A Mendelian randomization study of insulin therapy for type 1 diabetes increasing the potential risk of ovarian cancer

Xue Bai ¹ and Ling Zhang²

Ji Lin University

Abstract

Background: Type 1 diabetes (T1D) has been associated with a higher risk of Ovarian cancer (OC), albeit the mechanisms underlying this association remain elusive. A better understanding of the relationship between T1D and OC may contribute to improved primary prevention of OC. We aimed to investigate the putative causal role of T1D on OC, and to identify the potentially mediating effects of the usage of insulin product underlying this relationship.

Methods: We performed a two-sample Mendelian randomization (MR) analysis using genetic variants associated with T1D and OC from genome-wide association studies. Then, a multivariable MR analysis was conducted to investigate whether T1DM has an independent effect on OC after adjusting for potential confounders. Finally, the mediating role of insulin product was subsequently explored using mediation analysis via two-step MR.

Results: the MR estimated based on IVW method indicated a causal association between genetically determined T1D and Ovarian cancer (OC) (OR: 1.0006, 95% CI 1.0001–1.0011; P = 0.0164). After adjusting for body
mass index, Smoking, physical activity, age at menopause and age at menarche, respectively. We found that a causal relationship between T1DM and OC was still statistically significant (OR > 1, P < 0.05). The two-step MR analysis revealed that insulin product acted as a mediating moderator between the T1D and OC (mediated proportion, 1.07%).

**Conclusions:** Our findings suggest that T1D may confer a risk effect to OC, mediated in part by therapeutic insulin product. Therefore, precise dosage of insulin product or an alternative to insulin in T1D patients have a profound significance in terms of the prevention of OC.

**Keywords:** Type 1 diabetes (T1D), Insulin product, Ovarian cancer (OC), Mendelian randomization (MR)

**1. Introduction**

Ovarian cancer (OC) is defined as an extremely threatening gynecological malignancy with crucial global impact on women's well-being and economic burden. It stands as a dominating cause of female mortality worldwide, with a miserable escalation in mortality rates[1]. The absence of early clinical feature and the dearth of sensitive and specific diagnostic indicators system lead to the diagnosis of advanced-stage disease in over 70% of ovarian cancer patients[2]. Furthermore, it is worth notice that the 5-year survival rate for individuals diagnosed with advanced-stage ovarian cancer stands at approximately 25%, an extraordinarily lower figure compared to the 5-year survival rate of individuals with early-stage ovarian
cancer, which reaches 92% [3]. Despite the advancement of screening methods and treatment modalities, survival rates for patients with Ovarian cancer have not ameliorated. Consequently, exploring the etiology and higher risk factors of ovarian cancer are of great significance for prevention and treatment of disease actively and effectively.

The pathogenesis of ovarian cancer is still a blue ocean of medical research. Reported nosogenesis include genetic factors, endocrine factors, fertility factors, immune factors, chronic inflammation, oxidative stress, and environmental factors[4][5][6]. Type 1 diabetes (T1D) is an autoimmune and metabolic disease that results in pancreatic β-cell damage, giving rise to a decrease in insulin synthesis or a complete lack of insulin secretion in patients, which contributes to elevated blood glucose[7]. Over the years, countless scientists have been devoting themselves to explore the relationship between Type 1 diabetes and ovarian cancer. There is sufficient biological evidence that overexpression of IGF-I is demonstrated in ovarian cancer[8][9]. Besides, low serum sex hormone-binding globulin[10][11], inflammatory pathways[12][13], leptin overproduction[14][15], and microRNAs[16][17][18][19][20], may explain part of the mechanisms associating Type 1 Diabetes to Ovarian cancer. Epidemiologic evidence also suggests that people with Type 1 diabetes are at significantly higher risk for many forms of cancer[21][22][23][24][25]. Although there have been basic experimental
researches and epidemiological investigation [8-24], regarding the association between Type 1 diabetes and ovarian cancer, there are no randomized clinical trials. Meanwhile, many studies have a great deal of confounding factors, with a high degree of heterogeneity and a deficiency of causality inference. However, due to the profound impact of Ovarian cancer on women's health and quality of life, it is important to explore the causal association between Type 1 diabetes and ovarian cancer. When a causal relationship between Type 1 diabetes and ovarian cancer is established, it will provide a new clinical strategy for the management of ovarian cancer.

