

# Age at Menarche and Time Spent in Education: A Mendelian Randomization Study

D. Gill<sup>1,2</sup> · F. Del Greco M<sup>3</sup> · T. M. Rawson<sup>1</sup> · P. Sivakumaran<sup>1</sup> · A. Brown<sup>1</sup> · N. A. Sheehan<sup>4</sup> · C. Minelli<sup>1</sup>

Received: 14 November 2016 / Accepted: 13 July 2017 / Published online: 9 August 2017  
© The Author(s) 2017. This article is an open access publication

**Abstract** Menarche signifies the primary event in female puberty and is associated with changes in self-identity. It is not clear whether earlier puberty causes girls to spend less time in education. Observational studies on this topic are likely to be affected by confounding environmental factors. The Mendelian randomization (MR) approach addresses these issues by using genetic variants (such as single nucleotide polymorphisms, SNPs) as proxies for the risk factor of interest. We use this technique to explore whether there is a causal effect of age at menarche on time spent in education. Instruments and SNP-age at menarche estimates are identified from a Genome Wide Association Study (GWAS) meta-analysis of 182,416 women of European descent. The effects of instruments on time spent in education are estimated using a GWAS meta-analysis of 118,443 women performed by the Social Science Genetic Association Consortium (SSGAC). In our main analysis, we demonstrate a small but statistically significant causal effect of age at menarche on time spent in education: a 1 year increase in age at menarche is associated with 0.14 years (53 days) increase in time spent in education (95% CI

0.10–0.21 years,  $p=3.5\times 10^{-8}$ ). The causal effect is confirmed in sensitivity analyses. In identifying this positive causal effect of age at menarche on time spent in education, we offer further insight into the social effects of puberty in girls.

**Keywords** Menarche · Puberty · Education · Educational attainment · Mendelian randomization

## Introduction

Puberty is a time of physiological change in the human body, and its effects extend into the social domains of life (Stattin and Magnusson 1990). In girls, menarche signifies the primary event in puberty. The initiation of the menstrual cycle is associated with reorganization of the self-image, changes in peer relationships, and increased engagement in risk behaviours (Crosnoe 2000). In adolescent females, elevated levels of gonadal hormones follow menarche and influence behavioural development (Schulz and Sisk 2016). Heightened neural plasticity during puberty may predispose to a greater sensitivity to hormones (Piekarski et al. 2016). The significance of menarche as a life-course transition varies with its timing (Schulz and Sisk 2016), and previous observational work has suggested that girls with early puberty have a more difficult journey through school (Cavanagh et al. 2007). While the latter goes on to propose that earlier puberty might be causally associated with less time spent in education, this has not yet been demonstrated. Furthermore, the influence of potentially confounding variables must be reliably excluded to ensure that spurious associations are not interpreted as causal. Previous work has demonstrated age of menarche to be influenced by obesity (Dvornyk and Waqar ul 2012), family size and

Edited by Sarah Medland.

**Electronic supplementary material** The online version of this article (doi:10.1007/s10519-017-9862-2) contains supplementary material, which is available to authorized users.

✉ D. Gill  
dipender.gill@imperial.ac.uk

- <sup>1</sup> Imperial College London, London, UK
- <sup>2</sup> Imperial College Healthcare NHS Trust, London, UK
- <sup>3</sup> Institute for Biomedicine, Eurac Research, Bolzano, Italy
- <sup>4</sup> University of Leicester, Leicester, UK

socio-demographic factors (Chavarro et al. 2004), all of which are also associated with less time spent in education (Winding et al. 2013). Therefore, observational studies on this topic may be limited by confounding, making it difficult to decipher causal effects.

In such situations, the Mendelian randomization (MR) technique can often be used to overcome these limitations by using single nucleotide genetic polymorphisms (SNPs) as instrumental variables (IVs) to explore the direction and magnitude of any causal effect of age at menarche on time spent in education (Davey Smith and Ebrahim 2005). Genes are allocated randomly at the time of conception and are therefore independent of classical confounding. The demonstration that SNPs known to modify age at menarche also modify time spent in education can provide indirect evidence of a causal effect of age at menarche on time spent in education, provided that the necessary assumptions are satisfied (Sheehan et al. 2008). Indeed, such an approach has been previously used to show that earlier menarche causes a higher level of depressive symptoms at 14 years (Sequeira et al. 2016).

