Leisure time sedentary behaviour and risks of breast, colorectal, and prostate cancer: A Mendelian randomization analysis

Nikos Papadimitriou[1], Nabila Kazmi[2], Niki Dimou[1], Konstantinos K Tsilidis[3,4], Richard M Martin[2,5,6], Sarah J Lewis[5], Brigid M Lynch[7-9], Michael Hoffmeister[10], Sun-Seog Kweon[11,12], Li Li[13], Roger L Milne[7,8,14], Lori C Sakoda[15,16], Robert E Schoen[17], Amanda I Phipps[16,18], Jane C Figueiredo[19,20], Ulrike Peters[16,18], Suzanne C. Dixon-Suen[7,21], Marc J Gunter[1], Neil Murphy[1]

[1]. Nutrition and Metabolism Branch, International Agency for Research on Cancer, Lyon, France.

[2]. MRC Integrative Epidemiology Unit (IEU), Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK.

[3]. Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece.

[4]. Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK.

[5]. Bristol Medical School, Department of Population Health Sciences, University of Bristol, Bristol, UK.

[6]. National Institute for Health Research (NIHR) Bristol Biomedical Research Centre, University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol, Bristol, UK.

[7]. Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, Victoria, Australia.

[8]. Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia.

[9]. Physical Activity Laboratory, Baker Heart and Diabetes Institute, Melbourne, Victoria, Australia.

[10] Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany.

[11] Department of Preventive Medicine, Chonnam National University Medical School, Gwangju, Korea.

[12] Jeonnam Regional Cancer Center, Chonnam National University Hwasun Hospital, Hwasun, Korea.

[13] Department of Family Medicine, University of Virginia, Charlottesville, Virginia, USA.

[14] Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, Victoria, Australia.

[15] Division of Research, Kaiser Permanente Northern California, Oakland, California, USA.

[16] Public Health Sciences Division, Fred Hutchinson Cancer Center, Seattle, Washington, USA.

[17] Department of Medicine and Epidemiology, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA.

[18].Department of Epidemiology, University of Washington, Seattle, Washington, USA.

[19] Department of Medicine, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA.

[20] Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.

[21] Institute for Physical Activity and Nutrition, School of Exercise and Nutrition Sciences, Deakin University, Geelong, Victoria, Australia.

Corresponding author: Dr Nikos Papadimitriou, Branch of Nutrition and Metabolism, International Agency for Research on Cancer, 150 Cours Albert Thomas, 69008 Lyon, France, Tel.: +33 4 72738485, E-mail: papadimitrioun@iarc.who.int

Competing interests:

The authors declare no competing interests.

Disclaimer: Where authors are identified as personnel of the International Agency for Research on Cancer / World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer / World Health Organization.

Availability of data and materials

The summary statistics used in this study are outlined in the supplementary materials.

Funding

This work is supported by a WCRF grant (WCRF_2020_019). RMM is supported by a Cancer Research UK Programme Grant, the Integrative Cancer Epidemiology Programme (C18281/A29019). RMM is a member of the MRC IEU which is supported by the Medical Research Council and the University of Bristol (MC_UU_12013/1-9). RMM is supported by the National Institute for Health Research (NIHR) Bristol Biomedical Research Centre which is funded by the National Institute for Health Research and is a partnership between University Hospitals Bristol NHS Trust, Weston NHS Foundation Trust and the University of Bristol. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR or the UK Department of Health and Social Care.

BML is supported by the Victorian Cancer Agency (MCRF-18005).

SJL is supported by a WCRF grant (WCRF_2020_019).

Information on consortia funding is included in the supplementary material.

Acknowledgements

The breast cancer genome-wide association analyses for BCAC and CIMBA were supported by Cancer Research UK (PPRPGM-Nov20\100002, C1287/A10118, C1287/A16563, C1287/A10710, C12292/A20861, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692, C8197/A16565) and the Gray Foundation, The National Institutes of Health (CA128978, X01HG007492- the DRIVE consortium), the PERSPECTIVE project supported by the Government of Canada through Genome Canada and the Canadian Institutes of Health Research (grant GPH-129344) and the Ministère de l'Économie, Science et Innovation du Québec through Genome Québec and the PSRSIIRI-701 grant, the Quebec Breast Cancer Foundation, the European Community's Seventh Framework Programme under grant agreement n° 223175 (HEALTH-F2-2009-223175) (COGS), the European Union's Horizon 2020 Research and Innovation Programme (634935 and 633784), the Post-Cancer GWAS initiative (U19 CA148537, CA148065 and 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 (CRN-87521), the Komen Foundation for the Cure, the Breast Cancer Research Foundation and the Ovarian Cancer Research Fund. All studies and funders are listed in Zhang H et al (Nat Genet, 2020).

Additional information on consortia acknowledgements is included in the supplementary material

Abstract (332 words)

Introduction: Sedentary behaviours have been associated with increased risks of some common cancers in epidemiological studies; however, it is unclear if these associations are causal.

Methods: We used univariable and multivariable two-sample Mendelian randomization (MR) to examine potential causal relationships between sedentary behaviours and risks of breast, colorectal, and prostate cancer. Genetic variants associated with self-reported leisure television watching and computer use were identified from a recent genome-wide association study (GWAS). Data related to cancer risk were obtained from cancer GWAS consortia. A series of sensitivity analyses were applied to examine the robustness of the results to the presence of confounding.

Results:

In the univariable models, a 1-standard deviation (SD:1.5 hours/day) increment in hours of television watching increased risk of breast cancer (OR per 1-SD: 1.15, 95% confidence interval [CI]: 1.05, 1.26) and colorectal cancer (OR per 1-SD: 1.32, 95%CI: 1.16, 1.49) while there was little evidence of an association for prostate cancer risk (OR per 1-SD: 0.94, 95%CI: 0.84, 1.06). In multivariable MR models adjusted for years of education, the effect estimates for television watching were attenuated towards the null (breast cancer, OR per 1-SD: 1.08, 95%CI: 0.92, 1.27; colorectal cancer, OR per 1-SD: 1.08, 95%CI: 0.90, 1.31). However, in post-hoc analyses we found evidence of years of education having a possible confounding and mediating role in the association between television watching with breast and colorectal cancer. Consistent results were observed for each cancer site according to sex

(colorectal cancer), anatomical subsites, and cancer subtypes. There was little evidence of associations between genetically predicted computer use and risk of all three cancers.

Conclusion: We found evidence of positive associations between hours of television watching and risks of breast and colorectal cancer. However, these findings should be interpreted cautiously given the complex confounding and mediating role we found for years of education on the television watching and cancer relationship. Future studies using objective measures of exposure (e.g., accelerometers) can provide new insights into the possible role of sedentary behaviour in cancer development.

