Impact of glycemic traits, type 2 diabetes and metformin use on breast and prostate cancer risk: a Mendelian randomization study

Shiu Lun Au Yeung, Catherine Mary Schooling

ABSTRACT

Objectives Observational studies suggest glycemic traits and type 2 diabetes are positively associated, and metformin inversely associated with breast and prostate cancer risk. However, observational studies are susceptible to unmeasured confounding while studies of metformin use are also vulnerable to immortal time bias. The use of Mendelian randomization may reduce confounding due to random allocation of relevant genetic markers at birth, and may reduce immortal time bias (for metformin-related variants analysis) since the start of exposure is at birth.

Research design and methods We identified strong genetic predictors of fasting glucose, glycated hemoglobin, and type 2 diabetes from the Meta-Analyses of Glucose and Insulin-related traits Consortium and Diabetes Genetics Replication And Meta-analysis Consortium (n=140,595 for glucose; n=123,665 for HbA1c; n=898,130 for type 2 diabetes) and of AMPK-instrumented HbA1c reduction as a proxy of metformin and applied them to large genome-wide association studies of breast cancer (Breast Cancer Association Consortium; BCAC) and prostate cancer (Prostate Cancer Association Group to Investigate Cancer-Associated Alterations in the Genome; PRACTICAL). We used inverse variance weighting to obtain estimates. Sensitivity analyses included use of MR-Egger, weighted median, exclusion of pleiotropic instruments, and validation using UK Biobank (breast cancer only).

Results There was no association of fasting glucose (OR 1.03 per mmol/L, 95% CI 0.85 to 1.25), HbA1c (OR 1.02 per %, 95% CI 0.73 to 1.45), or type 2 diabetes (OR 0.98 per log odds, 95% CI 0.95 to 1.01) with breast cancer in BCAC, with similar findings from UK Biobank. There was no association of fasting glucose (OR 0.93 per mmol/L, 95% CI 0.73 to 1.17), HbA1c (OR 0.90 per %, 95% CI 0.58 to 1.40) or type 2 diabetes (OR 1.02 per log odds, 95% CI 0.97 to 1.07) with prostate cancer in PRACTICAL. No strong evidence was observed for AMPK-instrumented HbA1c reduction on cancer risk.

Conclusion Glycemic traits and type 2 diabetes unlikely cause breast and prostate cancer. Whether metformin can be repurposed for cancer prevention remains unclear.

INTRODUCTION

Previous studies have suggested a potential relation of glycemic traits with cancer, particularly hormone-related cancers, such as breast and prostate cancer. However, the positive associations are not always consistently observed, and hence may be indicative of potential biases or residual confounding in previous studies. Studies on metformin, the first line medication for type 2 diabetes which reduces hepatic glucose production, have also shown a potential inverse relation with cancers, such as breast and prostate cancer, although these studies could be vulnerable to immortal time bias due to misallocation of exposure time. Some studies which corrected for immortal time bias showed no relation of metformin with cancer risk, but these studies remain open to confounding and selection bias. Others argued the approach to correct for immortal...
time bias remains controversial and requires additional validations.\textsuperscript{10} Potentially stronger study designs, such as the marginal structural models applied to electronic health records with inverse probability weighting,\textsuperscript{11} and Mendelian randomization using biomarkers of metformin\textsuperscript{12} did not provide strong evidence for a causal role of metformin in breast cancer, with similar findings for prostate cancer in a recent meta-analysis.\textsuperscript{13} Given the relation between antidiabetes medications and breast cancer risk appears to vary across classes of medication with different modes of actions (eg, metformin vs sulfonylurea),\textsuperscript{14} this may imply glycemic traits per se are not the main factors accounting for any reduction in cancer risk among metformin users. However, whether glycemic traits are a potential target of intervention to reduce these cancers is unclear.

Mendelian randomization may provide more credible evidence concerning the role of glycemic traits in breast and prostate cancer. Mendelian randomization uses the random allocation of genotype during conception to infer the causal relations, and hence reduces confounding.\textsuperscript{15} A previous Mendelian randomization study suggested a potentially positive relation of 2-hour glucose and fasting insulin with breast cancer although the study may be biased by the choice of instruments.\textsuperscript{16,17} No Mendelian randomization studies of glycemic indicators have considered prostate cancer as the outcome. We conducted a Mendelian randomization study to address the causal role of glycemic traits (ie, fasting glucose and glycated hemoglobin (HbA1c)) and type 2 diabetes in breast and prostate cancer using relevant genetic consortia and UK Biobank. Lastly, to further address the potential for repurposing metformin to reduce risk of these cancers, we chose the genetic variants functionally relevant to metformin, that is, genes related to AMP-activated protein kinase (AMPK),\textsuperscript{18} in a separate analysis. The advantage of using genetics is it may circumvent immortal time bias given the start of higher/lower exposure is at birth for all participants.