Traditional observational studies may cause deflection and even misjudgment of the research results due to various observable and unobservable confounding factors, reverse causality and bias; meanwhile, they principally focus on the correlation between exposure factors and outcomes instead of the actual causality. Mendelian randomization (MR) is a new epidemiological method that imitates the design of randomized controlled studies[26]. It uses single nucleotide polymorphisms (SNPs) as instrumental variables (IVs) to infer causal relationship between the risk factors and interested outcomes. SNP is randomly assigned to individuals with gametes during meiosis [27], which is similar to the essential demands of randomized controlled trials. Meanwhile, genovariation precedes the occurrence of diseases, which avoids the potential influence of reverse
causality. Consequently, MR is an ideal way to explore the causal relationship between Type 1 diabetes and Ovarian cancer.

This study used a two-sample MR and Mediated Mendelian randomization analyses design to investigate whether Type 1 diabetes has a causal association with Ovarian cancer and estimate the mediating roles between them to provide scientific evidence for Ovarian cancer prevention and treatment.

2. Materials and methods

2.1 Experimental design

The current investigation selected Type 1 diabetes (T1D) as the exposure factor and extracted instrumental variables (IVs) from the exposure dataset in the shape of single nucleotide polymorphisms (SNPs) that revealed conspicuous associations with T1D. The outcome variable of interest was ovarian cancer (OC). MR analysis was conducted to investigate the causal relationship between the exposure and outcome variables. To evaluate heterogeneity and dispose potential issues of horizontal pleiotropy, Cochran’s Q test, MR Presso’s test and MR Egger’s test were preferred. Furthermore, sensitivity analysis was implemented to ascertain the robustness and reliability of the expected results [28]. The reliability of MR analysis is predicated on the implementation of three conditional assumptions [29]: (i) There exists a firm correlation between instrumental variables and exposure factors. (ii) No confounding factors
are existent that might affect the relationship between exposure and outcome variables, especially, no genetic polymorphisms.

(iii) Instrumental variables do not exert a direct influence on the outcome, but exclusively effect the outcome via the exposure factors. Refer to Figs 1 for a visual depiction of this assumptions.

2.2 Data sources

All data were independently acquired from the IEU GWAS database (https://gwas.mrcieu.ac.uk/terms/), founded by the University of Bristol, UK. Detailed information is provided in Table 1.
Table 1

An overview of the GWAS datasets used in the MR analysis

<table>
<thead>
<tr>
<th>Contribution</th>
<th>Traits</th>
<th>Sample size</th>
<th>Number of SNPs</th>
<th>Author</th>
<th>Population</th>
<th>GWAS ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposure</td>
<td>Type 1 diabetes</td>
<td>457,695</td>
<td>24,182,422</td>
<td>Sakaue</td>
<td>European</td>
<td>ebi-a-GCST90018925</td>
</tr>
<tr>
<td>Confounders</td>
<td>Body mass index</td>
<td>461,460</td>
<td>9,851,867</td>
<td>Ben Elsworth</td>
<td>European</td>
<td>ukb-b-19953</td>
</tr>
<tr>
<td></td>
<td>smoked</td>
<td>461,066</td>
<td>9,851,867</td>
<td>Ben Elsworth</td>
<td>European</td>
<td>ukb-b-20261</td>
</tr>
<tr>
<td></td>
<td>Age at menopause</td>
<td>143,819</td>
<td>9,851,867</td>
<td>Ben Elsworth</td>
<td>European</td>
<td>ukb-b-17422</td>
</tr>
<tr>
<td></td>
<td>Age at menarche</td>
<td>182,416</td>
<td>2,441,816</td>
<td>Perry JR</td>
<td>European</td>
<td>ieu-a-1095</td>
</tr>
<tr>
<td></td>
<td>physical activity</td>
<td>377,234</td>
<td>11,808,007</td>
<td>Klimentidis</td>
<td>European</td>
<td>ebi-a-GCST006097</td>
</tr>
<tr>
<td>Mediators</td>
<td>insulin product</td>
<td>462,933</td>
<td>9,851,867</td>
<td>Ben Elsworth</td>
<td>European</td>
<td>ukb-b-15445</td>
</tr>
<tr>
<td>Outcome</td>
<td>Ovarian cancer</td>
<td>199,741</td>
<td>9,822,229</td>
<td>Burrows</td>
<td>European</td>
<td>ieu-b-4963</td>
</tr>
</tbody>
</table>