Abbreviations

BMI: body mass index

IVW: inverse-variance weighted

GWAS: genome-wide association study

LD: linkage disequilibrium

MR: Mendelian randomization

PSA: prostate-specific antigen

OR: odds ratio

RCT: randomised control trial

SNP: single nucleotide polymorphism

SD: Standard deviation

Introduction (3988 words)

Breast, colorectal, and prostate cancer are three of the most common malignancies collectively accounting for an estimated 29% of new cancer cases in 2020 (1). Sedentary behaviour is defined as any waking behaviour characterized by energy expenditure ≤1.5 metabolic equivalents while in a sitting, reclining, or lying posture (2). The most common sedentary activities are television watching and computer use; these are more accurately recalled than total sedentary time and are therefore commonly used as surrogates of sedentary behaviour (3). A recent US study reported that approximately two-thirds of adults spent two or more hours each day watching television and around 50% spend more than one hour using their computer outside work (4). Studies in the UK and in the US estimated that adults on average spend five to six hours per day sitting (4, 5). Given such a high prevalence, sedentary behaviours represent an important public health challenge as they have been linked with multiple adverse health outcomes (6, 7).

Numerous observational studies have examined the associations between sedentary behaviours and the risks of breast, colorectal, and prostate cancer (8). A meta-analysis of case-control and cohort studies reported that sedentary behaviour was not associated with colorectal cancer risk (8). More recently, however, a UK Biobank analysis, found that greater volumes of television watching was associated with elevated colon cancer risk (9).The aforementioned meta-analysis did not observe any significant associations between sedentary behaviour and risk of prostate cancer (8). For breast cancer, when the meta-analysis included cohort studies only, sedentary behaviour was associated with a higher breast cancer risk (8). Clarifying causal associations from such observational evidence is hampered by inherent biases of the study design, such as residual confounding and reverse causality (10-12). Mendelian randomization (MR) is an alternative way to investigate potential causal associations. MR uses germline genetic variants as proxies (or instrumental variables) for

exposures of interest to make causal inferences between an exposure and an outcome (13). Unlike traditional observational epidemiology, MR can be largely free of conventional confounding owing to the random independent assignment of alleles during meiosis (14). In addition, multivariable MR methods have been developed to adjust for confounding if found to be present. There should be no reverse causation in MR studies, as germline genetic variants are fixed at conception and are consequently unaffected by the disease process (14). A recent MR analysis reported a positive effect estimate for television watching with lung cancer risk (15). However, similar analyses investigating possible causal effects of sedentary behaviours for other common cancers have not been conducted.

We used a two-sample MR framework to examine potential causal associations between self-reported sedentary behaviours and risks of breast, colorectal, and prostate cancer. Genetic variants associated with leisure television watching and computer use were identified from a recent genome-wide association study (GWAS) (16) and we then examined how these genetic variants related to risks of breast, colorectal, and prostate cancer using large-scale GWAS consortia data (17-19).

Methods

Data on leisure sedentary behaviours

Summary-level data on duration of leisure sedentary behaviours were obtained from a recently published GWAS conducted in 408,815 participants of European ancestry from the UK Biobank using BOLT-LMM v2.3beta2, using a mixed linear model correcting for population structure and cryptic relatedness (16). To ascertain the duration of the sedentary behaviours, participants within the UK Biobank were asked three questions, “In a typical DAY, how many hours do you spend watching television?”, “In a typical DAY, how many hours do you spend using the computer? (Do not include using a computer at work)” and “In

a typical DAY, how many hours do you spend driving?"(16). This GWAS identified 209 and 52 genome-wide-significant single nucleotide polymorphisms (SNPs) (P-value < 5×10^{-8}) for leisure television watching and computer use respectively using a linkage disequilibrium (LD) of $R^2 < 0.005$ within a five megabase window (Supplemental Tables 1-2). The GWAS also identified five genetic variants associated with driving; however, we did not include these instruments in our MR analyses due to low statistical power (see Statistical power, below). The 261 SNPs included in both instruments were identified in 204 loci demonstrating a partial overlap between the two phenotypes with 22 common loci. The selected SNPs explained approximately 2% and 0.5% of the variability in television watching and computer use respectively.

Data on breast, colorectal, and prostate cancer

Summary data for the associations of the above genetic variants with breast cancer were obtained from a GWAS of 247,173 women (133,384 breast cancer cases and 113,789 controls) of European ancestry from the Breast Cancer Association Consortium (19). We included six related outcomes in our analyses (overall, luminal A, luminal B, luminal B HER2 negative, HER2 enriched, and triple negative breast cancer).

For colorectal cancer, summary data from 98,715 participants (52,775 colorectal cancer cases and 45,940 controls) were drawn from a meta-analysis within the ColoRectal Transdisciplinary Study, the Colon Cancer Family Registry, and the Genetics and Epidemiology of Colorectal Cancer consortia (17). We included five outcomes in our analyses (overall colorectal cancer, colorectal cancer for men, colorectal cancer for women, colon cancer, and rectal cancer). The summary statistics did not include UK Biobank study to avoid potential overlap with the leisure sedentary behaviours GWAS.

For prostate cancer, summary data from a meta-analysis of 140,254 (79,148 prostate cancer cases and 61,106 controls) men of European ancestry in the Prostate Cancer Association Group to Investigate Cancer-Associated Alterations in the Genome and the Genetic Associations and Mechanisms in Oncology/Elucidating Loci Involved in Prostate Cancer Susceptibility consortia (18). The same consortia also conducted a GWAS of aggressive prostate cancer involving 15,167 cases and 58,308 controls, in which cancer cases were defined as aggressive based on the following characteristics: Gleason score ≥8, Prostate-Specific Antigen (PSA)>100 ng/mL, metastatic disease (M1) or death from prostate cancer (18).

All cancer estimates for the two exposures of interest are provided in Supplemental Tables 3-8. All participants provided written informed consent. Ethics were approved by respective institutional review boards.

Statistical power

The statistical power was calculated *a priori* using an online tool at http://cnsgenomics.com/shiny/mRnd/ (20). Under the scenario of a type 1 error of 5%, for leisure television use an expected OR per 1 standard deviation (SD) ≥ 1.09, ≥ 1.14 and ≥ 1.11 was needed to have adequate statistical power (> 80%) for overall breast, colorectal and prostate cancer respectively. Supplemental Table 9 presents the power estimates for the three exposures of interest by subtypes or subsites of breast, colorectal, and prostate cancer.

Statistical analysis

A two-sample MR approach using summary data and the fixed-effect IVW method was implemented. All results correspond to an OR per 1-SD increment in genetically-predicted hours of leisure sedentary behaviour (television watching: 1.5 hours/day; computer

use: 1.2 hours/day). The heterogeneity of the causal estimates by cancer subtype (breast cancer), subsite (colorectal cancer) and sex (colorectal cancer only) was investigated by calculating the I^2 metric using a fixed effect meta-analysis model (21).

