Evaluating the impact of AMPK activation, a target of metformin, on risk of cardiovascular diseases and cancer in the UK Biobank: A Mendelian randomization study

Running title: Impact of metformin on CVD and cancer

Shan Luo, MPH (0000-0003-3368-9935),[1] Catherine Mary Schooling, PhD (0000-0001-9933-5887),[1,2] Ian Chi Kei Wong, PhD (0000-0001-8242-0014),[3,4] Shiu Lun Au Yeung, PhD (0000-0001-6136-1836)[1]

Affiliations:

1 School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China
2 School of Public Health and Health Policy, City University of New York, New York, United States
3 Department of Pharmacology and Pharmacy, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China
4 Research Department of Practice and Policy, School of Pharmacy, University College London, United Kingdom

Correspondence to
Shiu Lun Au Yeung
1/F, Patrick Manson Building (North Wing), 7 Sassoon Road, Pokfulam, Hong Kong
Telephone: (+852) 3917 6740
Email: ayslryan@hku.hk

Tweet:
Our MR study suggested AMPK activation, a metformin target, may reduce type 2 diabetes, CAD, and overall cancer risk. Whether metformin can be repurposed for prevention of these diseases should be further explored. @Ryan_Au Yeung @hkumed

Word count: 3,479

Abstract

Aim Whether metformin reduces cardiovascular or cancer risk is unclear due to concerns over immortal time bias and confounding in observational studies. This study evaluated the effect of AMP-activated protein kinase (AMPK), the target of metformin, on risk of cardiovascular disease and cancer.

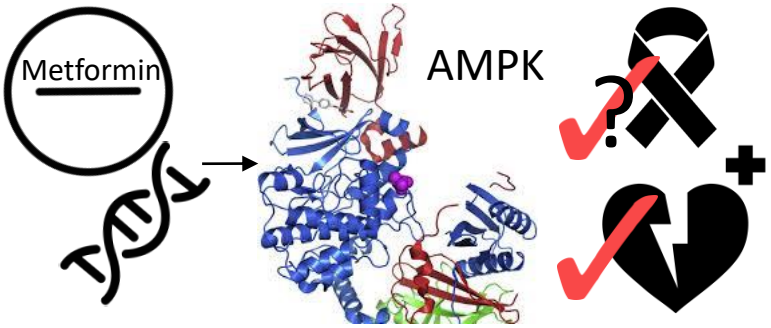
Methods This is a Mendelian randomization design, using AMPK, the pharmacologic target of metformin, to infer the AMPK pathway dependent effects of metformin on risk of cardiovascular disease and cancer in participants of white British ancestry in the UK Biobank.

Results A total of 391,199 participants were included (mean age 56.9 years; 54.1% women), of which 26,690 cases of type 2 diabetes, 38,098 cases of coronary artery disease and 80,941 cases of overall cancer. Genetically predicted reduction in HbA_{1c} (%) instrumented by AMPK variants was associated with a 61% reduction in risk of type 2 diabetes (odds ratio [OR] 0.39, 95% confidence interval [CI] 0.20 to 0.78, $P=7.69\times10^{-3}$), a 53% decrease in the risk of coronary artery disease (OR 0.47, 95% CI 0.26 to 0.84, P=0.01) and a 44% decrease in the risk of overall cancer (OR 0.56, 95% CI 0.36 to 0.85, $P=7.23\times10^{-3}$). Results were similar using median or quartiles of AMPK score, with dose-response effects (P for trend=4.18×10^{-3} for type 2 diabetes, 4.37×10^{-3} for coronary artery disease and 4.04×10^{-3} for overall cancer).

Conclusions This study provides some genetic evidence that AMPK activation by metformin may protect against cardiovascular disease and cancer, which needs to be confirmed by randomized controlled trials.

Key words Metformin, AMPK, type 2 diabetes, coronary artery disease, cancer, UK Biobank, Mendelian randomization

Graphical abstract:



Research in context

What is already know about this subject?

Metformin is the first-line pharmacologic treatment to manage hyperglycaemia in people with type 2 diabetes. Whether metformin reduces cardiovascular or cancer risk is unclear due to concerns over immortal time bias and confounding in observational studies.

What is the key question?

What is the association between AMP-activated protein kinase (AMPK), the pharmacological target of metformin, on lifetime risk of cardiovascular disease and cancer?

What are the new findings?

In Mendelian randomization analyses involving 391,199 participants of white British ancestry in the UK Biobank, reduction in HbA_{1c} instrumented by AMPK were associated with lower risk of coronary artery disease, and possibly cancer.

How might this impact on clinical practice in the foreseeable future?

Based on the genetics of AMPK, metformin use may protect against cardiovascular disease, and possibly cancer.

Introduction

Metformin is the first-line pharmacologic treatment to manage hyperglycaemia in people with type 2 diabetes, and is on the World Health Organization list of essential medicines [1]. Increasing evidence suggests that metformin may differ from other classes of anti-diabetic medications in having superior safety and lower risk of cardiovascular complications [2]. Furthermore, pharmaco-epidemiological studies have suggested metformin may reduce cardiovascular disease and cancer [3, 4], suggesting the possibility of its use for these diseases. Metformin not only impacts glycaemic traits but also other potentially relevant factors, such as growth differentiation factor 15 (GDF-15) and vascular endothelial growth factors [5]. However, pharmaco-epidemiological studies may be open to immortal-time bias and confounding, which may generate spurious protective effects of metformin, in particular for cancer related studies [6, 7]. To date, relevant randomized controlled trials of metformin in cardiovascular disease are not large enough to be definitive [8], whilst the impact of metformin on cancer has not been evaluated fully in a randomized controlled trial.

Mendelian randomization studies, which make use of the random allocation of genetic variants at conception, are less susceptible to confounding and time related biases than other observational studies, and are now increasingly used to infer health effects of medications, such as the use 3-hydroxy-3-methyl-glutaryl-coenzyme A reductase (*HMGCR*) variants to mimic the health effects of statins [9]. Previous Mendelian randomization studies have attempted to use genetics to infer health effects of metformin, but they were potentially underpowered and included non-specific instruments [10], or only evaluated health effects of metformin biomarkers instead of metformin itself [11]. To provide more definitive and direct evidence concerning the effect of metformin on cardiovascular disease and cancer risk, we conducted a Mendelian randomization study using AMP-activated protein kinase (AMPK),

the target of metformin, as a proxy of metformin use, in one of the largest prospective cohort studies globally.

