

## **Reverse causal effect of atrial fibrillation on 17 site-specific cancer risk: A Mendelian randomization study**

### **Authors**

Sehoon Park, MD;<sup>1,2</sup> Soojin Lee, MD;<sup>3,4</sup> Yaerim Kim, MD, PhD;<sup>5</sup> Semin Cho, MD,<sup>3,4</sup>  
Kwangsoo Kim, PhD;<sup>6</sup> Yong Chul Kim, MD, PhD;<sup>3</sup> Seung Seok Han, MD, PhD;<sup>3,7</sup>  
Hajeong Lee, MD, PhD;<sup>3</sup> Jung Pyo Lee, MD, PhD;<sup>3,7,8</sup> Kwon Wook Joo, MD, PhD;<sup>3,4,7</sup>  
Chun Soo Lim, MD, PhD;<sup>3,7,8</sup> Yon Su Kim, MD, PhD;<sup>1,3,4,7</sup> and Dong Ki Kim, MD,  
PhD<sup>3,4,7</sup>

### **Author affiliations**

<sup>1</sup>Department of Biomedical Sciences, Seoul National University College of Medicine,  
Seoul, Korea

<sup>2</sup>Department of Internal Medicine, Armed Forces Capital Hospital, Gyeonggi-do, Korea

<sup>3</sup>Department of Internal Medicine, Seoul National University Hospital, Seoul, Korea

<sup>4</sup>Department of Internal Medicine, Seoul National University College of Medicine, Seoul,  
Korea

<sup>5</sup>Department of Internal Medicine, Keimyung University School of Medicine, Daegu,  
Korea

<sup>6</sup>Transdisciplinary Department of Medicine & Advanced Technology, Seoul National  
University Hospital, Seoul, Korea

<sup>7</sup>Kidney Research Institute, Seoul National University, Seoul, Korea

<sup>8</sup>Department of Internal Medicine, Seoul National University Boramae Medical Center,  
Seoul, Korea

**Corresponding author**

Dong Ki Kim, MD, PhD

Professor

Department of Internal Medicine, Seoul National University College of Medicine

101 Daehak-ro, Jongno-gu, Seoul, 03080, Korea

Tel.: 82-2-2072-2303, Fax: 82-2-745-2264

Email: [dkkim73@gmail.com](mailto:dkkim73@gmail.com)

## **Abstract**

**Background:** Bidirectional association between atrial fibrillation (AF) and cancer was reported by observational investigations. Additional study is warranted to investigate the causal effects of AF on risk of cancer.

**Methods:** This study was a summary-level Mendelian randomization (MR) analysis. Genetic instrument for AF was developed from a genome-wide association study (GWAS) meta-analysis for AF including 537,409 European ancestry individuals including 55,144 AF cases. The outcome data for risk of 17 site-specific cancer from a previous GWAS meta-analysis of the UK Biobank (48,961/359,825 case/controls) and Genetic Epidemiology Research on Adult Health and Aging (16,001/50,525 case/controls) cohorts including European ancestry individuals was investigated. Inverse variance weighted method was the main MR method, supported by pleiotropy-robust sensitivity analysis including MR-Egger regression and penalized weighted median method.

**Results:** The causal estimates indicated that AF was causally linked to higher risk of cancers of lung, breast, cervix, endometrium, and melanoma. MR-Egger test for directional pleiotropy indicated absence of a pleiotropy in the identified causal estimates and MR-Egger regression and median-based methods provided similarly significant findings. On the other hand, the genetic predisposition of AF was significantly associated with lower risk of esophagus/gastric cancer, but possibility of a directional pleiotropy remained in the association.

**Conclusions:** AF is a causal factor for certain types of cancer. Appropriate cancer screening should be suggested in clinical guidelines for AF patients. Future trial is

necessary to confirm whether appropriate management of AF may reduce the risk of cancer which is a major cause of deaths in AF patients.

## Introduction

Atrial fibrillation (AF) is the most prevalent cardiac arrhythmia worldwide.<sup>1</sup> Socioeconomic burden related to AF is substantial and further project to increase in the future along with global ageing trends. AF is a direct etiology for thromboembolism, causes compromise in cardiac output, and is associated with various health outcomes including cardiovascular diseases and mortality.