METHODS

Study design
This is a two-sample Mendelian randomization design using summary statistics from genome-wide association studies (GWAS) and large biobanks.\textsuperscript{19} The assumptions of the Mendelian randomization design include: (1) instruments are related to exposures of interest; (2) instruments are not related to the confounders of the exposure-outcome relation; (3) instruments’ links with the outcome are only via the exposure of interest.\textsuperscript{19} Online supplementary figure 1A shows the framework for a Mendelian randomization design.

Genetic predictors of glycemic traits (fasting glucose, HbA1c) and type 2 diabetes

Fasting glucose and HbA1c
We extracted genetic instruments for fasting glucose and HbA1c from the Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC), which combined several GWAS to identify genetic loci related to glycemic traits.

For fasting glucose, we used the most recent GWAS, which included up to 140,595 participants (52% women). In brief, participants were excluded from the GWAS if they had physician-diagnosed diabetes, were on diabetes treatment or had fasting plasma glucose >7 mmol/L. Imputation on up to 78,739 participants was based on the HapMap II CEU reference panel while direct genotyping was performed using iSELECT Metabochip array (up to 61,422) and custom variant sets (up to 20,531). We only used the female-specific estimates (n=73,089) for the breast cancer, and the male-specific estimates (n=67,506) for the prostate cancer. The GWAS were adjusted for age, study site, and geographic covariates, and corrected for genomic control. Fasting glucose (mmol/L) was untransformed. Single-nucleotide polymorphisms (SNP) were not in high linkage disequilibrium (r²<0.001).

For HbA1c, we used the most recent GWAS of HbA1c, which included up to 159,940 participants without diabetes, imputed using the International HapMap Project Phase 2 reference panel.\textsuperscript{20} The GWAS was adjusted for age, sex, and study-specific covariates and corrected for genomic control. HbA1c (%) was untransformed. We only included loci that reached genome-wide significance among participants of European descent (up to 123,665), and were not in high linkage disequilibrium (r²<0.001). We also excluded rs1800562 which violated Hardy Weinberg equilibrium (p=0.0007) as previously.\textsuperscript{21,22}

Type 2 diabetes
We used the most recent GWAS of type 2 diabetes, which included up to 898,130 participants of European descent (74,124 cases and 82,4006 controls) from the DIAbetics Genetics Replication And Meta-analysis consortium.\textsuperscript{23} Imputation was performed using HaploType Reference Consortium (HRC) reference panel except deCODE GWAS, which was imputed using a population-specific reference panel. Type 2 diabetes was defined in various ways, including diagnostic fasting glucose ≥7.0 mmol/L, previous type 2 diabetes diagnosis, the use of antidiabetic medication, or from electronic health records. The genetic associations, apart from adjusting for study-specific covariates, controlled for population stratification and relatedness via adjustment for principal components after exclusion of related individuals or the use of a mixed model with random effects for kinship from a genetic relation matrix, with genomic control applied to association summary statistics. To avoid selection bias from stratiﬁying on a collider which may invalidate estimates, we used loci from the analyses without adjustment for body mass index. SNPs were not in high linkage disequilibrium (r²<0.001).
Genetic associations with breast cancer from the Breast Cancer Association Consortium

We used summary genetic associations from a large GWAS of breast cancer among women of European ancestry from Breast Cancer Association Consortium (BCAC; 122,977 cases and 105,974 controls).\(^2^4\) Imputation was based on the 1000 Genomes Project Phase 3 reference panel. The genetic analysis adjusted for study-specific covariates and up to 10 principal components to control for population stratification. To explore potential differences by breast cancer subtypes, we also examined the relation by estrogen receptor status (estrogen receptor positive and estrogen receptor negative).

Genetic associations with prostate cancer from the Prostate Cancer Association Group to Investigate Cancer-Associated Alterations in the Genome consortium

We used summary genetic associations from a large GWAS of prostate cancer from Prostate Cancer Association Group to Investigate Cancer-Associated Alterations in the Genome (PRACTICAL; 79,148 cases and 61,106 controls) of European descent,\(^2^5\) where majority of the cases were aged 60+ years. Imputation was based on 1000 Genomes Project Phase 3 reference panel. The OncoArray and each GWAS were adjusted for principal components and study-relevant covariates, while the overall analysis only adjusted for principal components.