Method

Study design

This is a Mendelian randomization design, using AMPK, the pharmacologic target of metformin, to infer the AMPK pathway dependent effects of metformin. The study design is depicted in Fig 1 [12].

Study population

The UK Biobank recruited ~500,000 participants intended to be aged 39-73 years between 2006 and 2010 from 22 recruitment centres across Scotland, Wales and England in the United Kingdom. Participants provided biological samples, completed questionnaires, including self-reported diseases and regular prescription medications, underwent assessments and had nurse-led interviews. A blood sample for standard haematological tests was collected by venepuncture in ethylenediaminetetraacetic acid tubes, and tested at the central processing laboratory in Stockport, within 24 hours of blood collection. HbA_{1c} was measured by high performance liquid chromatography on Bio-Rad Variant II Turbo analysers. Longitudinal follow-up via record linkage to all health service encounters and death is ongoing. Hospital inpatient data and cancer registries used International Classification of Diseases (ICD)-9 and ICD-10 codes and death registries used ICD-10 codes. Genotyping was undertaken with two similar arrays, the UK BiLEVE (Biobank Lung Exome Variant Evaluation) Axiom array (49,979 participants) and UK Biobank Axiom array (438,398 participants). Genotype imputation was based on the reference panel combining the UK10K haplotype and the Haplotype Reference Consortium reference panels. To reduce confounding by latent

population structure [13], we restricted the analysis to genetically verified white British participants and further excluded participants with 1) withdrawn consent, 2) sex-mismatch (genetic sex differs from reported sex), 3) putative sex chromosomes aneuploidy, 4) poor-quality genotyping (outliers in heterozygosity and missing rate > 1.5%), or 5) excessive relatedness (more than ten putative third-degree relatives), as per our previous study [14]. We used genotype and phenotype data from the UK Biobank provided in February 2020.

AMPK genetic score

We created a weighted AMPK genetic score to mimic the effects of AMPK activation by metformin use based on the strength of the association of genetic variants in the relevant gene regions with HbA_{1c} in MAGIC (Meta-Analyses of Glucose and Insulin-related traits Consortium), a genome wide association study (GWAS) of HbA_{1c}, with validation in the UK Biobank. Specifically, we selected genetic variants within 1-Megabase pairs downstream and upstream of each of the *PRKAA1*, *PRKAA2*, *PRKAB1*, *PRKAB2*, *PRKAG1*, *PRKAG2* and *PRKAG3* genes that encodes AMPK subunits [15]. We selected low linkage disequilibrium ($r^2 < 0.3$) variants associated with HbA_{1c} at a nominal level of statistical significance ($P \leq 0.05$) in MAGIC, restricted to people of European ancestry to minimize population stratification (n=123,665) [16]. We then validated the associations in the UK Biobank (using multivariable linear regression, adjusted for age, sex, age at recruitment, genotyping array and the first 20 principal components of genetic ancestry) and only retained variants also reaching statistical significance ($P \leq 0.05$) in the UK Biobank, which were used to construct the AMPK score. ESM Table 1 and ESM Fig. 1 show the details regarding the 44 variants used to construct the AMPK score. A weighted AMPK score was calculated for each participant by summing the number of HbA_{1c}-lowering alleles that a participant inherited at each variant included in the AMPK score, weighted by the effect of that variant on HbA_{1c}

measured in percentage (as estimated in MAGIC) [16]. This score was then considered as above and below the median to mimic metformin use and non-use. We also considered AMPK score in quartiles to assess whether our finding was robust to the way we consider AMPK scores.

Sensitivity analysis

As a sensitivity analysis, we used a stringent variant selection criteria by only using variants associated with HbA_{1c} at genome wide significance ($P \leq 5\times10^{-8}$) in both MAGIC and UK Biobank and not in linkage disequilibrium with the other variants ($r^2 < 0.01$), which gave rs2732480, ESM Table 1. Rs2732480 was associated with lower HbA_{1c} in both MAGIC ($P = 2\times10^{-9}$) and UK Biobank ($P = 1.07\times10^{-142}$). The effect allele was associated with lower HbA_{1c} % (beta -0.012, 95% confidence interval [CI] -0.016 to -0.008).

Study outcomes

The primary outcomes were coronary artery disease and overall cancer. The secondary outcomes were stroke and three main cancers, i.e. breast cancer, colorectal cancer and prostate cancer. Each disease outcome was defined based on self-report medical conditions at baseline, or subsequent primary and secondary diagnoses of hospital episodes (ICD-9 and ICD-10), or cancer register (ICD-9 and ICD-10), or underlying and contributing causes of death (ICD-10).

Positive control outcomes

We included type 2 diabetes and HbA_{1c} as positive control outcomes given these are the expected effects of metformin use. Type 2 diabetes was ascertained using a validated algorithm [17]. Specifically, the criteria included 1) self-reported type 2 diabetes at baseline,

2) indication of type 2 diabetes based on diagnostic codes (ICD-9 250 and ICD-10 E11), 3) diabetes medications (metformin, sodium-glucose cotransporter 2 inhibitors, glucagon-like peptide 1 receptor agonists, dipeptidyl peptidase 4 inhibitors, sulfonylureas and thiazolidinediones), and 4) hyperglycaemic blood result (either $HbA_{1c} \geq 6.5\%$ or 48 mmol/mol, or random glucose ≥ 11.1 mmol/L). The algorithmic definitions are described in ESM Table 2. HbA_{1c} was measured in mmol/mol, (IFCC unit, International Federation of Clinical Chemistry), and was converted to percentage (NGSP unit, National Glycohemoglobin Standardization Program) using the equation: NGSP = (0.09148×IFCC)+2.152 [18].