Cancer, another leading causes of mortality,<sup>2</sup> is commonly prevalent with AF in elderly individual and is a common cause of death in AF patients. Previous reports suggested that cancer increases the risk of AF through certain mechanisms.<sup>3</sup> However, a reverse-association has also been suggested, showing that presence of AF is an independent risk factor for future cancer.<sup>4-6</sup> However, as the findings are based on observational studies and delayed diagnosis of previous subclinical cancer is possible, the issue of reverse causation hinders the causal interpretation of the associations of AF and risk of cancer. Furthermore, as AF and cancer are both prevalent in elderly individuals commonly with multiple comorbidities, unmeasured confounding effect may bias the previous observational findings. Therefore, additional study is warranted to investigate the causal effects of AF on site-specific cancer risk, which would aid establishing new guidelines for cancer screenings in the emerging population with AF.<sup>4</sup>

Mendelian randomization (MR) is an analytic method to demonstrate causal effect by implementing genetic instruments.<sup>7</sup> As genetic instrument is determined before birth, causal estimates from genetic predisposition to an outcome are minimally affected by reverse causation or confounding effects. As those with a higher genetic predisposition for an exposure of interest would actually have higher prevalence of the exposure, a significant association between the genetic predisposition and outcome suggests a presence

of a causal effect. MR has been widely adopted in recent medical literature and identified important causal pathways between various exposures and complex health outcomes.<sup>8-10</sup>

In this study, we aimed to investigate the causal effect of AF on 17 site-specific cancer risks by MR analysis. We implemented large-scale genome-wide association study (GWAS) results for AF and various types of cancers and performed a large-scale summary-level MR analysis.

## **Methods**

### ***Ethical considerations***

The study was approved by the Institutional Review Boards of Seoul National University Hospital (IRB No. E-2012-004-1177). Our investigation for the UK Biobank data base has been approved by the consortium (application No. 53799). The requirement for informed consent was waived because the study analyzed public GWAS results and anonymous databases.

### ***Study setting***

The study was a summary-level MR. To limit the effects from different ethnic distributions in a genetic analysis, the studied data were restricted to those from European ancestry. The genetic instrument was developed from the previous GWAS meta-analysis for AF.<sup>11</sup> The individual-level data of the UK Biobank was implemented to identify genetic variants associated with potential confounders.<sup>12</sup> The outcome summary statistics for 17 site-specific cancer risks were provided from a previous GWAS meta-analysis, including UK Biobank and Genetic Epidemiology Research on Adult Health and Aging (GERA) cohorts (URL: [https://github.com/Wittelab/pancancer\\_pleiotropy](https://github.com/Wittelab/pancancer_pleiotropy)).<sup>13</sup>

### *Genetic instrument*

The previous GWAS meta-analysis identified single nucleotide polymorphisms (SNPs) showing genome-wide significant association with paroxysmal/persistent AF (Table 1).<sup>11</sup> The study meta-analyzed GWAS results from AFGen consortium Broad AF study, UK Biobank, and Biobank Japan. The SNPs reported in the study was relevant as being enriched in cardiac structural and electrophysiological pathways. Along with the findings from multi-ethnic population, the study reported 84 independent sentinel SNPs from the European ancestry individuals, including 537,409 European ancestry individuals including 55,144 AF cases. We constructed the genetic instrument from the 84 SNPs and Steiger filtering ensured that the genetic effects are from AF to outcome.<sup>14</sup> During trimming of the instrument, we considered the following core assumptions of MR.<sup>7</sup>

First, the relevance assumption means that the genetic instrument should be strongly associated with the exposure of interest. The previous GWAS meta-analysis already provided independent SNPs strongly associated with AF and we confirmed the association within the individual-level UK Biobank data by validating the explained variance by allele scores calculated from the genetic instrument.

Second, the independence assumption means that the genetic instrument should not be associated with confounders. We disregarded the SNPs that showed strong ( $P < 1 \times 10^{-5}$ ) association with phenotypical hypertension, diabetes mellitus, obesity, dyslipidemia, thyroid disorder,<sup>10</sup> and smoking by performing GWAS, with PLINK 2.0, with the individual-level UK Biobank data.<sup>15</sup> In the GWAS, the linear regression model was constructed adjusted for age, sex, age $\times$ sex, age<sup>2</sup>, and the first 10 principal components. Furthermore, we performed additional sensitivity MR analyses which are robust to pleiotropic effects and relax this second assumption for the genetic instruments.<sup>16,17</sup>

Third, the exclusion-restriction assumption means that the genetic effect should be through the exposure of interest. There is yet a formal test for this assumption, but median-based MR method, which was performed as a sensitivity MR analysis, can relax this assumption in up to half of the genetic instruments.<sup>17</sup>

### ***Individual-level UK Biobank data***

UK Biobank is a population-scale data including genetic and various phenotypical information. The UK Biobank included > 500,000 individuals aged 40-69 years across the United Kingdom, and the details have been published previously.<sup>12,18</sup> In this analysis, we included unrelated 337,138 white British UK Biobank participants after excluding those who were outliers due to heterozygosity or missing rates or had sex chromosome aneuploidy as the individual-level data.<sup>8</sup> The data has been utilized to identify confounder associated SNPs and to confirm whether the genetic instrument successfully predicted phenotypical AF. In the data, history of diabetes mellitus was self-reported, and hypertension or dyslipidemia history was identified through usage of relevant medication. Obesity was determined by baseline body mass index  $\geq 30$  kg/m<sup>2</sup>. Thyroid disease was identified by relevant diagnostic codes. The phenotypical AF was determined from the hospital records or death certificates including main causes of deaths, identified by an ICD-10 diagnostic code of I48 or ICD-9 diagnostic code of 4273.