2.3 Instrumental variables selection

Initially, SNPs revealing significant associations with exposures (selection criteria: P<5.0×10−8) were selected based on the 1,000 whole-genome European population. The linkage disequilibrium parameter (r2) was set at 0.001, and the genetic distance was founded in 10,000 kb to insure the independence of each instrumental variable (IV), excluding interference from other IVs. Subsequently, SNPs with strong linkage disequilibrium (r2 > 0.8) were applied to compensate for missing SNPs in the outcome dataset.

Finally, exposure and outcome data were combined while keeping the corresponding values for exposure and outcome with the same effect alleles[30]. The residual SNPs constitute the ultimate instrumental variables connected with the exposure. To consider the proportion of SNP phenotypic variation in the Mendelian randomization process, we calculated R2 values via using the formula $R^2 = 2 \times \beta^2 \times \text{EAF} \times (1-\text{EAF})/[2$
\[ \times \beta_2 \times \text{EAF} \times (1 - \text{EAF}) + \text{SE}_2 \times 2 \times N \times \text{EAF} \times (1 - \text{EAF}) \], where \( \beta \) is the effect estimate of the genetic variant in the exposure GWAS, EAF represents the Allele 1 frequency, SE represents the standard error, and N represents the sample size[31][32]. And then, the instrumental strength of SNPs for each socioeconomic trait was calculated using the F-statistic, evaluated as \( F = \left[ \frac{(N-k-1)}{k} \right] \times \left[ \frac{R^2}{1-R^2} \right] \), where N is the sample size, k represents the total number of SNPs, and R2 represents the total proportion of phenotypic variation interpreted by all the SNPs in the MR analysis. An F-statistic greater than 10 manifests that the combined SNPs serve as sufficiently strong instruments to illuminate phenotypic variation, while an F-statistic of 10 or less elucidates a weak instrument[32].

2.4 Mendelian randomization

MR analysis is a advantageous tool in epidemiological studies. The study began with a two-sample Mendelian randomization employing T1D and confounders (including body mass index, Smoke, physical activity, age at menopause, age at menarche) as exposures and OC as an outcome indicator severally. As T1D and confounders are correlated in clinical studies, multivariate Mendelian randomization was performed to correct for the results. Inverse variance weighted (IVW), weighted median estimator (WME) and MR-Egger regression were the main methods in Mendelian randomization.

2.5 Mediated Mendelian randomization
To explore the role of insulin product in the relationship between T1D and OC, we used a two-step MR approach, as shown in Fig. 2. The two-step approach is considered less tendency to the biases inherent in common multivariate methods [33]. In MVMR, the total effect of each exposure is accounted for by direct effect and an indirect effect. There is a mediator if the following conditions were met: 1) there was a correlation between T1D and the mediator (β1); 2) the mediator was correlated with OC (β2). 3) T1D was closely related to OC but not adjusted for the mediator (β3’). The mediation ratio was calculated as (β1 x β2)/(β3’). β1 x β2 is the indirect effect and the total effect (β3) is β3’ + β1 x β2.

Fig. 2 A diagram of mediator MR analysis. This diagram illustrates the potential causal pathway from T1D to OC, with Insulin Product serving as mediators.

