Epidemiology

Adult height is associated with increased risk of ovarian cancer: a Mendelian randomisation study

Suzanne C. Dixon-Suen et al.

BACKGROUND: Observational studies suggest greater height is associated with increased ovarian cancer risk, but cannot exclude bias and/or confounding as explanations for this. Mendelian randomisation (MR) can provide evidence which may be less prone to bias.

METHODS: We pooled data from 39 Ovarian Cancer Association Consortium studies (16,395 cases; 23,003 controls). We applied two-stage predictor-substitution MR, using a weighted genetic risk score combining 609 single-nucleotide polymorphisms. Study-specific odds ratios (OR) and 95% confidence intervals (CI) for the association between genetically predicted height and risk were pooled using random-effects meta-analysis.

RESULTS: Greater genetically predicted height was associated with increased ovarian cancer risk overall (pooled-OR (pOR) = 1.06; 95% CI: 1.01–1.11 per 5 cm increase in height), and separately for invasive (pOR = 1.06; 95% CI: 1.01–1.11) and borderline (pOR = 1.15; 95% CI: 1.02–1.29) tumours.

CONCLUSIONS: Women with a genetic propensity to being taller have increased risk of ovarian cancer. This suggests genes influencing height are involved in pathways promoting ovarian carcinogenesis.

INTRODUCTION

Observational studies have reported a positive association between adult height and ovarian cancer risk. However, these studies were subject to the biases inherent in conventional observational studies, including selection bias, differential and non-differential reporting bias and confounding. The degree to which these factors could account for the observed association is uncertain. Mendelian randomisation (MR) uses genetic markers as proxies for environmental exposures and, due to the singular qualities of genotype data, can provide complementary evidence by overcoming many biases affecting conventional studies. We used MR to examine the relationship between height and ovarian cancer risk in the Ovarian Cancer Association Consortium (OCAC), aiming to provide more certainty about the relationship between height and ovarian cancer risk. We hypothesised that greater genetically predicted height would be associated with increased risk.

MATERIALS AND METHODS

Study population and outcomes
We pooled data from 16,395 genetically European women with primary ovarian/fallopian tube/peritoneal cancer and 23,003 controls from 39 OCAC studies (Table 1; Supplementary Table 1). The data set and methods have been described previously. Participants were genotyped via the Collaborative Oncological Gene-Environment Study. Twenty-two studies provided height data (16 provided parity, oral contraceptive (OC) use, education and age at menarche information) for >50% of their participants. We first considered all cases, then stratified by tumour behaviour. Secondary analyses stratified by histologic subtype/behaviour.

Genetic risk score
The Genetic Investigation of ANthropometric Traits (GIANT) Consortium had previously identified 697 single-nucleotide polymorphisms (SNPs) significantly associated with height. In our sample, 92 of these SNPs had been genotyped and the remainder were imputed using 1000 Genome Project data. After excluding poorly-imputed SNPs (quality $R^2 < 0.6$), 609 remained (92 genotyped/517 imputed) (Supplementary Table 2). In controls, minor allele frequencies (MAFs) were >5% (except for 16 SNPs, MAFs 1.7–4.9%).

We constructed a weighted genetic risk score (GRS) for height by summing height-increasing alleles across the 609 SNPs ('GRS-609', the GRS), weighting alleles by $\beta$-coefficients for their association with height reported by GIANT. The score represents predicted additional height conferred by these variants, compared to having no height-increasing alleles. We report results for 5 cm increments.

Statistical analysis
Statistical methods have been described previously. Briefly, we used individual-level OCAC data for two-stage predictor-substitution MR. first, we predicted height from the weighted GRS for
all participants using coefficients from linear regression in 17,649 controls with height data; second, within each study, we used logistic regression to model disease status on GRS-predicted height. Models adjusted for age and five principal components for population structure. We combined study-specific estimates using meta-analysis, generating pooled odds ratios (pOR) and 95% confidence intervals (CI) for the trend in risk per 5 cm increase in predicted height. We had 97% power to detect an OR of 1.10 (mRND tool). Sensitivity analyses included removing 16 SNPs with MAFs <5%, and restricting to SNPs with imputation $r^2 > 0.9$ (GRS-363), SNPs representing distinct loci (GRS-377), and directly-genotyped
SNPs (‘GRS-92’). We examined whether potential confounders of the association in observational studies were associated with the GRS. To assess robustness to pleiotropy (where SNPs may influence risk via pathways not mediated through height), we conducted MR-Egger regression and assessed smaller GRSS excluding SNPs with the highest probability of acting via other pathways from GRS to outcome (SNPs associated with ovarian/other hormonal cancers (breast, prostate), hormone levels and in/ near tumour initiation/growth genes). We identified these potentially pleiotropic, pathway-specific SNPs via the NHGRI GWAS Catalog, the UCSC Genome/Table Browsers and from lists of SNPs nominated for iCOGS genotyping by ovarian, breast and prostate cancer researchers (to capture SNPs of interest unpublished at the time of analysis).