External validation

To validate our findings from the UK Biobank, we conducted an external validation study for the outcomes using summary statistics for type 2 diabetes, (12,171 cases and 56,862 controls) from the DIAGRAM (DIAbetes Genetics Replication And Meta-analysis) consortium [19], coronary artery disease, (60,801 cases and 123,504 controls) from the CARDIoGRAMplusC4D (Coronary ARtery DIsease Genome wide Replication and Meta-analysis plus The Coronary Artery Disease Genetics) consortium [20], stroke (40,585 cases of stroke and 406,111 controls) from the MEGASTROKE consortium [21]; breast cancer (122,977 cases and 105,974 controls) from the BCAC (Breast Cancer Association Consortium) [22]; and prostate cancer (79,148 cases and 61,106 controls) from the PRACTICAL (Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome) consortium [23]. All participants were of predominantly European ancestry and non-overlapping with the participants in the UK Biobank to avoid bias due to population structure and any potential bias due to participant overlap for a weak instrument [24].

Statistical analysis

To assess the assumption of independence of the genetic instruments (AMPK groups) from potential confounders, we assessed the association of AMPK groups with confounders (age at recruitment, body mass index [BMI], smoking status, alcohol drinking status, education level, Townsend deprivation index,) using chi-square tests or analysis of variance. To demonstrate AMPK had the expected effect on HbA_{1c}, we assessed the differences in HbA_{1c} and random glucose between each group using analysis of variance. We assessed the association of AMPK categories with HbA_{1c} using multivariable linear regression. We assessed the association of AMPK categories with risk of T2D, cardiovascular diseases and cancers using multivariable logistic regression. All regression analyses were adjusted for sex (if relevant), age at recruitment, genotyping array and the first 20 principal components of genetic ancestry. As per previous studies, we also assessed the impact of genetically predicted reduction in HbA_{1c} (%) instrumented by AMPK variants on risk of T2D, CAD and overall cancer [25].

For the external validation, we performed a standard Mendelian randomization analysis. We obtained the summary statistics of each variant included in the AMPK score on risk of T2D, coronary artery disease, stroke, breast cancer and prostate cancer as reported by each consortium. We obtained the Wald ratio for each variant (the ratio of the genetic association with outcome to the genetic association of exposure), and then combined them using weighted generalized linear regression in an inverse-variance weighted manner and accounted for the correlation between variants [26]. The correlations between variants were obtained in 503 participants of European ancestry from the 1000 Genomes Project (Phase 3). Since variants are from multiple gene regions that may have different mechanisms of effect, a

random-effects model was used [26]. We aligned the effect allele of each variant to the HbA_{1c} decreasing allele. We used the Cochran's Q statistic to assess heterogeneity of the Wald ratios [27] where high heterogeneity may indicate the presence of invalid genetic variants [28].

Exploring the association of HbA_{1c} with cardiovascular disease and cancer risk using Mendelian randomization

To preclude the possibility that the observed effects of AMPK activation, a target of metformin, are due to lowering HbA_{1c}, we also assessed the association of genetically predicted lower HbA_{1c} on cardiovascular disease and cancer risk. As previously, we obtained 38 independent genetic variants strongly related to HbA_{1c} ($P \leq 5\times10^{-8}$) from the MAGIC (ESM Table 3), and applied them to the relevant outcomes in the UK Biobank using inverse variance weighting, MR-Egger and weighted median method [14].

The AMPK score was generated using *PLINK 2.0*. Mendelian randomization analyses were performed with *MendelianRandomisation* package and all analyses were performed using R software, version 3.5.1 (R Foundation for Statistical Computing). A two-tailed *P* value less than 0.05 was considered statistically significant.

Results

Participant characteristics

A total of 391,199 participants were included in the main analysis (mean age, 56.9 years; 54.1% women). For T2D, there were 26,690 cases. For cardiovascular disease, there were 38,098 cases of coronary artery disease and 11,358 cases of stroke. For cancer, there were 80,941 cases of overall cancer, 9,251 cases of breast cancer, 5,861 cases of colorectal cancer,

and 8,970 cases of prostate cancer. In Table 1, $HbA_{1c,}$ glucose, insulin therapy users and metformin therapy users were significantly lower in the high AMPK group than in the low group. No other significant differences in baseline characteristics between the two groups were found.

Association of AMPK score with glycaemic traits and T2D

Compared with participants with a low AMPK score (below median favouring higher HbA_{1c}), participants with a high AMPK score (above median) had 0.032 % lower HbA_{1c}, 95% CI, 0.028 to 0.035, $P=2.34\times10^{-64}$, and 0.013 lower random blood glucose mmol/L (95% CI, 0.005 to 0.022; $P=1.09\times10^{-3}$). High AMPK score (above median) was also associated with a decreased risk of T2D (odds ratio (OR) 0.96, 95% CI 0.94 to 0.99; $P=4.16\times10^{-3}$), as shown in Fig. 2a. AMPK quartiles were associated with a stepwise decrease in HbA_{1c} (quartile 2, 0.011%, 95% CI 0.006 to 0.016, $P=2.70\times10^{-5}$; quartile 3, 0.028%, 95% CI 0.022 to 0.033, $P=2.02\times10^{-25}$ and quartile 4, 0.047%, 95% CI 0.042 to 0.052, $P=4.67\times10^{-70}$), and a corresponding stepwise decrease in the risk of T2D (P for trend=4.18×10^{-3}, Fig. 2a). Genetically predicted reduction in HbA_{1c} (%) instrumented by AMPK variants was associated with a 61% decrease in the risk of T2D (OR 0.39 per % reduction, 95% CI 0.20 to 0.78, $P=7.69\times10^{-3}$) (Fig. 3).

Association of AMPK with cardiovascular diseases

High AMPK score (above median) was associated with a 3% lower risk of coronary artery disease (OR 0.97, 95% CI 0.95 to 0.99; $P=5.69\times10^{-3}$, Fig. 2b), but not stroke (ESM Fig. 2a). AMPK quartile was associated with a stepwise decrease in the risk of coronary artery disease (P for trend=4.37×10^{-3}, Fig. 2b). Genetically predicted reduction in HbA_{1c} (%) instrumented

by AMPK variants was associated with a 53% decrease in the risk of coronary artery disease (OR 0.47 per % reduction, 95% CI 0.26 to 0.84, P=0.01) (Fig. 3).

