dramatic changes in testosterone and estrogen occur at puberty. In an RCT, estrogen increased sleep duration in postmenopausal women³⁴, but the effect could differ by baseline levels. Animal studies found estrogen suppresses sleep in female rats^{35,36}. As such, the mechanisms underlying the effects on sleep duration in women remain to be clarified. In men testosterone shortened sleep duration¹⁶. Second, earlier age of puberty was associated with higher risk of depressive symptoms only in girls in this cohort³⁷, which could account for the different estimates of earlier puberty with sleep duration by sex. Third, puberty-related sex differences in hypothalamic–pituitary–adrenal (HPA) axis activity may be another contributing factor, such as via corticosterone³⁸.

Age of menarche causally affects obesity¹³ and cancer³⁹. Age of menarche also affecting sleep duration may explain the observed, but possibly non-causal relation of sleep duration with obesity²⁸. Sex-specific associations are consistent with previous observational studies showing the association of earlier puberty with sleep duration was less clear in boys^{18,27}, which may partly be due to the lack of recordable and validated measures of pubertal timing compared with the more clear-cut milestone of menarche in girls. This means some estimates in boys could be biased towards the null by non-differential misclassification. For example, in the observational study, we had a larger sample with age of menarche ($n = 3,177$) than other pubertal markers. However, the same direction of effect estimates in women and men in the MR study suggests a role for pubertal timing in sleep duration in general rather than menarche specifically. Together with evidence of a similar causal effect of age of voice breaking on obesity and cardio metabolic traits²⁴, our results could indicate sleep duration is a marker, not a driver, of obesity⁴⁰ and possibly other chronic diseases.

Methods

Observational study. First, we used the “Children of 1997” birth cohort to conduct the observational study, which is a population-representative Chinese birth cohort ($n = 8,327$) that covered 88% of all births in Hong Kong from April 1 to May 31, 1997, described in detail elsewhere²². Baseline characteristics, including SEP (parental education and an indicator of parental migrant status) and birth characteristics were obtained from a self-administered questionnaire at recruitment⁴¹. Passive follow-up via record linkage was instituted in 2005 to obtain pubertal stage from the Student Health Service (SHS), based on an internal reference number, Department of Health, which provides free annual check-ups for all school students. Active follow-up via direct contact was instituted in 2007, with surveys conducted in 2008/9 (Survey I), 2010/12 (Survey II), 2011/12 (Survey III) and a “Children of 1997” Biobank Clinical follow-up in 2013–6.

Exposure – Age of puberty. Markers of puberty including breast/genitalia, pubic hair development and age of menarche/voice breaking were the exposures. Pubertal status were visually assessed by physicians at the SHS according to the criteria of Marshall and Tanner in grades 1, 3, 5, and 7 (usually at 6–7 years, 8–9 years, 10–11 years, and 12–13 years, respectively)^{42,43}. We defined the onset of puberty as onset of breast development for girls and genital development for boys, as measured by a change from Tanner stage I to stage II. Since the exact age of pubertal onset could not be precisely observed, we assumed the onset occurred midway between the latest time point when Tanner stage I was observed and the earliest time point when Tanner stage II was observed, assuming equal intervals between Tanner stages^{44,45}. The age of onset of pubic hair development was estimated in same way. Children with infeasible sequences of pubertal stages, such as Tanner stage II before stage I, were excluded ($n = 87$)³⁷. Age of menarche was self-reported at SHS clinics, in Survey III and in the “Children 1997” Biobank Clinical follow-up (for the Chinese birth cohort). Age of voice breaking was self-reported in Survey III and in the “Children 1997” Biobank Clinical follow-up.

Outcome – Sleep duration. The main outcome was sleep duration, which was assessed at three time points: Survey I (11.5 years), Survey II (13.5 years) and at the Biobank Clinical follow-up (17.5 years) in the Chinese “Children of 1997” birth cohort. Sleep duration was obtained in Survey I and II using a parent-reported questionnaire. Sleep duration was asked as “ ≤ 1 hours”, “2–4 hours”, “5–8 hours”, “9–12 hours” and “ ≥ 13 hours”, but almost all reported “5–8 hours” or “9–12 hours” sleep, so sleep duration was classified as < 9 hours and $9+$ hours. Sleep duration (in hours) was also obtained by self-report in the Biobank Clinical follow-up from the difference between bedtime and wake-up time, reported as the most common evening bedtime and wake-up time during the past month at about 17.5 years old. We also considered sleep duration as < 9 hours and $9+$ hours for consistency.

