

Review

Not peer-reviewed version

Using Genetics to Investigate Relationships Between Phenotypes: Application to Endometrial Cancer

Kelsie Bouttle , Nathan Ingold , [Tracy A. O'Mara](#) *

Posted Date: 1 July 2024

doi: 10.20944/preprints202406.1957.v1

Keywords: genome-wide association study; endometrial cancer; cross-phenotype; statistical genetics



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Review

Using Genetics to Investigate Relationships between Phenotypes: Application to Endometrial Cancer

Kelsie Bouttle, Nathan Ingold and Tracy A. O'Mara *

Cancer Research Program, QIMR Berghofer Medical Research Institute, Brisbane QLD 4006, Australia

* Correspondence: Tracy.OMara@qimrberghofer.edu.au; Tel.: +61-7-3362-0389

Abstract: Genome-wide association studies (GWAS) have accelerated the exploration of genotype-phenotype associations, facilitating the discovery of replicable genetic markers associated with specific traits or complex diseases. This narrative review explores the statistical methodologies developed using GWAS data to investigate relationships between various phenotypes, focusing on endometrial cancer, the most prevalent gynecological malignancy in developed nations. Advancements in analytical techniques such as genetic correlation, colocalization, cross-trait locus identification, and causal inference analyses have enabled deeper exploration of associations between different phenotypes, enhancing statistical power to uncover novel genetic risk regions. These analyses have unveiled shared genetic associations between endometrial cancer and many phenotypes, enabling identification of novel endometrial cancer risk loci and furthering our understanding of risk factors and biological processes underlying this disease. The current status of research in endometrial cancer is robust; however, this review demonstrates that further opportunities exist in statistical genetic that hold promise for advancing the understanding of endometrial cancer and other complex diseases.

Keywords: genome-wide association study; endometrial cancer; cross-phenotype; statistical genetics

1. Introduction

Genome-wide association studies (GWAS) play a pivotal role in advancing the identification of genotype-phenotype associations. GWAS entails a comprehensive examination of the entire genome for genetic markers, predominantly single nucleotide polymorphisms (SNPs), demonstrating statistically significant (P value $< 5 \times 10^{-8}$) associations with specific traits or complex diseases. Identification of genetic risk loci through GWAS allows for the discovery of interventions, identification of high genetic risk groups and guidance of treatment protocol [1].

Traits or diseases often present with shared clinical and epidemiological risk factors and can be linked through molecular, biological and population-based data. GWAS has facilitated the exploration of relationships between different diseases and phenotypes through analytical advancements in genetic association, correlation, and causal inference testing. These approaches can provide information on comorbidities for particular diseases, modifiable risk factors and enable cross-phenotype GWAS meta-analyses to identify new genetic loci.

This narrative review will present methods developed using GWAS data to explore relationships between different phenotypes (Table 1) and discuss their application to endometrial cancer, the most commonly diagnosed gynecological cancer in developed countries.

2. Endometrial Cancer

Endometrial cancer has witnessed noteworthy trends in incidence and mortality in recent years. The incidence has surged by over 132% in the past 30 years, with 417,000 new diagnoses worldwide in 2020 [2,3]. While the incidence rates vary globally, there has been a discernible increase in developed nations, likely attributed to factors such as ageing populations and the rising prevalence of obesity [2,4]. There is a well-established link between obesity and endometrial cancer, with excess

adiposity leading to increased levels of estrogen, thereby augmenting the risk of developing this malignancy [5]. Although endometrial cancer is typically diagnosed in post-menopausal women at around 60 years of age, up to 5 percent of cases occur in women under the age of 40 [4].

The leading theoretical pathogenetic pathway for endometrial carcinoma involves prolonged exposure to elevated estrogen levels, whether from exogenous or endogenous sources, which stimulates unopposed endometrial proliferation without adequate opposition by progestin [4,5]. In addition to obesity, other established risk factors include diabetes, polycystic ovary syndrome (PCOS), hypertension, and lifestyle factors like smoking and diet, which all may indirectly influence estrogen levels [6,7]. Direct estrogen-related factors, such as the use of combined oral contraceptive pills for risk reduction or unopposed estrogen replacement therapy for increased risk, further contribute to the disease's dynamics [8]. Reproductive factors like early menarche, late menopause, and nulliparity also play significant roles [8].

Epidemiological observations and family-based studies have demonstrated that the genetic effect of endometrial cancer is significant, with heritability estimates ranging from 27% to 52% and a two-to-threefold increased risk associated with a family history of endometrial cancer [9–12]. Rare pathogenic germline variants within mismatch-repair genes (i.e., *MLH1*, *MSH2*, *MSH6*, *PMS2*, and *EPCAM*) observed an initial genetic predisposition for endometrial cancer in women with Lynch syndrome [13,14]. Despite their rarity in the general population, estimates suggest that these high-risk germline variants contribute to 3% of endometrial cancer cases [15,16].

The genetics of endometrial cancer, particularly elucidated through GWAS, have significantly advanced our understanding of the disease's etiology, functional mechanisms, and translational implications, demonstrating the effects of common genetic variation (minor allele frequency > 1%) on endometrial cancer risk [17,18]. The largest endometrial cancer GWAS to date, conducted by the Endometrial Cancer Association Consortium (ECAC), used data from nearly 13,000 cases and identified 16 genetic loci associated with endometrial cancer risk [19]. Risk estimates for these common variants individually are expected to only slightly increase the risk associated with endometrial cancer; however, cumulatively they explain about 28% of the familial relative risk [17]. The identified 16 variants are estimated to account for only a quarter of the variance that genetics can explain, implicating a further potential for discovery in unidentified genetic loci.

Increases in sample size and availability of GWAS summary statistics, in addition to larger genetic cohorts, will improve power and facilitate the discovery of genetic risk regions. Patterns of shared genetic influence combined with existing epidemiological observations promise to elucidate functional pathways and contribute to our understanding of the biological underpinning and comorbidity of diseases.

3. Genetic Correlation

Genetic correlation is a critical metric to quantify the overall genetic similarity between complex traits irrespective of environmental confounders, which are prevalent and mostly unavoidable in conventional epidemiologic studies. Genetic correlation (denoted as r_g) ranges from 0 to 1 and describes the average effect of pleiotropy across all causal loci. It is frequently used in the initial identification of associations between traits of interest.

Table 1. Programs available for exploring relationships between traits using GWAS data.

Program	Ref	Description	Data Input ¹	Examples ²
Genetic correlation				
BOLT-REML	[20]	A Monte Carlo algorithm for variance component analysis to estimate genetic correlations and partition SNP heritability among multiple phenotypes. Computationally fast.	Individual	37340002 30929738
GenomicSEM	[21]	Synthesizes genetic correlations of multiple traits with unknown amounts of sample overlap	Summary	32606422 35513722
GCTA	[22]	Provides highly accurate estimates of genetic correlations between phenotypes while accounting for different genetic architectures.	Individual	24944428
PCGC-s	[23]	Estimates genetic correlation and partitioned heritability large datasets while accounting for case-control sampling and covariates.	Summary	31488892
LD Score Regression	[24]	Estimates genetic correlations across multiple phenotypes while accounting for cryptic relatedness and population stratification.	Summary	30093612 29608257 26414676
HDL	[25]	Highly powered and accurate estimates of genetic correlations fully account for whole genome LD and reduce the variance of genetic correlation estimates.	Summary	35492879 34997191
MTG2	[26]	Combines the average information algorithm used by REML with an eigen-decomposition of the genomic relationship matrix to estimate genetic variance.	Individual	29977057 35729236
GNOVA	[27]	Provides powerful statistical inference through annotation-stratified genetic covariance analysis that is robust to LD and sample overlap.	Summary	37034223 34634379
q-HESS	[28]	Localized and precise quantification of genetic correlation between pairs of traits due to small-region genetic variation. Accounts for LD and sample overlap while making no distributional assumptions on the causal effect size under a fixed effects model.	Summary	34355204 34561436
LAVA	[29]	Tests local genetic correlation between two phenotypes. Can also analyze local heritability and conditional genetic relationships between several phenotypes.	Summary	36471075 38637617
Colocalization				
COLOC	[30]	Bayesian statistical test to enable the computation of probabilities that two traits share a common genetic causal variant from single variant association p-values and MAFs. Locus-specific analysis.	Summary	34268601 34355204
GWAS-PW	[31]	Bayesian statistical test to enable the computation of probabilities that two traits share a common genetic causal variant from single variant association p-values and MAFs. Locus-specific analysis.	Summary	33144283 35178771 35851147
Cross-Trait Locus Identification				