Here we use MR to investigate the causal effect of age at menarche on time spent in education. By gaining insight into the effect of age at menarche on time spent in education, we hope to further our understanding of the social implications of this physiological and psychological transition.

## Methods

### SNP-age at menarche association estimates

SNPs for use as instruments in the MR analysis were identified from a GWAS meta-analysis of 57 studies in 182,416 women of European descent, where age at menarche was established by self-reporting, and analyses within each study were adjusted for birth year, to account for secular trends, and genomic control, to account for population stratification (Perry et al. 2014). This identified 122 independent SNPs at 106 genomic loci to be associated with age at menarche ( $p$  value  $< 5 \times 10^{-8}$ ). We measure the strength of the instruments using the F statistic, which is a function of the magnitude and precision of their genetic effects (Li and Martin 2002; Palmer et al. 2012).

### SNP-time spent in education association estimates

The effects of the 122 instruments on time spent in education were estimated using a GWAS meta-analysis of 118,443 women across 62 studies performed by the Social Science Genetic Association Consortium (SSGAC), the summary estimates for which can be downloaded from

<http://www.thessgac.org/data> (Okbay et al. 2016). The analysis was performed on women, aged 30 years or above, of European descent whose mother tongue was the same as the main language of the country in which they were educated (Okbay et al. 2016). Although study populations were heterogeneous in terms of their educational systems, with different survey questions and data registers used to evaluate time spent in education across studies, comparability was maximized by mapping each major educational qualification on to one of seven categories of the 1997 International Standard Classification of Education (ISCED) of the United Nations Educational, Scientific and Cultural Organization, and then imputing a time spent in education equivalent for each ISCED category (Okbay et al. 2016).

### Mendelian randomization estimates

Individual MR estimates for each of the 122 SNPs were derived using the Wald estimator, which is the ratio of the estimates of the two genetic associations (i.e. SNP-time spent in education estimate over SNP-age at menarche estimate) (Didelez et al. 2010), with standard error derived using the Delta method (Thompson et al. 2016). MR estimates across the individual SNPs were pooled using a fixed-effect inverse-variance weighted (IVW) meta-analysis. This approach assumes an additive model with no interactions for the SNP-age at menarche and SNP-time spent in education relationships.

### Sensitivity analyses

A critical assumption in MR is the absence of pleiotropy—that genetic instruments only modify time spent in education through age at menarche and not by any other independent pathways. In the absence of this condition, MR could produce biased estimates (Sheehan et al. 2008). In the meta-analysis of the 122 MR estimates, the  $I^2$  index (which we call  $I^2_{MR}$ ) describes the percentage of total variation in MR estimates across instruments that arises because of heterogeneity rather than chance, and can be used as a proxy for pleiotropy (Del Greco M et al. 2015). We define heterogeneity to be present if  $I^2_{MR} > 25\%$ . To address pleiotropy and other possible sources of bias in this work, further sensitivity analyses were performed:

1. *MR-Egger* This is an adaptation of Egger regression applied to the context of two-sample MR that uses multiple genetic variants (Bowden et al. 2015). The MR-Egger approach can be used to provide unbiased results in the presence of pleiotropic instruments under the assumption that the magnitude of pleiotropic effects is independent of the magnitude of the corresponding SNP-age at menarche effects (Bowden et al.