Sensitivity analyses

MR studies have three main assumptions that must be satisfied in order for their causal estimates to be valid, which in the context of this study are: 1) the genetic instrument is strongly associated with the levels of exposure (sedentary behaviour); 2) the genetic instrument is not associated with any potential confounder of the exposure (sedentary behaviour)—outcome (cancer) association; and 3) the genetic instrument does not affect the outcome (cancer) independently of the exposure (sedentary behaviour) (i.e. exclusion of horizontal pleiotropy). The strength of each genetic instrument can be evaluated through the F-statistic (provided by the initial GWAS) (16). Several sensitivity analyses were conducted to identify and correct for the presence of horizontal pleiotropy in the results from the main analysis. Cochran's Q was computed to quantify heterogeneity across the individual causal effects, with a P-value ≤ 0.05 indicating the presence of pleiotropy, and consequently, a random effects IVW MR analysis was used (21, 22). MR-Egger regression was performed in which the intercept term can deviate from zero allowing estimation of the causal effect even in the presence of invalid genetic variants. Large deviations from zero represent the presence of horizontal pleiotropic effects across the genetic variants. In such a case, the slope of the MR-Egger regression provides valid MR estimates when the pleiotropic effects of the genetic variants are independent from the genetic associations with the exposure (23, 24). Moreover, causal estimates were also computed using the weighted-median method that can give valid MR estimates under the presence of horizontal pleiotropy when up to 50% of the included instruments are invalid (25). The MR pleiotropy residual sum and outlier test (MR-PRESSO)

was also used to assess the presence of pleiotropy. The MR-PRESSO test relies on a regression framework to identify outlying genetic variants which may potentially be pleiotropic, we then reran the analysis after excluding these outlying variants (26). We also examined the selected genetic instruments and their proxies ($r^2 > 0.8$) and their associations with secondary phenotypes (P-value $< 5 \times 10{-}8$) in populations of European descent in Phenoscanner (http://www.phenoscanner.medschl.cam.ac.uk/) to explore potential pleiotropy of the included SNPs. Since several of the genetic variants were also associated with adiposity or education-related phenotypes - such as body mass index (BMI) and educational attainment - we performed multivariable MR to investigate whether any initial significant associations for sedentary behaviour are confounded by these two traits as well as additional secondary traits such as lifetime smoking and alcohol consumption which have previously been linked with cancer risk (27-29). For BMI, summary data from a GWAS meta-analysis of about 700,000 participants of European descent within the Genetic Investigation of ANthropometric Traits (GIANT) consortium and UK Biobank was obtained (30). For years of educational attainment, we obtained summary level data from a published GWAS of 1.1 million participants of European descent within the Social Science Genetic Association Consortium and which measured the number of completed years of schooling among those individuals (31). Data on alcohol consumption (drinks per week) was drawn from a GWAS of 1.2 million individuals (32). The data for lifetime smoking was obtained from a recent GWAS and MR study on causal effects of lifetime smoking on risk for depression and schizophrenia (33). In the current analysis we used data of 766,345 participants which is publicly available. All relevant summary statistics for the multivariable MR analyses is given in supplemental tables 10-17. For multivariable MR, we also calculated two variables: the conditional $F_{early\ life\ body\ size}$, $F_{adult\ body\ size}$ which can be used to examine how much variance the genetic variants explain on the main (sedentary behaviours) and secondary exposures (e.g.,

years of education); F values over 10 suggest little evidence of weak instrument bias (34). Finally, as a post-hoc analysis based on the results from the multivariable MR and trying to understand the observed attenuation, we also conducted a bidirectional MR study to examine the associations between sedentary behaviours and the four secondary traits (BMI, years of education, alcohol consumption, and lifetime smoking) (supplemental tables 18-21).

All the analyses were conducted using the MendelianRandomization and TwoSampleMR packages, while the LD clumping (LD < 0.001) in the multivariable MR analyses between SNPs of sedentary behaviour phenotypes with those for the secondary traits was done using the ieugwasr R package (https://mrcieu.github.io/ieugwasr/) and the R programming language (version 4.1.2) (35-37). Reporting guidelines for MR studies were followed (MR STROBE checklist).

Results

MR estimates for leisure television watching

A 1 SD (1.5 hours/day) increment in genetically-predicted duration of leisure television watching increased breast cancer risk (OR per 1 SD: 1.15, 95% confidence interval [CI]: 1.05, 1.26, P-value: 0.002) (Table 1). Similar magnitude positive effect estimates were found for all molecular subtypes of breast cancer (I^2 = 0%, P-heterogeneity=0.98) (Table 1).

A 1 SD increment in genetically-predicted duration of leisure television watching increased colorectal cancer risk (OR per 1 SD: 1.32, 95% CI: 1.16, 1.49, P-value: 2×10^{-5}) with similar significant estimates being observed for men and women (I^2 = 42%, P-heterogeneity=0.19) and by subsite (I^2 = 45%, P-heterogeneity=0.17) (Table 2).

There was little evidence that a 1 SD increment in genetically-predicted duration of leisure television watching was associated with risk of overall (OR per 1 SD: 0.94, 95% CI:

0.84, 1.06, P-value:0.34) or aggressive (OR per 1 SD: 0.95, 95% CI: 0.81, 1.13, P-value:0.59) (overall vs aggressive; I^2 = 0%, P-heterogeneity=0.92) prostate cancer (Table 3).

Based on the Cochran's Q values there was evidence of heterogeneity of SNP effects for most outcomes except for triple negative breast cancer (Tables 1-3). Scatter plots (with coloured lines representing the slopes of the different regression analyses) and funnel plots of the association between leisure television watching and the risk of breast, colorectal and prostate cancer risk are presented in Supplemental Figures 1-6.

The multivariable MR analysis adjusting for years of education led to the attenuation of all effect estimates between genetically-predicted television watching and the risk of breast (OR per 1 SD: 1.08, 95% CI: 0.92, 1.27) and colorectal cancer (OR per 1 SD: 1.08, 95% CI: 0.90, 1.31) (Figure 1, Supplemental Table 22). Additional attenuations were observed for the models adjusting for lifetime smoking. For women, risk estimates for colorectal cancer were attenuated towards the null in all multivariable MR models adjusting for each of the four secondary traits (Figure 1, Supplemental Table 22). Finally, genetically-predicted television watching was associated with HER2 negative, HER2 positive, and triple negative breast cancer after adjusting for BMI in the multivariable MR models with effect sizes ranging from 1.32 to 1.46 per SD (Figure 1).

MR estimates for leisure computer use

There was little evidence of any causal effect of longer duration of genetically predicted leisure computer use with overall breast, colorectal, and prostate cancer (Tables 1-3). Inverse effect-estimates were found for triple negative breast cancer (OR per 1 SD: 0.68, 95% CI: 0.50, 0.93, P-value: 0.02) and rectal cancer (OR per 1 SD: 0.66, 95% CI: 0.49, 0.89, P-value: 6×10^{-3}) (Tables 1,2). Despite this, little evidence of heterogeneity was found by breast cancer subtype (I^2 = 36%, P-heterogeneity=0.17), colorectal cancer subsite (I^2 = 45%,

P-heterogeneity=0.15), or by prostate cancer status (overall vs aggressive; I^2 = 0%, P-heterogeneity=0.34), or sex (colorectal cancer: I^2 = 31%, P-heterogeneity=0.23).

Based on Cochran's Q values, heterogeneity in SNP effects was found for overall breast cancer, luminal A breast cancer, luminal B breast cancer, and colorectal cancer. Scatter plots (with coloured lines representing the slopes of the different regression analyses) and funnel plots of the association between leisure computer use and risks of breast, colorectal and prostate cancer are presented in Supplemental Figures 7-12.