Program	Ref	Description	Data Input ¹	Examples ²
MTAG	[32]	Joint analysis of multiple traits to increase statistical power and account for sample overlap.	Summary	34268601 31488892 29292387
CPASSOC	[33]	Assess cross-phenotype associations for both continuous or binary traits.	Summary	37636041 31669095
MV-PLINK	[34]	Computationally fast implementation of canonical correlation analysis, including multiple phenotypes and uses	Individual	35278618
MultiPhen	[35]	Employs ordinal regression for joint multivariate modelling of multiple phenotypes, with increased statistical power and an appropriate type 1 error rate.	Individual	35701404 35680855
conjFDR	[36]	A model-free strategy for analysis that leverages genetic overlap between two phenotypes which boosts statistical power and identifies shared genomic association regardless of the cross-trait correlations.	Summary	37752828 31792363
bGWAS	[37]	A Bayesian method that leverages published studies for related risk factors to construct priors. Increase power to identify susceptibility variants and allows for assessment of posterior and direct effects.	Summary	37168552 35653391
RE2C	[38]	A generalized likelihood model that accounts for correlations of statistics and achieves optimal power under the condition of heterogeneity	Summary	35492870 3734002 35753705
MetABF	[39]	Employs a Bayesian framework using both an independent and fixed effect model to meta-analysis GWAS statistics. An efficient tool that allows the expected relationships between studies or traits to be encoded in the analysis.	Summary	36653479 35492870
Causal Inference				
LCVA	[40]	Distinguishes causal relationships among genetically correlated phenotypes such that a positive result is more likely to be the true causal effect.	Summary	36653534 36151087 31669095
MiXer	[41]	Applies a bivariate causal model to quantify and visualize polygenic overlap by estimating the total number of shared and trait-specific causal variants.	Summary	37752828 34761251
Mendelian randomization	[42]	Uses instrument variables in statistical models to identify causal relationships between an exposure and outcome. Various programs and techniques have been developed (see Table 2)	Both	34268601

¹GWAS data required for analysis: Individual-level genotypes, or summary statistics. ²Example publications (PubMed IDs) that have used these approaches, note this is not exhaustive.

Several methods have been developed (Table 1), the most commonly used being linkage disequilibrium (LD) score regression, mainly owing to its computational efficiency and use of GWAS summary statistics [24]. LD score regression has been used for an array of diseases and phenotypes, uncovering potentially shared genetic architecture between schizophrenia and a range of psychiatric, metabolite, personality, immune, cardiovascular, substance-related and anthropometric traits. [24,43–46] Interestingly, epidemiological studies have previously reported both opposing and direct comorbidity between schizophrenia and several cancer types.[47,48] LD score regression has also estimated a significant positive genetic correlation between schizophrenia and breast cancer, which may partly explain the epidemiological bidirectional association between the two traits and suggests shared biological mechanisms. [49]

The largest endometrial GWAS published to date determined genetic correlations between endometrial cancer and 224 non-cancer traits [19]. A significant positive genetic correlation, consistent with existing epidemiological observations, was found between endometrial cancer risk, type 2 diabetes, and anthropometric traits related to obesity (e.g., body mass index (BMI) and waist circumference) [19]. A significant negative correlation was found between years of schooling and the age of menarche, both of which negatively correlate with obesity-related traits [19].

LD score regression has explored the relationship between endometrial cancer and various cancers, finding a strong correlation with ovarian and ER-positive breast cancer [50,51]. Additionally, LD score regression has unveiled a potential shared genetic architecture with non-cancerous gynecological diseases, including polycystic ovarian syndrome (PCOS), uterine fibroids and endometriosis [52–54]. Unlike the genetic correlation between uterine fibroid and endometrial cancer, adjustment for genetically predicted BMI at least partly mediated the genetic correlation between PCOS and endometrial cancer [54]. Further research into cross-trait genetic correlation will enable a better understanding of endometrial cancer genetic predisposition.

While a powerful tool for understanding the overall genetic similarity of complex traits, genetic correlation analysis has several limitations to consider. LD score regression employs a polygenic model and is most effective when analyzing traits with a polygenic genetic architecture. However, when significant SNPs account for a sizable proportion of heritability, analyzing only those SNPs can prove more efficient [24]. It is crucial to recognize that genetic correlation analysis cannot establish causal relationships or determine the directionality of effects [24]. Any observed genetic correlation could result from a true direct relationship between two traits, or the genetic variant could be associated with an unknown risk factor, which also affects both traits. Therefore, while significant results may suggest shared genetic architecture, caution is needed in interpreting these findings without additional evidence from functional studies or experimental designs capable of elucidating causal relationships.

The absence of genome-wide genetic correlation does not overrule the occurrence of locus-specific genetic correlation. A correlation between two traits could result from multiple genetic variants, some of which may have opposing effects on the two traits; in these circumstances, local genetic correlations can be more insightful in identifying shared common causes [55]. Regional genetic correlations can quantify which genomic regions disproportionately contribute to the genome-wide correlation [28]. A large-scale cross-cancer study used ρ -HESS (Heritability Estimation using Summary Statistics) to successfully identify thirteen pairs of cancers with statistically significant local genetic correlations across eight distinct genomic regions [56]. This study found a positive genetic correlation between endometrial and prostate cancer at region 17q12 and a statistically significant local genetic correlation at 5p15.33 across six pairs of cancers, including ER-negative breast, pancreatic, glioma, melanoma, lung, pancreatic, prostate and colorectal cancer. Despite the varying direction of these correlations, the result postulates that the 5p15.33 region may harbor key genetic variants related to multiple cancer types supported by the number of susceptibility variants already reported in this region [56].

Pinpointing specific regions that may drive the global genetic correlation, as well as regions that might be neutral or antagonistic, can complement genome-wide analysis and deepen understanding of the genetic architecture of the traits. Such analysis has not been widely pursued in endometrial

cancer, thus presenting an area ripe for exploring locally shared genetic pathways that may go unnoticed in genome-wide genetic correlation analysis.

4. Colocalization

Colocalization refers to the identification of causal variants shared between different traits after controlling for the independent signals identified in individual analysis [57,58]. Multi-trait colocalization enhances the statistical power to identify shared variants across multiple traits and provides a more robust indication of the variant's potential causality [30]. Programs developed for colocalization analysis use Bayesian statistical tests to enable the computation of posterior probabilities that can disentangle whether the association signals across traits colocalize (colocalization) or are driven by distinct causal variants (pleiotropy) (Table 1) [30]. While programs such as COLOC [30] look at a specific queried locus, GWAS-PW [31] can perform genome-wide analysis, partitioning the genome into predefined regions.

Colocalization analysis in endometrial cancer research has unveiled intriguing insights into the shared genetic architecture between this malignancy and other traits or diseases. Colocalization analyses have indicated shared genetic variants between endometrial cancer and traits such as ovarian cancer [50] and COVID-19 phenotypes [59]. Colocalization analysis can also include expression quantitative trait loci (eQTLs) that have enabled the identification of several novel candidate endometrial cancer susceptibility genes [54,60–62].

Colocalization analysis faces several limitations; for example, in instances of very high linkage disequilibrium (LD), distinguishing between shared pleiotropic variants and those acting independently becomes challenging, leading to ambiguity in attributing shared genetic influences to specific traits. Additionally, limitations in statistical power may hinder the detection of true colocalization events, particularly in independent studies with small sample sizes or low genetic variant density. Like genetic correlation analysis, colocalization analysis does not provide information on the directionality of effects or causality, highlighting the need for complementary experimental approaches to validate findings and elucidate underlying biological mechanisms.

5. Cross-Trait Locus Identification

Cross-trait locus identification is the process of identifying genetic variants associated with multiple phenotypes or traits simultaneously. This approach uses a meta-analysis framework to combine summary or individual-level statistics of distinct or related traits to detect pleiotropic loci with shared associations. Cross-trait GWAS leverages power from multiple genetically correlated phenotypes to detect genetic loci that may not reach genome-wide statistical significance when considering each trait individually. Moreover, cross-trait analysis offers the benefit of utilizing existing datasets, obviating the necessity for additional genotyping efforts, thus presenting a considerable practical advantage.

Several methods have been developed for cross-trait variant identification (Table 1), with the most commonly used being Multi-Trait Analysis of GWAS (MTAG) due to its robustness to sample overlap and its ability to leverage GWAS summary statistics from multiple traits simultaneously [32]. MTAG can improve statistical power by borrowing information across traits and has been shown to increase the discovery of associated loci, especially for traits with smaller sample sizes [32]. Additional methods for cross-trait variant identification include Cross-Phenotype Association (CPASSOC), employing a Bayesian framework to account for correlation structures and offering valuable insights into shared genetic mechanisms [33]; Multivariate PLINK (MV-PLINK), which is particularly effective in extensive individual-level datasets, employing multivariate linear regression to detect common variants while controlling for covariates and population structure [34]; and Multi-Phenotype Conditional False Discovery Rate (MultiPhen conjFDR), regulating false discovery rates across diverse traits and highlighting variants with the strongest evidence of shared associations [35,36].

Cross-trait variant identification has shown promise in advancing our understanding of endometrial cancer by uncovering novel risk regions replicable in larger endometrial cancer GWAS.