Genetic variants functionally relevant to metformin

Although the mechanisms by which metformin confers beneficial effects on glucose metabolism remains unclear, \(\text{AMPK}\) likely has a role.\(^1^8\) We chose genetic variants covering the subunits of \(\text{AMPK}\), including \(\text{PRKAA1}, \text{PRKAA2}, \text{PRKAB1}, \text{PRKAB2}, \text{PRKAG1}, \text{PRKAG2},\) and \(\text{PRKAG3}\).\(^2^6\) Specifically, we included all SNPs in relevant gene regions, as identified by the corresponding positions (hg19 build) based on PubMed (https://www.ncbi.nlm.nih.gov/gene/). Following a previous study,\(^2^7\) we checked the validity of these genetic predictors by (1) checking each SNP’s relation with HbA1c (p<0.05) using MAGIC,\(^2^8\) and (2) checking whether the SNP relates to relevant gene expression using the Genotype-Tissue Expression (GTex) project data (https://gtexportal.org/home/, version 7), which gives expression quantitative trait loci analyses for 48 tissues among 620 donors (p<0.05).\(^2^9\) Specifically, we checked if the expression increasing allele (association with the lowest p value if expression was reported across tissues) corresponded to the HbA1c decreasing allele in MAGIC. Online supplementary figure 1B shows the framework of a Mendelian randomization design, with respect to exploring the impact of metformin using \(\text{AMPK}\) variants.

Outcomes

The outcomes were breast and prostate cancer risk.

Statistical analyses for glycemic traits and type 2 diabetes

We approximated the \(F\) statistic of each instrument for fasting glucose, HbA1c and type 2 diabetes to indicate instrument strength, that is, the square of the gene–exposure association divided by the square of the corresponding SE of the association,\(^2^9\) where higher \(F\) statistic indicates lower risk of weak instrument bias. We used inverse variance weighting (IVW) with multiplicative random effects to meta-analyze the Mendelian randomization estimates for each SNP, so as to derive an overall causal estimate for glycemic traits and type 2 diabetes on breast and prostate cancer. We used \(I^2\) to assess the consistency of the Mendelian randomization estimates, where high \(I^2\) may indicate the presence of invalid instruments. Given IVW assumes balanced pleiotropy and may be invalid in the presence of invalid instruments, we also conducted additional sensitivity analyses which rely on other assumptions.\(^3^0\) Consistency across sensitivity analyses further strengthens the estimates,\(^3^1\) an approach used in previous studies.\(^3^2\)\(^3^3\)

MR-Egger regression

MR-Egger regression is similar to IVW, except the intercept is not fixed to zero, and hence allows for horizontal pleiotropy. The intercept also captures the degree of horizontal pleiotropy, and a p<0.05 indicates the presence of such pleiotropy, which in turn suggests the IVW estimate may be invalid.\(^3^4\) MR-Egger requires the Instrument Strength Independent of Direct Effect assumption and has lower statistical power.

Weighted median

The weighted median produces valid estimates as long as 50% of the information is derived from valid instruments.\(^3^5\)

Exclusion of pleiotropic instruments

We searched for pleiotropic effects of the instruments via PhenoScanner (but excluding UK Biobank).\(^3^6\) which may also be a cause of breast cancer, such as sex hormone-binding globulin,\(^3^7\) estrogen and pubertal stage,\(^3^8\)\(^3^9\) height,\(^4^0\) lipids,\(^4^1\) obesity,\(^4^2\) and sleep,\(^4^3\) and repeated the analyses without these instruments. For prostate cancer, we considered Microseminoprotein-beta,\(^4^4\) and testosterone and pubertal stage\(^4^5\) as potential causes. We assume these are horizontal, not vertical, pleiotropic effects.\(^4^6\)

Replication in the UK Biobank (breast cancer only)

We repeated the analyses using the UK Biobank, which had a different sampling strategy than BCAC,\(^4^4\) and hence can be used to test for robustness of results due to potential biases introduced by study design. We used the summary genetic associations from UK Biobank generated by the Scalable and Accurate Implementation of Generalized (SAIGE) mixed model, restricted to British
White of European ancestry. Imputation was based on the HRC reference panel. Breast cancer (12671 female cases and 388549 controls) was defined using the corresponding PheWAS code (174 for breast cancer), which corresponds to the relevant International Classification of Diseases codes. SAIGE controls for sample relatedness. Mediation analysis was adjusted for sex, birth year, and the first four principal components.

Statistical analyses for AMPK-instrumented HbA1c to mimic metformin use

We used IVW with fixed effects to meta-analyze the Mendelian randomization estimates for each SNP, so as to derive an overall causal estimate for AMPK-instrumented HbA1c reduction, as proxy of metformin use, on breast and prostate cancer. Given these SNPs were correlated, we used a correlation matrix, obtained from 1000 Genomes via MR Base in the IVW analysis, as per our previous study.