2.6 Sensitivity analysis
This study implemented a 'leave-one-out' sensitivity analysis, whereby individual SNPs were removed one at a time, to assess whether the observed variation was dominant the association between the exposure and
outcome variables. Secondly, to confirm the presence of horizontal pleiotropy in this MR analysis, the MR-Egger intercept test was absorbed. If the intercept term in the MR-Egger intercept analysis did not prove statistically significant, there is no horizontal pleiotropy within the study [34]. Finally, Cochran’s Q statistic was implemented in this study to ascertain heterogeneity. A significant level of heterogeneity within the analysis was manifested by a statistically significant result from the Cochran’s Q statistic test. However, when the number of SNPS is greater than 3, IVW may have certain heterogeneity tolerate.

2.7 Statistical analysis

In this study, the "TwoSampleMR" and "Mendelian randomization" [35]. Packages accessible within RStudio software were used to carry out the exposure and outcome analyses. The outcomes from the Mendelian randomization (MR) analysis were defined as beta (β) values, indicating the impact of T1D and insulin product on OC. Furthermore, the associated 95% confidence intervals (CI) were recorded for all causal estimates. A significance threshold of p < 0.05 was applied to discern statistical significance.

3 Results

3.1 Two-sample Mendelian randomization

In the IVW models, compelling evidence indicated a significant association between Type 1 Diabetes and the risk of ovarian cancer (OR:
1.0006, 95% CI 1.0001–1.0011; \(P = 0.0164\)). Additionally, we assessed the association of the 17-SNP instrument with OC using MR-Egger, weighted median. MR-Egger regression showed an OR of 1.0009 (95% CI 1.0001–1.0011; \(P = 0.0492\)), while weighted median (OR 1.0009, 95% CI 1.0002–1.0016; \(P = 0.0087\)) techniques provided directionally consistent results. These findings suggest that the results remain robust even in the face of potential violations of MR assumptions. Furthermore, no evidence of pleiotropy or heterogeneity was detected in the MR-PRESSO global test, MR-Egger intercept test, and Cochran’s Q test. The data is summarized in Supplementary Tables S1-2. To obtain MR estimates for each individual SNP, we conducted the analysis multiple times for each exposure-outcome combination. In each iteration, a different single SNP was used for the analysis. The scatterplots depicted in Fig.3 illustrate that SNPs exhibiting a larger effect on Type 1 Diabetes also exert a greater impact on the risk of ovarian cancer. Each method is expressed by a different colored line, with the slope of the line manifesting the estimated causal effect. A leave-one-out analysis is feasible, wherein the MR analysis is repeated while excluding each SNP independently. This allows us to explicit that a majority of the associated signals were not primarily impacted by a single genetic marker, as is demonstrated in Fig.4.
Table 2

Two-sample MR analysis results under different methods

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Outcome</th>
<th>Method</th>
<th>nSNPs</th>
<th>Beta</th>
<th>SE</th>
<th>P-value</th>
<th>OR</th>
<th>OR_lci95</th>
<th>OR_Uci95</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1D</td>
<td>OC</td>
<td>Inverse variance weighted</td>
<td>17</td>
<td>0.0006</td>
<td>0.0003</td>
<td>0.0164</td>
<td>1.0006</td>
<td>1.0001</td>
<td>1.0011</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weighted median</td>
<td>17</td>
<td>0.0009</td>
<td>0.0003</td>
<td>0.0087</td>
<td>1.0009</td>
<td>1.0002</td>
<td>1.0016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MR Egger</td>
<td>17</td>
<td>0.0009</td>
<td>0.0004</td>
<td>0.0492</td>
<td>1.0009</td>
<td>1.0001</td>
<td>1.0018</td>
</tr>
</tbody>
</table>

SNP, single nucleotide polymorphism; SE, standard error; OR, odds ratio; CI, confidence interval.