In the multivariable MR analysis for triple negative breast cancer, after adjusting for years of education, alcohol, or BMI the inverse effect estimates for genetically-predicted computer use found in the univariable MR analysis were no longer statistically significant with the new attenuated effect sizes ranging from 0.73 to 1.06 per SD (Figure 1, Supplemental Table 22). Similarly, the inverse effect estimates for rectal cancer observed in the univariable analysis were attenuated after adjusting for years of education or alcohol consumption (Figure 1, Supplemental Table 22).

Evaluation of assumptions and sensitivity analyses

The strength of the genetic instruments according to the F-statistic was ≥10 for both exposures of interest and ranged between 23 and 164 (Supplemental Tables 1-3). Little evidence of directional pleiotropy was observed based on the MR-Egger's test (MR-Egger intercept P-values > 0.05) (Tables 1-3). The effect estimates from MR Egger regression models were generally in the same direction with those from the main analysis but with wider confidence intervals (Tables 1-3). Similarly, the weighted-median approach effect estimates were consistent in direction and magnitude to the IVW models (Tables 1-3). The MR-PRESSO analysis identified several (10 in total) outlying SNPs (Supplemental Table 23); however, no major differences were observed when these outlying genetic variants were

excluded from the analyses (Tables 1-3). After examining Phenoscanner, we found that several of the genetic variants were also associated with adiposity or education-related phenotypes, such as BMI and highest qualification (Supplemental Table 24). In the multivariable MR framework, the conditional F statistics were in general above 10 (indicating little evidence of weak instrument bias) for both our exposures of interest and the adjusting factors with a few exceptions. For models including television watching and years of education, conditional F statistics for both variables were below 10. Also, adjusting for BMI or years of education resulted in low F statistics (<10) for computer use.

MR estimates for the bidirectional MR

In post-hoc analyses, inverse bidirectional associations were observed between the genetically-predicted duration of leisure television watching and years of education. A one SD increase in genetically-predicted duration of leisure television watching reduced years of education by 0.54 SD (95% CI: -0.58 to -0.49). Similarly, a one SD increase in genetically-predicted years of education reduced duration of leisure television watching by 0.63 SD (95% CI: -0.66 to -0.59) (Figure 2, Supplemental Tables 25,26). These observations taken together with the inverse effect estimate found for years of years of education with breast and colorectal cancer (Supplemental Table 27) point to education having a complex dual confounding and mediating role in the association between television watching with breast and colorectal cancer risk. Contrary to this, positive bidirectional associations were observed for genetically-predicted duration of leisure computer use ($beta_{computer\ use \rightarrow education}$0.59; 95% CI: 0.48 to 0.70 and $beta_{education \rightarrow computer\ use}$0.34; 95% CI: 0.30 to 0.37). Additionally, positive bidirectional associations were observed between the genetically-predicted duration of leisure television watching with BMI and smoking status while, inverse bidirectional associations were observed between the genetically-predicted duration of leisure

computer use and smoking status. Finally, alcohol consumption was inversely associated with computer use (Figure 2, Supplemental Tables 25,26).

Discussion

In this MR analysis, a high level of genetically-predicted television watching increased risks of breast and colorectal cancer. The effect estimates for television watching were robust according to most of the univariable sensitivity analyses conducted to assess the influence of pleiotropy. After multivariable MR adjustment for years of education, the positive effects were attenuated; however, our post-hoc analyses suggest that education has a complex dual confounding and mediating role in the association between television watching with these cancers and adjustment for years of education is not appropriate. We found little evidence that genetically-predicted leisure computer use was associated with breast, colorectal, and prostate cancer.

Inconsistent results have been reported in prospective cohort studies that have examined the association between sedentary behaviours and breast cancer risk. A recent meta-analysis reported a statistically significant 10% higher risk for the highest sedentary behaviour group when compared with the lowest group (8). However, a recent study in UK Biobank found little evidence of any association between hours spent watching television and the risk of breast cancer (OR per 1 hour increase: 1.01, 95% CI: 0.99, 1.03) (9). In our analysis we initially observed positive associations between hours of television watching and the risk of breast cancer. However, these positive effect estimates were attenuated towards the null in our multivariable MR models adjusting for other risk factors, particularly years of education.

Numerous observational studies have investigated the associations between sedentary behaviours and colorectal cancer risk. Results from the most recent meta-analysis of case-control and cohort studies reported a non-significant 10% risk increase for colorectal cancer

for the highest sedentary behaviour group when compared with the lowest group (RR=1.10, 95% CI: 0.96–1.26) (8). Television viewing time has been the most investigated sedentary behaviour trait and positive associations have been found with colon cancer (9, 38). A recent UK Biobank analysis reported that higher levels of television watching time were associated with greater colon cancer risk (HR per 1-hour increase, 1.04, 95% CI: 1.01–1.07; P-value=0.016), but not rectal cancer (9). The same UK Biobank study found no association between leisure computer use and colorectal cancer risk (9). Results from our univariable MR analyses were generally consistent with this prior observational evidence, with positive effect estimates found for television watching, and little evidence of an association between computer use and colorectal cancer risk, except of rectal cancer. However, like our breast results, these associations attenuated towards the null in multivariable MR models adjusted for years of education and smoking (colorectal; television watching) or alcohol (rectal; computer use).

We found little evidence of any associations between sedentary behaviours and prostate cancer risk, consistent with prior observational evidence (9, 38). The null effects we found were similar for overall and aggressive prostate cancer risk.

Strong genetic correlations have been reported between each of television watching (inverse) and computer use (positive) and years of education (r_g^{TV}= -0.79 and r_g^{PC}= 0.53) (16). The low conditional F statistics in our multivariable models including the sedentary behaviour traits with years of education provided a further indicator of strong correlations. A recent MR study reported an inverse association between years of education and breast (OR, 0.89, 95% CI: 0.83, 0.96; P-value = 0.001) and a positive association for prostate cancer (OR, 1.10, 95% CI: 1.01, 1.21; P-value = 0.035) (39). In agreement with that, we observed inverse effect estimates for years of education in our multivariable models for breast and also for colorectal cancer. An additional MR study found that higher educational attainment levels were further

inversely associated with smoking, BMI, and sedentary behaviours, and positively with vigorous physical activity levels and alcohol consumption (40). Therefore, education may be a proxy for overall lifestyle, with higher educated individuals practising healthier lifestyle behaviours and actively participating in screening programs that lower their risk of developing cancer (39). Additionally, traits like sedentary behaviours, education, smoking, alcohol consumption, and obesity are correlated and it is therefore difficult to disentangle their complex interrelationships. As an example, in our post-hoc analyses we found evidence of education having a dual confounding and mediating role in the association between television watching with breast and colorectal cancers.