Mendelian randomization. Second, in order to valid the observational results, we used summary genetic associations from 2 different genome-wide association studies (GWAS) to test each association. We obtained genetic predictors of age of menarche from a GWAS of a combined study including ReproGen, 23andMe and UK Biobank²³, of age of voice breaking from 23andMe²⁴, and of Tanner stage from EGG²⁵. We obtained genetic associations with sleep duration from UK Biobank²⁶, restricted to participants of European descent. We obtained SNPs strongly (p -value $< 5 \times 10^{-8}$) associated with age of puberty from the largest and most recent genome-wide association studies (GWAS)^{23–25}. Linkage disequilibrium between these SNPs was identified using the “Clumping” function of MR-base⁴⁶. We used UK Biobank data to check for any associations at Bonferroni corrected significance of the selected SNPs with potential confounders, such as education, smoking, physical activity and alcohol use⁴⁷, using the UK Biobank, a large cohort study accessible to researchers worldwide²⁶. We repeated the

analysis after removing these SNPs in sensitivity analysis. Potentially pleiotropic effects (linked to the outcome other than via sleep) of the chosen SNPs were obtained from comprehensive curated genotype to phenotype cross-references, Ensembl and PhenoScanner⁴⁸. As sensitivity analysis we also repeated the analysis after excluding potentially pleiotropic SNPs. To identify any unknown pleiotropic effects, statistically, we used MR-Egger and the Mendelian randomization pleiotropy residual sum and outlier (MR-PRESSO) test⁴⁹.

These genetic instruments were applied to the largest publicly available GWAS of sleep duration from the UK Biobank, as described above. The GWAS includes 335,410 unrelated individuals of white British ancestry and provides sex-specific and overall genetic associations [35], adjusted, where appropriate, for sex, age, age squared, the interaction of age and sex, of sex and age squared and the first 20 principal components⁴⁷. Sleep duration was considered in three ordered categories: <7 hours, ≥7 and <8 hours and ≥8 hour.

Statistical analysis. In the “Children of 1997” birth cohort we used chi-squared tests and Cohen effect sizes to compare confounders for children with and without information about pubertal status. The association of pubertal timing with sleep were assessed using multivariable linear regression. Confounders were selected as likely common causes of sleep duration and pubertal timing⁵⁰, including parents’ place of birth, highest parental occupation, household income per head, highest parental education level⁴¹. Multicollinearity was assessed by variance inflation factor (VIF). A VIF of 10 or above suggests that interpretation of the relevant coefficients could be problematic⁵¹.

To account for loss to follow-up, we used a combination of multiple imputation (MI) and IPW⁵². First, we used multiple imputation to predict missing confounders and exposures⁵³. Second, we estimated IPWs using logistic regression to retrieve the original sample⁵⁴. Third, Rubin’s Rules were used to combine each IPW effect estimator with its corresponding sandwich variance estimator.

In the two-sample MR study, the causal associations of age of menarche with sleep duration were obtained using instrumental variable analysis. The F statistic for each SNP and overall was calculated to evaluate the strength of the instrument^{55,56}. We pooled Wald estimates (SNP on outcome/SNP on exposure)⁵⁷ for independent SNPs ($R^2 < 0.01$) using IVW meta-analysis with multiplicative random effects which assumes balanced pleiotropy. To assess heterogeneity, we used the I^2 statistic, where a higher value indicates more pleiotropy⁵⁸. However, given the possibility of unknown unbalanced pleiotropy, IVW could be invalid. When an exposure had at least 3 genetic predictors, we used the weighted median (WM) and MR-Egger with more relaxed assumptions. The WM gives robust estimates as long as valid SNPs contribute >50% of the information. MR-Egger with wide confidence intervals provides valid estimates and checks for potentially unknown horizontal pleiotropy (the SNPs affect the outcomes via mechanisms other than sleep duration) through non-null intercept, but it requires that the Instrument Strength Independent of Direct effect (INSIDE) assumption is satisfied. Since MR-Egger could not detect outliers and has limited statistical power⁵⁹, MR-PRESSO was used to identify horizontal pleiotropic outliers and if necessary to correct for pleiotropy via outlier removal⁴⁹. Using 100,000 simulations, we obtained the empirical p -value for the MR-PRESSO global test. The Mendelian randomization estimate is valid if the global test is non-significant ($p > 0.05$). We harmonized the effect allele for exposure and outcomes on the effect allele letter for non-palindromic SNPs, and confirmed the same effect allele for palindromic SNPs, i.e., coded (A/T or C/G) from the effect allele frequency and the coding used (forward or reverse). Palindromic SNPs which could not be aligned unequivocally were replaced by proxy SNPs.