Secondary analyses defined cases by histologic subtype/behaviour. Among 16 studies with height/confounder data, we conducted conventional analysis (adjusted for parity, OC use, education, menarche age; stratified by study, 5-year age group) and compared results with MR-estimates among the same women.

Analyses were performed using SAS 9.2 (SAS Institute Inc., Cary, NC) and Stata 13.0 (StataCorp LP, College Station, TX). This work and each contributing study was approved by the appropriate institutional review board/ethics committee. All participants provided informed consent.

RESULTS
Population characteristics
We included 16,395 cases (14,549 invasive tumours, 1691 borderline, 155 of unknown behaviour) and 23,003 controls (Table 1). The median diagnosis year was 2003, with 74% diagnosed post-2000. Participants were aged 18–94 (median 56 years) at diagnosis/interview. Mean height ranged from 159 to 167 cm across 22 studies with data, and was 163 (standard error (SE) = 0.06) cm for controls and 164 (SE = 0.05) cm for cases (p < 0.0001).
Genetic risk score characteristics
The GRS-609 was normally distributed in controls, ranging from 15.45 to 18.99 (median = 17.23; interquartile range = 16.92–17.54). It explained 13% of variance in height, 17% after adjusting for age and principal components (partial-$R^2 = 12$%; first-stage regression partial-$F$-statistic = 2505.8 (df = 1), $p < 0.001$). A 1-unit GRS-609 increase was associated with 5.2 cm greater height. Average height was 6.2 cm greater in the highest vs. lowest GRS quartile.

Cochran’s $I^2$ and $p$-values for heterogeneity showed no evidence of inter-study heterogeneity in the relationship between either the GRS-609 ($I^2 = 34$%, $p$-heterogeneity = 0.07) or the simplified GRS-363 ($I^2 = 32$%, $p$-heterogeneity = 0.08) and height among controls (Supplementary Figure 1a, b). The GRS-609 was not associated with most potential confounders of the height-ovarian cancer association in observational studies, including age, parity, OC use and education (Supplementary Table 3). The GRS was marginally positively associated with age at menarche ($p = 0.03$), consistent with known genetic overlap between these traits.

Primary outcomes
Women with greater genetically predicted height had a modestly increased risk of developing ovarian cancer ($p$ = 1.06, 95% CI: 1.01–1.11 per 5 cm) (Fig. 1a; Table 2) with a greater magnitude of association for borderline ($p$ = 1.15; 95% CI: 1.02–1.29) than invasive tumours ($p$ = 1.06; 95% CI: 1.01–1.11; Fig. 1b, c; Table 2). No significant inter-study heterogeneity was noted (Fig. 1a–c). GRS-363 ($p$ = 1.06, 95% CI: 1.00–1.11, all tumours) and GRS-377 (OR = 1.07; 95% CI: 1.01–1.12) results were similar to the GRS-609. The association was stronger when we restricted to 92 genotyped SNPs ($p$ = 1.14; 95% CI: 1.04–1.25). Estimates from analyses excluding low-MAF SNPs, excluding case-only studies, or adjusting for age at menarche, were similar to primary analyses. When we sequentially excluded SNPs associated with ovarian or other hormonal cancers, hormone levels and tumour development, estimates were similar or stronger than GRS-609 results. MR-Egger suggested minimal bias from pleiotropy ($p = 0.1$; MR-Egger beta = 0.163 corresponded to an OR per 5 cm of 1.13 (95% CI: 1.02–1.25), confirming a significant positive association).

In contrast, for women with height and confounder data (16 studies), the conventional analysis suggested no association (adjusted-OR = 1.01, 95% CI: 0.99–1.04 per 5 cm). Conducting MR within the same 16 studies yielded results similar to overall analyses (OR = 1.06, 95% CI: 1.00–1.13) (Supplementary Table 4).