The main strength of the current study is the use of large-scale summary genetic data from consortia and the UK Biobank that allowed us to investigate the role of leisure sedentary behaviours on risk of developing breast, colorectal, and prostate cancer. A limitation of our study is that leisure sedentary behaviours were derived from self-reported questionnaires that are prone to measurement error (41, 42). An alternative approach is to use genetic instruments derived from objectively measured levels of physical activity using accelerometer data from the UK Biobank. However, a current limitation is that the number of genetic instruments is small as the GWAS on accelerometer data was analysed in a subset of 90,000 participants. Analysing two highly correlated phenotypes together, like sedentary behaviours and years of education may have introduced collinearity which leads to greater imprecision and possible bias. Furthermore, caution is needed regarding the results from the analyses for leisure computer use as the genetic instruments explained a very small proportion of the phenotypic variance resulting in a low powered analysis. Also, our analyses focused solely on leisure sedentary behaviours. The genetic correlation between television watching and objectively measured sedentary behaviour in UK Biobank was weak while, the correlation for computer use was higher (r_g^{TV}= 0.14 and r_g^{PC}= 0.46) (16). This can be at least partially explained from

the fact that accelerometers measure total but not domain-specific sedentary time (e.g., television watching) and that has been observed in previous observational studies (3, 43). Therefore, our results cannot be generalised to overall sedentary behaviour. Finally, the results cannot be generalised to diverse populations due to the lack of ancestral diversity in UK Biobank.

In conclusion, we found that higher genetically predicted television watching time increased risks of breast and colorectal cancer in univariable models. When we adjusted for years of education in multivariable MR models, these positive effect estimates were no longer present. However, these multivariable results should be interpreted cautiously as we detected evidence of education having a dual confounding and mediating role in the associations between television watching with risks of breast and colorectal cancer. Future analyses utilising objective measures of exposure (e.g., accelerometers) and novel analytic frameworks (e.g., target trial emulation) are required to provide new insights into the possible role of sedentary behaviour in cancer development.

References

1. Sung H, Ferlay J, Siegel RL, *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021;71(3):209-249.
2. Tremblay MS, Aubert S, Barnes JD, *et al.* Sedentary Behavior Research Network (SBRN) - Terminology Consensus Project process and outcome. Int J Behav Nutr Phys Act 2017;14(1):75.
3. Prince SA, LeBlanc AG, Colley RC, *et al.* Measurement of sedentary behaviour in population health surveys: a review and recommendations. PeerJ 2017;5:e4130.
4. Yang L, Cao C, Kantor ED, *et al.* Trends in Sedentary Behavior Among the US Population, 2001-2016. JAMA 2019;321(16):1587-1597.
5. British heart foundation. *Physical Inactivity and Sedentary Behaviour Report 2017*. https://www.bhf.org.uk/informationsupport/publications/statistics/physical-inactivity-report-2017.
6. Biswas A, Oh PI, Faulkner GE, *et al.* Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis. Ann Intern Med 2015;162(2):123-32.
7. Patterson R, McNamara E, Tainio M, *et al.* Sedentary behaviour and risk of all-cause, cardiovascular and cancer mortality, and incident type 2 diabetes: a systematic review and dose response meta-analysis. Eur J Epidemiol 2018;33(9):811-829.
8. Lynch BM, Mahmood S, Boyle T. Sedentary Behaviour and Cancer. In: Leitzmann MF, Jochem C, Schmid D, (eds). *Sedentary Behaviour Epidemiology*. Cham: Springer International Publishing; 2018, 245-298.
9. Hunter RF, Murray JM, Coleman HG. The association between recreational screen time and cancer risk: findings from the UK Biobank, a large prospective cohort study. Int J Behav Nutr Phys Act 2020;17(1):97.
10. Lawlor DA, Davey Smith G, Kundu D, *et al.* Those confounded vitamins: what can we learn from the differences between observational versus randomised trial evidence? Lancet 2004;363(9422):1724-7.
11. Davey Smith G, Ebrahim S. Epidemiology--is it time to call it a day? Int J Epidemiol 2001;30(1):1-11.
12. Prince SA, Reed JL, McFetridge C, *et al.* Correlates of sedentary behaviour in adults: a systematic review. Obes Rev 2017;18(8):915-935.
13. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol 2003;32(1):1-22.
14. Lawlor DA, Harbord RM, Sterne JA, *et al.* Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med 2008;27(8):1133-63.
15. Gao Y, Mi J, Liu Z, *et al.* Leisure Sedentary Behavior and Risk of Lung Cancer: A Two-Sample Mendelian Randomization Study and Mediation Analysis. Front Genet 2021;12:763626.
16. van de Vegte YJ, Said MA, Rienstra M, *et al.* Genome-wide association studies and Mendelian randomization analyses for leisure sedentary behaviours. Nat Commun 2020;11(1):1770.
17. Huyghe JR, Bien SA, Harrison TA, *et al.* Discovery of common and rare genetic risk variants for colorectal cancer. Nat Genet 2019;51(1):76-87.
18. Schumacher FR, Al Olama AA, Berndt SI, *et al.* Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. Nat Genet 2018;50(7):928-936.
19. Zhang H, Ahearn TU, Lecarpentier J, *et al.* Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. Nat Genet 2020;52(6):572-581.
20. Brion MJ, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. Int J Epidemiol 2013;42(5):1497-501.

21. Higgins JP, Thompson SG, Deeks JJ*, et al.* Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557-60.
22. Bowden J, Del Greco MF, Minelli C*, et al.* A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. Stat Med 2017;36(11):1783-1802.
23. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44(2):512-25.
24. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. Eur J Epidemiol 2017;32(5):377-389.
25. Bowden J, Davey Smith G, Haycock PC*, et al.* Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol 2016;40(4):304-14.
26. Verbanck M, Chen CY, Neale B*, et al.* Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet 2018;50(5):693-698.
27. Botteri E, Borroni E, Sloan EK*, et al.* Smoking and Colorectal Cancer Risk, Overall and by Molecular Subtypes: A Meta-Analysis. Am J Gastroenterol 2020;115(12):1940-1949.
28. Macacu A, Autier P, Boniol M*, et al.* Active and passive smoking and risk of breast cancer: a meta-analysis. Breast Cancer Res Treat 2015;154(2):213-24.
29. Bagnardi V, Rota M, Botteri E*, et al.* Alcohol consumption and site-specific cancer risk: a comprehensive dose-response meta-analysis. Br J Cancer 2015;112(3):580-93.
30. Pulit SL, Stoneman C, Morris AP*, et al.* Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. Hum Mol Genet 2019;28(1):166-174.
31. Lee JJ, Wedow R, Okbay A*, et al.* Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat Genet 2018;50(8):1112-1121.
32. Liu M, Jiang Y, Wedow R*, et al.* Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. Nat Genet 2019;51(2):237-244.
33. Wootton RE, Richmond RC, Stuijfzand BG*, et al.* Evidence for causal effects of lifetime smoking on risk for depression and schizophrenia: a Mendelian randomisation study. Psychol Med 2020;50(14):2435-2443.
34. Sanderson E, Davey Smith G, Windmeijer F*, et al.* An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. Int J Epidemiol 2019;48(3):713-727.
35. Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. Int J Epidemiol 2017;46(6):1734-1739.
36. Hemani G, Zheng J, Elsworth B*, et al.* The MR-Base platform supports systematic causal inference across the human phenome. Elife 2018;7.
37. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. PLoS Genet 2017;13(11):e1007081.
38. Schmid D, Leitzmann MF. Television viewing and time spent sedentary in relation to cancer risk: a meta-analysis. J Natl Cancer Inst 2014;106(7).
39. Yuan S, Xiong Y, Michaelsson M*, et al.* Genetically predicted education attainment in relation to somatic and mental health. Sci Rep 2021;11(1):4296.
40. Davies NM, Hill WD, Anderson EL*, et al.* Multivariable two-sample Mendelian randomization estimates of the effects of intelligence and education on health. Elife 2019;8.
41. Sargan JD. The Estimation of Economic Relationships using Instrumental Variables. Econometrica 1958;26(3):393-415.
42. Urda JL, Larouere B, Verba SD*, et al.* Comparison of subjective and objective measures of office workers' sedentary time. Prev Med Rep 2017;8:163-168.

