Association between genetic polymorphisms and endometrial cancer risk: a systematic review

Cemsel Bafligil, Deborah J Thompson, Artitaya Lophatnananon, Miriam J Smith, Neil Aj Ryan, Anie Naqvi, D Gareth Evans, Emma J Crosbie

Department of Obstetrics, University of Manchester, Manchester, UK;
Division of Population Sciences, University of Manchester, Manchester, UK;
Division of Evolution and Genomic Sciences, University of Manchester, Manchester, UK;
University of Manchester Medical School, Manchester, UK;
Department of Obstetrics and Gynaecology, Manchester University NHS Foundation Trust, Manchester, UK

ABSTRACT

Introduction Endometrial cancer is one of the most commonly diagnosed cancers in women. Although there is a hereditary component to endometrial cancer, most cases are thought to be sporadic and lifestyle related. The aim of this study was to systematically review prospective and retrospective case–control studies, meta-analyses and genome-wide association studies to identify genomic variants that may be associated with endometrial cancer risk.

Methods We searched MEDLINE, Embase and Cinahl from 2007 to 2019 without restrictions. We followed PRISMA 2009 guidelines. The search yielded 3015 hits in total. Following duplicate exclusion, 2674 abstracts were screened and 453 full-texts evaluated based on our pre-defined screening criteria. 149 articles were eligible for inclusion.

Results We found that single nucleotide polymorphisms (SNPs) in HNF1B, KLF, EIF2AK, CYP19A1, SOX4 and MYC were strongly associated with incident endometrial cancer. Nineteen variants were reported with genome-wide significance and a further five with suggestive significance. No convincing evidence was found for the widely studied MDM2 variant rs2279744. Publication bias and false discovery rates were noted throughout the literature.

Conclusion Endometrial cancer risk may be influenced by SNPs in genes involved in cell survival, oestrogen metabolism and transcriptional control. Larger cohorts are needed to identify more variants with genome-wide significance.

INTRODUCTION

Endometrial cancer is the most common gynaecological malignancy in the developed world.1 Its incidence has risen over the last two decades as a consequence of the ageing population, fewer hysterectomies for benign disease and the obesity epidemic. In the USA, it is estimated that women have a 1 in 35 lifetime risk of endometrial cancer, and in contrast to cancers of most other sites, cancer-specific mortality has risen by approximately 2% every year since 2008 related to the rapidly rising incidence.2

Endometrial cancer has traditionally been classified into type I and type II based on morphology.3 The more common subtype, type I, is mostly comprised of endometrioid tumours and is oestrogen-driven, arises from a hyperplastic endometrium, presents at an early stage and has an excellent 5-year survival rate.4 By contrast, type II includes non-endometrioid tumours, specifically serous, carcinoma and clear cell subtypes, which are biologically aggressive tumours with a poor prognosis that are often diagnosed at an advanced stage.5 Recent efforts have focused on a molecular classification system for more accurate categorisation of endometrial tumours into four groups with distinct prognostic profiles.6 7

The majority of endometrial cancers arise through the interplay of familial, genetic and lifestyle factors. Two inherited cancer predisposition syndromes, Lynch syndrome and the much rarer Cowden syndrome, substantially increase the lifetime risk of endometrial cancer, but these only account for around 3–5% of cases.8 9 Having first or second degree relative(s) with endometrial or colorectal cancer increases endometrial cancer risk, although a large European twin study failed to demonstrate a strong heritable link.10 The authors failed to show that there was greater concordance in monozygotic than dizygotic twins, but the study was based on relatively small numbers of endometrial cancers. Lu and colleagues reported an association between common single nucleotide polymorphisms (SNPs) and endometrial cancer risk, revealing the potential role of SNPs in explaining part of the risk in both the familial and general populations.12 Thus far, many SNPs have been reported to modify susceptibility to endometrial cancer; however, much of this work predates genome wide association studies and is of variable quality. Understanding genetic predisposition to endometrial cancer could facilitate personalised risk assessment with a view to targeted prevention and screening interventions.13

This emerged as the most important unanswered research question in endometrial cancer according to patients, carers and healthcare professionals in our recently completed James Lind Womb Cancer Alliance Priority Setting Partnership.14 It would be particularly useful for non-endometrioid endometrial cancers, for which advancing age is so far the only predictor.15

We therefore conducted a comprehensive systematic review of the literature to provide an overview of the relationship between SNPs and endometrial cancer risk. We compiled a list of the most robust endometrial cancer-associated SNPs. We assessed the applicability of this panel of SNPs with a theoretical polygenic risk score (PRS) calculation. We also critically appraised the meta-analyses investigating the
most frequently reported SNPs in MDM2. Finally, we described all SNPs reported within genes and pathways that are likely involved in endometrial carcinogenesis and metastasis.

METHODS
Our systematic review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) collaboration 2009 recommendations. The registered protocol is available through PROSPERO (CRD42018091907).16

Search strategy
We searched Embase, MEDLINE and Cumulative Index to Nursing and Allied Health Literature (CINAHL) databases via the Healthcare Databases Advanced Search (HDAS) platform, from 2007 to 2018, to identify studies reporting associations between polymorphisms and endometrial cancer risk. Key words including MeSH (Medical Subject Heading) terms and free-text words were searched in both titles and abstracts. The following terms were used: “endomet*”, “uter*”, “womb”, “cancer(s)”, “neoplasm(s)”, “endometrium tumour”, “carcinoma”, “adenocarcinoma”, “clear cell carcinoma”, “carcinosarcoma”, “SNP”, “single nucleotide polymorphism”, “GWAS”, and “genome-wide association study/ies”. No other restrictions were applied. The search was repeated with time restrictions between 2018 and June 2019 to capture any recent publications.

Eligibility criteria
Studies were selected for full-text evaluation if they were primary articles investigating a relationship between endometrial cancer and SNPs. Study outcome was either the increased or decreased risk of endometrial cancer relative to controls reported as an odds ratio (OR) with corresponding 95% confidence intervals (95% CIs).

Study selection
Three independent reviewers screened all articles uploaded to a screening spreadsheet developed by Helena VonVille.17 Disagreements were resolved by discussion. Chronbach’s α score was calculated between reviewers and indicated high consistency at 0.92. Case–control, prospective and retrospective studies, genome-wide association studies (GWAS), and both discovery and validation studies were selected for full-text evaluation. Non-English articles, editorials, conference abstracts and proceedings, letters and correspondence, case reports and review articles were excluded.