Secondary outcomes
After stratifying by subtype/behaviour, the strongest associations were seen for clear cell (OR = 1.20, 95% CI: 1.04–1.38) and low-grade/borderline serous cancers (OR = 1.15, 95% CI: 1.01–1.30) (Table 2). However, CIs were wide and overlapping due to lower statistical power in these subgroup analyses. The estimate for clear cell cancers was also significantly elevated in our conventional analyses (Supplementary Table 4).

**DISCUSSION**
We used a 609-SNP GRS to examine the relationship between height and ovarian cancer risk for women of European ancestry. Our data indicate a modest positive association between genetically predicted height and ovarian cancer risk, which may be stronger for borderline cancers. Height may be relevant to cancer risk as a marker for lifetime growth-factor levels (e.g. IGF-1) and/or early-life exposures (socio-economic/environmental/nutritional).3, 21, 22

Observational studies are subject to biases (reverse causality, selection bias, differential/non-differential reporting, confounding) which cannot be ruled out as possible explanations for observed associations. By using genotype, the MR technique can overcome some of these biases, given three assumptions. We confirmed the two verifiable assumptions: the GRS was associated with height, and not with most known confounders. The GRS-menarche age association is unlikely to explain the observed association, because age at menarche is only weakly associated with ovarian cancer, and women with later menarche have if anything lower ovarian cancer risk, so if this affected our results, we would expect the true effect to be at least as strong as the reported association. Also, removing hormone-related SNPs, or adjusting for menarche age, did not attenuate estimates. Owing to the limited current biological understanding of all 609 SNPs, we could not

### Table 2. Association between increasing height (per 5 cm)—predicted by a weighted 609-locus genetic risk score—and risk of ovarian cancer, stratified by study

| Histologic subtypea | N studies | N controls | N cases | Odds ratios (95% CI)b |
|---------------------|-----------|------------|---------|-----------------------|
| **Primary outcomes** |           |            |         |                       |
| All ovarian cancers  | 39        | 23,003     | 16,395  | 1.06 (1.01–1.11)       |
| Invasive             | 39        | 23,003     | 14,549  | 1.06 (1.01–1.11)       |
| Borderlinec          | 20        | 16,463     | 1680    | 1.15 (1.02–1.29)       |
| **Secondary outcomes, by histologic subtype and behaviour** | | | | |
| Serous               |           |            |         |                       |
| High-grade4          | 39        | 23,003     | 7933    | 1.05 (0.99–1.11)       |
| Invasive low-grade and borderline | 32       | 21,131     | 1408    | 1.15 (1.01–1.30)       |
| Mucinous (invasive and borderline) | 38       | 22,410     | 1567    | 1.08 (0.96–1.21)       |
| Endometrioid (invasive) | 39       | 23,003     | 2059    | 1.05 (0.95–1.16)       |
| Clear cell (invasive) | 35        | 22,051     | 948     | 1.20 (1.04–1.38)       |

Weights applied were β-coefficients for the relationship between each SNP and height as reported in the meta-analysis of genome-wide association studies conducted by the Genetic Investigation of ANthropometric Traits (GIANT) Consortium. On the basis of the additive SNP effects suggested by GIANT, the score summed alleles across the 609 SNPs. For the 92 genotyped SNPs, where values were missing (<2.5% per SNP), we used imputed probabilities. Includes studies with >5 cases. *Pooled study-specific odds ratios are reported for primary outcomes; odds ratios stratified by study are reported for secondary outcomes (secondary analyses used single models stratified by study to maximise power). Of the 1691 borderline tumours included in the all-case analysis, 1680 were from 20 studies with >5 cases each. Includes all invasive serous cancers except low-grade (G1) (n = 469) as well as invasive serous cancers of unknown grade (n = 1957) and primary peritoneal cancers of unknown behaviour (n = 44), because in both instances the majority would be high-grade serous. CI confidence interval.
conclusively exclude the presence of alternate pathways from height genes to ovarian cancer (assumption three). However, MR-Egger and sensitivity analyses excluding pathway-specific SNPs provided some evidence for their absence, minimising the likelihood of our observed association is explained by pathways separate from height/growth. Although height data were not available for the entire population, this is unlikely to have affected our results as we used these data only to refine the height predictions from the GR, and there is no reason to believe the GRs-height relationship would be different for women with and without height data. Further strengths of our analysis include the large number of SNPs and power to detect modest differences.