43. Clark BK, Healy GN, Winkler EA*, et al.* Relationship of television time with accelerometer-derived sedentary time: NHANES. Med Sci Sports Exerc 2011;43(5):822-8.

Figure 1: Associations of leisure time television watching and computer use with breast and colorectal cancer after adjusting for the four secondary traits. The black dot corresponds to the 1-SD odds ratio and the corresponding error bar to the 95% confidence interval. Abbreviations: BMI: Body mass index; IVW: inverse variance weighting; SD: standard deviation

Figure 2: Bidirectional associations of leisure time television watching and computer use with the four secondary traits: BMI, years of education, smoking, and alcohol. The solid lines correspond to the effects of time television watching and computer use on the four secondary traits while the dashed lines correspond to the effects of the four secondary traits on time television watching and computer use. The black colour corresponds to statistically significant associations and the grey colour to non-significant. All the results, odds ratios and 95% confidence intervals, correspond to a 1-SD change in the levels of the variables. Abbreviations: BMI: Body mass index

Table1: Mendelian Randomization estimates for sedentary behaviour and breast cancer risk

Methods	Leisure television watching				Leisure computer use			
	Estimates (OR)*	95% CI	P-value	P-value for pleiotropy† or heterogeneity‡	Estimates (OR)*	95% CI	P-value	P-value for pleiotropy† or heterogeneity‡
Breast cancer								
Inverse-variance weighted	1.15	1.05, 1.26	0.002	1×10^{-17}	1.01	0.84, 1.23	0.89	1×10^{-9}
MR-Egger	1.48	0.98, 2.23	0.06	0.22	0.69	0.19, 2.48	0.57	0.55
Weighted median	1.16	1.05, 1.27	0.003		1.06	0.87, 1.28	0.57	
MR-PRESSO	1.12	1.03, 1.20	0.008	3×10^{-8}	1.04	0.88, 1.23	0.62	8×10^{-4}
Luminal A breast cancer								
Inverse-variance weighted	1.20	1.06, 1.35	0.002	6×10^{-19}	1.06	0.84, 1.34	0.62	4×10^{-6}
MR-Egger	1.55	0.90, 2.69	0.11	0.34	1.58	0.35, 7.10	0.55	0.60
Weighted median	1.15	1.01, 1.31	0.03		1.06	0.83, 1.35	0.66	
MR-PRESSO	1.14	1.03, 1.26	0.01	3×10^{-7}	1.06	0.87, 1.31	0.54	0.003
Luminal B breast cancer								
Inverse-variance weighted	1.14	0.94, 1.38	0.19	0.03	0.89	0.58, 1.36	0.58	0.02
MR-Egger	1.16	0.47, 2.89	0.74	0.96	1.95	0.12, 30.3	0.63	0.57
Weighted median	1.13	0.86, 1.48	0.40		0.97	0.57, 1.67	0.92	
MR-PRESSO					0.82	0.57, 1.17	0.28	0.11
Luminal B HER2 negative breast cancer								
Inverse-variance weighted	1.14	0.96, 1.36	0.13	0.004	1.03	0.76, 1.40	0.84	0.19
MR-Egger	1.07	0.48, 2.39	0.86	0.88	0.27	0.04, 2.25	0.23	0.22
Weighted median	1.30	1.03, 1.63	0.03		1.15	0.76, 1.75	0.52	
MR-PRESSO								
HER2 enriched breast cancer								
Inverse-variance weighted	1.21	0.91, 1.60	0.19	0.02	0.67	0.40, 1.13	0.13	0.69
MR-Egger	1.31	0.35, 4.95	0.68	0.90	0.08	0.00, 2.16	0.13	0.20
Weighted median	1.25	0.84, 1.86	0.28		0.65	0.31, 1.35	0.25	
MR-PRESSO								
Triple negative breast cancer								
Inverse-variance weighted	1.16	0.99, 1.35	0.06	0.10	0.68	0.50, 0.93	0.02	0.24
MR-Egger	1.54	0.72, 3.29	0.27	0.45	0.41	0.05, 3.35	0.40	0.63
Weighted median	1.31	1.04, 1.67	0.02		0.73	0.47, 1.14	0.16	
MR-PRESSO								

Abbreviations: CI, confidence intervals; MR: Mendelian Randomization; OR: odds ratio; MR-PRESSO: MR pleiotropy residual sum and outlier test

* The estimates correspond to a standard deviation increase in duration of sedentary activity

† P-value or pleiotropy based on MR-Egger intercept

‡ P-value for heterogeneity based on Q statistic

Table2: Mendelian Randomization estimates for sedentary behaviour and colorectal cancer risk

Methods	Leisure television watching				Leisure computer use			
	Estimates (OR)*	95% CI	P-value	P-value for pleiotropy† or heterogeneity‡	Estimates (OR)*	95% CI	P-value	P-value for pleiotropy† or heterogeneity‡
Colorectal cancer								
Inverse-variance weighted	1.32	1.16, 1.49	2×10^{-5}	9×10^{-9}	0.90	0.70, 1.13	0.33	0.02
MR-Egger	1.35	0.76, 2.39	0.31	0.94	0.35	0.08, 1.55	0.17	0.21
Weighted median	1.40	1.20, 1.63	2×10^{-5}		1.08	0.81, 1.45	0.59	
MR-PRESSO								
Colorectal cancer in men								
Inverse-variance weighted	1.45	1.23, 1.67	5×10^{-6}	3×10^{-3}	0.79	0.61, 1.04	0.10	0.2
MR-Egger	1.72	0.84, 3.53	0.14	0.63	0.61	0.09, 4.06	0.61	0.79
Weighted median	1.52	1.23, 1.88	9×10^{-5}		0.76	0.51, 1.13	0.17	
MR-PRESSO								
Colorectal cancer in women								
Inverse-variance weighted	1.25	1.06, 1.46	0.007	0.003	1.02	0.74, 1.40	0.89	0.05
MR-Egger	1.02	0.50, 2.08	0.96	0.57	0.31	0.04,2.29	0.25	0.24
Weighted median	1.25	1.01, 1.54	0.04		1.20	0.81, 1.79	0.36	
MR-PRESSO					1.08	0.83, 1.42	0.58	0.27
Colon cancer								
Inverse-variance weighted	1.36	1.19, 1.57	2×10^{-5}	5×10^{-5}	0.90	0.72, 1.14	0.42	0.06
MR-Egger	1.48	0.78, 2.80	0.24	0.80	0.26	0.05, 1.42	0.12	0.14
Weighted median	1.49	1.25, 1.79	2×10^{-5}		0.96	0.68, 1.34	0.82	
MR-PRESSO								
Rectal cancer								
Inverse-variance weighted	1.60	1.32, 1.93	2×10^{-6}	8×10^{-7}	0.66	0.49, 0.89	0.006	0.57
MR-Egger	1.97	0.82, 4.71	0.13	0.63	0.88	0.13, 6.05	0.90	0.76
Weighted median	1.86	1.48, 2.36	3×10^{-7}		0.81	0.53, 1.25	0.34	
MR-PRESSO								