Candidate-gene studies with at least 100 women and GWAS with at least 1000 women in the case arm were selected to ensure reliability of the results, as explained by Spencer et al.18 To construct a panel of up to 30 SNPs with the strongest evidence of association, those with the strongest p values were selected. For the purpose of an SNP panel, articles utilising broad European or multi-ethnic cohorts were selected. Where overlapping populations were identified, the most comprehensive study was included.

Data extraction and synthesis
For each study, the following data were extracted: SNP ID, nearby gene(s)/chromosome location, OR (95% CI), p value, minor or effect allele frequency (MAF/EAF), EA (effect allele) and OA (other allele), adjustment, ethnicity and ancestry, number of cases and controls, endometrial cancer type, and study type including discovery or validation study and meta-analysis. For risk estimates, a preference towards most adjusted results was applied. For candidate-gene studies, a standard p value of <0.05 was applied and for GWAS a p value of <5×10^{-8}, indicating genome-wide significance, was accepted as statistically significant. However, due to the limited number of SNPs with p values reaching genome-wide significance, this threshold was then lowered to <1×10^{-5}, allowing for marginally significant SNPs to be included. As shown by Mavaddat et al,19 for breast cancer, SNPs that fall below genome-wide significance may still be useful for generating a PRS and improving the models.

We estimated the potential value of a PRS based on the most significant SNPs by comparing the predicted risk for a woman with a risk score in the top 1% of the distribution to the mean predicted risk. Per-allele ORs and MAFs were taken from the publications and standard errors (SEs) for the lnORs were derived from published 95% CIs. The PRS was assumed to have a Normal distribution, with mean 2∑pIβIpI and SE, σ, equal to √2∑pIβ2IpI(1−pI), according to the binomial distribution, where the summation is over all SNPs in the risk score. Hence the relative risk (RR) comparing the top 1% of the distribution to the mean is given by exp(Zα0.01σ), where Z is the inverse of the standard normal cumulative distribution.

RESULTS
The flow chart of study selection is illustrated in figure 1. In total, 453 text articles were evaluated and, of those, 149 articles met our inclusion criteria. One study was excluded from table 1, for having an Asian-only population, as this would make it harder to compare with the rest of the results which were all either multi-ethnic or Caucasian cohorts, as stated in our inclusion criteria for the SNP panel.20 Any SNPs without 95% CIs were also excluded from any downstream analysis. Additionally,
Table 1  List of top SNPs most likely to contribute to endometrial cancer risk identified through systematic review of recent literature21-25

| Reference | SNP ID | Nearby gene(s) | Location | OR (95% CI) | UCI (95% CI) | P | EAF | EA | QQ | Ethnicity | Cases (n) | Controls (n) | EC type | Position | Datasets* |
|-----------|--------|----------------|----------|-------------|--------------|---|-----|----|----|-----------|----------|--------------|--------|----------|-----------|
| O’Mara et al, 201821 | n11263761 | HNF1B | 17q12 | 1.15 | 1.12 | 1.19 | 3.20e-20 | 0.52 | A | G | EUR | 12906 | 108979 | All | Intrinsic | NSECG, UK1-CORGI, SEARCH, WTCCC, ANECs, QIMR, HCS, EZC2, BCCS, HIECS, ICOGS, BBCC, BSUCH, ESTHER, GC-HBC, GENICA, MARIE, MoMaTec, NCBS, BCCS, SBCS, UKBGS, NECS, ABCFS, ABCB1, BCCES, MCCS, LCS, LES, LBCC, BECS, GESBC, HABCs, CAHRES, RENDOCAS, MISS, PKARMA, SMC, CBIR, STUDY98, MECS, MCBCC, MMBH, WH, UKBB |
| O’Mara et al, 201821 | n7981863 | KLFS, KLFI 2 | 13q22.1 | 1.16 | 1.12 | 1.20 | 2.70e-17 | 0.72 | C | T | EUR | 12906 | 108979 | All | Intrinsic | NSECG, UK1-CORGI, SEARCH, WTCCC, ANECs, QIMR, HCS, EZC2, BCCS, HIECS, ICOGS, BBCC, BSUCH, ESTHER, GC-HBC, GENICA, MARIE, MoMaTec, NCBS, BCCS, SBCS, UKBGS, NECS, ABCFS, ABCB1, BCCES, MCCS, LCS, LES, LBCC, BECS, GESBC, HABCs, CAHRES, RENDOCAS, MISS, PKARMA, SMC, CBIR, STUDY98, MECS, MCBCC, MMBH, WH, UKBB |
| O’Mara et al, 201821 | n1740828 | SOX4 | 6q22.3 | 1.15 | 1.11 | 1.19 | 4.20e-16 | 0.52 | G | A | EUR | 12906 | 108979 | All | Regulatory region | NSECG, UK1-CORGI, SEARCH, WTCCC, ANECs, QIMR, HCS, EZC2, BCCS, HIECS, ICOGS, BBCC, BSUCH, ESTHER, GC-HBC, GENICA, MARIE, MoMaTec, NCBS, BCCS, SBCS, UKBGS, NECS, ABCFS, ABCB1, BCCES, MCCS, LCS, LES, LBCC, BECS, GESBC, HABCs, CAHRES, RENDOCAS, MISS, PKARMA, SMC, CBIR, STUDY98, MECS, MCBCC, MMBH, WH, UKBB |
| Chen et al, 201426 | n1953358 | UNCO0520 | 14q22.3 | 1.36 | 1.20 | 1.53 | 4.76e-07 | 0.49 | G | A | ME | 1055 | 1778 | Endometrioid | Intrinsic | AHS, EDGE, FHCR, MECS |
| Zhan et al, 201136 | n8718643 | PROS1 | 17q12 | 1.10 | 1.06 | 1.13 | 4.30e-08 | 0.38 | G | A | EUR | 12906 | 108979 | All | Regulatory region | NSECG, UK1-CORGI, SEARCH, WTCCC, ANECs, QIMR, HCS, EZC2, BCCS, HIECS, ICOGS, BBCC, BSUCH, ESTHER, GC-HBC, GENICA, MARIE, MoMaTec, NCBS, BCCS, SBCS, UKBGS, NECS, ABCFS, ABCB1, BCCES, MCCS, LCS, LES, LBCC, BECS, GESBC, HABCs, CAHRES, RENDOCAS, MISS, PKARMA, SMC, CBIR, STUDY98, MECS, MCBCC, MMBH, WH, UKBB |
| Smeenk et al, 201437 | n9399640 | N/A | 6q16.3 | 1.33 | 1.18 | 1.49 | 3.01e-06 | 0.53 | T | C | ME | 1055 | 1778 | Endometrioid | Intrinsic | AHS, EDGE, FHCR, MECS |

Continued
SNPs in linkage disequilibrium ($r^2 > 0.2$) with each other were examined, and of those in linkage disequilibrium, the SNP with strongest association was reported. Per allele ORs were used unless stated otherwise.