All statistical analysis was performed using Stata version 13.1 (StataCorp LP, College Station, TX) and R version 3.3.3 (R Foundation for Statistical Computing, Vienna, Austria), and the “MendelianRandomization”, “TwoSampleMR” and “MRPRESSO” packages.

Ethics approval and consent to participate. Ethical approval for the study, including comprehensive health related analyses, was obtained from Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (HKU/HA HKW IRB). The Mendelian randomization study is an analysis of publicly available summary data that does not require ethical approval.

Data availability

The datasets for MR analysis of this article are available from two published genome-wide association studies: <https://www.nature.com/articles/ng.3841> and <http://www.nealelab.is/uk-biobank>. The “Children of 1997” data were available on request after approval. <https://aprmay97.sph.hku.hk>. The “Children of 1997” data and sample access committee is responsible for responding to birth cohort data access requests by bona fide researchers, and the relevant procedures involved (e.g., application and suggested revisions). The committee will also ensure access to the data is in line with the interests of the birth cohort participants, and access is in accordance with said protocol (i.e. measures to avoid confidentiality breaches, and notification of relevant publications arising from the requested data).

Received: 1 September 2019; Accepted: 17 January 2020;

Published online: 21 February 2020

References

1. Keyes, K. M., Maslowsky, J., Hamilton, A. & Schulenberg, J. The great sleep recession: changes in sleep duration among US adolescents, 1991–2012. *Pediatrics* **135**, 460–468, <https://doi.org/10.1542/peds.2014-2707> (2015).
2. Mindell, J. A., Meltzer, L. J., Carskadon, M. A. & Chervin, R. D. Developmental aspects of sleep hygiene: Findings from the 2004 National Sleep Foundation Sleep in America Poll. *Sleep Medicine* **10**, 771–779, <https://doi.org/10.1016/j.sleep.2008.07.016>.
3. Meerlo, P., Sgoifo, A. & Suchecki, D. Restricted and disrupted sleep: effects on autonomic function, neuroendocrine stress systems and stress responsivity. *Sleep. Med. Rev.* **12**, 197–210, <https://doi.org/10.1016/j.smrv.2007.07.007> (2008).