Association of AMPK with cancer

High AMPK score (above median) was associated with lower risk of overall cancer (OR 0.98, 95% CI 0.96 to 1.00, P=0.01 and P for trend=4.04×10^{-3}, Fig. 2c), but not with prostate cancer, breast cancer or colorectal cancer (ESM Fig. 2b to d). AMPK quartile was associated with prostate cancer (Quartile 2, OR 0.91, 95% CI 0.85 to 0.96, P=1.61×10^{-3}, Quartile 3, OR 0.93, 95% CI 0.88 to 0.99, P=0.02) although the dose response was unclear (ESM Fig. 2b). Genetically predicted reduction in HbA_{1c} (%) instrumented by AMPK variants was associated with a 44% decrease in the risk of overall cancer (OR 0.56 per % reduction, 95% CI 0.36 to 0.85, P=7.23×10^{-3}) (Fig. 3).

Sensitivity analysis by using a more stringent variant selection criteria

One % reduction in HbA_{1c} instrumented by rs2732480 was associated with a decreased risk of T2D (OR 0.11, 95% CI 0.02 to 0.50, P=4.08×10^{-3}), coronary artery disease (OR 0.22, 95% CI 0.06 to 0.81, P=0.02). The direction with overall cancer was consistent with the main analysis but with wider CI (OR 0.45, 95% CI 0.17 to 1.14, P=0.09), ESM Table 4.

External validation

In external replication analyses, genetically predicted lower HbA_{1c} instrumented by AMPK variants was associated with decreased risk of T2D (OR 0.11 per % reduction, 95% CI 0.04 to 0.35; P=1.78×10^{-4}) and coronary artery disease (OR 0.48 per % reduction, 95 CI 0.33 to 0.72; P=2.89×10^{-4}), but not with stroke, breast cancer or prostate cancer (ESM Table 5 and

ESM Fig. 3a to e). The Q statistic suggested possible heterogeneity for the association with T2D, coronary artery disease, stroke and breast cancer.

Association of HbA_{1c} with cardiovascular disease and cancer risk in the UK Biobank using Mendelian randomization

ESM Table 6 shows that genetically predicted higher HbA_{1c} was associated with higher risk of coronary artery disease (OR 1.41, 95% CI 1.03 to 1.93, P=0.03), and possibly with lower risk of overall cancer (OR 0.84, 95% CI 0.70 to 1.01, P=0.07), but not for stroke, or any cancer subtype.

Discussion

To the best of our knowledge, this is one of the first Mendelian randomization studies to ascertain the effects of metformin, based on AMPK variants, on cardiovascular diseases and cancer. Using a design more robust to immortal time biases and confounding, our study is consistent with previous pharmaco-epidemiological studies suggesting that metformin use may reduce coronary artery disease and overall cancer risk. We added some genetic evidence that the putative cancer-protective effect of metformin via AMPK pathways is unlikely by glycaemic control.

A protective effect of metformin on cardiovascular health was observed in small randomized controlled trials using surrogate outcomes [29], the UK Prospective Diabetes Study post trial analysis [30], and a recent meta-analysis [8], which were consistent with our findings. Although a genetically predicted reduction in HbA_{1c} is protective against coronary artery disease [14], it is apparent that metformin's protective effect is not solely due to its improvement in glycaemic profile given these benefits are not clearly observed for all other

classes of anti-diabetic medications [8], such as sulfonylureas and insulin [2, 31]. Metformin increases in GDF-15, a stress responsive cytokine which suppresses appetite and promoting weight loss [32], and hence provides a potential mechanistic pathway by which metformin reduces cardiovascular disease risk. However, changes in GDF-15 were not clearly associated with coronary artery disease risk based on our previous Mendelian randomization study [11]. On the contrary, sulfonylureas and insulin may lead to cardiotoxicity via weight gain, hypoglycaemia [31], or alteration of hormone levels [33].

The relation of metformin use with cancer risk is more controversial given the concern over immortal time bias [6]. Our study, where the start of "exposure" is at birth, effectively removes this bias. As such, our study adds by showing that immortal time bias alone may not have explained the inverse relation of metformin use with cancer risk. Given previous studies generally have ruled out the causal role of glycaemic traits in cancer risk [10, 34], possible mechanisms underlying the anti-cancer property of metformin is likely via pre-cursors of glycaemic traits or of glycaemic-independent pathways [35]. People without growth hormone appear to be protected against both diabetes and cancer [36]. This may suggest a possible pathway via growth hormone or the closely related insulin like growth factor-1 [37]. Glycaemic-independent pathways may include inhibition of tumour-mesothelial cell interaction by suppressing hypoxia-inducible factor 1α and transforming growth factor TGF-β signalling [38], immune-mediated via metabolic reprogramming of tumour-specific T cells [39], and GDF-15 overexpressing fibroblasts promote the growth of tumour xenografts [40]. The examination of these potential mechanisms can be explored in further studies. Big data approaches, such as metabolomics, may also be warranted to better understand the full spectrum of effects of metformin, and hence help identify the main pathways in which metformin confers the additional benefits on cardiovascular disease and cancer [41].

Although our study is more robust to confounding and immortal time bias than previous observational studies, there are limitations. First, whilst our study suggested AMPK activation by metformin may protect against coronary artery disease, and possibly cancer, the estimates from this study cannot be used directly to infer the health impact of metformin given the differences in exposure time where randomized controlled trials often consider short-term pharmacologic treatment in contrast to the effect of lifelong exposures estimated by Mendelian randomization [12]. Moreover, our study using AMPK variants may only predict the effect of metformin which acts on the AMPK activation pathways, and metformin may also have AMPK-independent pathways and could be explored in additional studies to fully capture the overall effect of metformin on cardiovascular disease and cancer [42]. Second, we used a lower threshold than genome-wide statistical significance to select AMPK variants as proxy of metformin use to maximize total prediction of AMPK function by the genetic score. We reduced the possibility of false positives by cross checking the variants' association with HbA_{1c} in two independent studies. We also repeated the analysis with stringent variant selection criteria which gave a consistent conclusion. However, this may compromise the generalizability of the genetic score in other studies [43]. Third, we cannot rule out selection bias resulting from the recruitment of generally healthier participants and survivors in the UK Biobank, which may bias the estimate towards null. We also cannot rule out selection bias from competing risk before recruitment for diseases which share risk factors with other diseases that typically occur at younger ages, which could have biased estimates for stroke and prostate cancer to the null. Forth, the Q statistic suggested possible heterogeneity in some analyses. These heterogeneity may imply multiple gene regions encoding subunits of AMPK may have different mechanisms of influencing the outcomes and should be explored in future studies [28]. Fifth, given the pleiotropic effects of metformin and

its association with multiple non-glycemic makers [5], it would be difficult to identify a suitable negative control outcome. Nevertheless, we also assessed the impact of HbA_{1c} on these outcomes and found that HbA_{1c} unlikely explained all the observed associations related to AMPK. Lastly, we could not exclude the possibility that metformin may reduce sub-types of cancer as the number of cases was not large enough for adequate statistical power although the direction of effect for some cancer subtypes are similar to the overall cancer. Few AMPK genetic variants were available for prostate cancer and breast cancer in those consortia and we were unable to create an overall genetic score in the associated analyses to increase statistical power. Together with possible selection biases embedded in these GWAS [44], these might explain the discrepancy between the estimates from the UK Biobank and external consortia. Further investigations in large consortia on specific cancers may help verify the potential anti-cancer property of metformin.