### Top SNPs associated with endometrial cancer risk

Following careful interpretation of the data, 24 independent SNPs with the lowest $p$ values that showed the strongest association with endometrial cancer were obtained (table 1).21–25 These SNPs are located in or around genes coding for transcription factors, cell growth and apoptosis regulators, and enzymes involved in the steroidogenesis pathway. All the SNPs presented here were reported on the basis of a GWAS or in one case, an exome-wide association study, and hence no SNPs from candidate-gene studies made it to the list. This is partly due to the nature of larger GWAS providing more comprehensive and powered results as opposed to candidate gene studies. Additionally, a vast majority of SNPs reported by candidate-gene studies were later refuted by large-scale GWAS such as in the case of TERT and MDM2 variants.26–27 The exception to this is the CYP19 gene, where candidate-gene studies reported an association between variants in this gene with endometrial cancer in both Asian and broad European populations, and this association was more recently confirmed by large-scale GWAS.21 26–30 Moreover, a recent article authored by O’Mara and colleagues reviewed the GWAS that identified most of the currently known SNPs associated with endometrial cancer.31

Most of the studies represented in table 1 are GWAS and the majority of these involved broad European populations. Those having a multi-ethnic cohort also consisted primarily of broad European populations. Only four of the variants in table 1 are located in coding regions of a gene, or in regulatory flanking regions around the gene. Thus, most of these variants would not be expected to cause any functional effects on the gene or the resulting protein. An eQTL search using GTEx Portal showed that some of the SNPs are significantly associated ($p<0.05$) with modified transcription levels of the respective genes in various tissues such as prostate (rs11263761), thyroid (rs9668337), pituitary (rs2747716), breast mammary (rs882380) and testicular (rs2498794), as summarised in table 2.

The only variant for which there was an indication of a specific association with non-endometrioid endometrial cancer was rs148261157 near the BCL11A gene. The A allele of this SNP had a moderately higher association in the non-endometrioid arm (OR 1.64, 95%CI 1.32 to 2.04; $p=9.6 \times 10^{-5}$) compared with the endometrioid arm (OR 1.25, 95%CI 1.14 to 1.38; $p=4.7 \times 10^{-5}$).21

Oestrogen receptors $\alpha$ and $\beta$ encoded by ESR1 and ESR2, respectively, have been extensively studied due to the assumed role of oestrogens in the development of endometrial cancer. O’Mara et al reported a lead SNP (rs7957594) in the ESR1 region that was associated with endometrial cancer ($p=1.86 \times 10^{-5}$).23 However, this SNP did not reach genome-wide significance in a more recent large GWAS.21 No statistically significant associations have been reported between endometrial cancer and SNPs in the ESR2 gene region.

AKT is an oncogene linked to endometrial carcinogenesis. It is involved in the PI3K/AKT/mTOR pro-proliferative signalling pathway to inactivate apoptosis and allow cell survival. The A allele of rs2494737 and G allele of rs2498796 were reported to be associated with increased and decreased risk of endometrial cancer in 2016, respectively.22 30 However, this association was not replicated in a larger GWAS in 2018.21 Nevertheless, given...
Table 2  List of eQTL hits for the selected panel of SNPs