- 2015). The degree of heterogeneity in the SNP-age at menarche estimates generated by the different instruments, as measured using the  $I^2$  statistic (called  $I^2_{GX}$  here), is used to quantify any potential bias arising in the MR-Egger analysis due to measurement error. An  $I^2_{GX}$  estimate close to 100% would suggest that such a phenomenon is not creating bias, as greater heterogeneity reduces regression dilution with MR-Egger (Bowden et al. 2016b).
2. *Weighted median estimator* Used as a further sensitivity analysis here, this approach orders the MR estimates generated using each instrument separately by the inverse of their variances; selecting the median result provides a single MR estimate, with confidence intervals generated using a parametric bootstrap method (Bowden et al. 2016a). The weighted median estimator does not require that the magnitude of any pleiotropic effects of the instruments are uncorrelated to their effects on the intermediate phenotype, as MR-Egger does, but instead assumes that at least half of the instruments are valid (Bowden et al. 2016b).
  3. *Exclusion of instruments also associated with body mass index (BMI)* Age at menarche has been previously demonstrated to be associated with obesity (Dvornyk and Waqar ul 2012). If BMI is also associated with time spent in education, then any instruments for age at menarche that are also associated with BMI might be introducing pleiotropic effects by this mechanism. Sensitivity analysis was therefore also performed by repeating the MR-analysis using the fixed-effect IVW meta-analysis approach with the exclusion of SNPs also associated with BMI at genome-wide significance level (Supplementary Table 1).
  4. *Unweighted allele score* The use of SNP-age at menarche estimates generated from the GWAS discovery analysis rather than the replication analysis, which had a sample size 20 times smaller (8689 vs. 182,416) (Perry et al. 2014), may result in the possible upward bias that is typical of discovery stage results (Ioannidis et al. 2001). MR analysis using a fixed-effect IVW meta-analysis of SNP-time spent in education association estimates across the 122 SNPs, which is equivalent to an “unweighted allele score” (Charoen et al. 2016), is not affected by this form of bias and is therefore used here as a sensitivity analysis.
  5. *SNP-time spent in education estimates for men using an unweighted allele score* As a check that the SNP-time spent in education association observed in women is indeed driven by mediation through age at menarche and not via alternative pathways, as implied by our instrumental variable assumptions, we estimate this association in men using an unweighted allele score for the 122 age at menarche instruments. Since men

do not undergo menarche, the exposure under investigation, lack of any association in men would provide further evidence that an association in women is due to a causal effect of age at menarche on time spent in education. These SNP-time spent in education estimates were obtained from a GWAS meta-analysis of 147,474 men also performed by the SSGAC, the summary estimates for which can be downloaded from <http://www.thessgac.org/data> (Okbay et al. 2016).

All analyses were performed using Stata 14 (StataCorp LP) and R version 3.3.2 (R Core Team).

## Results

Supplementary Tables 2 and 3 report individual SNP estimates of the per-allele effects on age at menarche (years) and time spent in education for women (standard deviation change in time spent in education, measured in years), respectively, while Supplementary Table 4 reports individual SNP MR estimates for the causal effect of age at menarche on time spent in education (standard deviation change in time spent in education, measured in years, per year increase in age at menarche). The considered SNPs are all strong instruments for age at menarche, with F statistics ranging from 25 to 576 (Supplementary Table 2), which are all greater than the recommended threshold of 10 (Lawlor et al. 2008).

The fixed-effect IVW meta-analysis of all 122 SNPs shows a statistically significant causal effect of age at menarche on time spent in education: a 1 year increase in age at menarche is associated with a 0.04 standard deviation units increase in time spent in education (95% CI 0.03–0.06), with a p value of  $3.5 \times 10^{-8}$  (Supplementary Fig. 1). With the standard deviation of time in education reported as 3.6 years (Okbay et al. 2016), this equates to 0.14 years (53 days, 95% CI 0.10–0.21 years). However, there is evidence of pleiotropy among instruments, with a between-instrument  $I^2_{MR}$  of 48% (95% CI 36–58%). Further sensitivity analyses were performed, and MR-Egger regression analysis estimated that a 1 year increase in age at menarche is associated with 0.10 standard deviation increase in time spent in education (95% CI 0.03–0.18,  $p=0.01$ ) (Supplementary Fig. 1). The  $I^2_{GX}$  statistic is 85%, suggesting that there is no major evidence of measurement error biasing MR-Egger analysis (Bowden et al. 2016b). The MR-Egger intercept is  $-0.002$  (95% CI  $-0.006$  to  $0.001$ ,  $p=0.139$ ), suggesting no evidence of directional pleiotropy (Bowden et al. 2015). Supplementary Fig. 2 shows the funnel plot of the minor allele frequency corrected GX estimates by the GY/GX estimates; there is no major asymmetry around the fixed-effect IVW

meta-analysis causal estimate (dashed red line) to suggest directional pleiotropy.

The weighted median approach estimates that a 1 year increase in age at menarche is associated with 0.05 standard deviation increase in time spent in education (95% CI 0.03–0.07,  $p=7.34\times 10^{-5}$ ) (Supplementary Fig. 1). Fixed-effect IVW meta-analysis after excluding the 12 SNPs associated with BMI (Supplementary Table 1) also shows a statistically significant causal effect of age at menarche on time spent in education: a 1 year increase in age at menarche is associated with 0.05 standard deviation increase in time spent in education (95% CI 0.03–0.06,  $p=1.74\times 10^{-8}$ ), although evidence of pleiotropy persists ( $I^2_{MR}$  47%, 95% CI 34–58%). Thus, removing these 12 SNPs that are associated with BMI did not have any major effect on the results obtained, and for this reason, the results of the original IVW approach are reported as the main analysis.