Conclusion

This Mendelian randomization study provides some genetic evidence that AMPK activation by metformin may reduce coronary artery disease risk, and possibly overall cancer risk. Whether metformin can be repurposed for coronary artery disease and cancer should be explored in large randomized controlled trials.

Contribution statement: SL and SLAY designed the study, wrote the research plan and interpreted the results. SL undertook analyses with feedback from SLAY, CMS and ICKW. SL and SLAY wrote the manuscript with critical comments from CMS and ICKW. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. All authors gave final approval of the version to be published. SL is the guarantor of this work, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Funding: This research is funded by Seed Fund for Basic Research, The University of Hong Kong (#201811159054 to SLAY). SL is supported by Bau Tsu Zung Bau Kwan Yeu Hing Research and Clinical Fellowship (*200008682.920006.20006.400.01), The University of Hong Kong. ICKW received research funding outside the submitted work from Amgen, Bristol-Myers Squibb, Pfizer, Janssen, Bayer, GSK, Novartis, the Hong Kong Research Grant Council, and the Hong Kong Health and Medical Research Fund, National Institute for Health Research in England, European Commission, National Health and Medical Research Council in Australia, and also received speaker fees from Janssen and Medice in the previous 3 years. The study sponsor/funder was not involved in the design of the study; the collection, analysis, and interpretation of data; writing the report; and did not impose any restrictions regarding the publication of the report.

Authors' relationships and activities: The authors declare that there are no relationships or activities that might bias, or be perceived to bias, their work.

Acknowledgements: This research has been conducted using the UK Biobank Resource (www.ukbiobank.ac.uk) under application number 51001. Summary Data on type 2 diabetes

have been contributed by DIAGRAM investigators and have been downloaded from http://diagram-consortium.org/. Summary data on coronary artery disease have been contributed by CARDIoGRAMplusC4D investigators and have been downloaded from www.CARDIOGRAMPLUSC4D.ORG. Summary data on stroke have been contributed by the MEGASTROKE investigators and have been downloaded from http://www.megastroke.org/. The MEGASTROKE project received funding from sources specified at http://www.megastroke.org/acknowledgments.html. Summary data on breast cancer have been contributed by BCAC investigators and have been downloaded from http://bcac.ccge. medschl.cam.ac.uk/bcacdata/oncoarray/gwas-icogs-and-oncoarraysummary-results/. 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-SIIRI-701, The National Institutes of Health (U19 CA148065, X01HG007492),Cancer Research UK (C1287/A10118, C1287/A16563, C1287/A10710) and The European Union (HEALTH-F2-2009-223175 and H2020 633784 and 634935). All studies and funders are listed in Michailidou et al (2017). Summary data on prostate cancer have been contributed by The PRACTICAL consortium, CRUK, BPC3, CAPS, PEGASUS. The Prostate cancer genome-wide association analyses are supported by the Canadian Institutes of Health Research, European Commission's Seventh Framework Programme grant agreement n° 223175 (HEALTH-F2-2009-223175), Cancer Research UK Grants C5047/A7357, C1287/A10118, C1287/A16563, C5047/A3354, C5047/A10692, C16913/A6135, and The National Institute of Health (NIH) Cancer Post-Cancer GWAS initiative grant: No. 1 U19 CA 148537-01 (the GAME-ON initiative). We would also like to thank the following for funding support: The Institute of Cancer Research and The Everyman Campaign, The Prostate Cancer Research Foundation, Prostate Research Campaign UK (now

PCUK), The Orchid Cancer Appeal, Rosetrees Trust, The National Cancer Research Network UK, The National Cancer Research Institute (NCRI) UK. We are grateful for support of 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 acknowledge grant support from The National Health and Medical Research Council, Australia (126402, 209057, 251533, , 396414, 450104, 504700, 504702, 504715, 623204, 940394, 614296,), VicHealth, Cancer Council Victoria, The Prostate Cancer Foundation of Australia, The Whitten Foundation, PricewaterhouseCoopers, and Tattersall's. EAO, DMK, and EMK acknowledge the Intramural Program of the National Human Genome Research Institute for their support. Genotyping of the OncoArray was funded by the US National Institutes of Health (NIH) [U19 CA 148537 for ELucidating Loci Involved in Prostate cancer SuscEptibility (ELLIPSE) project and X01HG007492 to the Center for Inherited Disease Research (CIDR) under contract number HHSN268201200008I] and by Cancer Research UK grant A8197/A16565. Additional analytic support was provided by NIH NCI U01 CA188392 (PI: Schumacher). Funding for the iCOGS infrastructure came from: the European Community's Seventh Framework Programme under grant agreement n° 223175 (HEALTH-F2-2009-223175) (COGS), Cancer Research UK (C1287/A10118, C1287/A 10710, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692, C8197/A16565), the National Institutes of Health (CA128978) and Post-Cancer GWAS initiative (1U19 CA148537, 1U19 CA148065 and 1U19 CA148112 – the GAME-ON initiative), the Department of Defence (W81XWH-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 U.S. National Institutes of Health, National Cancer Institute (cooperative agreements U01-CA98233 to D.J.H., U01-CA98710 to S.M.G.,

U01-CA98216 toE.R., and U01-CA98758 to B.E.H., and Intramural Research Program of NIH/National Cancer Institute, Division of Cancer Epidemiology and Genetics). CAPS GWAS study was supported by the Swedish Cancer Foundation (grant no 09-0677, 11-484, 12-823), the Cancer Risk Prediction Center (CRisP; www.crispcenter.org), a Linneus Centre (Contract ID 70867902) financed by the Swedish Research Council, Swedish Research Council (grant no K2010-70X-20430-04-3, 2014-2269). PEGASUS was supported by the Intramural Research Program, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health.