Abbreviations: CI, confidence intervals; MR: Mendelian Randomization; OR: odds ratio; MR-PRESSO: MR pleiotropy residual sum and outlier test

* The estimates correspond to a standard deviation increase in duration of sedentary activity

† P-value or pleiotropy based on MR-Egger intercept

‡ P-value for heterogeneity based on Q statistic

Table3: Mendelian Randomization estimates for sedentary behaviour and prostate cancer risk

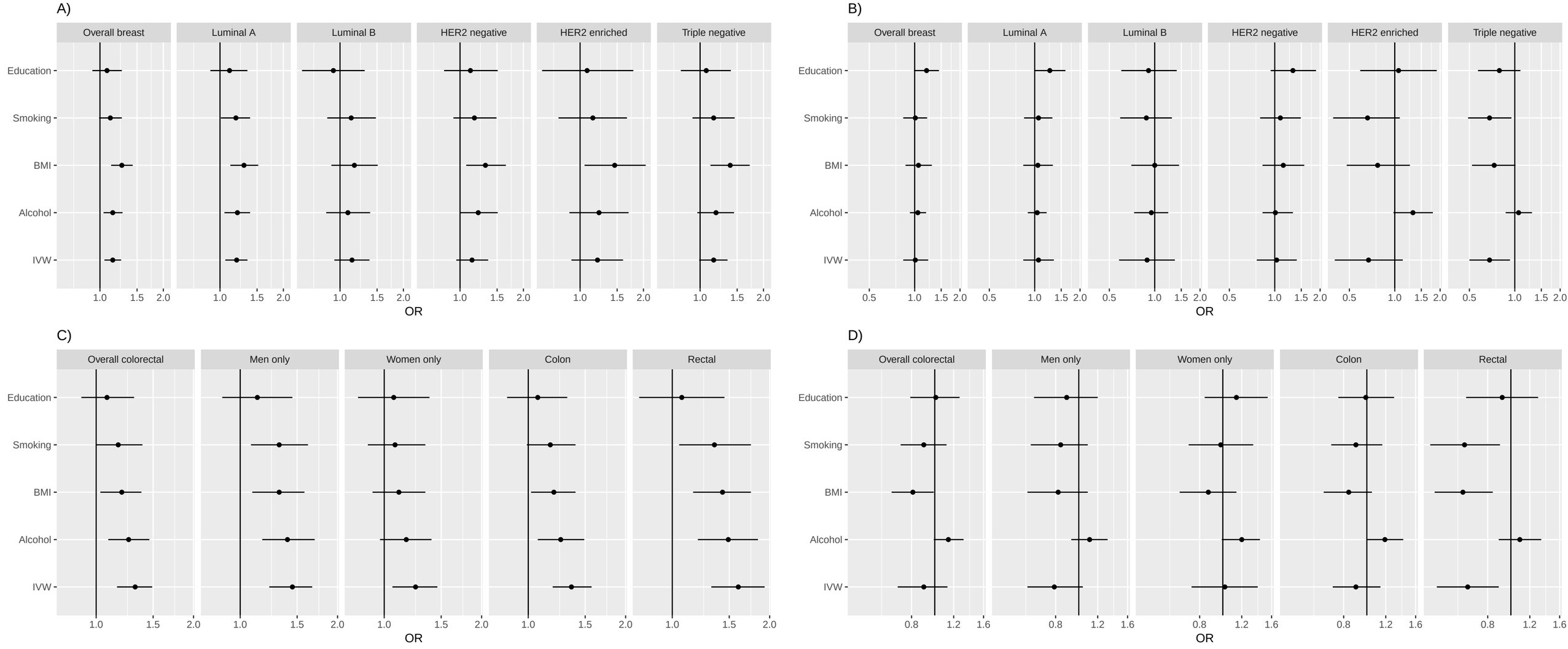
Methods	Leisure television watching				Leisure computer use			
	Estimates (OR)*	**95% CI**	**P-value**	**P-value for pleiotropy† or heterogeneity‡**	**Estimates (OR)***	**95% CI**	**P-value**	**P-value for pleiotropy† or heterogeneity‡**
Prostate cancer								
Inverse-variance weighted	0.94	0.84, 1.06	0.34	3×10^{-12}	1.08	0.89, 1.34	0.42	0.01
MR-Egger	1.19	0.71, 1.99	0.51	0.37	0.70	0.19, 2.56	0.59	0.5
Weighted median	0.94	0.83, 1.08	0.41		1.13	0.88, 1.46	0.33	
MR-PRESSO	0.92	0.84, 1.02	0.13	1×10^{-5}	1.14	0.96, 1.35	0.13	0.09
Advanced prostate cancer								
Inverse-variance weighted	0.95	0.81, 1.13	0.59	3×10^{-4}	0.91	0.69, 1.22	0.54	0.1
MR-Egger	1.46	0.68, 3.16	0.33	0.26	1.05	0.14, 8.17	0.96	0.89
Weighted median	0.82	0.66, 1.02	0.07		0.96	0.62, 11.48	0.84	
MR-PRESSO								