In past studies, joint analyses aimed at identifying cross-trait variants associated with endometrial cancer have employed MTAG, incorporating genetically correlated traits such as ovarian cancer, PCOS, and uterine fibroids[50,54]. Additionally, joint analyses with endometriosis used a fixed-effects GWAS meta-analysis for cross-trait variant identification [53]. Preliminary findings using RE2C [38] supported the identification of five novel risk loci for endometrial cancer in a joint analysis with three hormone-related cancers (breast, prostate, and ovarian) [63]. Other cross-trait analyses of COVID-19 and endometrial cancer employed CPASSOC and determined five independent pleiotropic SNPs for endometrial cancer located at three previously identified loci (9q34.2, 2p16.1 and 15q21.2)[59]. Using a Bayesian framework (bGWAS)[37], a multi-trait GWAS analysis included several endometrial cancer risk factors and identified a novel risk locus at 7q22.1, replicated in an independent cohort [61]. This newly identified locus at 7q22.1 encompasses CYP3A7, which encodes an enzyme that metabolizes testosterone and synthesizes estrogen. Exposure to estrogen is suspected to elevate the risk of endometrial carcinogenesis. Despite these insightful findings, there is ample room for more research in cross-trait variant identification, including a broader range of traits, to improve our understanding of the underlying genetic architecture of endometrial cancer and shed light on potential biological pathways and mechanisms driving its development.

Despite its potential benefits, cross-trait variant identification also presents several limitations. Firstly, a shared causal variant may exhibit varying biological impacts among different traits, leading to heightened heterogeneity in reported effect sizes. Secondly, a genetic variant might solely influence a subset of study traits. Methods aimed at identifying cross-trait variants typically report an overall association with a genetic locus without attributing the association to a specific trait. Occasionally, the same genetic variant can exert opposing effects, increasing the risk of one disease while decreasing the risk of another. In such instances, these methods may fail to detect an association when a genetic locus exhibits antagonistic effects on the traits under investigation. Lastly, discriminating between heterogeneous effects and statistical noise can prove challenging when assessing multiple traits of differing power and study design. These complexities underscore the need for robust methodologies and cautious interpretation in cross-trait variant identification analyses to ensure reliable insights into shared genetic architecture across diverse traits and diseases.

6. Causal Inference Analysis

Mendelian randomization (MR) is the most commonly used genetic approach for causal inference analysis, a statistical method that uses genetic variants as instrumental variables (IVs) to infer causal relationships of observed associations between an exposure trait and an outcome trait [42]. MR is analogous to a conventional randomized controlled trial; however, unlike random intervention allocation, MR relies on the premise that the germline alleles of these variants are randomly distributed (due to Mendel's Law of Independent Assortment), thereby establishing an unconfounded relationship between exposure and outcome [42,64]. When the exposure traits have a substantial genetic component, the IV explains a larger proportion of variation in the risk factor and improves power. As GWAS samples become larger and more SNP associations are identified, the IVs will only become a stronger proxy for the exposure trait, increasing statistical power in MR studies [55].

Several publications have described the underlying statistical framework and assumptions required for MR [42]. Briefly, for the validity of causal effect estimates, MR analysis relies on three primary assumptions that must be satisfied (Figure 1): the relevance assumption, which asserts that genetic variants are robustly associated with the risk factor; the independence assumption, which stipulates that there are no external confounding influences affecting the association between the genetic variants, the risk-factor and the outcome; and the exclusion restriction assumption, which posits that the genetic variants affect only the outcome through the risk-factor (also known as the pleiotropy assumption); this assume there is no pleiotropic effect influencing the gene-outcome association other than that of the vertical pleiotropy implicating the causal pathway between the risk-factor and the outcome. MR methods should be combined with robust estimation methods (Table 2) to tackle bias introduced due to pleiotropy [65].

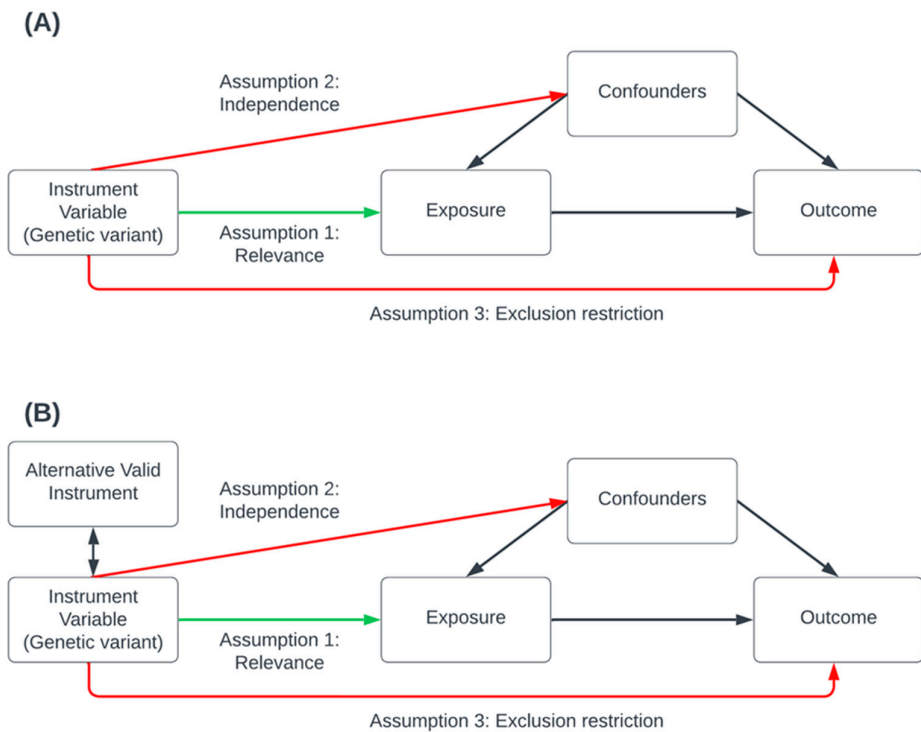


Figure 1. Conceptual visualization of the key genetic instrumental variable assumptions of Mendelian Randomization. In this valid Mendelian randomization (MR) simulation, a genetic variant causally affects the exposure, which may, in turn, causally affect the outcome while accounting for confounders that influence both the exposure and the outcome. Causal effects are depicted using one-sided arrows indicating the direction of causation. Panel (A) illustrates the three key assumptions: the green arrow represents the relevance assumption, indicating a valid causal pathway, while the red arrows represent causal pathways explicitly excluded by the independence and exclusion restriction assumptions. Panel (B) depicts a scenario in which a valid alternative instrument, though not causal, is in linkage disequilibrium with a causal variant introducing a bias due to pleiotropy. This is tackled by applying a range of MR sensitivity analysis methods and recognition of assumptions underlying the analysis when interpreting results.

Table 2. Mendelian randomization software packages and consistency assumption approaches.

Methods	Ref	Description
Packages		
MR-Base	[66]	A web platform housing GWAS summary statistics that can perform two-sample MR analyses
MendelianRandomization	[67]	R software package that implements several methods for MR analyses based on summary statistics including multivariable MR
CAUSE	[68]	R software package for MR analysis accounting for both uncorrelated and correlated horizontal pleiotropy effects
TwoSampleMR	[69]	R software package to perform a range of two-sample MR analyses using GWAS summary data from two independent exposure and outcome cohorts.
OneSampleMR	[70]	R software package to perform a range of one-sample MR analyses using GWAS data from a single cohort (individual-level data)
Consistency Assumption: Instrument Strength Independent of Direct Effect		
MR-Egger	[71]	A sensitivity analysis tool used to test for bias from pleiotropy caused by some genetic variants having multiple proximal phenotypic

Methods	Ref	Description
correlations, making them invalid instrumental variables. Egger’s test provides a valid causal effect estimate when some or all the genetic variants are invalid instrumental variables.		
Consistency Assumption: Majority Valid		
Weight-median	[72]	A sensitivity analysis tool that uses GWAS summary data for MR with multiple genetic variants. Provides a consistent causal effect estimate even when up to 50% of the information comes from invalid instrumental variables.
Consistency Assumption: Plurality Valid		
Weighted-MBE	[73]	A sensitivity analysis tool using summary data that is robust to horizontal pleiotropy. Provides a consistent causal estimate when the largest number of similar individual-instrument causal effect estimates comes from valid instruments, even if the majority of instruments are invalid.
Consistency Assumption: Horizontal pleiotropy around 0		
MR-LDP	[74]	An efficient variational Bayesian expectation-maximization algorithm using GWAS summary statistics to estimate the causal effects of complex traits that have multiple instrumental variants within LD. The random component eliminates the impact of horizontal pleiotropy.
MR-RAPS	[75]	Uses GWAS summary data under a random-effect model to estimate the causal effects of genetic variants while accounting[68] for pleiotropy. It is robust to outliers but sensitive to the assumption that pleiotropy is normally distributed around 0.
Consistency Assumption: Outlier-robust		
GSMR + HEIDI	[76]	Uses summary GWAS data to perform MR analysis by accounting for LD between the variants, thereby improving statistical power. Detects and accounts for outliers that could violate MR assumptions.
MR-GRAPPLE	[77]	Uses GWAS summary statistics to identify multiple pleiotropic pathways and determine the causal effect, under a likelihood model pervasive pleiotropy accounted for as long as the InSIDE assumption holds for all genetic instruments.
MR-Lasso	[65]	Extension of the IVW-MR framework by adding an intercept term for each genetic variant and a lasso penalty term for identification. Aim removes a potential source of bias (outliers) that could violate the analysis's assumptions.
MR-Robust	[65]	IVW method is performed by regression resulting in MM-estimation (robust against influential points) and Tukey's loss function (robust against outliers). Aims to downweigh outliers which could cause a violation of the assumptions underlying the analysis.
MR-PRESSO	[78]	Uses summary-level data to test and correct for horizontal pleiotropic outliers using a regression framework to “leave-one-out” approach which detects and removes outliers from the analysis to determine which SNP is driving the difference in computed residual sum of squares.