| SNP ID   | Significant eQTL for | P        | Tissue               | Other gene(s)          | Other tissue(s)                                      |
|----------|----------------------|----------|----------------------|------------------------|-----------------------------------------------------|
| rs17601876 | GLDN                 | 1.2e-08  | Adipose – subcutaneous | SPPL2A, DMXL2          | Skin – sun exposed (lower leg); colon – sigmoid; cells – cultured fibroblasts; muscle – skeletal; spleen; skin – not sun exposed (suprapubic); nerve – tibial |
|          | CYP19A1              | 3.4e-07  | Whole blood           |                        |                                                     |
|          | CYP19A1              | 5.8e-06  | Adipose – subcutaneous |                        |                                                     |
| rs3184504  | TME1116              | 1.7e-04  | Adipose – subcutaneous | ALDH2, LINC01405, ADAM1B | Oesophagus – mucosa; skin – not sun exposed (suprapubic); skin – not exposed (lower leg); muscle – skeletal; artery – aorta; heart – atrial appendage; artery – tibial; colon – sigmoid; brain – nucleus accumbens (basal ganglia) |
|          | MAPKAPK5             | 2.6e-04  | Adipose – subcutaneous |                        |                                                     |
| rs2747716  | RP11-62M4.81         | 4.2e-11  | Pituitary             | HDDD2                  | Artery – tibial; pancreas; thyroid; brain – nucleus accumbens (basal ganglia); brain – substantia nigra; oesophagus – muscularis; nerve – tibial; Brain – cauda (basal ganglia); adipose – vesical (omentum); brain – spinal cord (cervical c-1); artery – aorta; artery – tibial; heart – hypothalamus; muscle – skeletal; brain – cerebellum; heart – left ventricle; brain – putamen (basal ganglia); brain – frontal cortex (BA9); brain – cerebellar hemisphere |
|          | RP11-62M4.81         | 8.2e-11  | Adipose – subcutaneous |                        |                                                     |
|          | HEY2                 | 9.7e-10  | Testis                |                        |                                                     |
|          | HEY2                 | 2.1e-09  | Ovary                 |                        |                                                     |
|          | RP11-62M4.81         | 1.7e-07  | Breast – mammary tissue |                        |                                                     |
|          | RP11-62M4.81         | 1.3e-06  | Ovary                 |                        |                                                     |
| rs9668337  | BHLHE41              | 9.0e-17  | Thyroid               | RP11-283G6.3           | Cells – cultured fibroblasts                        |
|          | SSPN                 | 1.1e-04  | Thyroid               |                        |                                                     |
| rs882380   | SNX11                | 3.1e-25  | Adipose – subcutaneous | RPS-890E16.5, CBX1, LRRC46, MRPL10, RP11-6N17.4, CDK5RAP3, SP6, PRR15L, RP1-890E16.2, PNPFR, RP11-6N17.3, HOXB1, HOXB-AS1, NFEL21 | Skin – sun exposed (lower leg); cells – cultured fibroblasts; adipose – vesical (omentum); lung; skin – not sun exposed (suprapubic); pancreas; spleen; oesophagus – muscularis; artery – aorta; heart – atrial appendage; liver; colon – transverse; thyroid; artery – tibial; colon – sigmoid; oesophagus – gastro-oesophageal junction; stomach; muscle – skeletal; small intestine – terminal ileum; prostate; brain – cerebellum; brain – cerebellar hemisphere; minor salivary gland; adrenal gland; oesophagus – mucosa |
|          | SNX11                | 1.0e-21  | Whole blood           |                        |                                                     |
|          | SNX11                | 1.2e-13  | Breast – mammary tissue |                        |                                                     |
|          | COP22                | 9.3e-12  | Testis                |                        |                                                     |
|          | SKAP1                | 3.3e-08  | Whole blood           |                        |                                                     |
|          | HOXB2                | 2.6e-05  | Adipose – subcutaneous |                        |                                                     |
| rs937213   | EIF2AK4              | 4.7e-11  | Adipose – visceral (omentum) | SRP14                  | Thyroid; oesophagus – mucosa; skin – sun exposed (lower leg); stomach; oesophagus – muscularis; pancreas; skin – not sun exposed (suprapubic); colon – transverse; adipose – subcutaneous; lung; colon – sigmoid; muscle – skeletal; nerve – tibial; whole blood; oesophagus – gastro-oesophageal junction; artery – tibial; adrenal gland; spleen; heart – left ventricle; heart – atrial appendage |
|          | EIF2AK4              | 3.4e-08  | Breast – mammary tissue | N/A                    |                                                     |
|          | RP11-521C20.5        | 5.4e-07  | Testis                | N/A                    |                                                     |
|          | RP11-521C20.5        | 7.4e-07  | Prostate              | N/A                    |                                                     |
| rs2498794  | AKT1                 | 1.7e-30  | Thyroid               | ZBTB42                 | Oesophagus – mucosa; artery – tibial; oesophagus – muscularis; skin – sun exposed (lower leg); skin – not sun exposed (suprapubic); cells – cultured fibroblasts; artery – aorta; oesophagus – gastro-oesophageal junction; adipose – subcutaneous; colon – sigmoid; colon – transverse; heart – atrial appendage |
|          | ADSSL1               | 5.5e-25  | Testis                | N/A                    |                                                     |
|          | SIVA1                | 1.8e-07  | Adipose – visceral (omentum) |                        |                                                     |
|          | ADSSL1               | 2.6e-05  | Ovary                 | N/A                    |                                                     |
|          | SIVA1                | 4.4e-05  | Breast – mammary tissue | N/A                    |                                                     |
| rs10835920 | WT1-A5               | 5.5e-06  | Spleen                | N/A                    | Oesophagus – muscularis                             |
| rs148261157 | KIAA1841            | 1.3e-05  | Oesophagus – muscularis | N/A                    |                                                     |
| rs113998067 | RSP01               | 2.7e-10  | Artery – tibial       | EPHA10, HFL3, DNAL1    | Nerve – tibial; artery – aorta; colon – transverse |
| rs1129506  | EVI2A                | 4.3e-20  | Whole blood           | OMG, RAB11FIP4         | Spleen; oesophagus – mucosa; artery – tibial; lung; artery – aorta; skin – exposed (lower leg); nerve – tibial; heart – atrial appendage; adipose – visceral (omentum); cells – cultured fibroblasts; liver; stomach; brain – amygdala; skin – not sun exposed (suprapubic); brain – cauda (basal ganglia); muscle – skeletal; colon – sigmoid |
|          | NF1                  | 3.5e-09  | Adipose – subcutaneous |                        |                                                     |
|          | NF1                  | 2.2e-07  | Thyroid               |                        |                                                     |
|          | NF1                  | 3.7e-07  | Testis                |                        |                                                     |
| rs673604   | ZMYM1                | 7.0e-07  | Adipose – subcutaneous | RP4-66SN4.8, ZMYM4, KIAA0319L, TFAP2E | Skin – sun exposed (lower leg); oesophagus – muscularis; cells – EBV-transformed lymphocytes; oesophagus – mucosa; nerve – tibial; brain – cerebellum |
|          | MAP7D1               | 1.0e-05  | Whole blood           |                        |                                                     |
| rs1953358  | LINC00520            | 1.5e-05  | Skin – not sun exposed (suprapubic) | N/A                    |                                                     |
| rs8178648  | PROS1                | 3.0e-04  | Skin – sun exposed (lower leg) | N/A                    |                                                     |

Top significant eQTL hits from different tissues are shown in the table. There were no significant hits reported for some SNPs which are hence not included in this table.

EBV, Epstein-Barr virus; SNP, single nucleotide polymorphism.

the previous strong indications, and biological basis that could explain endometrial carcinogenesis, we decided to include an AKT1 variant (rs2498794) in our results.

PTEN is a multi-functional tumour suppressor gene that regulates the AKT/PKB signalling pathway and is commonly mutated in many cancers including endometrial cancer. Loss-of-function germline mutations in PTEN are responsible for Cowden syndrome, which exerts a lifetime risk of endometrial cancer of up to 28%. Lacey and colleagues studied SNPs in the PTEN gene region; however, none showed significant differences in frequency between 447 endometrial cancer cases and 439 controls of European ancestry.
Cancer genetics

(transcriptional activator). Three SNPs have been identified in or around KLF5 that are associated with endometrial cancer. The G allele of rs1841589 (OR 1.15, 95% CI 1.11 to 1.21; p=4.83×10⁻¹¹), the A allele of rs9600103 (OR 1.23, 95% CI 1.16 to 1.30; p=3.76×10⁻¹⁰) and C allele of rs7981863 (OR 1.16, 95% CI 1.12 to 1.20; p=2.70×10⁻¹⁵) have all been found to be associated with an increased likelihood of endometrial cancer in large European cohorts.²¹ ²³ ²⁴ ²⁵ It is worth noting that these SNPs are not independent, and hence they quite possibly tag the same causal variant.