Competing interests: none.

Ethical approval: The UK Biobank received ethical approval from the North West Multi-centre Research Ethics Committee (11/NW/0382), and all participants provided written informed consent. No ethics approval was acquired for the analyses using summary statistics. The contributing studies to the consortium received ethical approval from their specific institutional review boards, and written informed consent was obtained from all participants.

Data availability

The data generated and analysed during the current study are available from the corresponding author on reasonable request.

Table 1: Baseline characteristics of the participants in the UK Biobank

Baseline characteristics	**AMPK score < median (n = 194634)**	**AMPK score ≥ median (n = 196565)**	**P value**
Age at recruitment (yr)	56.9±8.0	56.9±8.0	0.85
Female, No. (%)	105048 (26.9)	106632 (27.3)	0.08
Current smoker, No. (%)	19869 (5.1)	19689 (5.0)	0.09
Current alcohol drinker, No. (%)	181839 (46.5)	183332 (46.9)	0.22
Blood pressure (mmHg)			
Systolic	138.3±18.7	138.3±18.6	0.73
Diastolic	82.3±10.1	82.3±10.1	0.34
Body mass index (kg/m^2)	27.4±4.8	27.4±4.8	0.61
Education level Degree, No. (%)	88727 (22.7)	89717 (22.9)	0.94
Townsend deprivation index	-1.55±2.94	-1.56±2.93	0.21
HbA_{1c} (mmol/mol)	36.14±6.49	35.8±6.33	*<0.001*
HbA_{1c} (%)	5.46±0.59	5.43±0.58	*<0.001*
Random glucose (mmol/L)	5.13±1.23	5.11±1.20	*<0.001*
Current treatment, No. (%)			
Antihypertensive therapy	43400 (11.1)	43564 (11.1)	0.31
Insulin therapy	2197 (0.6)	2077 (0.5)	*0.03*
Metformin therapy	5465 (1.4)	5224 (1.3)	*0.004*

Values are means ± standard deviations. To convert values for HbA_{1c} (mmol/mol, IFCC unit) to percentage (NGSP unit), with master equation NGSP = (0.09148 * IFCC) + 2.152; To convert glucose from mmol/L to mg/dL, multiply by 18.

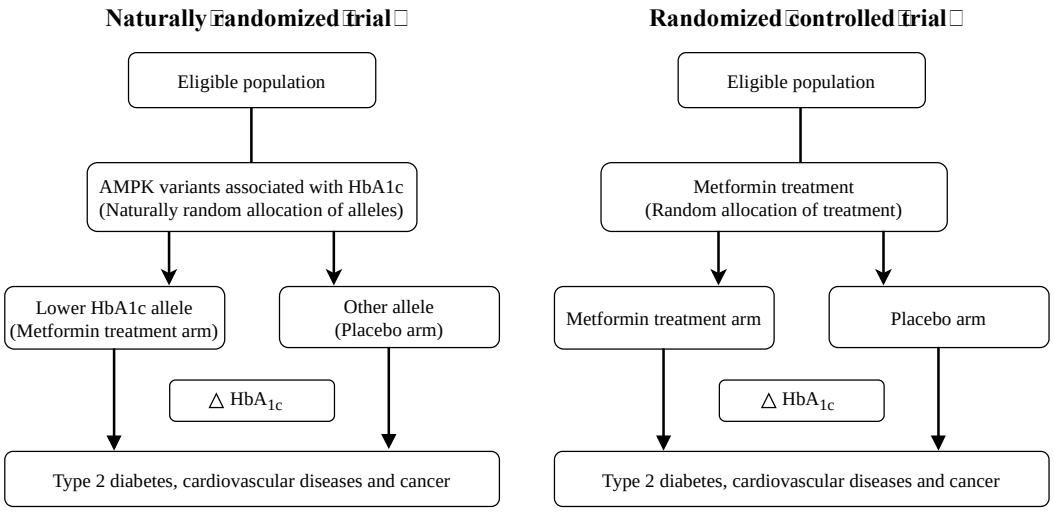
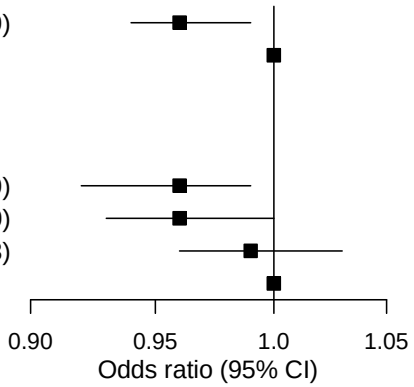


Fig 1. Study design of this Mendelian randomization study and its comparison to randomized controlled trial.

Fig 2. Association of AMP-activated protein kinase score with risk of type 2 diabetes, coronary heart disease and overall cancer in the UK Biobank

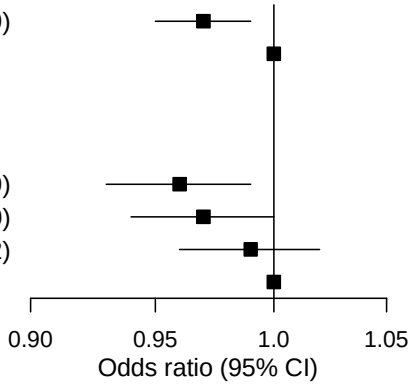
a) Type 2 diabetes

Group	No. of cases (%)	OR (95% CI)
Primary analysis		
AMPK score above median	13,167 (3.4)	0.96 (0.94, 0.99)
AMPK score below median	13,523 (3.5)	Reference
Dose−response analysis		
Quartile of AMPK score		
4	6613 (1.7)	0.96 (0.92, 0.99)
3	6554 (1.7)	0.96 (0.93, 1.00)
2	6711 (1.7)	0.99 (0.96, 1.03)
1	6812 (1.7)	Reference