MR analyses can clarify the causal nature between putative risk factors and endometrial cancer susceptibility [18]. A recent study analyzed all known and suspected risk factors for endometrial cancer risk using MR and then performed a multivariable analysis to distil risk factors into five independent factors: waist circumference (in a module with BMI), age of menarche, age of menopause, SHBG levels and testosterone levels [61]. Factors such as cigarette smoking, dietary

factors, cannabis consumption, gut microbiota, sedentary behavior, and caffeine consumption have all received attention in MR research [79–86]. When interpreting results from published studies, it is crucial to consider the quality of the MR analysis as the quality of evidence provided relies on the satisfaction of instrumental variable assumptions. The accessibility of summary-level data for MR has contributed to a surge in the tool's popularity, often leading to its opportunistic use without substantiated biological or functional relevance. The risk factor should only be considered a common genetic predictor in these instances. While this represents a weaker claim, it still holds its place in scientific literature. Causation claims in MR analysis should be entirely for genetic variants with well-established biological and statistical relevance. For example, the established relationship between obesity and endometrial cancer risk means obesity-related factors have received extensive focus from MR research. Several studies leveraging GWAS [87,88] data have substantiated the causal link between higher BMI and increased risk of endometrial cancer across multiple European and Japanese cohorts [87–92] and histological subtypes [19,87].

MR approaches can unravel the associations between endometrial cancer and a trait that may be affected by their relationships with BMI. For example, childhood adiposity is an apparent health problem epidemiologically associated with endometrial cancer risk [93,94]. However, the causal nature of this association and whether it represents a direct or indirect effect mediated by adult obesity remain unclear. MR studies have disentangled the relationship between child and adult adiposity and subtype-specific endometrial cancer risk, identifying direct independent effects of childhood adiposity on the risk of non-endometrioid endometrial cancer but a minimal indirect effect that adult adiposity mediates on endometrioid EC risk [93]. These novel findings shed light on the critical role of targeting adiposity at different life stages to limit subtype-specific endometrial cancer risk. Similarly, MR analyses have determined molecular mediators underlying endometrial cancer risk [90]. This study identified two sex-steroid hormones (bioavailable testosterone and SHBG), and fasting insulin strongly mediate the relationship between excess adiposity and endometrial carcinogenesis [90]. These findings suggest that in the future, there is scope to investigate targeting these hormone-related and insulin-related traits for endometrial cancer prevention.

With the rapid advancement of MR in endometrial cancer research, awareness of the limitations of this approach is essential for the correct interpretation of results. As mentioned, traditional approaches rely on the core assumptions being met for accurate causal effect estimates [95]. Bias in MR studies can arise from confounding of genetic-intermediate phenotype-disease associations, trait heterogeneity, and linkage disequilibrium, which may obscure causal effect estimates [76,96–98]. Sensitivity analyses and careful selection of instrumental variables are essential to address these issues and ensure reliable causal inferences [98,99]. Additionally, canalization or developmental compensation processes can distort MR estimates by mitigating the effects of genetic variants on phenotype expression, further complicating interpretation [98–100]. Other limitations include confounding due to population stratification, dynastic effects, assortative mating, selection bias, and collider bias, all of which can introduce systematic errors and undermine the validity of MR findings [98,101,102].

7. Conclusions

In summary, this review underscores the crucial role of cross-trait GWAS in elucidating genotype-phenotype associations and advancing our understanding of complex diseases such as endometrial cancer. It provides an updated synopsis of the genetic architecture of endometrial cancer by comprehensively detailing related genetic studies in the field. Through leveraging large-scale publicly available data, joint analysis has effectively highlighted the interplay between genetic susceptibility in different phenotypes, offering insights into comorbidities, modifiable risk factors, and genetic predisposition. While primary analytic methods have been instrumental in unveiling significant associations with shared common genetic variants in endometrial cancer, these variants only explain a fraction of the expected risk variance, suggesting the potential for introducing new statistical tools to identify novel risk loci. Despite the potential benefits of cross-trait GWAS, they also

present methodological challenges, underscoring the need for robust methodologies and cautious interpretation to ensure reliable insights.

Furthermore, this review delineates significant findings achieved through various stages of genome-wide cross-trait analysis, including genetic correlation, colocalization analysis, cross-trait meta-analysis, and Mendelian randomization. It emphasizes that investigation of shared genetic factors in endometrial cancer is still in its infancy, offering numerous promising avenues for future exploration. The expanding data repositories and innovative analytical methodologies enhance the capacity for identifying risk loci. Integrating multi-omics datasets can also deepen our understanding of the molecular mechanisms underlying disease susceptibility and progression, potentially unveiling novel biomarkers. These advancements broaden the scope for downstream analyses focused on discovering new biological pathways and therapeutic targets, with the ultimate goal of clinical translation of multiple diseases. Cross-trait GWAS may facilitate the development of polygenic risk scores and predictive models to assess individual risk profiles and guide personalized prevention and intervention strategies. Moreover, expanding GWAS data to include under-represented populations, thus better reflecting the global community and increasing emphasis on cross-population analyses, may offer novel insights into disease etiology and pathogenesis. The field holds promise for further elucidating the genetic basis of endometrial cancer and other complex diseases, offering multiple avenues for improved prevention, diagnosis, and treatment strategies that can impact on the global community.

Author Contributions: Conceptualization, T.A.O'M.; investigation, K.B., N.I.; data curation, K.B.; writing—original draft preparation, K.B.; writing—review and editing, N.I., T.A.O'M.; visualization, K.B.; supervision, T.A.O'M.; funding acquisition, T.A.O'M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by an Investigator Grant from the National Health and Medical Research Council of Australia, grant number APP1173170.

Institutional Review Board Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the review; in the collection, or interpretation of data; in the writing of the manuscript; or in the decision to publish.

References

1. Uffelmann, E.; Huang, Q.Q.; Munung, N.S.; de Vries, J.; Okada, Y.; Martin, A.R.; Martin, H.C.; Lappalainen, T. Genome-Wide Association Studies. *Nature Reviews Methods Primers* **2021**, *1*, 1–21, doi:10.1038/s43586-021-00056-9.
2. Gu, B.; Shang, X.; Yan, M.; Li, X.; Wang, W.; Wang, Q.; Zhang, C. Variations in Incidence and Mortality Rates of Endometrial Cancer at the Global, Regional, and National Levels, 1990–2019. *Gynecol. Oncol.* **2021**, *161*, 573–580, doi:10.1016/j.ygyno.2021.01.036.
3. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249, doi:10.3322/caac.21660.
4. Crosbie, E.J.; Kitson, S.J.; McAlpine, J.N.; Mukhopadhyay, A.; Powell, M.E.; Singh, N. Endometrial Cancer. *Lancet* **2022**, *399*, 1412–1428, doi:10.1016/S0140-6736(22)00323-3.
5. Kitson, S.J.; Crosbie, E.J. Endometrial Cancer and Obesity. *Obstet. Gynaecol.* **2019**, *21*, 237–245, doi:10.1111/tog.12601.
6. Setiawan, V.W.; Yang, H.P.; Pike, M.C.; McCann, S.E.; Yu, H.; Xiang, Y.-B.; Wolk, A.; Wentzensen, N.; Weiss, N.S.; Webb, P.M.; et al. Type I and II Endometrial Cancers: Have They Different Risk Factors? *J. Clin. Oncol.* **2013**, *31*, 2607–2618, doi:10.1200/JCO.2012.48.2596.
7. Raglan, O.; Kalliala, I.; Markozannes, G.; Cividini, S.; Gunter, M.J.; Nautiyal, J.; Gabra, H.; Paraskevidis, E.; Martin-Hirsch, P.; Tsilidis, K.K.; et al. Risk Factors for Endometrial Cancer: An Umbrella Review of the Literature. *Int. J. Cancer* **2019**, *145*, 1719–1730, doi:10.1002/ijc.31961.
8. Kaaks, R.; Lukanova, A.; Kurzer, M.S. Obesity, Endogenous Hormones, and Endometrial Cancer Risk: A Synthetic Review. *Cancer Epidemiol. Biomarkers Prev.* **2002**, *11*, 1531–1543.
9. Mucci, L.A.; Hjelmberg, J.B.; Harris, J.R.; Czene, K.; Havelick, D.J.; Scheike, T.; Graff, R.E.; Holst, K.; Möller, S.; Unger, R.H.; et al. Familial Risk and Heritability of Cancer Among Twins in Nordic Countries. *JAMA* **2016**, *315*, 68–76, doi:10.1001/jama.2015.17703.