All analyses were performed using R V.3.5.2 (R Development Core Team, Vienna, Austria) using R packages ‘TwoSampleMR’ and ‘MendelianRandomization’.46 47

RESULTS

The genetic instruments (up to 21 for fasting glucose, up to 34 for HbA1c, up to 159 for type 2 diabetes) used in this study are shown in online supplementary appendices 1–3. All SNP-specific F statistics for all exposures were >10, indicating that weak instrument bias was unlikely. The mean F statistics for fasting glucose, HbA1c, and type 2 diabetes were 98.7, 75.6, and 84.7, respectively, in the breast cancer analyses. The mean F statistics for fasting glucose, HbA1c, and type 2 diabetes were 79.8, 60.5, and 72.4, respectively, in the prostate cancer analyses. Based on PhenoScanner, the majority of the pleiotropic effects for breast cancer were adiposity-related traits such as fat and muscle mass. For prostate cancer, only rs1260326 and rs780093 were considered as pleiotropic because these SNPs relate to testosterone. Among the 428 SNPs in the genes related to AMPK, 49 were associated with HbA1c (p<0.05), with 43 SNPs within the PRKAG2 region. Further screening based on the GTEx suggested 27 SNPs were related to the gene expression. We removed an additional five SNPs because the expression-increasing allele did not correspond to HbA1c-reducing allele. Nine SNPs were retained after excluding highly correlated SNPs (r2≥0.8) (online supplementary appendix 4).

Table 1 shows that there was no association of fasting glucose (OR 1.03 per mmol/L, 95% CI 0.85 to 1.25), HbA1c (OR 1.02 per %, 95% CI 0.73 to 1.45), or type 2 diabetes (OR 0.98 per log odds, 95% CI 0.95 to 1.01) with overall breast cancer in BCAC. Table 2 shows glycemic traits and type 2 diabetes were similarly not associated with estrogen receptor positive or negative breast cancer risk. Exclusion of potential pleiotropic SNPs did not change the conclusion. Table 3 shows that glycemic traits and type 2 diabetes were not associated with overall breast cancer risk in UK Biobank, with similar associations with or without pleiotropic SNPs. Exclusion of pleiotropic SNPs generally reduced I2 (tables 1–3).

Table 4 shows that there was no association of fasting glucose (OR 0.93 per mmol/L, 95% CI 0.73 to 1.17), HbA1c (OR 0.90 per %, 95% CI 0.58 to 1.40), or type 2 diabetes (OR 1.02 per log odds, 95% CI 0.97 to 1.07) with prostate cancer in PRACTICAL, with similar observations in the sensitivity analyses. Exclusion of pleiotropic SNPs did not change the conclusion. Exclusion of pleiotropic SNPs generally reduced I2.

Table 5 shows that AMPK-instrumented HbA1c reduction was not associated with breast cancer in BCAC or UK Biobank, nor with prostate cancer in PRACTICAL although the estimates had wide CIs.

DISCUSSION

This Mendelian randomization study did not support a causal role of fasting glucose, HbA1c, or type 2 diabetes, with or without pleiotropic associations, on breast or prostate cancer risk. The estimated I2 was 22% for glycemic traits and 50% for type 2 diabetes. The mean F statistics for all exposures were >10, indicating that weak instrument bias was unlikely. The mean F statistics for fasting glucose, HbA1c, and type 2 diabetes were 98.7, 75.6, and 84.7, respectively, in the breast cancer analyses. The mean F statistics for fasting glucose, HbA1c, and type 2 diabetes were 79.8, 60.5, and 72.4, respectively, in the prostate cancer analyses. Based on PhenoScanner, the majority of the pleiotropic effects for breast cancer were adiposity-related traits such as fat and muscle mass. For prostate cancer, only rs1260326 and rs780093 were considered as pleiotropic because these SNPs relate to testosterone. Among the 428 SNPs in the genes related to AMPK, 49 were associated with HbA1c (p<0.05), with 43 SNPs within the PRKAG2 region. Further screening based on the GTEx suggested 27 SNPs were related to the gene expression. We removed an additional five SNPs because the expression-increasing allele did not correspond to HbA1c-reducing allele. Nine SNPs were retained after excluding highly correlated SNPs (r2≥0.8) (online supplementary appendix 4).

Table 1 shows that there was no association of fasting glucose (OR 1.03 per mmol/L, 95% CI 0.85 to 1.25), HbA1c (OR 1.02 per %, 95% CI 0.73 to 1.45), or type 2 diabetes (OR 0.98 per log odds, 95% CI 0.95 to 1.01) with overall breast cancer in BCAC. Table 2 shows glycemic traits and type 2 diabetes were similarly not associated with estrogen receptor positive or negative breast cancer risk. Exclusion of potential pleiotropic SNPs did not change the conclusion. Table 3 shows that glycemic traits and type 2 diabetes were not associated with overall breast cancer risk in UK Biobank, with similar associations with or without pleiotropic SNPs. Exclusion of pleiotropic SNPs generally reduced I2 (tables 1–3).