4. Luntamo, T. *et al.* Psychosocial determinants of headache, abdominal pain, and sleep problems in a community sample of Finnish adolescents. *Eur. child. Adolesc. psychiatry* **21**, 301–313, <https://doi.org/10.1007/s00787-012-0261-1> (2012).
5. Bertisch, S. M. *et al.* Insomnia with Objective Short Sleep Duration and Risk of Incident Cardiovascular Disease and All-Cause Mortality: Sleep Heart Health Study. *Sleep*, <https://doi.org/10.1093/sleep/zsy047> (2018).
6. Shan, Z. *et al.* Sleep duration and risk of type 2 diabetes: a meta-analysis of prospective studies. *Diabetes Care* **38**, 529–537, <https://doi.org/10.2337/dc14-2073> (2015).
7. Lu, C. *et al.* Long-Term Sleep Duration as a Risk Factor for Breast Cancer: Evidence from a Systematic Review and Dose-Response Meta-Analysis. *Biomed. Res. Int.* **2017**, 4845059, <https://doi.org/10.1155/2017/4845059> (2017).
8. Twenge, J. M., Krizan, Z. & Hisler, G. Decreases in self-reported sleep duration among U.S. adolescents 2009–2015 and association with new media screen time. *Sleep. Med.* **39**, 47–53, <https://doi.org/10.1016/j.sleep.2017.08.013> (2017).
9. Matthews, K. A., Jennings, J. R. & Lee, L. Socioeconomic status in childhood predicts sleep continuity in adult Black and White men. *Sleep. health* **4**, 49–55, <https://doi.org/10.1016/j.sleh.2017.09.008> (2018).
10. Lin, Y. *et al.* Temporal and bi-directional associations between sleep duration and physical activity/sedentary time in children: An international comparison. *Prev. Med.*, <https://doi.org/10.1016/j.ypmed.2017.12.006> (2017).
11. Owens, J. A. & Weiss, M. R. Insufficient sleep in adolescents: causes and consequences. *Minerva pediatrica* **69**, 326–336, <https://doi.org/10.23736/s0026-4946.17.04914-3> (2017).
12. Hoyt, L. T. *et al.* Girls' Sleep Trajectories Across the Pubertal Transition: Emerging Racial/Ethnic Differences. *J. Adolesc. Health* **62**, 496–503, <https://doi.org/10.1016/j.jadohealth.2017.10.014> (2018).
13. Gill, D. *et al.* Age at menarche and adult body mass index: a Mendelian randomization study. *International journal of obesity (2005)*, <https://doi.org/10.1038/s41366-018-0048-7> (2018).
14. Dashti, H. S. *et al.* Genome-wide association study identifies genetic loci for self-reported habitual sleep duration supported by accelerometer-derived estimates. *Nat. Commun.* **10**, 1100, <https://doi.org/10.1038/s41467-019-08917-4> (2019).
15. Wells, J. C. K., Nesse, R. M., Sear, R., Johnstone, R. A. & Stearns, S. C. Evolutionary public health: introducing the concept. *Lancet* **390**, 500–509, [https://doi.org/10.1016/s0140-6736\(17\)30572-x](https://doi.org/10.1016/s0140-6736(17)30572-x) (2017).
16. Liu, P. Y. *et al.* The short-term effects of high-dose testosterone on sleep, breathing, and function in older men. *J. Clin. Endocrinol. Metab.* **88**, 3605–3613, <https://doi.org/10.1210/jc.2003-030236> (2003).
17. Hays, J. *et al.* Effects of Estrogen plus Progestin on Health-Related Quality of Life. *N. Engl. J. Med.* **348**, 1839–1854, <https://doi.org/10.1056/NEJMoa030311> (2003).
18. Johnson, E. O., Roth, T., Schultz, L. & Breslau, N. Epidemiology of DSM-IV insomnia in adolescence: lifetime prevalence, chronicity, and an emergent gender difference. *Pediatrics* **117**, e247–256, <https://doi.org/10.1542/peds.2004-2629> (2006).
19. Liu, X., Chen, H., Liu, Z.-Z., Fan, F. & Jia, C.-X. Early Menarche and Menstrual Problems Are Associated with Sleep Disturbance in a Large Sample of Chinese Adolescent Girls. *Sleep* **40**, zsx107-zsx107, <https://doi.org/10.1093/sleep/zsx107> (2017).
20. Sequeira, M. E., Lewis, S. J., Bonilla, C., Smith, G. D. & Joinson, C. Association of timing of menarche with depressive symptoms and depression in adolescence: Mendelian randomisation study. *Br. J. psychiatry: J. Ment. Sci.* **210**, 39–46, <https://doi.org/10.1192/bjp.bp.115.168617> (2017).
21. Gill, D. *et al.* Age at Menarche and Time Spent in Education: A Mendelian Randomization Study. *Behav. Genet.* **47**, 480–485, <https://doi.org/10.1007/s10519-017-9862-2> (2017).
22. Schooling, C. M., Hui, L. L., Ho, L. M., Lam, T. H. & Leung, G. M. Cohort profile: 'children of 1997': a Hong Kong Chinese birth cohort. *Int. J. Epidemiol.* **41**, 611–620, <https://doi.org/10.1093/ije/dyq243> (2012).
23. Day, F. R. *et al.* Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk. *Nature Genetics* **49**, 834, <https://doi.org/10.1038/ng.3841>, <https://www.nature.com/articles/ng.3841#supplementary-information> (2017).
24. Day, F. R. *et al.* Shared genetic aetiology of puberty timing between sexes and with health-related outcomes. *Nature communications* **6**, <https://doi.org/10.1038/ncomms9842> (2015).
25. Cousminer, D. L. *et al.* Genome-wide association study of sexual maturation in males and females highlights a role for body mass and menarche loci in male puberty. *Hum. Mol. Genet.* **23**, 4452–4464, <https://doi.org/10.1093/hmg/ddu150> (2014).
26. Allen, N. E., Sudlow, C., Peakman, T. & Collins, R. UK Biobank Data: Come and Get It. *Science translational medicine* **6**, 224ed224–224ed224, <https://doi.org/10.1126/scitranslmed.3008601> (2014).
27. Zhang, J. *et al.* Emergence of Sex Differences in Insomnia Symptoms in Adolescents: A Large-Scale School-Based Study. *Sleep*. **39**, 1563–1570, <https://doi.org/10.5665/sleep.6022> (2016).
28. Jones, S. E. *et al.* Genome-Wide Association Analyses in 128,266 Individuals Identifies New Morningness and Sleep Duration Loci. *PLoS Genet.* **12**, e1006125, <https://doi.org/10.1371/journal.pgen.1006125> (2016).
29. Cespedes, E. M. *et al.* Comparison of Self-Reported Sleep Duration With Actigraphy: Results From the Hispanic Community Health Study/Study of Latinos Sueño Ancillary Study. *Am. J. Epidemiol.* **183**, 561–573, <https://doi.org/10.1093/aje/kwv251> (2016).
30. Hirshkowitz, M. *et al.* National Sleep Foundation's updated sleep duration recommendations: final report. *Sleep. Health: J. Natl Sleep Found.* **1**, 233–243, <https://doi.org/10.1016/j.sleh.2015.10.004> (2015).
31. Abreu, A. P. *et al.* Central precocious puberty caused by mutations in the imprinted gene MKRN3. *N. Engl. J. Med.* **368**, 2467–2475, <https://doi.org/10.1056/NEJMoa1302160> (2013).
32. Lopez, P. M., Subramanian, S. V. & Schooling, C. M. Effect measure modification conceptualized using selection diagrams as mediation by mechanisms of varying population-level relevance. *J. Clin. Epidemiol.* **113**, 123–128, <https://doi.org/10.1016/j.jclinepi.2019.05.005> (2019).
33. Taylor, A. E. *et al.* Mendelian randomization in health research: using appropriate genetic variants and avoiding biased estimates. *Econ. Hum. Biol.* **13**, 99–106, <https://doi.org/10.1016/j.ehb.2013.12.002> (2014).
34. Maki, P. M., Gast, M. J., Vieweg, A. J., Burriss, S. W. & Yaffe, K. Hormone therapy in menopausal women with cognitive complaints: a randomized, double-blind trial. *Neurol.* **69**, 1322–1330, <https://doi.org/10.1212/01.wnl.0000277275.42504.93> (2007).
35. Hadjimarkou, M. M., Benham, R., Schwarz, J. M., Holder, M. K. & Mong, J. A. Estradiol suppresses rapid eye movement sleep and activation of sleep-active neurons in the ventrolateral preoptic area. *Eur. J. Neurosci.* **27**, 1780–1792, <https://doi.org/10.1111/j.1460-9568.2008.06142.x> (2008).
36. Schwartz, M. D. & Mong, J. A. Estradiol suppresses recovery of REM sleep following sleep deprivation in ovariectomized female rats. *Physiol. Behav.* **104**, 962–971, <https://doi.org/10.1016/j.physbeh.2011.06.016> (2011).
37. Wang, H., Lin, S. L., Leung, G. M. & Schooling, C. M. Age at Onset of Puberty and Adolescent Depression: "Children of 1997" Birth Cohort. *Pediatrics* **137**, <https://doi.org/10.1542/peds.2015-3231> (2016).
38. Andersen, M. L., Martins, P. J. F., D'Almeida, V., Bignotto, M. & Tufik, S. Endocrinological and catecholaminergic alterations during sleep deprivation and recovery in male rats. *J. sleep. Res.* **14**, 83–90, <https://doi.org/10.1111/j.1365-2869.2004.00428.x> (2005).
39. Bonilla, C. *et al.* Pubertal development and prostate cancer risk: Mendelian randomization study in a population-based cohort. *BMC Med.* **14**, 66, <https://doi.org/10.1186/s12916-016-0602-x> (2016).
40. Wang, J., Li, A. M. & Lam, H. Sleep Duration and Adiposity in Children and Adults: Observational and Mendelian Randomization. *Studies.* **27**, 1013–1022, <https://doi.org/10.1002/oby.22469> (2019).
41. Wang, J. *et al.* Sleep duration and risk of diabetes: Observational and Mendelian randomization studies. *Prev. Med.* **119**, 24–30, <https://doi.org/10.1016/j.ypmed.2018.11.019> (2019).
42. Marshall, W. A. & Tanner, J. M. Variations in the pattern of pubertal changes in boys. *Arch. Dis. Child.* **45**, 13–23 (1970).