Despite potential limitations of conventional observational studies, our MR-estimate is almost identical to previously reported associations, suggesting previous estimates were not appreciably biased. The World Cancer Research Fund/American Institute for Cancer Research meta-analysis of 24 prospective studies, and a study pooling 47 prospective/case–control studies, both reported a significant 7–8% increase in risk (combining invasive/borderline cancers) per 5 cm height increase.3, 4 The lack of association seen in the OCAC conventional height analysis reflects the greater potential for bias in case–control studies and illustrates the value of MR in overcoming these biases. Few previous studies have examined borderline cancers separately, a strength of our analysis. Previous observational studies have not reported consistent patterns by histologic subtype2, 4, 23; our secondary analyses of MR estimates for invasive and borderline cancers separately confirm a GRS-height relationship would be different for women with and without height data. Further strengths of our analysis include the large number of SNPs and power to detect modest differences.

Using MR, we have established that the previously observed association between height and ovarian cancer risk is unlikely to have been explained by bias, and that genetic factors influencing height play roles in ovarian cancer development. Height could therefore be used, with other risk factors, to identify women at elevated risk. Further research should continue to explore mechanisms underpinning this association.

ACKNOWLEDGEMENTS

We thank all the individuals who took part in this study and all the researchers, clinicians, and technical and administrative staff who have made possible many of the new insights contributing to this work. In particular, for their contribution to the design and conduct of the individual studies that contributed to the analysis, we thank: D. Bowtell, A. deFazio, D. Gertig, A. Green, P. Parsons, N. Hayward and D. Whiteman (AUS); G. Peuteman, T. Van Brussel and D. Smeets (BEL); U. Elieber (GER); S. Reckemeyer, A. Korotkaia and S. Reina-Campanon (HJO); J. Meyer, C. Hilker, S. Reckemeyer, A. Korotkaia and S. Reina-Campanon (HJO); the SEARCH team, C. Cuccorini, C. Baynes, and D. Conroy (SEA); the Scottish Gynaecological Clinical Trials group and SCOTROC1 investigators (SRO); I. Jacobs, M. Widschwendter, E. Wozniak, A. Ryan, J. Ford, N. Balogun and C. Karpinski (UKO); and C. Pye (UKR).

ADDITIONAL INFORMATION

Supplementary information is available for this paper at https://doi.org/10.1038/s41416-018-0011-3.

Competing interests: The authors declare no competing financial interests.