Use of an unweighted allele score for the 122 instruments in women shows a statistically significant positive association with time spent in education ( $p=3.61\times 10^{-7}$ ). This is reassuring that the causal effect of age at menarche on time spent in education shown by our main analysis is not attributable to bias due to use of SNP-age at menarche estimates from discovery stage results. The unweighted allele score is only used here to test for a causal effect of age at menarche on time spent in education and not to estimate the magnitude of this effect.

Supplementary Table 5 reports individual SNP estimates for the per-allele effect of the 122 age at menarche instruments on time spent in education for men. The unweighted allele score for the 122 age at menarche instruments using SNP-time spent in education estimates for men is not significant ( $p=0.72$ ), thus strengthening our belief that the observed association in women is driven by mediation through age at menarche.

In summary, all sensitivity analyses support our findings of a statistically significant causal effect of age at menarche on time spent in education.

## Discussion

We have used MR to investigate the causal effect of age at menarche on time spent in education. Under the required assumptions, this technique circumvents the classical confounding seen in observational studies, and our main (fixed-effect IVW meta-analysis of all 122 SNPs) analysis suggests that for every year increase in age at menarche, women spend an extra 53 days in education on average (95% CI 37–78 days). The main limitation of the MR approach is possible bias due to pleiotropic instruments,

and we have addressed this using several sensitivity analyses.

The physical, behavioural and cognitive aspects of development that are associated with puberty vary in their timing. A lower age at menarche has been hypothesized to result in earlier physical and physiological development, but without matching levels of cognitive and behavioural development. This delay can lead to inadequate coping strategies, greater risk-taking behaviour, lower social competence, and higher rates of internalizing and affective disorders (Brooks-Gunn and Warren 1989; Stice et al. 2004; Westling et al. 2012), all of which may culminate in less time spent in education.

Previous work has highlighted the adverse effects of early puberty on self-perception, peer-relationships and risk-taking behaviours in girls, with consequent effects on performance in school (Correll 2001; Crosnoe 2000; Graber 2013; Mendle et al. 2007). However, there has been comparatively little work directly exploring the effect of age at menarche on time spent in education. One survey-based observational study performed by Koivusilta et al. investigating how age at menarche predicts time spent in education separately considered samples of 903, 1430 and 1584 Finnish girls aged 12 years, 14 years and 16 years respectively (Koivusilta and Rimpela 2004). This work measured the timing of menarche as early (age 11 years or younger), average (age 12 or 13 years) or late (age 14 years or older), with time spent in education divided into categories of 9–10, 11–12, 13–15 and 16–18 years. While ordinal logistic regression did not identify any effect of age at menarche on time spent in education, it is possible that use of ordered categorical variables, rather than continuous ones, might have resulted in loss of power to detect the small effect size identified in our analysis. Furthermore, such observational work is also susceptible to the effects of confounding, such as from socio-demographic factors (Koivusilta and Rimpela 2004), the direction of which is hard to predict (away from the null or towards the null).

There are a number of possible sources of bias in our work. The age at menarche estimates used were self-reported and are therefore susceptible to recall bias (Perry et al. 2014). Furthermore, the different studies used to generate SNP-time spent in education association results were spread out over decades, with birth years ranging from 1901 to 1989 (Okbay et al. 2016). The social environment is likely to have changed over this period, making the MR analysis susceptible to varying SNP-environment interactions (Brennan 2004). Finally, although the use of SNP-age at menarche estimates generated from the GWAS discovery analysis may result in upward bias (“winner’s curse”), we have shown that a causal effect remains when testing it using an unweighted allele score, which is not affected by this “winner’s curse”.

In summary, we have used the MR approach to tackle traditional confounding in investigating the effect of age at menarche on time spent in education. We demonstrate a small positive causal effect, which offers further insight into the effects of puberty in girls. Given the significance of education on future life course, to include effects on health, this finding provides further insight into the social, psychological and physiological factors that determine time in education (Brennan 2004; Kingston et al. 2003; Li and Powdthavee 2015).