10. Lu, Y.; Ek, W.E.; Whiteman, D.; Vaughan, T.L.; Spurdle, A.B.; Easton, D.F.; Pharoah, P.D.; Thompson, D.J.; Dunning, A.M.; Hayward, N.K.; et al. Most Common “sporadic” Cancers Have a Significant Germline Genetic Component. *Hum. Mol. Genet.* **2014**, *23*, 6112–6118, doi:10.1093/hmg/ddu312.
11. Johnatty, S.E.; Tan, Y.Y.; Buchanan, D.D.; Bowman, M.; Walters, R.J.; Obermair, A.; Quinn, M.A.; Blomfield, P.B.; Brand, A.; Leung, Y.; et al. Family History of Cancer Predicts Endometrial Cancer Risk Independently of Lynch Syndrome: Implications for Genetic Counselling. *Gynecol. Oncol.* **2017**, *147*, 381–387, doi:10.1016/j.ygyno.2017.08.011.
12. Lichtenstein Paul; Holm Niels V.; Verkasalo Pia K.; Iliadou Anastasia; Kaprio Jaakko; Koskenvuo Markku; Pukkala Eero; Skytthe Axel; Hemminki Kari Environmental and Heritable Factors in the Causation of Cancer — Analyses of Cohorts of Twins from Sweden, Denmark, and Finland. *N. Engl. J. Med.* **343**, 78–85, doi:10.1056/NEJM200007133430201.
13. Spurdle, A.B.; Bowman, M.A.; Shamsani, J.; Kirk, J. Endometrial Cancer Gene Panels: Clinical Diagnostic vs Research Germline DNA Testing. *Mod. Pathol.* **2017**, *30*, 1048–1068, doi:10.1038/modpathol.2017.20.
14. Dörk, T.; Hillemanns, P.; Tempfer, C.; Breu, J.; Fleisch, M.C. Genetic Susceptibility to Endometrial Cancer: Risk Factors and Clinical Management. *Cancers* **2020**, *12*, doi:10.3390/cancers12092407.
15. Ryan, N.A.J.; Glaire, M.A.; Blake, D.; Cabrera-Dandy, M.; Evans, D.G.; Crosbie, E.J. The Proportion of Endometrial Cancers Associated with Lynch Syndrome: A Systematic Review of the Literature and Meta-Analysis. *Genet. Med.* **2019**, *21*, 2167–2180, doi:10.1038/s41436-019-0536-8.
16. Buchanan, D.D.; Tan, Y.Y.; Walsh, M.D.; Clendenning, M.; Metcalf, A.M.; Ferguson, K.; Arnold, S.T.; Thompson, B.A.; Lose, F.A.; Parsons, M.T.; et al. Tumor Mismatch Repair Immunohistochemistry and DNA MLH1 Methylation Testing of Patients with Endometrial Cancer Diagnosed at Age Younger than 60 Years Optimizes Triage for Population-Level Germline Mismatch Repair Gene Mutation Testing. *J. Clin. Oncol.* **2014**, *32*, 90–100, doi:10.1200/JCO.2013.51.2129.
17. O’Mara, T.A.; Glubb, D.M.; Kho, P.F.; Thompson, D.J.; Spurdle, A.B. Genome-Wide Association Studies of Endometrial Cancer: Latest Developments and Future Directions. *Cancer Epidemiol. Biomarkers Prev.* **2019**, *28*, 1095–1102, doi:10.1158/1055-9965.EPI-18-1031.
18. Wang, X.; Glubb, D.M.; O’Mara, T.A. 10 Years of GWAS Discovery in Endometrial Cancer: Aetiology, Function and Translation. *EBioMedicine* **2022**, *77*, 103895, doi:10.1016/j.ebiom.2022.103895.
19. O’Mara, T.A.; Glubb, D.M.; Amant, F.; Annibaldi, D.; Ashton, K.; Attia, J.; Auer, P.L.; Beckmann, M.W.; Black, A.; Bolla, M.K.; et al. Identification of Nine New Susceptibility Loci for Endometrial Cancer. *Nat. Commun.* **2018**, *9*, 3166, doi:10.1038/s41467-018-05427-7.
20. Loh, P.-R.; Bhatia, G.; Gusev, A.; Finucane, H.K.; Bulik-Sullivan, B.K.; Pollack, S.J.; Schizophrenia Working Group of Psychiatric Genomics Consortium; de Candia, T.R.; Lee, S.H.; Wray, N.R.; et al. Contrasting Genetic Architectures of Schizophrenia and Other Complex Diseases Using Fast Variance-Components Analysis. *Nat. Genet.* **2015**, *47*, 1385–1392, doi:10.1038/ng.3431.
21. Grotzinger, A.D.; Rhemtulla, M.; de Vlaming, R.; Ritchie, S.J.; Mallard, T.T.; Hill, W.D.; Ip, H.F.; Marioni, R.E.; McIntosh, A.M.; Deary, I.J.; et al. Genomic Structural Equation Modelling Provides Insights into the Multivariate Genetic Architecture of Complex Traits. *Nat. Hum. Behav.* **2019**, *3*, 513–525, doi:10.1038/s41562-019-0566-x.
22. Yang, J.; Lee, S.H.; Goddard, M.E.; Visscher, P.M. GCTA: A Tool for Genome-Wide Complex Trait Analysis. *Am. J. Hum. Genet.* **2011**, *88*, 76–82, doi:10.1016/j.ajhg.2010.11.011.
23. Weissbrod, O.; Flint, J.; Rosset, S. Estimating SNP-Based Heritability and Genetic Correlation in Case-Control Studies Directly and with Summary Statistics. *Am. J. Hum. Genet.* **2018**, *103*, 89–99, doi:10.1016/j.ajhg.2018.06.002.
24. Bulik-Sullivan, B.; Finucane, H.K.; Anttila, V.; Gusev, A.; Day, F.R.; Loh, P.-R.; ReproGen Consortium; Psychiatric Genomics Consortium; Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control Consortium 3; Duncan, L.; et al. An Atlas of Genetic Correlations across Human Diseases and Traits. *Nat. Genet.* **2015**, *47*, 1236–1241, doi:10.1038/ng.3406.
25. Ning, Z.; Pawitan, Y.; Shen, X. High-Definition Likelihood Inference of Genetic Correlations across Human Complex Traits. *Nat. Genet.* **2020**, *52*, 859–864, doi:10.1038/s41588-020-0653-y.
26. Lee, S.H.; van der Werf, J.H.J. MTG2: An Efficient Algorithm for Multivariate Linear Mixed Model Analysis Based on Genomic Information. *Bioinformatics* **2016**, *32*, 1420–1422, doi:10.1093/bioinformatics/btw012.
27. Lu, Q.; Li, B.; Ou, D.; Erlendsdottir, M.; Powles, R.L.; Jiang, T.; Hu, Y.; Chang, D.; Jin, C.; Dai, W.; et al. A Powerful Approach to Estimating Annotation-Stratified Genetic Covariance via GWAS Summary Statistics. *Am. J. Hum. Genet.* **2017**, *101*, 939–964, doi:10.1016/j.ajhg.2017.11.001.
28. Shi, H.; Mancuso, N.; Spendlove, S.; Pasaniuc, B. Local Genetic Correlation Gives Insights into the Shared Genetic Architecture of Complex Traits. *Am. J. Hum. Genet.* **2017**, *101*, 737–751, doi:10.1016/j.ajhg.2017.09.022.
29. Werme, J.; van der Sluis, S.; Posthuma, D.; de Leeuw, C.A. An Integrated Framework for Local Genetic Correlation Analysis. *Nat. Genet.* **2022**, *54*, 274–282, doi:10.1038/s41588-022-01017-y.