Table 4 shows that there was no association of fasting glucose (OR 0.93 per mmol/L, 95% CI 0.73 to 1.17), HbA1c (OR 0.90 per %, 95% CI 0.58 to 1.40), or type 2 diabetes (OR 1.02 per log odds, 95% CI 0.97 to 1.07) with prostate cancer in PRACTICAL, with similar observations in the sensitivity analyses. Exclusion of pleiotropic SNPs did not change the conclusion. Exclusion of pleiotropic SNPs generally reduced I2.

Table 5 shows that AMPK-instrumented HbA1c reduction was not associated with breast cancer in BCAC or UK Biobank, nor with prostate cancer in PRACTICAL although the estimates had wide CIs.

Table 1 The impact of glycemic traits on breast cancer risk in Breast Cancer Association Consortium using Mendelian randomization

<table>
<thead>
<tr>
<th>All SNPs</th>
<th>Inverse variance weighting with multiplicative random effects</th>
<th>MR-Egger</th>
<th>Weighted median</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR</td>
<td>95% CI</td>
<td>I2</td>
<td>OR</td>
</tr>
<tr>
<td>Glucose (mmol/L) (n=21)</td>
<td>1.03</td>
<td>0.85 to 1.25</td>
<td>0.763</td>
</tr>
<tr>
<td>HbA1c (%) (n=32)</td>
<td>1.02</td>
<td>0.73 to 1.45</td>
<td>0.843</td>
</tr>
<tr>
<td>Type 2 diabetes (n=147)</td>
<td>0.98</td>
<td>0.95 to 1.01</td>
<td>0.707</td>
</tr>
</tbody>
</table>

Excluding pleiotropic SNPs

| Glucose (mmol/L) (n=11) | 1.00 | 0.81 to 1.23 | 0.72 | 1.09 | 0.69 to 1.71 | -0.004 | 0.68 | 0.94 | 0.81 to 1.10 |
| HbA1c (%) (n=19) | 1.06 | 0.91 to 1.23 | 0.00 | 1.12 | 0.85 to 1.49 | -0.002 | 0.62 | 1.06 | 0.85 to 1.32 |
| Type 2 diabetes (n=108) | 0.97 | 0.94 to 1.01 | 0.585 | 0.90 | 0.80 to 1.01 | 0.005 | 0.18 | 0.97 | 0.93 to 1.01 |

SNP, single-nucleotide polymorphism.
Table 2  The impact of glycemic traits on breast cancer risk in Breast Cancer Association Consortium using Mendelian randomization, stratified by estrogen receptor subtypes

<table>
<thead>
<tr>
<th>All SNPs</th>
<th>Inverse variance weighting with multiplicative random effects</th>
<th>MR-Egger</th>
<th>Weighted median</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>95% CI</td>
<td>I²</td>
</tr>
<tr>
<td>Estrogen receptor (+ve)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucose (mmol/L) (n=21)</td>
<td>1.03</td>
<td>0.85 to 1.24</td>
<td>0.638</td>
</tr>
<tr>
<td>HbA1c (%) (n=32)</td>
<td>1.06</td>
<td>0.75 to 1.50</td>
<td>0.776</td>
</tr>
<tr>
<td>Type 2 diabetes (n=147)</td>
<td>0.98</td>
<td>0.95 to 1.02</td>
<td>0.651</td>
</tr>
<tr>
<td>Excluding pleiotropic SNPs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucose (mmol/L) (n=11)</td>
<td>1.00</td>
<td>0.81 to 1.24</td>
<td>0.603</td>
</tr>
<tr>
<td>HbA1c (%) (n=19)</td>
<td>1.06</td>
<td>0.89 to 1.27</td>
<td>0.00</td>
</tr>
<tr>
<td>Type 2 diabetes (n=108)</td>
<td>0.98</td>
<td>0.93 to 1.02</td>
<td>0.555</td>
</tr>
<tr>
<td>Estrogen receptor (-ve)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucose (mmol/L) (n=21)</td>
<td>1.06</td>
<td>0.82 to 1.38</td>
<td>0.577</td>
</tr>
<tr>
<td>HbA1c (%) (n=32)</td>
<td>1.12</td>
<td>0.68 to 1.84</td>
<td>0.751</td>
</tr>
<tr>
<td>Type 2 diabetes (n=147)</td>
<td>0.97</td>
<td>0.93 to 1.01</td>
<td>0.467</td>
</tr>
<tr>
<td>Excluding pleiotropic SNPs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucose (mmol/L) (n=11)</td>
<td>1.01</td>
<td>0.80 to 1.26</td>
<td>0.200</td>
</tr>
<tr>
<td>HbA1c (%) (n=19)</td>
<td>1.31</td>
<td>0.87 to 1.96</td>
<td>0.547</td>
</tr>
<tr>
<td>Type 2 diabetes (n=108)</td>
<td>0.95</td>
<td>0.89 to 1.00</td>
<td>0.342</td>
</tr>
</tbody>
</table>

SNP, single-nucleotide polymorphism.

diabetes in breast cancer, with similar observations by estrogen receptor subtype. Similarly, we did not find strong evidence for a role of glycemic traits in prostate cancer, which has not been investigated previously. As such, results from previous observational studies may be confounded by lifestyle factors, socioeconomic position or closely correlated traits.