The MYC family of proto-oncogenes encode transcription factors that regulate cell proliferation, which can contribute to cancer development if dysregulated. The recent GWAS by O’Mara et al reported three SNPs within the MYC region that reached genome-wide significance with conditional p values reaching at least 5×10⁻⁷.²⁶

To test the utility of these SNPs as predictive markers, we devised a theoretical PRS calculation using the log ORs and EAFs per SNP from the published data. The results were very encouraging with an RR of 3.16 for the top 1% versus the mean, using all the top SNPs presented in table 1 and 2.09 when using only the SNPs that reached genome-wide significance (including AKT1).

Controversy surrounding MDM2 variant SNP309

MDM2 negatively regulates tumour suppressor gene TP53, and as such, has been extensively studied in relation to its potential role in predisposition to endometrial cancer. Our search identified six original studies of the association between MDM2 SNP rs2279744 (also referred to as SNP309) and endometrial cancer, all of which found a statistically significant increased risk per copy of the G allele. Two more original studies were included through our full-text evaluation; however, these were not included here as they did not meet our inclusion criteria—one due to small sample size, the other due to studying rs2279744 status dependent on another SNP.³⁶ ³⁷ Even so, the two studies were described in multiple meta-analyses that are listed in table 3. Different permutations of these eight original studies appear in at least eight published meta-analyses. However, even the largest meta-analysis contained <2000 cases (table 3)³⁸

In comparison, a GWAS including nearly 13 000 cases found no evidence of an association with OR and corresponding 95% CI of 1.00 (0.97 to 1.03) and a p value of 0.93 (personal communication).³¹ ³² Nevertheless, we cannot completely rule out a role for MDM2 variants in endometrial cancer predisposition as the candidate-gene studies reported larger effects in Asians, whereas the GWAS primarily contained participants of European ancestry. There is also some suggestion that the SNP309 variant is in linkage disequilibrium with another variant, SNP285, which confers an opposite effect.

It is worth noting that the SNP285C/SNP309G haplotype frequency was observed in up to 8% of Europeans, thus requiring correction for the confounding effect of SNP285C in European studies.³⁹ However, aside from one study conducted by Knappskog et al, no other study including the meta-analyses corrected for the confounding effect of SNP285.⁴₀ Among the studies presented in table 3, Knappskog et al (2012) reported that after correcting for SNP285, the OR for association of this haplotype with endometrial cancer was much lower, though still significant. Unfortunately, the meta-analyses which synthesised Knappskog et al (2012) as part of their analysis, did not correct for SNP285C in the European-based studies they included.³⁸ ⁴¹ ⁴² It is also concerning that two meta-analyses using the same primary articles failed to report the same result, in two instances.³⁸ ⁴²–⁴⁴

Table 3 Characteristics of studies that examined MDM2 SNP rs2279744

| Reference       | OR (95% CI) | P values | EAF | Ancestry | Cases (n) | Controls (n) | EC type | Dataset(s) |
|-----------------|------------|----------|-----|----------|-----------|--------------|---------|------------|
| Terry 2008      | 1.32 (1.11 to 1.56) | 0.002 | N/A | European | 591       | 1543         | N/A     | NHS, WHS   |
| Ashton 2009     | 1.37 (1.06 to 1.79) | 0.067 | Caucasian | 191     | 291       | All          | Hospital based |
| Nunobiki 2009   | 2.28 (2.02 to 2.54) | 0.030 | 0.49 | Japanese | 102       | 95           | All      | Hospital based |
| Ueda 2009       | 1.91 (1.5 to 3.47)  | 0.035 | 0.51 | Japanese | 119       | 108          | All      | Hospital based |
| Wan 2011        | 1.54 (1.21 to 1.94) | 0.000 | N/A | N/A      | N/A       | N/A          | N/A     | Walsh 2007, Terry 2008, Ashton 2009, Nunobiki 2009, Ueda 2009 |
| Li 2011         | 1.75 (1.16 to 2.63) | 0.007 | N/A | European, Asian | 1001 | 1889        | N/A     | Walsh 2007, Terry 2008, Ashton 2009, Nunobiki 2009, Ueda 2009 |
| Knappskog 2012 | 1.22 (1.03 to 1.44) | 0.038 | 0.36 | European | 392       | 956          | N/A     | Hospital based |
| Zajac 2012      | 1.33 (1.12 to 1.58) | 0.001 | N/A | European | 152       | 100          | N/A     | Hospital based |
| Yoneda 2013     | 1.64 (0.81 to 3.28) | 0.450 | 0.45 | Asian     | 125       | 200          | All     | Population based |
| Peng 2013       | 1.6 (1.21 to 2.13)  | 0.001 | N/A | European, Asian | 2069 | 4546        | N/A     | Walsh 2007, Terry 2008, Ashton 2009, Nunobiki 2009, Knappskog 2012, Yoneda 2013 |
| Zhao 2014       | 1.41 (1.04 to 1.92) | 0.030 | N/A | European, Asian | 1278 | 2189        | N/A     | Walsh 2007, Terry 2008, Ashton 2009, Ueda 2009, Zajac 2012, Yoneda 2013 |
| Wang 2014       | 1.32 (1.06 to 1.64) | 0.010 | N/A | European, Asian | 1967 | 4460        | N/A     | Walsh 2007, Terry 2008, Ashton 2009, Nunobiki 2009, Ueda 2009, Zajac 2012, Knappskog 2012, Yoneda 2013 |
| Xue 2014        | 1.14 (0.79 to 1.65) | 0.490 | N/A | European | 1769      | 4172        | N/A     | Walsh 2007, Terry 2008, Ashton 2009, Nunobiki 2009, Ueda 2009, Zajac 2012, Knappskog 2012, Yoneda 2013 |
| Zhang 2018      | 1.46 (1.25 to 1.72) | 0.035 | N/A | European, Asian | 1690 | 4151        | N/A     | Walsh 2007, Terry 2008, Ashton 2009, Nunobiki 2009, Ueda 2009, Zajac 2012, Knappskog 2012, Yoneda 2013 |
| Zou 2018        | 1.23 (1.06 to 1.41) | 0.005 | N/A | European, Asian | 3535 | 6476        | All     | Walsh 2007, Terry 2008, Ashton 2009, Nunobiki 2009, Zajac 2012, Zajac 2012, Okamoto 2015, Gansmo 2017 |