**Acknowledgements** The authors would like to thank Prof. Chris McManus and Dr. Katherine Woolf for their advice on this work.

**Funding** No funding was received for this work.

#### Compliance with ethical standards

**Conflict of interest** D. Gill, F. Del Greco M, T. M. Rawson, P. Sivakumaran, A. Brown, N. A. Sheehan and C. Minelli declare that they have no conflicts of interest.

**Ethical approval** This work used publicly available summary GWAS meta-analysis results, and therefore ethical approval was not required. All procedures were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

**Informed consent** For this type of study formal consent is not required.

**Statement of Human and Animal Rights** This article does not contain any studies with animals performed by any of the authors.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided 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.

## References

- Bowden J, Davey Smith G, Burgess S (2015) Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 44(2):512–525. doi:10.1093/ije/dyv080
- Bowden J, Davey Smith G, Haycock PC, Burgess S (2016a) Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol* 40(4):304–314. doi:10.1002/gepi.21965
- Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan NA, Thompson JR (2016b) Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I<sup>2</sup> statistic. *Int J Epidemiol*. doi:10.1093/ije/dyw220
- Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan N, Thompson J (2017) A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Stat Med*. doi:10.1002/sim.7221
- Brennan P (2004) Commentary: Mendelian randomization and gene-environment interaction. *Int J Epidemiol* 33(1):17–21. doi:10.1093/ije/dyh033
- Brooks-Gunn J, Warren MP (1989) Biological and social contributions to negative affect in young adolescent girls. *Child Dev* 60(1):40–55
- Cavanagh SE, Riegle-Crumb C, Crosnoe R (2007) Puberty and the education of girls. *Soc Psychol Q* 70(2):186–198
- Charoen P, Nitsch D, Engmann J, Shah T, White J, Zabaneh D, Jefferis B, Wannamethee G, Whincup P, Cassidy AM, Gaunt T (2016) Mendelian randomisation study of the influence of eGFR on coronary heart disease. *Sci Rep* 6:28514. doi:10.1038/srep28514
- Chavarro J, Villamor E, Narvaez J, Hoyos A (2004) Socio-demographic predictors of age at menarche in a group of Colombian university women. *Ann Hum Biol* 31(2):245–257. doi:10.1080/03014460310001652239
- Correll SJ (2001) Gender and the career choice process: the role of biased self-assessments. *Am J Sociol* 106(6):1691–1730. doi:10.1086/321299
- Crosnoe R (2000) Friendships in childhood and adolescence: the life course and new directions. *Soc Psychol Q* 63(4):377–391. doi:10.2307/2695847
- Davey Smith G, Ebrahim S (2005) What can Mendelian randomisation tell us about modifiable behavioural and environmental exposures? *BMJ* 330(7499):1076–1079. doi:10.1136/bmj.330.7499.1076
- Del Greco M F, Minelli C, Sheehan NA, Thompson JR (2015) Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Stat Med*. doi:10.1002/sim.6522
- Didelez V, Meng S, Sheehan NA (2010) Assumptions of IV methods for observational epidemiology. *Stat Sci* 25(1):22–40. doi:10.1214/09-STS316
- Dvornyk V, Waqar ul H (2012) Genetics of age at menarche: a systematic review. *Hum Reprod Update* 18(2):198–210. doi:10.1093/humupd/dmr050
- Graber JA (2013) Pubertal timing and the development of psychopathology in adolescence and beyond. *Horm Behav* 64(2):262–269. doi:10.1016/j.yhbeh.2013.04.003
- Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG (2001) Replication validity of genetic association studies. *Nat Genet* 29(3):306–309. doi:10.1038/ng749
- Kingston PW, Hubbard R, Lapp B, Schroeder P, Wilson J (2003) Why education matters. *Sociol Educ* 76(1):53–70. doi:10.2307/3090261
- Koivusilta L, Rimpela A (2004) Pubertal timing and educational careers: a longitudinal study. *Ann Hum Biol* 31(4):446–465. doi:10.1080/03014460412331281719
- Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G (2008) Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* 27(8):1133–1163. doi:10.1002/sim.3034
- Li B, Martin EB (2002) An approximation to the F distribution using the chi-square distribution. *Comput Stat Data Anal* 40:21–26
- Li J, Powdthavee N (2015) Does more education lead to better health habits? Evidence from the school reforms in Australia. *Soc Sci Med* 127:83–91. doi:10.1016/j.socscimed.2014.07.021
- Mendle J, Turkheimer E, Emery RE (2007) Detrimental psychological outcomes associated with early pubertal timing in adolescent girls. *Dev Rev* 27(2):151–171. doi:10.1016/j.dr.2006.11.001
- Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA, Turley P, Chen GB, Emilsson V, Meddens SF, Oskarsson S (2016) Genome-wide association study identifies 74 loci