Funding: This work was supported by: the National Cancer Institute at the U.S. National Institutes of Health [RO1 CA069566, RO1 CA08066, RO1 CA130407, RO1 CA138563, RO1-CN25403, RO1-CN55242, RO1-PC67001, RO1-PC67010, RO1-CA17054, P30-CA072720, P30-CA08748, P30-CA14089, P30-CA15083, P50-CA15009, P50-CA136393, P50-CA159981, R15A035860, R01 CA063876, R01 CA063682, R01 CA092204, R01-CA095523, R01-CA10655, R01-CA54419, R01-CA58598, R01-CA61107, R01-CA61132, R01-CA76016, R01-CA83918, R01-CA87538, R01-CA112523, R01-CA122443, R03-CA113148, R03-CA15195, U01-CA69417, U01-CA71966 and Intramural Research funds]; the European Commission’s Seventh Framework Programme (agreement number 223175 HEALTH F2 2009-223175); Cancer Research UK [C490/A16561, C536/A10806, C536/A6689, C2187/A1011, C2187/A 1070, C2129/A11174, C2182/A12014, C5047/A8384, C5047/A15007, C5047/A10692 and CB197/A16565]; a National Institutes of Health (CA128978), Post-Cancer GWAS Initiative [1U19 CA148537, 1U19 CA148065 and 1U19 CA14812—The Genetic Associations and Mechanisms in Oncology (GAME-ON) initiative]; the U.S. Department of Defense [DAMD17-02-1-0669, W81XWH-07-0449, DAMD17-02-1-0666, W81XWH-10-1-0280 and W81XWH-10-1-0341]; the Canadian Institutes of Health Research (CIHR) [MOP-86727 and MOP-87734 to L.E.K.] and the CIHR Team in Familial Risks of Breast Cancer; the Komen Foundation for the Cure; the Breast Cancer Research Foundation; the Ovarian Cancer Research Fund (thanks to donations by the family and friends of Kathryn Sladek Smith); the U.S. Army Medical Research and Materiel Command [DAMD17-01-1-0729 and DAMD17-02-1-0669]; the National Health and Medical Research Council of Australia [199600, 209057, 251553, 402081, 504711, 1073898 and fellowships to G.C.T. and P.M.W.]; Cancer Councils of Queensland, Victoria, New South Wales, South Australia and Tasmania and the Cancer Foundation of Western Australia [Multi-State Grant Applications 191, 211 and 182]; VicHealth; the ELAN Program of the University of Erlangen-Nuremberg; the National Kankerplan of Belgium; the German Federal Ministry of Education and Research of Germany Programme of Clinical Biomedical Research [01 GB 9401]; the German Cancer Research Center; the Roswell Park Cancer Institute Alliance Foundation [3P0 CA016056]; the Rudolf-Bartling Foundation; the Helsinki University Central Hospital Research Fund; the National Institutes of Health/National Center for Research Resources/General Clinical Research Center [MO1 RR000056]; an American Cancer Society Early Detection Professorship [SIOP-06-258-01-COUN to B.Y.K.]; the National Center for Advancing Translational Sciences (NCATS) [UL1TR000124 to B.Y.K.]; the Danish Cancer Society [94-222-52]; the Mermaid I project; the Mayo Foundation; the Minnesota Ovarian Cancer Alliance; the Fred C. and Katherine B. Andersen Foundation; the Cancer Foundation; the New Institute of New Jersey; Helse Vest; the Norwegian Cancer Society: the Research Council of Norway; Radboud University Medical Centre; the Oregon Health and Science University (OHSU) Foundation; Pomeranian Medical University; the UK National Institute for Health Research Biomedical Research Centres at the University of Cambridge, University College London Hospital, and the Royal Marsden Hospital; the Imperial Experimental Cancer Research Centre [C1312/A15389]; the U.S. Public Health Service (PSA-04205); the Lon V. Smith Foundation [LVS-39420]; The Eve Appeal; The Oak Foundation; the California Cancer Research Program [00-01389V-20170 and 201020]; the Polish Ministry of Science and Higher Education [4 POSC 028 14 and 2 POSA 068 27]; and the Maria Skłodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw.

Note: This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution 4.0 International licence (CC BY 4.0).

REFERENCES

1. Emerging Risk Factors Collaboration. Adult height and the risk of cause-specific death and vascular morbidity in 1 million people: individual participant meta-analysis. Int. J. Epidemiol. 41, 1419–1433 (2012).
2. Schouten, L. J. et al. Height, body mass index, and ovarian cancer: a pooled analysis of 12 cohort studies. Cancer Epidemiol. Biomark. Prev. 17, 902–912 (2008).
3. Collaborative Group on Epidemiological Studies of Ovarian Cancer. Ovarian cancer and body size: individual participant meta-analysis including 25,157 women with ovarian cancer from 47 epidemiological studies. PLoS Med. 9, e1001200 (2012).
4. Davey Smith, G. & Ebrahim, S. ‘Mendelian randomization’: can genetic epidemiology contribute to understanding environmental determinants of disease? Int. J. Epidemiol. 32, 1–22 (2003).
5. Gayther, S. A. et al. Tagging single nucleotide polymorphisms in cell cycle control genes and susceptibility to invasive epithelial ovarian cancer. Cancer Res. 67, 3027–3035 (2007).
6. Dixon, S. C. et al. Adult body mass index and risk of ovarian cancer by subtype: a Mendelian randomization study. Int. J. Epidemiol. 45, 884–895 (2016).
7. Pharoah, P. D. P. et al. GWAS meta-analysis and replication identifies three new susceptibility loci for ovarian cancer. Nat. Genet. 45, 362–370 (2013).
8. Wood, A. R. et al. Defining the role of common variation in the genomic and biological architecture of adult human height. Nat. Genet. 46, 1173–1186 (2014).
9. Abecasis, G. R. et al. A map of human genome variation from population-scale sequencing. Nature 467, 1061–1073 (2010).
10. Burgess, S. Identifying the odds ratio estimated by a two-stage instrumental variable analysis with a logistic regression model. Stat. Med. 32, 4726–4747 (2013).
11. Didelot, V., Meng, S. & Sheehan, N. A. Assumptions of IV methods for observational epidemiology. Stat. Sci. 25, 22–40 (2010).