Fig 3. Scatter plot of single SNP with T1D as the exposure and OC as the outcome.
3.2 Multivariate Mendelian randomization

To adjust for potential pleiotropic pathways that might confound the relationship between T1D and OC, an MVMR model was utilized. In this model, the combined effect of T1D on body mass index, Smoking, physical activity, age at menopause and age at menarche respectively was used as an exposure for OC outcomes. The results of the IVW analysis revealed that the previously significant correlation between T1D and OC was not attenuated in the MVMR model and was still statistically significant, detailed results are presented in Table 3. Furthermore, there is no pleiotropy or heterogeneity in the MR-PRESSO global test, MR-Egger intercept test, and Cochran’s Q test (Supplementary Tables S1-2).
Table 3.

Forest plot of the MVMR analysis. The direct causal effect of T1D on OC risk was determined by adjusting for traits such as Body mass index, Physical activity, Smoked, Age at menarche, Age at menopause.

<table>
<thead>
<tr>
<th>Exposure</th>
<th>MVMR</th>
<th>Outcome</th>
<th>nsnp</th>
<th>method</th>
<th>OR(95%CI)</th>
<th>Pval</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1D</td>
<td></td>
<td>OC</td>
<td>17</td>
<td>IVW</td>
<td>1.00062(1.0001to1.001)</td>
<td>0.016</td>
</tr>
<tr>
<td>Adjusting Body mass index</td>
<td>7</td>
<td>IVW</td>
<td>1.00071(1.0002to1.001)</td>
<td>0.012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusting Physical activity</td>
<td>13</td>
<td>IVW</td>
<td>1.00072(1.0002to1.001)</td>
<td>0.011</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusting Smoked</td>
<td>14</td>
<td>IVW</td>
<td>1.00070(1.0002to1.001)</td>
<td>0.009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusting age at menarche</td>
<td>9</td>
<td>IVW</td>
<td>1.00087(1.0003to1.001)</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusting age at menopause</td>
<td>11</td>
<td>IVW</td>
<td>1.00066(1.0001to1.001)</td>
<td>0.024</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.3 Mediated Mendelian randomization

In order to investigate the mediating role of insulin product in the relationship between T1D and OC, a mediated MR analysis was conducted. The results of this study are detailed in Table 4. Our study revealed that insulin product serves as a mediator in the relationship between T1D and OC, with the mediating effect estimated to be ($\beta = 6.63 \times 10^{-4}$, 95% CI $2.10 \times 10^{-4}$ to $1.12 \times 10^{-3}$, $P = 0.04$).
Table 4.
The mediation effect of mediators on the causal effect of T1D and OC

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Mediators</th>
<th>Outcome</th>
<th>Direct effect A (β1, 95% CI)</th>
<th>Direct effect B (β2, 95% CI)</th>
<th>Mediation effect (β, 95% CI)</th>
<th>P</th>
<th>Mediate proportion (β, 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1D</td>
<td>Insulin product</td>
<td>OC</td>
<td>8.66e-03 (7.5)</td>
<td>7.72e-02 (2.5)</td>
<td>6.63e-04 (2.1)</td>
<td>0.004</td>
<td>1.067 (0.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2e-03 (981e-03)</td>
<td>5e-02 (0.013)</td>
<td>e04,112e (0.0)</td>
<td></td>
<td>38,1794</td>
</tr>
</tbody>
</table>

4. Discussion

Ovarian cancer and Type 1 diabetes bring tremendous security challenges to public health. The relationship between T1D and OC remains a subject of debate, with the mechanism of this association not fully revealed. In this study, we utilized GWAS database and employed two-sample MR to examine the causal relationship between T1D and OC. The findings indicate that T1D may potentially accelerate the progression of OC. The interconnectedness between T1D and OC was mediated by insulin product. Insulin product emerges as a risk factor for OC, underlining the precise control of insulin medicine to mitigate the risk of OC.