- associated with educational attainment. *Nature* 533(7604):539–542. doi:[10.1038/nature17671](https://doi.org/10.1038/nature17671)
- Palmer TM, Lawlor DA, Harbord RM, Sheehan NA, Tobias JH, Timpson NJ, Smith GD, Sterne JA (2012) Using multiple genetic variants as instrumental variables for modifiable risk factors. *Stat Methods Med Res* 21(3):223–242. doi:[10.1177/0962280210394459](https://doi.org/10.1177/0962280210394459)
- Perry JR, Day F, Elks CE, Sulem P, Thompson DJ, Ferreira T, He C, Chasman DI, Esko T, Thorleifsson G, Albrecht E (2014) Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. *Nature* 514(7520):92–97. doi:[10.1038/nature13545](https://doi.org/10.1038/nature13545)
- Piekarski DJ, Johnson CM, Boivin JR, Thomas AW, Lin WC, Delovich K, Galarce EM, Wilbrecht L (2016) Does puberty mark a transition in sensitive periods for plasticity in the associative neocortex? *Brain Res*. doi:[10.1016/j.brainres.2016.08.042](https://doi.org/10.1016/j.brainres.2016.08.042)
- Schulz KM, Sisk CL (2016) The organizing actions of adolescent gonadal steroid hormones on brain and behavioral development. *Neurosci Biobehav Rev*. doi:[10.1016/j.neubiorev.2016.07.036](https://doi.org/10.1016/j.neubiorev.2016.07.036)
- Sequeira ME, Lewis SJ, Bonilla C, Davey Smith G, Joinson C (2016) Association of timing of menarche with depressive symptoms and depression in adolescence: Mendelian randomisation study. *Br J Psychiatry*. doi:[10.1192/bjp.bp.115.168617](https://doi.org/10.1192/bjp.bp.115.168617)
- Sheehan NA, Didelez V, Burton PR, Tobin MD (2008) Mendelian randomisation and causal inference in observational epidemiology. *PLoS Med* 5(8):e177. doi:[10.1371/journal.pmed.0050177](https://doi.org/10.1371/journal.pmed.0050177)
- Staley JR, Blackshaw J, Kamat MA, Ellis S, Surendran P, Sun BB, Paul DS, Freitag D, Burgess S, Danesh J, Young R (2016) PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics* 32(20):3207–3209. doi:[10.1093/bioinformatics/btw373](https://doi.org/10.1093/bioinformatics/btw373)
- Stattin H, Magnusson D (1990) Pubertal maturation in female development. Lawrence Erlbaum Associates, Hillsdale
- Stice E, Burton EM, Shaw H (2004) Prospective relations between bulimic pathology, depression, and substance abuse: unpacking comorbidity in adolescent girls. *J Consult Clin Psychol* 72(1):62–71. doi:[10.1037/0022-006X.72.1.62](https://doi.org/10.1037/0022-006X.72.1.62)
- Thompson JR, Minelli C, Del Greco M F (2016) Mendelian randomization using public data from genetic consortia. *Int J Biostat*. doi:[10.1515/ijb-2015-0074](https://doi.org/10.1515/ijb-2015-0074)
- Westling E, Andrews JA, Peterson M (2012) Gender differences in pubertal timing, social competence, and cigarette use: a test of the early maturation hypothesis. *J Adolesc Health* 51(2):150–155. doi:[10.1016/j.jadohealth.2011.11.021](https://doi.org/10.1016/j.jadohealth.2011.11.021)
- Winding TN, Nohr EA, Labriola M, Biering K, Andersen JH (2013) Personal predictors of educational attainment after compulsory school: influence of measures of vulnerability, health, and school performance. *Scand J Public Health* 41(1):92–101. doi:[10.1177/1403494812467713](https://doi.org/10.1177/1403494812467713)