*Walsh et al 2007 and Gansmo et al 2017 did not meet eligibility criteria for us to include in our evaluation. EAF, effect allele frequency; EC, endometrial cancer; SNP, single nucleotide polymorphism.*

Bafligil C, et al. J Med Genet 2020;0:1–10. doi:10.1136/jmedgenet-2019-106529
DISCUSSION

This article represents the most comprehensive systematic review to date, regarding critical appraisal of the available evidence of common low-penetrance variants implicated in predisposition to endometrial cancer. We have identified the most robust SNPs in the context of endometrial cancer risk. Of those, only 19 were significant at genome-wide level and a further five were considered marginally significant. The largest GWAS conducted in this field was the discovery- and meta-GWAS by O’Mara et al., which utilised 120,968 cases and 108,979 controls. Despite the inclusion of all published GWAS and around 5000 newly genotyped cases, the total number did not reach anywhere near what is currently available for other common cancers such as breast cancer. For instance, BCAC (Breast Cancer Association Consortium) stands at well over 200,000 individuals with more than half being cases, and resulted in identification of ~170 SNPs in relation to breast cancer. A total of 313 SNPs including imputations were then used to derive a PRS for breast cancer. Therefore, further efforts should be directed to recruit more patients, with deep phenotypic clinical data to allow for relevant adjustments and subgroup analyses to be conducted for better precision.

A recent pre-print study by Zhang and colleagues examined the polygenicity and potential for SNP-based risk prediction for 14 common cancers, including endometrial cancer, using available summary-level data from European-ancestry datasets. They estimated that there are just over 1000 independent endometrial cancer susceptibility SNPs, and that a PRS comprising all such SNPs would have an area under the receiver-operator curve of 0.64, similar to that predicted for ovarian cancer, but lower than that for the other cancers in the study. The modelling in the paper suggests that an endometrial cancer GWAS double the size of the current largest study would be able to identify susceptibility SNPs together explaining 40% of the genetic variance, but that in order to explain 75% of the genetic variance it would be necessary to have a GWAS comprising close to 150,000 cases and controls, far in excess of what is currently feasible.

We found that the literature consists mainly of candidate-gene studies with small sample sizes, meta-analyses reporting conflicting results despite using the same set of primary articles, and multiple reports of significant SNPs that have not been validated by any larger GWAS. The candidate-gene studies were indeed the most useful and cheaper technique available until the mid to late 2000s. However, a lack of reproducibility (particularly due to population stratification and reporting bias), uncertainty of reported associations, and considerably high false discovery rates make these studies much less appropriate in the post-GWAS era. Unlike the candidate-gene approach, GWAS do not require prior knowledge, selection of genes or SNPs, and provide vast amounts of data. Furthermore, both the genotyping process and data analysis phases have become cheaper, the latter particularly due to faster and open-access pre-phasing and imputation tools being made available.

It is clear from table 1 that some SNPs were reported with wide 95% CI, which can be directly attributed to small sample sizes particularly when restricting the cases to non-endometrioid histology only, low EAF or poor imputation quality. Thus, these should be interpreted with caution. Additionally, most of the SNPs reported by candidate-gene studies were not detected by the largest GWAS to date conducted by O’Mara et al. However, this does not necessarily mean that the possibility of those SNPs being relevant should be completely dismissed. Moreover, meta-analyses were attempted for other variants; however, these showed no statistically significant association and many presented with high heterogeneity between the respective studies (data not shown). Furthermore, as many studies utilised the same set of cases and/or controls, conducting a meta-analysis was not possible for a good number of SNPs. It is therefore unequivocal that the literature is crowded with numerous small candidate-gene studies and conflicting data. This makes it particularly hard to detect novel SNPs and conduct meaningful meta-analyses.

We found convincing evidence for 19 variants that indicated the strongest association with endometrial cancer, as shown in table 1. The associations between endometrial cancer and variants in or around HNF1B, CYP19A1, SOX4, MYC, KLF and EIF2AK found in earlier GWAS were then replicated in the latest and largest GWAS. These SNPs showed promising potential in a theoretical PRS we devised based on published data. Using all 24 or genome-wide significant SNPs only, women with a PRS in the top 1% of the distribution would be predicted to have a risk of endometrial cancer 3.16 and 2.09 times higher than the mean risk, respectively.

However, the importance of these variants and relevance of the proximate genes in a functional or biological context is challenging to evaluate. Long distance promoter regulation by enhancers may disguise the genuine target gene. In addition, enhancers often do not loop to the nearest gene, further complicating the relevance of nearby gene(s) to a GWAS hit. In order to elucidate biologically relevant candidate target genes in endometrial cancer, O’Mara et al looked into promoter-associated chromatin looping using a modern HiChIP approach. The authors utilised normal and tumoural endometrial cell lines for this analysis which showed significant enrichment for endometrial cancer heritability, with 103 candidate target genes identified across the 13 risk loci identified by the largest ECAC GWAS. Notable genes identified here were CDKN2A and WTI, and their antisense counterparts. The former was reported to be nearby of rs1679014 and the latter of rs10835920, as shown in table 1. Moreover, of the 36 candidate target genes, 17 were found to be downregulated while 19 were upregulated in endometrial tumours.