In recent decades, several observational studies have sought the link between T1D and OC. Nonetheless, since there is inherent nature of recall and select bias in observational study, certain biases might contribute to inaccurate reporting of causal relationship[36]. Several cohorts [37][38][39][40]and case–control [41] studies have been reported that a history of T1D is associated with an augmented risk of OC, however, other
relevant studies found a negative significant association [42][43][44][45]. Besides, a subgroup meta-analysis based on DM types indicated that the risk of OC in T1DM group (44%) is higher than in T2DM group (17%) [46]. However, MR analysis overcomes the limitations of observational studies, which unambiguously interpret the causal relationship between T1D and OC. Moreover, the positive link was even more prominent in studies that adjusted for confounding factors (such as, body mass index, smoking, physical activity, age at menopause and age at menarche) than these for unadjusted analysis.

The underlying carcinogenesis effect of T1D to OC was not completely revealed at present, but several plausible mechanisms have been used to explain the links between them. Previous studies have shown that the neoplastic process has been considered to influenced by T1D through these mechanisms, mainly including hyperglycemia, adipokine imbalances, and chronic inflammation[47][48][49]. Because of a prolonged exposure to inflammation and hyperglycemic condition, the repair cycles which are associated with incessant ovulation process could be slowed down, thus, resulting in an underlying risk of OC[50]. In addition, hyperglycemia may lead to endothelial dysfunction, endothelial cell death and aberrant neoangiogenesis, ultimately, contributing to cancer progression[51][52]. Studies have shown that adipose metabolic dysregulation as an hallmark of T1D, can result in high levels of inflammatory cytokines, including IL-6.
and TNF-α [49] These cytokines stimulate cell proliferation, invasion, and evasion of antitumor immunity pathways [13].

Mediated Mendelian randomization was employed to delve into the potential mechanism underlying this relationship. We find that therapeutic insulin product as a mediator could explain parts of this relationship. This is consistent with the results of some observational studies. Since 1972 in the animal experimental researches[53][54], it has been recognized that the insulin has promoted cancer growth due to the mitogenic effect of insulin by alteration in the PI3K/AKT pathway and mitotic kinase pathway [55].

Type 1 diabetes patients have widely used insulin replacement therapy to strictly control the metabolic status and prevent the progression of long-term complications associated with persistent hyperglycemia [56]. Although there are structural similarities between exogenous and endogenous insulin, the insulin replacement therapy is non-physiologic, and thus, a supraphysiologic dose of insulin needs to be injected in order to achieve the Physiological needs. This contributes to an inevitable therapeutic increase in the dose of insulin therapy, leading to exogenous systemic hyperinsulinemia [57]. Hyperinsulinemia increases the levels of estradiol and testosterone, which may result in cancer initiation and progression in patients [58][59].

The characteristic of our study is consistent with all MR assumptions.

We integrated data from GWAS of both Type 1 diabetes and ovarian cancer.
This robust methodology, coupled with extensive data sources, supports our causal inference by alleviating biases stemming from confounding and reverse causation. Furthermore, our findings provide a reliable theoretical framework for future investigations focused on improving therapeutic approaches for Type 1 diabetes, such as insulin preparation replacement reagent. This potential improvement may reduce the risk of ovarian cancer, carrying an important directive function for clinical management. However, there are several limitations in our research. On one hand, this investigation is the lack of comprehensive GWAS in non-European ancestries. Therefore, the assessment of these causal pathways in diverse ethnic groups should be paid to in the future studies. On the other hand, the utilization of pooled data from genome-wide association studies was short of individual-specific information, eliminating subgroup analyses for variables such as age, disease duration, treatment method, disease classification, and so on. This limitation makes it impossible to compare potential differences in causal effects among these subgroups.

5. Conclusion
In conclusion, our study provides evidence supporting a causal relationship between T1D and OC. Furthermore, therapeutic insulin product mediated the effect between T1D and OC. Therefore, precise dosage of insulin product or an alternative to insulin in T1D patients have a profound significance in terms of the prevention of OC.
References

in ovarian cancer cells via AKT activation. Arch Gynecol Obstet. 2015;292:1109-16.


[57] Shuwen Chen, Zaixin Guo, and Qi Yu. Genetic evidence for the causal association between type 1 diabetes and the risk of polycystic ovary syndrome; Hum Genomics. 2023; 17: 100.