43. Marshall, W. A. & Tanner, J. M. Variations in pattern of pubertal changes in girls. *Arch. Dis. Child.* **44**, 291–303 (1969).
44. Christensen, K. Y. *et al.* Characterization of the correlation between ages at entry into breast and pubic hair development. *Ann. Epidemiol.* **20**, 405–408, <https://doi.org/10.1016/j.annepidem.2010.02.005> (2010).
45. Hui, L. L., Lam, H. S., Leung, G. M. & Schooling, C. M. Duration of puberty in preterm girls. *Am J. Hum. Biol.* **29**, <https://doi.org/10.1002/ajhb.22963> (2017).
46. Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the human phenome. *eLife* **7**, <https://doi.org/10.7554/eLife.34408> (2018).
47. Benjiamin, N. <http://www.nealelab.is/blog/2017/9/11/details-and-considerations-of-the-uk-biobank-gwas> (update to Apr 2018).
48. Staley, J. R. *et al.* PhenoScanner: a database of human genotype-phenotype associations. *Bioinforma.* **32**, 3207–3209, <https://doi.org/10.1093/bioinformatics/btw373> (2016).
49. Verbanck, M., Chen, C. Y., Neale, B. & Do, R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat. Genet.* **50**, 693–698, <https://doi.org/10.1038/s41588-018-0099-7> (2018).
50. VanderWeele, T. J. & Shpitser, I. A new criterion for confounder selection. *Biometrics* **67**, 1406–1413, <https://doi.org/10.1111/j.1541-0420.2011.01619.x> (2011).
51. Vatcheva, K. P., Lee, M., McCormick, J. B. & Rahbar, M. H. Multicollinearity in Regression Analyses Conducted in Epidemiologic Studies. *Epidemiol.* **6**, 227, <https://doi.org/10.4172/2161-1165.1000227> (2016).
52. Seaman, S. R., White, I. R., Copas, A. J. & Li, L. Combining multiple imputation and inverse-probability weighting. *Biometrics* **68**, 129–137, <https://doi.org/10.1111/j.1541-0420.2011.01666.x> (2012).
53. Schafer, J. L. Multiple imputation: a primer. *Stat. Methods Med. Res.* **8**, 3–15, <https://doi.org/10.1177/096228029900800102> (1999).
54. Shaun, R. S. & Ian, R. W. Review of inverse probability weighting for dealing with missing data. *Stat. Methods Med. Res.* **22**, 278–295, <https://doi.org/10.1177/0962280210395740> (2011).
55. Li, B. & Martin, E. B. An approximation to the F distribution using the chi-square distribution. *Computational Stat. Data Anal.* **40**, 21–26, [https://doi.org/10.1016/S0167-9473\(01\)00097-4](https://doi.org/10.1016/S0167-9473(01)00097-4) (2002).
56. Palmer, T. M. *et al.* Using multiple genetic variants as instrumental variables for modifiable risk factors. *Stat. methods Med. Res.* **21**, 223–242, <https://doi.org/10.1177/0962280210394459> (2012).
57. Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet. Epidemiol.* **40**, 304–314, <https://doi.org/10.1002/gepi.21965> (2016).
58. Higgins, J. P., Thompson, S. G., Deeks, J. J. & Altman, D. G. Measuring inconsistency in meta-analyses. *Bmj* **327**, 557–560, <https://doi.org/10.1136/bmj.327.7414.557> (2003).
59. Bowden, J. *et al.* Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I² statistic. *Int. J. Epidemiol.* **45**, 1961–1974, <https://doi.org/10.1093/ije/dyw220> (2016).