Our study is consistent with a previous Mendelian randomization study, which showed no relation of fasting glucose with breast cancer.17 Our study adds by showing that the null findings were unlikely due to the use of sex-adjusted instruments given we had used instruments for glucose in women only from the MAGIC. However, the previous study suggested a potential positive relation of 2-hour glucose and insulin (when pleiotropic SNPs were removed) with breast cancer.17 Such differences suggest that factors related to glycemic control (eg, insulin and glucagon) may be more relevant than the glycemic traits themselves to cancer development and would deserve more investigations in future studies.

Pharmacological studies have suggested a potential inverse relation of metformin with breast and prostate

Table 3  The impact of glycemic traits on breast cancer risk in UK Biobank using Mendelian randomization

<table>
<thead>
<tr>
<th>All SNPs</th>
<th>Inverse variance weighting with multiplicative random effects</th>
<th>MR-Egger</th>
<th>Weighted median</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>95% CI</td>
<td>I²</td>
</tr>
<tr>
<td>Glucose (mmol/L) (n=21)</td>
<td>1.05</td>
<td>0.82 to 1.33</td>
<td>0.365</td>
</tr>
<tr>
<td>HbA1c (%) (n=34)</td>
<td>1.08</td>
<td>0.64 to 1.81</td>
<td>0.719</td>
</tr>
<tr>
<td>Type 2 diabetes (n=159)</td>
<td>0.98</td>
<td>0.95 to 1.03</td>
<td>0.292</td>
</tr>
<tr>
<td>Excluding pleiotropic SNPs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucose (mmol/L) (n=11)</td>
<td>1.02</td>
<td>0.77 to 1.35</td>
<td>0.35</td>
</tr>
<tr>
<td>HbA1c (%) (n=21)</td>
<td>1.15</td>
<td>0.71 to 1.86</td>
<td>0.608</td>
</tr>
<tr>
<td>Type 2 diabetes (n=120)</td>
<td>1.00</td>
<td>0.95 to 1.06</td>
<td>0.238</td>
</tr>
</tbody>
</table>

SNP, single-nucleotide polymorphism.
cancer.\textsuperscript{46} Metformin reduces glycemic traits suggesting a role of glycemic traits in these cancers. However, not all antidiabetic medications show an inverse relation. For example, use of insulin and glucagon-like peptide-1 analog is not related to breast cancer.\textsuperscript{48, 49} This discrepancy suggests that metformin’s potential anticancer property could be due to other pathways rather than its glycemic property. Baseline analysis of a randomized controlled trial suggested metformin may have pleiotropic effects, among which growth differentiation factor 15 (GDF-15) was the strongest biomarker\textsuperscript{50} although we previously showed that GDF-15 did not appear to be a cause of breast cancer in BCAC.\textsuperscript{12} Our findings on AMPK-instrumented HbA1c reduction also suggest no strong evidence for metformin use affecting cancer outcomes although they were wide CIs, and hence it would be premature to conclude there is no effect of metformin use on cancer risk. For example, AMPK-instrumented HbA1c reduction may increase breast cancer risk (OR 1.28, 95% CI 0.21 to 7.87) although the findings were not consistent in BCAC (OR 0.70, 95% CI 0.29 to 1.69). Such differences could be driven by differences in biases stemming from study design, such as differing levels of selection bias induced by inevitably only selected survivors of the exposure and competing risk. As such, it remains inconclusive whether the observed inverse relation between metformin and various cancers is a reflection of immortal time bias, notwithstanding the observations concerning metformin and prostate cancer are unclear.\textsuperscript{13} Results from randomized controlled trials and Mendelian randomization studies using genetic risk scores for AMPK in large well-designed biobanks would provide a more definitive answer to this important clinical question.

Strengths of this study included the use of Mendelian randomization, which is more robust to confounding than observational studies, and the use of validation data sets. Nevertheless, there are some limitations. First, the Mendelian randomization design has its own assumptions. However, we used genetic loci identified from GWAS. Genetic variants should be unconfounded. Sensitivity analyses with different assumptions about pleiotropy gave consistent estimates. Second, we were only able to use sex-specific genetic estimates for fasting glucose due to data availability and hence we could not ascertain whether the null findings for HbA1c and type 2 diabetes were driven by the use of sex-adjusted genetic estimates.