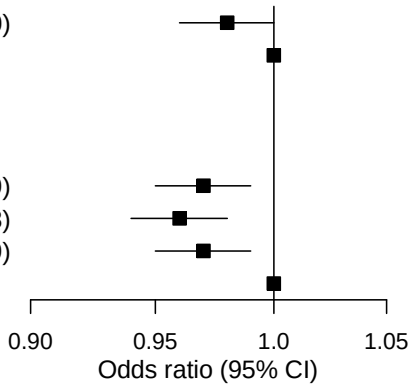
b) Coronary artery disease

Group	No. of cases (%)	OR (95% CI)
Primary analysis		
AMPK score above median	18,882 (4.8)	0.97 (0.95, 0.99)
AMPK score below median	19,216 (4.9)	Reference
Dose−response analysis		
Quartile of AMPK score		
4	9530 (2.4)	0.96 (0.93, 0.99)
3	9352 (2.4)	0.97 (0.94, 1.00)
2	9521 (2.4)	0.99 (0.96, 1.02)
1	9695 (2.5)	Reference



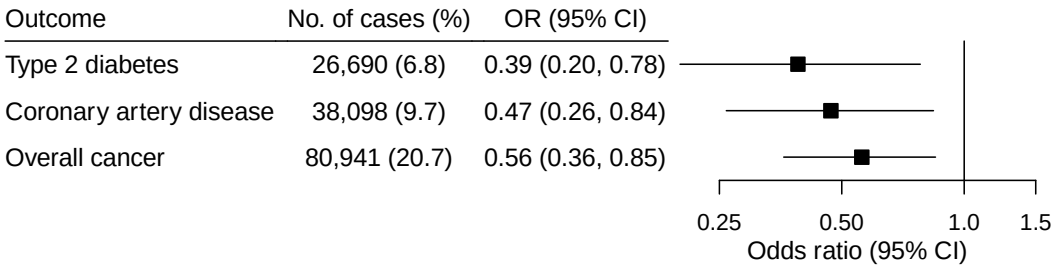
c) Overall cancer

Group	No. of cases (%)	OR (95% CI)
Primary analysis		
AMPK score above median	40,383 (10.3)	0.98 (0.96, 1.00)
AMPK score below median	40,558 (10.4)	Reference
Dose−response analysis		
Quartile of AMPK score		
4	20,478 (5.2)	0.97 (0.95, 0.99)
3	19,905 (5.1)	0.96 (0.94, 0.98)
2	19,999 (5.1)	0.97 (0.95, 0.99)
1	20,559 (5.3)	Reference



Boxes represent odds ratios (OR) and lines represent 95% confidence intervals (CI).

Fig 3. The impact of genetically predicted reduction in HbA_{1c} (%) instrumented by AMPK variants on risk of type 2 diabetes, coronary artery disease and overall cancer in the UK Biobank



Boxes represent odds ratio (OR) and lines represent 95% confidence intervals (CI).

References

[1] Bailey CJ (2017) Metformin: historical overview. Diabetologia 60(9): 1566-1576. 10.1007/s00125-017-4318-z

[2] Douros A, Dell'Aniello S, Yu OHY, Filion KB, Azoulay L, Suissa S (2018) Sulfonylureas as second line drugs in type 2 diabetes and the risk of cardiovascular and hypoglycaemic events: population based cohort study. BMJ 362: k2693. 10.1136/bmj.k2693

[3] Heckman-Stoddard BM, DeCensi A, Sahasrabuddhe VV, Ford LG (2017) Repurposing metformin for the prevention of cancer and cancer recurrence. Diabetologia 60(9): 1639-1647. 10.1007/s00125-017-4372-6

[4] Rena G, Lang CC (2018) Repurposing Metformin for Cardiovascular Disease. Circulation 137(5): 422-424. 10.1161/CIRCULATIONAHA.117.031735

[5] Gerstein HC, Pare G, Hess S, et al. (2017) Growth Differentiation Factor 15 as a Novel Biomarker for Metformin. Diabetes Care 40(2): 280-283. 10.2337/dc16-1682

[6] Suissa S, Azoulay L (2012) Metformin and the risk of cancer: time-related biases in observational studies. Diabetes Care 35(12): 2665-2673. 10.2337/dc12-0788

[7] Farmer RE, Ford D, Mathur R, et al. (2019) Metformin use and risk of cancer in patients with type 2 diabetes: a cohort study of primary care records using inverse probability weighting of marginal structural models. International Journal of Epidemiology 48(2): 527-537. 10.1093/ije/dyz005

[8] Zhu J, Yu X, Zheng Y, et al. (2020) Association of glucose-lowering medications with cardiovascular outcomes: an umbrella review and evidence map. Lancet Diabetes Endocrinol. 10.1016/S2213-8587(19)30422-X

[9] Ference BA, Robinson JG, Brook RD, et al. (2016) Variation in PCSK9 and HMGCR and Risk of Cardiovascular Disease and Diabetes. N Engl J Med 375(22): 2144-2153. 10.1056/NEJMoa1604304

[10] Au Yeung SL, Schooling CM (2019) Impact of glycemic traits, type 2 diabetes and metformin use on breast and prostate cancer risk: a Mendelian randomization study. BMJ Open Diabetes Res Care 7(1): e000872. 10.1136/bmjdrc-2019-000872

[11] Au Yeung SL, Luo S, Schooling CM (2019) The impact of GDF-15, a biomarker for metformin, on the risk of coronary artery disease, breast and colorectal cancer, and type 2 diabetes and metabolic traits: a Mendelian randomisation study. Diabetologia 62(9): 1638-1646. 10.1007/s00125-019-4913-2

[12] Ference BA (2018) How to use Mendelian randomization to anticipate the results of randomized trials. European Heart Journal 39(5): 360-362. 10.1093/eurheartj/ehx462

[13] Haworth S, Mitchell R, Corbin L, et al. (2019) Apparent latent structure within the UK Biobank sample has implications for epidemiological analysis. Nature Communications 10. ARTN 33310.1038/s41467-018-08219-1

[14] Au Yeung SL, Luo S, Schooling CM (2018) The Impact of Glycated Hemoglobin (HbA1c) on Cardiovascular Disease Risk: A Mendelian Randomization Study Using UK Biobank. Diabetes Care 41(9): 1991-1997. 10.2337/dc18-0289

[15] Xiao B, Heath R, Saiu P, et al. (2007) Structural basis for AMP binding to mammalian AMP-activated protein kinase. Nature 449(7161): 496-500. 10.1038/nature06161