30. Giambartolomei, C.; Vukcevic, D.; Schadt, E.E.; Franke, L.; Hingorani, A.D.; Wallace, C.; Plagnol, V. Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. *PLoS Genet.* **2014**, *10*, e1004383, doi:10.1371/journal.pgen.1004383.
31. Pickrell, J.K.; Berisa, T.; Liu, J.Z.; Séguérel, L.; Tung, J.Y.; Hinds, D.A. Detection and Interpretation of Shared Genetic Influences on 42 Human Traits. *Nat. Genet.* **2016**, *48*, 709–717, doi:10.1038/ng.3570.
32. Turley, P.; Walters, R.K.; Maghzian, O.; Okbay, A.; Lee, J.J.; Fontana, M.A.; Nguyen-Viet, T.A.; Wedow, R.; Zacher, M.; Furlotte, N.A.; et al. Multi-Trait Analysis of Genome-Wide Association Summary Statistics Using MTAG. *Nat. Genet.* **2018**, *50*, 229–237, doi:10.1038/s41588-017-0009-4.
33. Zhu, X.; Feng, T.; Tayo, B.O.; Liang, J.; Young, J.H.; Franceschini, N.; Smith, J.A.; Yanek, L.R.; Sun, Y.V.; Edwards, T.L.; et al. Meta-Analysis of Correlated Traits via Summary Statistics from GWASs with an Application in Hypertension. *Am. J. Hum. Genet.* **2015**, *96*, 21–36, doi:10.1016/j.ajhg.2014.11.011.
34. Ferreira, M.A.R.; Purcell, S.M. A Multivariate Test of Association. *Bioinformatics* **2009**, *25*, 132–133, doi:10.1093/bioinformatics/btn563.
35. O'Reilly, P.F.; Hoggart, C.J.; Pomyen, Y.; Calboli, F.C.F.; Elliott, P.; Jarvelin, M.-R.; Coin, L.J.M. MultiPhen: Joint Model of Multiple Phenotypes Can Increase Discovery in GWAS. *PLoS One* **2012**, *7*, e34861, doi:10.1371/journal.pone.0034861.
36. Smeland, O.B.; Frei, O.; Shadrin, A.; O'Connell, K.; Fan, C.-C.; Bahrami, S.; Holland, D.; Djurovic, S.; Thompson, W.K.; Dale, A.M.; et al. Discovery of Shared Genomic Loci Using the Conditional False Discovery Rate Approach. *Hum. Genet.* **2020**, *139*, 85–94, doi:10.1007/s00439-019-02060-2.
37. Mounier, N.; Kutalik, Z. BGWAS: An R Package to Perform Bayesian Genome Wide Association Studies. *Bioinformatics* **2020**, *36*, 4374–4376, doi:10.1093/bioinformatics/btaa549.
38. Lee, C.H.; Eskin, E.; Han, B. Increasing the Power of Meta-Analysis of Genome-Wide Association Studies to Detect Heterogeneous Effects. *Bioinformatics* **2017**, *33*, i379–i388, doi:10.1093/bioinformatics/btx242.
39. Trochet, H.; Pirinen, M.; Band, G.; Jostins, L.; McVean, G.; Spencer, C.C.A. Bayesian Meta-Analysis across Genome-Wide Association Studies of Diverse Phenotypes. *Genet. Epidemiol.* **2019**, *43*, 532–547, doi:10.1002/gepi.22202.
40. O'Connor, L.J.; Price, A.L. Distinguishing Genetic Correlation from Causation across 52 Diseases and Complex Traits. *Nat. Genet.* **2018**, *50*, 1728–1734, doi:10.1038/s41588-018-0255-0.
41. Frei, O.; Holland, D.; Smeland, O.B.; Shadrin, A.A.; Fan, C.C.; Maeland, S.; O'Connell, K.S.; Wang, Y.; Djurovic, S.; Thompson, W.K.; et al. Bivariate Causal Mixture Model Quantifies Polygenic Overlap between Complex Traits beyond Genetic Correlation. *Nat. Commun.* **2019**, *10*, 2417, doi:10.1038/s41467-019-10310-0.
42. Sanderson, E.; Glymour, M.M.; Holmes, M.V.; Kang, H.; Morrison, J.; Munafò, M.R.; Palmer, T.; Schooling, C.M.; Wallace, C.; Zhao, Q.; et al. Mendelian Randomization. *Nat Rev Methods Primers* **2022**, *2*, doi:10.1038/s43586-021-00092-5.
43. Rødevand, L.; Rahman, Z.; Hindley, G.F.L.; Smeland, O.B.; Frei, O.; Tekin, T.F.; Kutrolli, G.; Bahrami, S.; Hoseth, E.Z.; Shadrin, A.; et al. Characterizing the Shared Genetic Underpinnings of Schizophrenia and Cardiovascular Disease Risk Factors. *Am. J. Psychiatry* **2023**, appiajp20220660, doi:10.1176/appi.ajp.20220660.
44. Ohi, K.; Sumiyoshi, C.; Fujino, H.; Yasuda, Y.; Yamamori, H.; Fujimoto, M.; Shiino, T.; Sumiyoshi, T.; Hashimoto, R. Genetic Overlap between General Cognitive Function and Schizophrenia: A Review of Cognitive GWASs. *Int. J. Mol. Sci.* **2018**, *19*, doi:10.3390/ijms19123822.
45. Romero, C.; Werme, J.; Jansen, P.R.; Gelemtier, J.; Stein, M.B.; Levey, D.; Polimanti, R.; de Leeuw, C.; Posthuma, D.; Nagel, M.; et al. Exploring the Genetic Overlap between Twelve Psychiatric Disorders. *Nat. Genet.* **2022**, *54*, 1795–1802, doi:10.1038/s41588-022-01245-2.
46. Duncan, L.E.; Shen, H.; Ballon, J.S.; Hardy, K.V.; Noordsy, D.L.; Levinson, D.F. Genetic Correlation Profile of Schizophrenia Mirrors Epidemiological Results and Suggests Link Between Polygenic and Rare Variant (22q11.2) Cases of Schizophrenia. *Schizophr. Bull.* **2018**, *44*, 1350–1361, doi:10.1093/schbul/sbx174.
47. Catalá-López, F.; Suárez-Pinilla, M.; Suárez-Pinilla, P.; Valderas, J.M.; Gómez-Beneyto, M.; Martínez, S.; Balanzá-Martínez, V.; Climent, J.; Valencia, A.; McGrath, J.; et al. Inverse and Direct Cancer Comorbidity in People with Central Nervous System Disorders: A Meta-Analysis of Cancer Incidence in 577,013 Participants of 50 Observational Studies. *Psychother. Psychosom.* **2014**, *83*, 89–105, doi:10.1159/000356498.
48. Tran, E.; Rouillon, F.; Loze, J.-Y.; Casadebaig, F.; Philippe, A.; Vitry, F.; Limosin, F. Cancer Mortality in Patients with Schizophrenia: An 11-Year Prospective Cohort Study. *Cancer* **2009**, *115*, 3555–3562, doi:10.1002/cncr.24383.
49. Lu, D.; Song, J.; Lu, Y.; Fall, K.; Chen, X.; Fang, F.; Landén, M.; Hultman, C.M.; Czene, K.; Sullivan, P.; et al. A Shared Genetic Contribution to Breast Cancer and Schizophrenia. *Nat. Commun.* **2020**, *11*, 4637, doi:10.1038/s41467-020-18492-8.
50. Glubb, D.M.; Thompson, D.J.; Aben, K.K.H.; Alsulimani, A.; Amant, F.; Annibali, D.; Attia, J.; Barricarte, A.; Beckmann, M.W.; Berchuck, A.; et al. Cross-Cancer Genome-Wide Association Study of Endometrial Cancer and Epithelial Ovarian Cancer Identifies Genetic Risk Regions Associated with Risk of Both Cancers. *Cancer Epidemiol. Biomarkers Prev.* **2021**, *30*, 217–228, doi:10.1158/1055-9965.EPI-20-0739.