### ***Outcome GWAS summary statistics for site-specific cancer***

The previous GWAS meta-analysis for pan-cancer risk was performed with European ancestry individuals in the UK Biobank and GERA cohort (Table 1).<sup>12,13,19</sup> The study investigated 48,961/359,825 case/controls in the UK Biobank data and 16,001/50,525 case/controls in the GERA data. The study provided summary statistics for cancers of



oropharynx, thyroid, lung, breast, pancreas, esophagus/stomach, colon, rectum, prostate, kidney, bladder, cervix, endometrium, ovary, lymphocytic leukemia, melanoma, and non-Hodgkin lymphoma. In the UK Biobank data, cancer histories were identified through national cancer registries in the United Kingdom collecting cancer cases with medical records, death certificates, and other sources from early 1970 to 2015. In the GERA data, the participants were receiving Kaiser Permanente Northern (KPNC) health care plan and the registry collected treatment/diagnosis history of cancer from 1988 to 2016. In the meta-analysis, the first cancer diagnosis determined the cancer case, and the individuals without any cancer history constructed the control group. Those with a cancer history which was not the type of interest in the meta-analysis were excluded. Along with novel SNPs predicting site-specific cancer risks, the study had strength that previously reported SNPs in the literature were replicated in the results.

### ***MR methods***

We performed summary-level MR analysis. During harmonization of the data, overlapping SNPs between the genetic instrument and the outcome data were utilized, and we conservatively excluded the palindromic SNPs because whether the effect sizes were from the alleles in the same DNA strand may be questioned for such variants. We performed multiplicative random effect inverse variance weighted method as the main MR method, which allows balanced pleiotropic effect when heterogeneity among the instrument is present.<sup>20</sup> Next, we performed MR-Egger regression analysis with bootstrapped standard error to yield pleiotropy-robust causal estimates.<sup>16</sup> The MR-Egger regression allows pleiotropic effect for all of the genetic instruments but still can calculate valid causal estimates. In addition, MR-Egger intercept reflects the presence of directional pleiotropy, thus can be implemented as a formal test for a pleiotropic effect. The weakness

of the MR-Egger regression method is the weak statistical power and the method can be biased when group of instruments act through a same pleiotropic pathway [violation of the instrument Strength Independent of Direct Effect (InSIDE) assumption],<sup>21</sup> thus additional median-based sensitivity analysis is encouraged to validate the results. Therefore, we performed penalized weighted median method,<sup>17</sup> which provides valid causal estimates even up to half of the genetic instruments are invalid. The method has particular strength as it can relax the untestable exclusion-restriction assumption in 50% of the genetic instruments and less biased by violation of the InSIDE assumption. Then, we performed MR-PRESSO analysis, which detects and corrects the effects from outliers.<sup>22</sup> The causal estimates were calculated by MR-PRESSO when the analysis identified a correctable outlier effect. Within instrument heterogeneity was identified through Cochran's Q statistics P value.

Lastly, additional sensitivity analysis was performed considering the sample overlap (UK Biobank data) in the GWAS for genetic instrument develop and for for derivation of outcome summary statistics.<sup>23</sup> Such overlap may cause bias towards confounding effects, particularly when the instrumental power is weak. Following the previous literature,<sup>23</sup> we performed sensitivity analysis by including fewer SNPs but with stronger association strength, by including SNPs with  $P < 1 \times 10^{-10}$  association with AF as the genetic instrument.

A significant causal effect was addressed only when the inverse variance weighted method and all sensitivity MR methods provided consistently significant (two-sided P value  $< 0.05$ ) causal estimates. The summary-level MR was performed by the TwoSampleMR package in R (version 3.6.2, the R foundation).<sup>24</sup>

## Results

### ***Genetic instrument construction***

Of the 84 SNPs, there were 4 SNPs showing strong association with the potential confounders. After additionally disregarding non-overlapping or palindromic SNPs, 67 SNPs constructed the genetic instrument for AF. The explained variance for phenotypical AF by the allele scores for AF in the individual-level UK Biobank data was 2.4% (McFadden's pseudo-R square).