Table 4: The impact of glycemic traits on prostate cancer risk in Prostate Cancer Association Group to Investigate Cancer-Associated Alterations in the Genome consortium using Mendelian randomization

<table>
<thead>
<tr>
<th>All SNPs</th>
<th>Inverse variance weighting with multiplicative random effects</th>
<th>MR-Egger</th>
<th>Weighted median</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>95% CI</td>
<td>I²</td>
</tr>
<tr>
<td>Glucose (mmol/L) (n=11)</td>
<td>0.93</td>
<td>0.73 to 1.17</td>
<td>0.464</td>
</tr>
<tr>
<td>HbA1c (%) (n=19)</td>
<td>0.90</td>
<td>0.58 to 1.40</td>
<td>0.575</td>
</tr>
<tr>
<td>Type 2 diabetes (n=112)</td>
<td>1.02</td>
<td>0.97 to 1.07</td>
<td>0.690</td>
</tr>
<tr>
<td>Excluding pleiotropic SNPs*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucose (mmol/L) (n=10)</td>
<td>1.01</td>
<td>0.81 to 1.24</td>
<td>0.263</td>
</tr>
<tr>
<td>Type 2 diabetes (n=111)</td>
<td>1.02</td>
<td>0.97 to 1.07</td>
<td>0.686</td>
</tr>
</tbody>
</table>

*No pleiotropic SNPs identified for HbA1c.

Table 5: The impact of AMP-activated protein kinase (AMPK)-instrumented HbA1c reduction (%) on breast and prostate cancer using Mendelian randomization

<table>
<thead>
<tr>
<th>Nine single-nucleotide polymorphisms</th>
<th>Data sets</th>
<th>Inverse variance weighting with fixed effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>95% CI</td>
</tr>
<tr>
<td>Breast cancer (overall)</td>
<td>Breast Cancer Association Consortium</td>
<td>0.70</td>
</tr>
<tr>
<td>Breast cancer (estrogen receptor +ve)</td>
<td>Breast Cancer Association Consortium</td>
<td>0.77</td>
</tr>
<tr>
<td>Breast cancer (estrogen receptor −ve)</td>
<td>Breast Cancer Association Consortium</td>
<td>1.69</td>
</tr>
<tr>
<td>Breast cancer (overall)</td>
<td>UK Biobank</td>
<td>1.28</td>
</tr>
<tr>
<td>Prostate cancer</td>
<td>Prostate Cancer Association Group to Investigate Cancer-Associated Alterations in the Genome consortium</td>
<td>0.60</td>
</tr>
</tbody>
</table>
Notwithstanding this possibility, the GWAS on glucose suggested a lack of heterogeneity by sex regarding the SNPs predicting glucose (p≥0.001). Third, we relied on the summary statistics within the genes concerning AMPK to estimate the impact of metformin use on cancer risk. Such approach may have less statistical power compared with using genetic risk scores. Furthermore, we could not rule out the existence of other targets by which metformin may confer its potential anticancer properties. Lastly, we were unable to assess potential non-linearity between these traits and breast cancer risk as we did not have access to individual-level data. However, the possibility of non-linear relation is unlikely. Furthermore, we were unable to explore other cancers given lack of relevant summary statistics.

This Mendelian randomization study suggested a lack of relation of fasting glucose, HbA1c and type 2 diabetes with breast and prostate cancer while it remains unclear whether metformin reduces cancer risk. The putative anticancer property of metformin, if causal, is likely via glycaemia-independent pathways. Further research with better designs may help elucidate the underlying mechanisms, and hence identify new targets of intervention for cancer prevention.

Acknowledgements This research has been conducted using the UK Biobank Resource under application number 24460 (for breast cancer-related analyses). Data on sex-specific glucose have been contributed by MAGIC and have been downloaded from https://www.magicinvestigators.org/ (Vasiliki Lagou, Reedeik Miagi, Jouke-Jan J Hottenga, et al. Fasting glucose and insulin variability: sex-dimorphic genetic effects and novel loci (2019) In Preparation). Data on breast cancer have been contributed by BCAC investigators and have been downloaded from http://bcac.cge.medschl.cam.ac.uk/bcacdata/oncoarray/gwas-icogs-and-oncoarray-summary-results/. All studies and funders are listed in the study of Michailidou et al. Data on prostate cancer have been contributed by PRACTICAL investigators and have been downloaded from http://practical.iarc.fr/blog/?page_id=8164. We also thank the following for funding support: the Institute of Cancer Research and the Eversen Campaign, the Prostate Cancer Research Foundation, Prostate Research Campaign UK (now PCUK), the Orchid Cancer Appeal, RoseTrees Trust, the National Cancer Research Network UK, and the National Cancer Research Institute (NCRI) UK. We are grateful for support of the NIHR funding to the NIHR Biomedical Research Centre at the Institute of Cancer Research and the Royal Marsden NHS Foundation Trust. The Prostate Cancer Program of Cancer Council Victoria also acknowledges grant support from the National Health and Medical Research Council Australia, VicHealth, Cancer Council Victoria, the Prostate Cancer Foundation of Australia, the Whitten Foundation, PricewaterhouseCoopers, and Tattersall’s. EAO, DMK, and EMK from the original paper (Schumacher et al.) acknowledge the Intramural Program of the National Human Genome Research Institute for their support. The data used for the analyses described in this manuscript were obtained from the GTEx Portal on 15 August 2019.