[16] Wheeler E, Leong A, Liu CT, et al. (2017) Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transethnic genome-wide meta-analysis. PLoS Med 14(9): e1002383. 10.1371/journal.pmed.1002383

[17] Eastwood SV, Mathur R, Atkinson M, et al. (2016) Algorithms for the Capture and Adjudication of Prevalent and Incident Diabetes in UK Biobank. PLoS One 11(9): e0162388. 10.1371/journal.pone.0162388

[18] English E, Lenters-Westra E (2018) HbA1c method performance: The great success story of global standardization. Crit Rev Cl Lab Sci 55(6): 408-419. 10.1080/10408363.2018.1480591

[19] Morris AP, Voight BF, Teslovich TM, et al. (2012) Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nature Genetics 44(9): 981-+. 10.1038/ng.2383

[20] Nikpay M, Goel A, Won HH, et al. (2015) A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet 47(10): 1121-1130. 10.1038/ng.3396

[21] Malik R, Chauhan G, Traylor M, et al. (2018) Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nat Genet 50(4): 524-537. 10.1038/s41588-018-0058-3

[22] Michailidou K, Lindstrom S, Dennis J, et al. (2017) Association analysis identifies 65 new breast cancer risk loci. Nature 551(7678): 92-+. 10.1038/nature24284

[23] Schumacher FR, Al Olama AA, Berndt SI, et al. (2018) Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. Nature Genetics 50(7): 928-+. 10.1038/s41588-018-0142-8

[24] Burgess S, Davies NM, Thompson SG (2016) Bias due to participant overlap in two-sample Mendelian randomization. Genetic Epidemiology 40(7): 597-608. 10.1002/gepi.21998

[25] Yarmolinsky J, Bull CJ, Vincent EE, et al. (2020) Association Between Genetically Proxied Inhibition of HMG-CoA Reductase and Epithelial Ovarian Cancer. JAMA 323(7): 646-655. 10.1001/jama.2020.0150

[26] Burgess S, Dudbridge F, Thompson SG (2016) Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods. Statistics in Medicine 35(11): 1880-1906. 10.1002/sim.6835

[27] Greco MF, Minelli C, Sheehan NA, Thompson JR (2015) Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. Stat Med 34(21): 2926-2940. 10.1002/sim.6522

[28] Burgess S, Butterworth A, Thompson SG (2013) Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 37(7): 658-665. 10.1002/gepi.21758

[29] Larsen AH, Jessen N, Norrelund H, et al. (2019) A randomised, double-blind, placebo-controlled trial of metformin on myocardial efficiency in insulin-resistant chronic heart failure patients without diabetes. Eur J Heart Fail. 10.1002/ejhf.1656

[30] Holman RR, Paul SK, Bethel MA, Matthews DR, Neil HA (2008) 10-year follow-up of intensive glucose control in type 2 diabetes. N Engl J Med 359(15): 1577-1589. 10.1056/NEJMoa0806470

[31] Ferrannini E, DeFronzo RA (2015) Impact of glucose-lowering drugs on cardiovascular disease in type 2 diabetes. Eur Heart J 36(34): 2288-2296. 10.1093/eurheartj/ehv239

[32] Coll AP, Chen M, Taskar P, et al. (2019) GDF15 mediates the effects of metformin on body weight and energy balance. Nature. 10.1038/s41586-019-1911-y

[33] Zhao JV, Luo S, Schooling CM (2019) Sex-specific Mendelian randomization study of genetically predicted insulin and cardiovascular events in the UK Biobank. Commun Biol 2(1): 332. 10.1038/s42003-019-0579-z

[34] Shu X, Wu L, Khankari NK, et al. (2019) Associations of obesity and circulating insulin and glucose with breast cancer risk: a Mendelian randomization analysis. Int J Epidemiol 48(3): 795-806. 10.1093/ije/dyy201

[35] Vancura A, Bu PL, Bhagwat M, Zeng J, Vancurova I (2018) Metformin as an Anticancer Agent. Trends Pharmacol Sci 39(10): 867-878. 10.1016/j.tips.2018.07.006

[36] Guevara-Aguirre J, Rosenbloom AL (2015) Obesity, diabetes and cancer: insight into the relationship from a cohort with growth hormone receptor deficiency. Diabetologia 58(1): 37-42. 10.1007/s00125-014-3397-3

[37] Laron Z (2008) The GH-IGF1 axis and longevity. The paradigm of IGF1 deficiency. Hormones (Athens) 7(1): 24-27. 10.14310/horm.2002.1111034

[38] Hart PC, Kenny HA, Grassl N, et al. (2019) Mesothelial Cell HIF1alpha Expression Is Metabolically Downregulated by Metformin to Prevent Oncogenic Tumor-Stromal Crosstalk. Cell Rep 29(12): 4086-4098 e4086. 10.1016/j.celrep.2019.11.079

[39] Bahrambeigi S, Shafiei-Irannejad V (2019) Immune-mediated anti-tumor effects of metformin; targeting metabolic reprogramming of T cells as a new possible mechanism for anti-cancer effects of metformin. Biochem Pharmacol: 113787. 10.1016/j.bcp.2019.113787

[40] Bruzzese F, Hagglof C, Leone A, et al. (2014) Local and Systemic Protumorigenic Effects of Cancer-Associated Fibroblast-Derived GDF15. Cancer Res 74(13): 3408-3417. 10.1158/0008-5472.Can-13-2259

[41] Wurtz P, Wang Q, Soininen P, et al. (2016) Metabolomic Profiling of Statin Use and Genetic Inhibition of HMG-CoA Reductase. J Am Coll Cardiol 67(10): 1200-1210. 10.1016/j.jacc.2015.12.060

[42] Rena G, Hardie DG, Pearson ER (2017) The mechanisms of action of metformin. Diabetologia 60(9): 1577-1585. 10.1007/s00125-017-4342-z

[43] Torkamani A, Wineinger NE, Topol EJ (2018) The personal and clinical utility of polygenic risk scores. Nat Rev Genet 19(9): 581-590. 10.1038/s41576-018-0018-x

[44] CM Schooling, P Lopez, Z Yang, SL Au Yeung, JV Huang. Bias from competing risk before recruitment in Mendelian Randomization studies of conditions with shared etiology. Available https://www.biorxiv.org/content/10.1101/716621v3.