The authors also investigated overlap between the 13 endometrial cancer risk loci and top eQTL variants for each target gene. In whole blood, of the two particular lead SNPs, rs8822380 at 17q21.32 was a top eQTL for SNX11 and HOXB2, whereas rs937213 at 15q15.1 was a top eQTL for SRP14. In endometrial tumour, rs7579014 at 2p16.1 was found to be a top eQTL for BCL11A. This is particularly interesting because BCL11A was the only nearby/candidate gene that had a GWAS association reported in both endometrioid and non-endometrioid subtypes. The study looked at protein–protein interactions between endometrial cancer drivers and candidate target gene products. Significant interactions were observed with TP53 (most significant), AKT, PTEN, ESR1 and KRAS, among others. Finally, when 103 target candidate genes and 387 proteins were combined together, 462 pathways were found to be significantly enriched. Many of these are related to gene regulation, cancer, obesity, insulinemia and oestrogen exposure. This study clearly showed a potential biological relevance for some of the SNPs reported by ECAC GWAS in 2018.

Most of the larger included studies used cohorts primarily composed of women of broad European descent. Hence, there are negligible data available for other ethnicities, particularly African women. This is compounded by the lack of reference genotype data available for comparative analysis, making it harder for research to be conducted in ethnicities other than Europeans. This poses a problem for developing risk prediction
models that are equally valuable and predictive across populations. Thus, our results also are of limited applicability to non-European populations.

Furthermore, considering that non-endometrioid cases comprise a small proportion (~20%) of all endometrial cancer cases, much larger cohort sizes are needed to detect any genuine signals for non-endometrioid tumours. Most of the evaluated studies looked at either overall/mixed endometrial cancer subtypes or endometrioid histology, and those that looked at variant associations with non-endometrioid histology were unlikely to have enough power to detect any signal with statistical significance. This is particularly concerning because non-endometrioid subtypes are biologically aggressive tumours with a much poorer prognosis that contribute disproportionately to mortality from endometrial cancer. It is particularly important that attempts to improve early detection and prevention of endometrial cancer focus primarily on improving outcomes from these subtypes. It is also worth noting that, despite the current shift towards a molecular classification of endometrial cancer, most studies used the overarching classical Bokhman’s classification system, type I versus type II, or no histological classification system at all. Therefore, it is important to create and follow a standardised and comprehensive classification system for reporting tumour subtypes for future studies.

This study compiled and presented available information for an extensively studied, yet unproven in large datasets, SNP309 variant in MDM2. Currently, there is no convincing evidence for an association between this variant and endometrial cancer risk. Additionally, of all the studies, only one accounted for the opposing effect of a nearby variant SNP285 in their analyses. Thus, we conclude that until confirmed by a sufficiently large dataset, the top 24 SNPs for endometrial cancer susceptibility variants, emphasising the ability to detect genuine signals.

Thus, we conclude that until confirmed by a sufficiently large dataset, there is no convincing evidence that any of the SNPs evaluated in this study for non-endometrioid endometrial cancer are genuine signals for non-endometrioid cases, much larger cohort sizes are needed to detect any genuine signals for non-endometrioid cases. Therefore, it is important to create and follow a standardised and comprehensive classification system for reporting tumour subtypes for future studies.

This systematic review presents the most up-to-date evidence for endometrial cancer susceptibility variants, emphasising the need for further large-scale studies to identify these variants and validation of these associations. Until data from larger and more diverse cohorts are available, the top 24 SNPs presented here are the most robust common genetic variants that affect endometrial cancer risk. The multiplicative effects of these SNPs could be used in a PRS to personalise risk prediction models to be developed for targeted screening and prevention interventions for women at greatest risk of endometrial cancer.

Twitter Emma J Crosbie @DeEmmaCrosbie

Contributors CB planned the study, did the systematic review, analysed the data and wrote the manuscript. DIT and AL supervised the study and provided statistical support for the analysis. MS supervised the study. NAIR and AN supported data acquisition. DGE and EJC designed and planned the study, provided supervision and wrote the manuscript. EJC provided funding for the study. All authors reviewed and approved the final manuscript.

Funding CB, DGE, AL, MS and EJC are all supported by the National Institute for Health Research (NIHR) Manchester Biomedical Research Centre (IS-BRC-1215-20007). NAIR was supported through a Medical Research Council (MRC) Doctoral Research Fellowship (MR/M018433/1). EJC through an NIHR Clinician Scientist Fellowship (NIHR-CS-012-009) and DGE is an NIHR Senior Investigator (NF-SI-0513-10076). This research was funded by the NIHR Manchester BRC.

Disclaimer The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

The protocol for this systematic review was published at PROSPERO and the data that inform this manuscript are available upon reasonable request from the corresponding author.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: https://creativecommons.org/licenses/by/4.0/.

ORCID iDs
Cemsel Balligır https://orcid.org/0000-0002-2365-1194
Miriam J Smith http://orcid.org/0000-0002-3184-0817
D Gareth Evans http://orcid.org/0000-0002-8482-5784
Emma J Crosbie http://orcid.org/0000-0003-0284-8630

REFERENCES

1. Sundar S, Balega J, Crosbie E, Drake A, Edmondson R, Fotopoulou C, Gallos I, Ganesan R, Gupta J, Johnson N, Kitson S, Mackintosh M, Martin-Hirsch P, Miles T, Raji S, Reed N, Rolland P, Singh K, Sivalingam V, Walther A. BCOS uterine cancer guidelines: recommendations for practice. Eur J Obstet Gynecol Reprod Biol 2017;213:71-97.

2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin 2019;69:7-34.

3. Monte R, Leary A, Creutzberg C, Abu-Rustum N, Daraei E. Endometrial cancer. Lancet 2016;387:1094–108.

4. Tzur T, Kessous R, Weintalay AV. Current strategies in the diagnosis of endometrial cancer. Arch Gynecol Obstet 2017;296:5–14.

5. Clarke MA, Devesa SS, Harvey SV, Wentzensen N. Hysterec-tomy-corrected uterine corpus cancer incidence trends and differences in relative survival reveal racial disparities and rising rates of nondenominant cancers. J Clin Oncol 2019;37:1895–908.