51. Lindström, S.; Wang, L.; Feng, H.; Majumdar, A.; Huo, S.; Macdonald, J.; Harrison, T.; Turman, C.; Chen, H.; Mancuso, N.; et al. Genome-Wide Analyses Characterize Shared Heritability among Cancers and Identify Novel Cancer Susceptibility Regions. *J. Natl. Cancer Inst.* **2023**, *115*, 712–732, doi:10.1093/jnci/djad043.
52. Masuda, T.; Low, S.-K.; Akiyama, M.; Hirata, M.; Ueda, Y.; Matsuda, K.; Kimura, T.; Murakami, Y.; Kubo, M.; Kamatani, Y.; et al. GWAS of Five Gynecologic Diseases and Cross-Trait Analysis in Japanese. *Eur. J. Hum. Genet.* **2020**, *28*, 95–107, doi:10.1038/s41431-019-0495-1.
53. Painter, J.N.; O'Mara, T.A.; Morris, A.P.; Cheng, T.H.T.; Gorman, M.; Martin, L.; Hodson, S.; Jones, A.; Martin, N.G.; Gordon, S.; et al. Genetic Overlap between Endometriosis and Endometrial Cancer: Evidence from Cross-Disease Genetic Correlation and GWAS Meta-Analyses. *Cancer Med.* **2018**, *7*, 1978–1987, doi:10.1002/cam4.1445.
54. Kho, P.F.; Mortlock, S.; Amant, F.; Annibali, D.; Ashton, K.; Attia, J.; Auer, P.L.; Beckmann, M.W.; Black, A.; Brinton, L.; et al. Genetic Analyses of Gynecological Disease Identify Genetic Relationships between Uterine Fibroids and Endometrial Cancer, and a Novel Endometrial Cancer Genetic Risk Region at the WNT4 1p36.12 Locus. *Hum. Genet.* **2021**, *140*, 1353–1365, doi:10.1007/s00439-021-02312-0.
55. Kraft, P.; Chen, H.; Lindström, S. The Use Of Genetic Correlation And Mendelian Randomization Studies To Increase Our Understanding of Relationships Between Complex Traits. *Curr Epidemiol Rep* **2020**, *7*, 104–112, doi:10.1007/s40471-020-00233-6.
56. Chen, H.; Majumdar, A.; Wang, L.; Kar, S.; Brown, K.M.; Feng, H.; Turman, C.; Dennis, J.; Easton, D.; Michailidou, K.; et al. Large-Scale Cross-Cancer Fine-Mapping of the 5p15.33 Region Reveals Multiple Independent Signals. *HGG Adv* **2021**, *2*, 100041, doi:10.1016/j.xhgg.2021.100041.
57. Guo, H.; Fortune, M.D.; Burren, O.S.; Schofield, E.; Todd, J.A.; Wallace, C. Integration of Disease Association and eQTL Data Using a Bayesian Colocalisation Approach Highlights Six Candidate Causal Genes in Immune-Mediated Diseases. *Hum. Mol. Genet.* **2015**, *24*, 3305–3313, doi:10.1093/hmg/ddv077.
58. Jaffe, A.E.; Gao, Y.; Deep-Soboslay, A.; Tao, R.; Hyde, T.M.; Weinberger, D.R.; Kleinman, J.E. Mapping DNA Methylation across Development, Genotype and Schizophrenia in the Human Frontal Cortex. *Nat. Neurosci.* **2016**, *19*, 40–47, doi:10.1038/nn.4181.
59. Zhao, X.; Wu, X.; Xiao, J.; Zhang, L.; Hao, Y.; Xiao, C.; Zhang, B.; Li, J.; Jiang, X. A Large-Scale Genome-Wide Cross-Trait Analysis for the Effect of COVID-19 on Female-Specific Cancers. *iScience* **2023**, *26*, 107497, doi:10.1016/j.isci.2023.107497.
60. Kho, P.F.; Wang, X.; Cuéllar-Partida, G.; Dörk, T.; Goode, E.L.; Lambrechts, D.; Scott, R.J.; Spurdle, A.B.; O'Mara, T.A.; Glubb, D.M. Multi-Tissue Transcriptome-Wide Association Study Identifies Eight Candidate Genes and Tissue-Specific Gene Expression Underlying Endometrial Cancer Susceptibility. *Commun Biol* **2021**, *4*, 1211, doi:10.1038/s42003-021-02745-3.
61. Wang, X.; Kho, P.F.; Ramachandran, D.; Bafligil, C.; Amant, F.; Goode, E.L.; Scott, R.J.; Tomlinson, I.; Evans, D.G.; Crosbie, E.J.; et al. Multi-Trait Genome-Wide Association Study Identifies a Novel Endometrial Cancer Risk Locus That Associates with Testosterone Levels. *iScience* **2023**, *26*, 106590, doi:10.1016/j.isci.2023.106590.
62. Zhou, Y.; Zhou, X.; Sun, J.; Wang, L.; Zhao, J.; Chen, J.; Yuan, S.; He, Y.; Timofeeva, M.; Spiliopoulou, A.; et al. Exploring the Cross-Cancer Effect of Smoking and Its Fingerprints in Blood DNA Methylation on Multiple Cancers: A Mendelian Randomization Study. *Int. J. Cancer* **2023**, *153*, 1477–1486, doi:10.1002/ijc.34656.
63. Kar, S.P.; Lindström, S.; Hung, R.J.; Lawrenson, K.; Schmidt, M.K.; O'Mara, T.A.; Glubb, D.M.; Tyrer, J.P.; Schildkraut, J.M.; Chang-Claude, J.; et al. Combining Genome-Wide Studies of Breast, Prostate, Ovarian and Endometrial Cancers Maps Cross-Cancer Susceptibility Loci and Identifies New Genetic Associations. *bioRxiv* **2020**, 2020.06.16.146803.
64. Davies, N.M.; Holmes, M.V.; Davey Smith, G. Reading Mendelian Randomisation Studies: A Guide, Glossary, and Checklist for Clinicians. *BMJ* **2018**, *362*, k601, doi:10.1136/bmj.k601.
65. Slob, E.A.W.; Burgess, S. A Comparison of Robust Mendelian Randomization Methods Using Summary Data. *Genet. Epidemiol.* **2020**, *44*, 313–329, doi:10.1002/gepi.22295.
66. Hemani, G.; Zheng, J.; Elsworth, B.; Wade, K.H.; Haberland, V.; Baird, D.; Laurin, C.; Burgess, S.; Bowden, J.; Langdon, R.; et al. The MR-Base Platform Supports Systematic Causal Inference across the Human Phenome. *Elife* **2018**, *7*, doi:10.7554/eLife.34408.
67. Patel, A.; Ye, T.; Xue, H.; Lin, Z.; Xu, S.; Woolf, B.; Mason, A.M.; Burgess, S. MendelianRandomization v0.9.0: Updates to an R Package for Performing Mendelian Randomization Analyses Using Summarized Data. *Wellcome Open Res.* **2023**, *8*, 449, doi:10.12688/wellcomeopenres.19995.2.
68. Morrison, J.; Knoblach, N.; Marcus, J.H.; Stephens, M.; He, X. Mendelian Randomization Accounting for Correlated and Uncorrelated Pleiotropic Effects Using Genome-Wide Summary Statistics. *Nat. Genet.* **2020**, *52*, 740–747, doi:10.1038/s41588-020-0631-4.
69. TwoSampleMR: R Package for Performing 2-Sample MR Using MR-Base Database; Github;