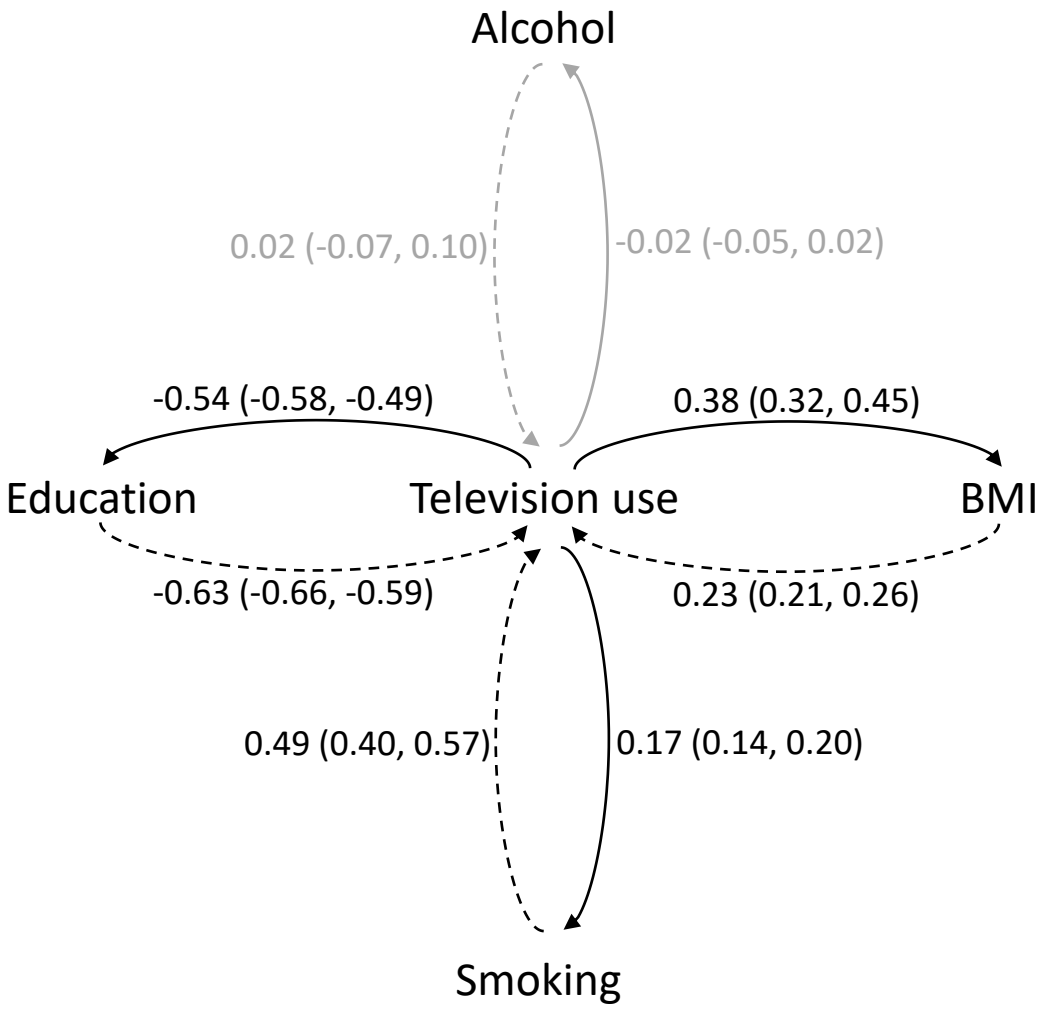
Abbreviations: CI, confidence intervals; MR: Mendelian Randomization; OR: odds ratio; MR-PRESSO: MR pleiotropy residual sum and outlier test

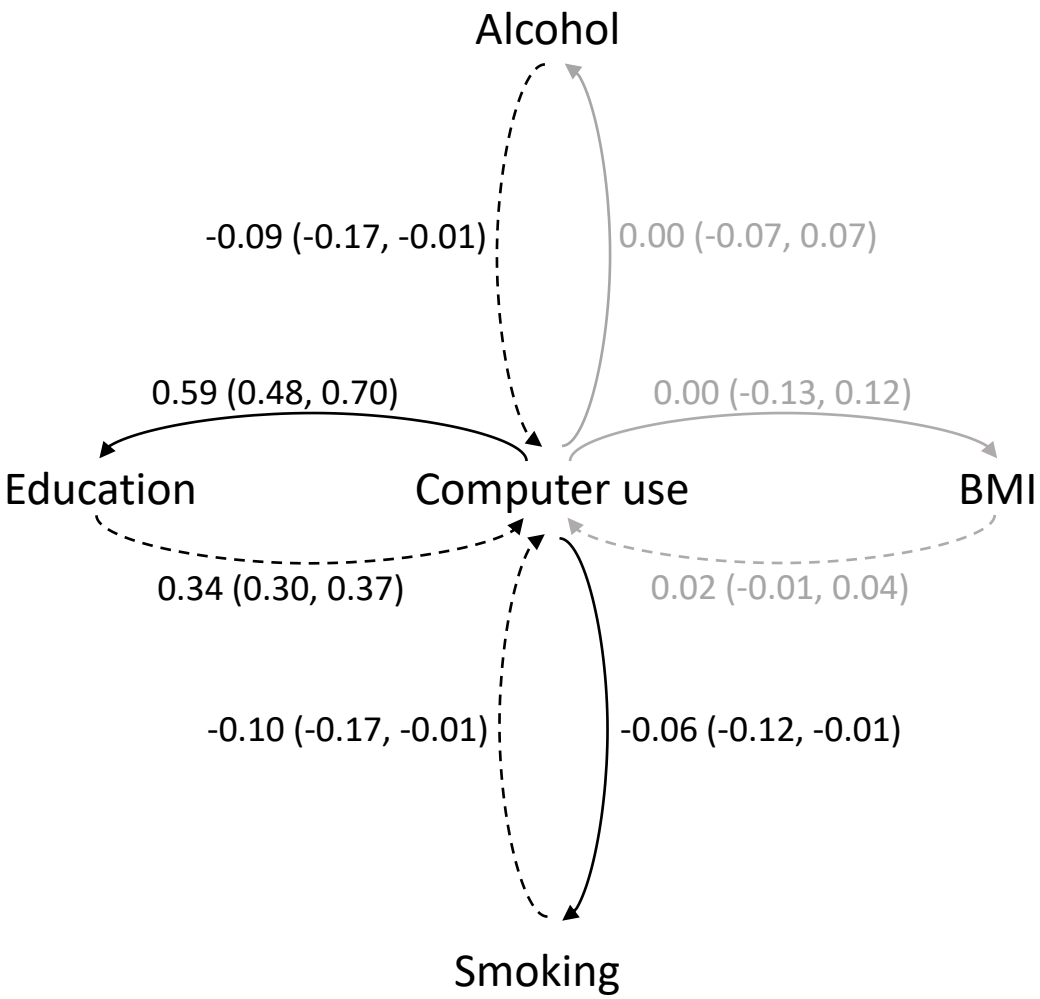
* The estimates correspond to a standard deviation increase in duration of sedentary activity

† P-value or pleiotropy based on MR-Egger intercept

‡ P-value for heterogeneity based on Q statistic


A)
Overall breast
Luminal A
Luminal B
HER2 negative
HER2 enriched
Triple negative
Education
Smoking
BMI
Alcohol
IVW
1.0 1.5 2.0
OR
B)
Overall breast
Luminal A
Luminal B
HER2 negative
HER2 enriched
Triple negative
Education
Smoking
BMI
Alcohol
IVW
0.5 1.0 1.5 2.0
OR
C)
Overall colorectal
Men only
Women only
Colon
Rectal
Education
Smoking
BMI
Alcohol
IVW
1.0 1.5 2.0
OR
D)
Overall colorectal
Men only
Women only
Colon
Rectal
Education
Smoking
BMI
Alcohol
IVW
0.8 1.2 1.6
OR


Alcohol
0.02 (-0.07, 0.10)
-0.02 (-0.05, 0.02)
-0.54 (-0.58, -0.49)
0.38 (0.32, 0.45)
Education
Television use
BMI
-0.63 (-0.66, -0.59)
0.23 (0.21, 0.26)
0.49 (0.40, 0.57)
0.17 (0.14, 0.20)
Smoking


Alcohol
-0.09 (-0.17, -0.01)
0.00 (-0.07, 0.07)
0.59 (0.48, 0.70)
0.00 (-0.13, 0.12)
Education
Computer use
BMI
0.34 (0.30, 0.37)
0.02 (-0.01, 0.04)
-0.10 (-0.17, -0.01)
-0.06 (-0.12, -0.01)
Smoking