6. Kandoth C, Schultz N, Chenniah AK, Akbari R, Liu Y, Shen H, Robertson AG, Pashanl S, Shen R, Benz CC, Yau C, Laird PW, Ding L, Zhang W, Mills GB, Kucherlapati R, Mardis ER, Levine DA, Getz G, Gabriel SB, Cibulskis K, Lander E, Savchenko A, Sougnez C, Lawrence M, Dooling D, Fulton R,ulton L, Kalicki-Verheijer J, McElhaney M, O’Laughlin M, Schmidt H, Wilson RK, Ye K, Alya A, Balaasendramadu M, Birol I, Butterfield YSN, Carlsen R, Carter C, Chu A, Chuah S, Chu HN, Dhalla N, Guin R, Hirst C, Holt RA, Jones SM, Lee D, Li H, Marra MA, Mayo M, Moore RA, Mungall AJ, Ptietrovich P, Schein JE, Spaihhimalani P, Tam A, Vorhal RJ, Saksena G, Orofino RC, Schumacher SE, Tabak B, Carter SL, Hernandez B, Gentry J, Salvesen HB, Ardile K, Winckler W, Beroukhim R, Meyerson M, Hadijanayis A, Lee S, Mahadawshy HS, Park P, Protopopov A, Ren XJ, Seth S, Song XZ, Tang JB, RB Y, Yang LX, Zeng D, Chin L, Zhang JH, Auman JT, Balu S, Bodenheimer T, Buda E, Hayes DN, Hoyle AP, Jeffreys SR, Jones CD, Meng SW, Mieczkowski PA, Mose LE, Parker JS, Perou CM, Roach J, Shi Y, Simons JV, Soloway MG, Tan DH, Topal MD, Waring S, WJ Y, Woadley K, Baylin SB, Botvinnik MAI, Lai PH, Camilla TI, Van Der Berg D, Weisheiler DJ, Cho J, DiCara D, Frazer S, Heiman D, Jing R, Lin P, Mallard W, Stojanov P, Voet D, Zhang HL, Zou LH, Nobile M, Reynolds SM, Shumilevich L, Aksay BA, Antipin Y, Ciriello G, Dresnog J, Gao JJ, Gross B, Jacobsen A, Ladanyi M, Reva B, Sander C, Sinha R, Sumer TO, Taylor BS, Cerami E, Weinhold N, Benz S, Goldstein T, Haussler D, Ng S, Saro C, Stuart J, Annala M, Broom BM, Cassavens TD, ZL, Li H, Li Y, Lurah AK, Wakefield C, Weinstein JD, Ziang NX, Liu XG, Broadhurst R, Adams C, Barr T, Black AD, Bowen J, Deardourff J, Frick J, Gaster-Foster JM, Grossman T, Harper HA, Hart-Kothari M, Helvel C, Hobensack A, Kuck H, Kneile L, Keraas L, Kichtenberg TM, Almclister A, Pyatt R, Ramiz NC, Tabler TR, Vanhoucke N, White P, Wise L, Zmeda E, Barnabas N, Berry-Green C, Blanc V, Boice L, Button M, Farkas A, Green A, Mackenzie J, Nicholson D, Kallergos SE, Gibbs CK, Karlan BY, Lester J, Onsur S, Borland PW, Cadungog M, Czerwinski C, Huelas-Egork-Orli L, Jacoco M, Petrelli N, Winkin G, Nemicovich-Danchenko E, Potopova Q, Rotin D, Berchuck C, Birrer M, DiSala P, Monovich L, Curley E, Gardner J, Mallory D, PENNY R, Dowdy SC, Winterhalter B, Dalso L, Gostout B, Meurer A, Teoman A, Dao F, Olvera N, Bogomolny F, Gark K, Gosalis ARA, Abramov M, Bartlett JM, Kodeeswaran S, Partan M, Joosekken C, Clarke BA, Goodman MT, Carney ME, Matsuno RK, Fisher J, Huang M, Ratamakwill W, Thorne L, Van DR, Edwards R, Elisafoua E, Zorn K, Goodfellow PJ, Muthz D, Kahn AB, Bell DW, Pollock PM, Wang C, Wheeler DA, Shiront E, Ayala B, CHU AL, Jensen MA, Kothiyil P, Pihl TD, Pontius J, Pot DA, Sneeder EE, Srinivasa D, Shaw KRT, Sheth M, Davidens T, Ellert E, Ferguson ML, Demchok JA, Yang LM, Geyer MS, Ozenberger BA, Sofia HJ, Shen R, Cancer Genome Atlas Research Network. Integrated genomic characterization of endometrial carcinoma. Nature 2013;497:67–73.

7. Stello E, Bosse T, Nout RA, Mackay H, Church DN, NJMwan HW, Leary A, Edmondson RJ, Powell ME, Crosbie EJ, Kitchener HC, Milleskhin L, Pollock PM, Smelt VT, Creutzberg CL. Refining prognostic and identifying targetable pathways for high-risk endometrial cancer: a TRANSPORTE initiative. Mod Pathol 2015;28:836–44.

8. Win AK, Reece JC, Ryan S. Family history and risk of endometrial cancer: a systematic review and meta-analysis. Obstet Gynecol 2015;125:89–98.
Cancer genetics

Michalidou K, Tyrer JP, Wang Q, Ahmed S, Healey CS, Shah M, Annibali D, Depreeuw J, Al-Tassan NA, Harris R, Meyer BE, Whiffin N, Hosking FJ, Kinnersley B, Farrington SM, Timofeeva M, Tenesa A, Campbell H, Haile RW, Hodgson S, Carvalio-Carmona L, Cheadle JP, Easton DF, Dunlop M, Houlston R, Spurdle A, Tomlinson I. Meta-analysis of genome-wide association studies identifies common susceptibility polymorphisms for colorectal and endometrial cancer. Nat Genet. 2012;44:455–52.

Zajac A, Stachowiak G, Petyriski T, Romanowicz H, Wilczyński J, Smolarek B. Association between MDM2 SNP309 polymorphism and endometrial cancer risk in Polish women. PloS 2012;7:278–83.

Dunning Wendy 2014;40:1957–67.

Michailidou K, Tyrer JP, Easton DF, Houlston R, Spurdle A, Tomlinson I. Meta-analysis of genome-wide association studies identifies common susceptibility polymorphisms for colorectal and endometrial cancer. Nat Genet. 2012;44:455–52.

Zajac A, Stachowiak G, Petyriski T, Romanowicz H, Wilczyński J, Smolarek B. Association between MDM2 SNP309 polymorphism and endometrial cancer risk in Polish women. PloS 2012;7:278–83.

Dunning Wendy 2014;40:1957–67.