70. Palmer, T. *OneSampleMR: R Package of Useful Functions for One-Sample Mendelian Randomization and Instrumental Variable Analyses*; Github;
71. Bowden, J.; Davey Smith, G.; Burgess, S. Mendelian Randomization with Invalid Instruments: Effect Estimation and Bias Detection through Egger Regression. *Int. J. Epidemiol.* **2015**, *44*, 512–525, doi:10.1093/ije/dyv080.
72. Bowden, J.; Davey Smith, G.; Haycock, P.C.; Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet. Epidemiol.* **2016**, *40*, 304–314, doi:10.1002/gepi.21965.
73. Hartwig, F.P.; Davey Smith, G.; Bowden, J. Robust Inference in Summary Data Mendelian Randomization via the Zero Modal Pleiotropy Assumption. *Int. J. Epidemiol.* **2017**, *46*, 1985–1998, doi:10.1093/ije/dyx102.
74. Cheng, Q.; Yang, Y.; Shi, X.; Yeung, K.-F.; Yang, C.; Peng, H.; Liu, J. MR-LDP: A Two-Sample Mendelian Randomization for GWAS Summary Statistics Accounting for Linkage Disequilibrium and Horizontal Pleiotropy. *NAR Genom. Bioinform.* **2020**, *2*, lqaa028, doi:10.1093/nargab/lqaa028.
75. Zhao, Q.; Chen, Y.; Wang, J.; Small, D.S. Powerful Three-Sample Genome-Wide Design and Robust Statistical Inference in Summary-Data Mendelian Randomization. *Int. J. Epidemiol.* **2019**, *48*, 1478–1492, doi:10.1093/ije/dyz142.
76. Zhu, Z.; Zheng, Z.; Zhang, F.; Wu, Y.; Trzaskowski, M.; Maier, R.; Robinson, M.R.; McGrath, J.J.; Visscher, P.M.; Wray, N.R.; et al. Causal Associations between Risk Factors and Common Diseases Inferred from GWAS Summary Data. *Nat. Commun.* **2018**, *9*, 224, doi:10.1038/s41467-017-02317-2.
77. Wang, J.; Zhao, Q.; Bowden, J.; Hemani, G.; Davey Smith, G.; Small, D.S.; Zhang, N.R. Causal Inference for Heritable Phenotypic Risk Factors Using Heterogeneous Genetic Instruments. *PLoS Genet.* **2021**, *17*, e1009575, doi:10.1371/journal.pgen.1009575.
78. Burgess, S.; Foley, C.N.; Allara, E.; Staley, J.R.; Howson, J.M.M. A Robust and Efficient Method for Mendelian Randomization with Hundreds of Genetic Variants. *Nat. Commun.* **2020**, *11*, 376, doi:10.1038/s41467-019-14156-4.
79. Dimou, N.; Omiyale, W.; Biessy, C.; Viallon, V.; Kaaks, R.; O'Mara, T.A.; Aglago, E.K.; Ardanaz, E.; Bergmann, M.M.; Bondonno, N.P.; et al. Cigarette Smoking and Endometrial Cancer Risk: Observational and Mendelian Randomization Analyses. *Cancer Epidemiol. Biomarkers Prev.* **2022**, *31*, 1839–1848, doi:10.1158/1055-9965.EPI-21-1176.
80. Wang, X.; Glubb, D.M.; O'Mara, T.A. Dietary Factors and Endometrial Cancer Risk: A Mendelian Randomization Study. *Nutrients* **2023**, *15*, doi:10.3390/nu15030603.
81. Niu, D.; Li, C.; Qu, H.; Zheng, Y. Does Cannabis Elevate Cancer Risk? : Evidence from Mendelian Randomization. *Wien. Klin. Wochenschr.* **2023**, doi:10.1007/s00508-023-02303-3.
82. Wei, Z.; Yang, B.; Tang, T.; Xiao, Z.; Ye, F.; Li, X.; Wu, S.; Huang, J.-G.; Jiang, S. Gut Microbiota and Risk of Five Common Cancers: A Univariable and Multivariable Mendelian Randomization Study. *Cancer Med.* **2023**, *12*, 10393–10405, doi:10.1002/cam4.5772.
83. Long, Y.; Tang, L.; Zhou, Y.; Zhao, S.; Zhu, H. Causal Relationship between Gut Microbiota and Cancers: A Two-Sample Mendelian Randomisation Study. *BMC Med.* **2023**, *21*, 66, doi:10.1186/s12916-023-02761-6.
84. Chen, Z.; Liu, C.; Wu, J.; Kong, F. Association of Coffee and Caffeine Consumption with Risk and Prognosis of Endometrial Cancer and Its Subgroups: A Mendelian Randomization. *Front Nutr* **2023**, *10*, 1291355, doi:10.3389/fnut.2023.1291355.
85. Chen, J.; Yang, K.; Qiu, Y.; Lai, W.; Qi, S.; Wang, G.; Chen, L.; Li, K.; Zhou, D.; Liu, Q.; et al. Genetic Associations of Leisure Sedentary Behaviors and the Risk of 15 Site-Specific Cancers: A Mendelian Randomization Study. *Cancer Med.* **2023**, *12*, 13623–13636, doi:10.1002/cam4.5974.
86. Larsson, S.C.; Lee, W.-H.; Kar, S.; Burgess, S.; Allara, E. Assessing the Role of Cortisol in Cancer: A Wide-Ranged Mendelian Randomisation Study. *Br. J. Cancer* **2021**, *125*, 1025–1029, doi:10.1038/s41416-021-01505-8.
87. Painter, J.N.; O'Mara, T.A.; Marquart, L.; Webb, P.M.; Attia, J.; Medland, S.E.; Cheng, T.; Dennis, J.; Holliday, E.G.; McEvoy, M.; et al. Genetic Risk Score Mendelian Randomization Shows That Obesity Measured as Body Mass Index, but Not Waist:Hip Ratio, Is Causal for Endometrial Cancer. *Cancer Epidemiol. Biomarkers Prev.* **2016**, *25*, 1503–1510, doi:10.1158/1055-9965.EPI-16-0147.
88. Prescott, J.; Setiawan, V.W.; Wentzensen, N.; Schumacher, F.; Yu, H.; Delahanty, R.; Bernstein, L.; Chanock, S.J.; Chen, C.; Cook, L.S.; et al. Body Mass Index Genetic Risk Score and Endometrial Cancer Risk. *PLoS One* **2015**, *10*, e0143256, doi:10.1371/journal.pone.0143256.
89. Nead, K.T.; Sharp, S.J.; Thompson, D.J.; Painter, J.N.; Savage, D.B.; Semple, R.K.; Barker, A.; Australian National Endometrial Cancer Study Group (ANECS); Perry, J.R.B.; Attia, J.; et al. Evidence of a Causal Association Between Insulinemia and Endometrial Cancer: A Mendelian Randomization Analysis. *J. Natl. Cancer Inst.* **2015**, *107*, doi:10.1093/jnci/djv178.
90. Hazelwood, E.; Sanderson, E.; Tan, V.Y.; Ruth, K.S.; Frayling, T.M.; Dimou, N.; Gunter, M.J.; Dossus, L.; Newton, C.; Ryan, N.; et al. Identifying Molecular Mediators of the Relationship between Body Mass Index

- and Endometrial Cancer Risk: A Mendelian Randomization Analysis. *BMC Med.* **2022**, *20*, 125, doi:10.1186/s12916-022-02322-3.
91. Masuda, T.; Ogawa, K.; Kamatani, Y.; Murakami, Y.; Kimura, T.; Okada, Y. A Mendelian Randomization Study Identified Obesity as a Causal Risk Factor of Uterine Endometrial Cancer in Japanese. *Cancer Sci.* **2020**, *111*, 4646–4651, doi:10.1111/cas.14667.
 92. Freuer, D.; Linseisen, J.; O'Mara, T.A.; Leitzmann, M.; Baurecht, H.; Baumeister, S.-E.; Meisinger, C. Body Fat Distribution and Risk of Breast, Endometrial, and Ovarian Cancer: A Two-Sample Mendelian Randomization Study. *Cancers* **2021**, *13*, doi:10.3390/cancers13205053.
 93. Kennedy, O.J.; Bafilogil, C.; O'Mara, T.A.; Wang, X.; Evans, D.G.; Kar, S.; Crosbie, E.J. Child and Adult Adiposity and Subtype-Specific Endometrial Cancer Risk: A Multivariable Mendelian Randomisation Study. *Int. J. Obes.* **2023**, *47*, 87–90, doi:10.1038/s41366-022-01231-y.
 94. Mariosa, D.; Smith-Byrne, K.; Richardson, T.G.; Ferrari, P.; Gunter, M.J.; Papadimitriou, N.; Murphy, N.; Christakoudi, S.; Tsilidis, K.K.; Riboli, E.; et al. Body Size at Different Ages and Risk of 6 Cancers: A Mendelian Randomization and Prospective Cohort Study. *J. Natl. Cancer Inst.* **2022**, *114*, 1296–1300, doi:10.1093/jnci/djac061.
 95. de Leeuw, C.; Savage, J.; Bucur, I.G.; Heskes, T.; Posthuma, D. Understanding the Assumptions Underlying Mendelian Randomization. *Eur. J. Hum. Genet.* **2022**, *30*, 653–660, doi:10.1038/s41431-022-01038-5.
 96. Bowden, J.; Del Greco M, F.; Minelli, C.; Davey Smith, G.; Sheehan, N.; Thompson, J. A Framework for the Investigation of Pleiotropy in Two-Sample Summary Data Mendelian Randomization. *Stat. Med.* **2017**, *36*, 1783–1802, doi:10.1002/sim.7221.
 97. Verbanck, M.; Chen, C.-Y.; Neale, B.; Do, R. Detection of Widespread Horizontal Pleiotropy in Causal Relationships Inferred from Mendelian Randomization between Complex Traits and Diseases. *Nat. Genet.* **2018**, *50*, 693–698, doi:10.1038/s41588-018-0099-7.
 98. Zheng, J.; Baird, D.; Borges, M.-C.; Bowden, J.; Hemani, G.; Haycock, P.; Evans, D.M.; Smith, G.D. Recent Developments in Mendelian Randomization Studies. *Curr Epidemiol Rep* **2017**, *4*, 330–345, doi:10.1007/s40471-017-0128-6.
 99. Smith, G.D.; Ebrahim, S. "Mendelian Randomization": Can Genetic Epidemiology Contribute to Understanding Environmental Determinants of Disease? *Int. J. Epidemiol.* **2003**, *32*, 1–22, doi:10.1093/ije/dyg070.
 100. Debat, V.; David, P. Mapping Phenotypes: Canalization, Plasticity and Developmental Stability. *Trends Ecol. Evol.* **2001**, *16*, 555–561, doi:10.1016/S0169-5347(01)02266-2.
 101. Hughes, R.A.; Davies, N.M.; Davey Smith, G.; Tilling, K. Selection Bias When Estimating Average Treatment Effects Using One-Sample Instrumental Variable Analysis. *Epidemiology* **2019**, *30*, 350–357, doi:10.1097/EDE.0000000000000972.
 102. Griffith, G.J.; Morris, T.T.; Tudball, M.J.; Herbert, A.; Mancano, G.; Pike, L.; Sharp, G.C.; Sterne, J.; Palmer, T.M.; Davey Smith, G.; et al. Collider Bias Undermines Our Understanding of COVID-19 Disease Risk and Severity. *Nat. Commun.* **2020**, *11*, 5749, doi:10.1038/s41467-020-19478-2.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.