Acknowledgements

The authors thank colleagues at the Student Health Service and Family Health Service of the Department of Health for their assistance and collaboration. The authors thank the UK Biobank consortium investigators. This work is a sub-study of the “Children of 1997 birth cohort”, which was initially supported by the Health Care and Promotion Fund, Health and Welfare Bureau, Government of the Hong Kong SAR [HCPF Grant # 216106] and re-established in 2005 with support from the Health and Health Services Research Fund [HHSRF Grant # 03040771, 05060671, 07080751, 07080841] and Research Fund for Control of Infectious Diseases (RFCID grant # 04050172, 06060592), and the University Research Committee Strategic Research Theme (SRT) of Public Health, The University of Hong Kong. The Biobank Clinical Follow-up was partly funded by the WYNG Foundation. The authors have no financial relationships relevant to this article to disclose.

Author contributions

J.W. conducted the analyses and wrote the first draft of the manuscript. C.M.S. revised drafts of the manuscript critically and supervised the study from conception to completion. M.K.K. and S.L.A.Y. helped to run the follow-up of the birth cohort and reviewed the manuscript. J.Z. checked the analysis and reviewed the manuscript. G.M.L., A.M.L. and H.S.L. provided content matter expertise and helped to collect the birth cohort data. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-59811-9>.

Correspondence and requests for materials should be addressed to C.M.S.

Reprints and permissions information is available at www.nature.com/reprints.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020