Contributors SLAY designed the study and wrote the analysis plan. SLAY undertook analyses with feedback from CMS, and all authors interpreted the results in the study. SLAY wrote the first draft of the manuscript with critical feedback and revisions from CMS. All authors gave final approval of the version to be published. SLAY is the guarantor of this work.

Funding The breast cancer genome-wide association analyses were supported by the Government of Canada through Genome Canada and the Canadian Institutes of Health Research, the ‘Ministère de l’Économie, de la Science et de l’Innovation du Québec’ through Genome Québec and grant PSR-SII-NI-701, the National Institutes of Health (U19 CA148065, 1K12HL705472), Cancer Research UK (C1287/A10118, C1287/A10563, C1287/A10710) and the European Union (HEALTH-F2-2009-223175 and H2020 633784 and 634935). The prostate cancer genome-wide association analyses were supported by the Canadian Institutes of Health Research, European Commission’s Seventh Framework Programme grant agreement number 223175 (HEALTH-F2-2009-223175), Cancer Research UK grants CS047/AT35, C1287/A10118, C1287/A10563, CS047 /A33547, CS047/A10692, and C16913/A6135, and the National Institute of Health (NIH) Cancer Post-Cancer GWAS initiative grant: U19 CA 145357-01 (the GAME-ON initiative). Grant support was granted by the National Health and Medical Research Council Australia (126402, 209057, 251533, 396414, 451040, 504700, 504702, 504715, 623204, 940394, 614296), VicHealth, Cancer Council Victoria, the Prostate Cancer Foundation of Australia, the Whitten Foundation, PricewaterhouseCoopers, and Tattersall’s. Genotyping of the OncoArray was funded by the US National Institutes of Health (U19 CA148065 for ELucidating Loci Involved in Prostate cancer SucEpiduity (ELLIPSE) project and X01HG007492 to the Center for Inherited Disease Research (CIDR) under contract number HSNS208201200008) and by Cancer Research UK grant A8197/A16565. Additional analytic support was provided by NIH NCi U01 CA183392 (Pl. Schumacher). Funding for the iCOGS infrastructure came from the European Community’s Seventh Framework Programme under grant agreement number 223175 (HEALTH-F2-2009-223175) (COSG), Cancer Research UK (C1287/A10118, C1287/A10710, C12292/A11174, C121/A12014, C5047/A3583, C5047/A15007, C5047/A16062, C1917/A16565), the National Institutes of Health (CA128978) and Post-Cancer GWAS initiative (U19 CA145357, U19 CA148065 and U19 CA148112—the GAME-ON initiative), the Department of Defence (WB1XWH-10-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, Komen Foundation for the Cure, the Breast Cancer Research Foundation, and the ovarian Cancer Research Fund. The BPC3 was supported by the US National Institutes of Health, National Cancer Institute (cooperative agreements U01-CAB8233 to DJ Hunter, U01-CAB8710 to SM Gapstur, U01-CAB8216 to E Riboli, and U01-CAB8758 to BE Henderson, and Intramural Research Program of NH/National Cancer Institute, Division of Cancer Epidemiology and Genetics). GWAS study was supported by the Swedish Cancer Foundation (grant numbers 09-0677, 11-484, 12-823), the Cancer Risk Prediction Center (CiRiP; www.criscenter.org), and a Linneus Centre (Contract ID 7086702) financed by the Swedish Research Council (grant numbers K2010-70X-20430-04-3, 2014-2269). PEGASUS was supported by the Intramural Research Program, Division of Cancer Epidemiology and Genetics, National Cancer Institute, and National Institutes of Health. The Genotype-Tissue Expression (GTEx) project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by ND, NIHBI, NHIL, NIDA, NIMH, and NINDS.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This study used previously available data and hence no ethics approval was required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs Shiu Lun Au Yeung http://orcid.org/0000-0001-6136-1836 Catherine Mary Schooling http://orcid.org/0000-0001-9933-5887

REFERENCES
Epidemiology/Health Services Research