### ***Summary-level MR results***

The summary-level MR results are graphically shown in Figure 2 and the numbers are present in Table 2. A genetically predicted AF was significantly associated with higher risk of cancers in lung, breast, cervix, endometrium, and melanoma, supported by all sensitivity analysis results. In the causal estimates for the cancers significantly affected from genetic predisposition of AF, MR-Egger test for directional pleiotropy or Cochran's Q statistics did not identify a significant pleiotropic effect or a heterogeneity. Even when we reduced the number of SNPs as the genetic instruments by including only 47 SNPs with  $P < 1 \times 10^{-10}$  association with AF, the findings were consistent both by the inverse variance weighted method and the pleiotropy-robust MR analyses (Supplemental Table 1). The MR-PRESSO analysis did not detect correctable outlier effects, thus, calculating the causal estimates from MR-PRESSO was unnecessary.

The causal estimates for pancreas cancer risk indicated that genetical predisposition for AF may be significantly associated with higher risk of pancreas cancer. However, the findings were non-significant in the sensitivity analysis reducing the number of SNPs by applying more stringent threshold.

On the other hand, higher genetic predisposition for AF was significantly associated with lower risk of esophagus/gastric cancer. The results were consistent throughout the

performed sensitivity analyses. However, MR-Egger test for directional pleiotropy indicated that significant pleiotropic effect was detected in the causal estimate for the risk of esophagus or stomach cancer.

The genetically predicted AF was significantly associated with non-Hodgkin lymphoma in the main analysis, but the significance was attenuated when we included fewer but stronger SNPs as the genetic instrument considering the overlapping sample issue.

## **Discussion**

In this MR analysis, we identified that AF may cause cancers of lung, breast, cervix, endometrium, and melanoma. The genetic predisposition of AF was significantly associated with lower risk of esophagus/gastric cancer, but possibility of a directional pleiotropy remained in the association. The findings suggest that AF is a causative factor for certain site-specific cancer risks.

The previous observational findings suggested bidirectional association between AF and cancer. The studies reporting that AF was higher incident cancer risks raised question whether AF may have causal effect on certain risks of malignancies.<sup>5,25</sup> However, as delayed diagnosis for AF and cancer is possible and the diseases share common risk factors, the presence of causal effect of AF on cancer was only suspected but yet confirmed.<sup>26</sup> In this MR analysis, we constructed genetic instrument for AF by implementing the largest GWAS result to date and applied to one of the largest GWAS meta-analyses for site-specific cancer risks by the summary-level MR. With our efforts to attain the key assumptions of MR, we identified that AF causally affects risks of certain site-specific cancers, which has been newly identified by this methodology.

Our findings may serve as an evidence for cancer screening guideline for AF patients. In AF patients, additional screening may be necessary for lung, breast, cervix, endometrium, and melanoma cancers because AF causally elevates the cancer risks. Or, screening of AF may be considered in the individuals with such cancers, as subclinical AF is common and may affect cancer progression considering its' present effect on cancer development. As non-cardiovascular mortality in AF patients is common and largely attributed by malignancies,<sup>27</sup> such strategy may ameliorate the burden of cancer-related complications and risk of mortality in AF patients by promoting early diagnosis.<sup>26,28</sup>

Our study encourages future studies for the mechanism regarding the identified causal effects. It is well described that AF causes systemic hypercoagulability and directly, formation of atrial thrombosis causing arterial thromboembolism.<sup>29,30</sup> As patients with primary venous thromboembolism showed higher risk of cancers without evidence of occult malignancy, vascular thrombosis has been suspected to have causal effect on cancer development.<sup>31,32</sup> In addition, two studies reported that not only venous, but also arterial thrombosis was associated with increased risk of incident cancer identification, from 746862 cancer patients in the United States<sup>33</sup> and from 6600 Danish patients diagnosed for low limb arterial thrombosis.<sup>34</sup> Although the possibility of delayed-diagnosis for occult malignancy after identification of arterial thrombosis cannot be disregarded, certain portion of the incident cancer events might have been related to causal effects from impaired circulation due to vascular thrombosis. Such hypothesis was supported by the previous observational findings suggesting anti-tumor effect of anti-thrombotic drugs.<sup>26,28,35</sup> Linkage between hypercoagulability state and cancer risk has been recently suggested.<sup>36</sup> Therefore, that AF provoking vascular thrombosis, hypercoagulability, and impaired circulation may be the vertical pathway of the causal effect from AF on risk of cancer development. A future study targeting the thrombogenic pathway as an etiology of

cancer development is warranted. In addition, the mechanism that the causal effects were apparent for specific cancer types should be studied. Furthermore, if the biologic pathway is relevant, additional trial may test the efficacy of appropriate usage of preventive anticoagulative therapy or prompt rhythm control for reducing the cancer risks in AF patients.<sup>28,37</sup> On the other hand, AF may have causal effect on risk of certain types of cancer through undetermined pathway, and hemodynamic effect or neurohormonal response related to AF may be investigated as a potential mechanism.

That genetic predisposition for AF was associated with lower risk of gastric/esophageal cancer needs to be explained, although the identification for direct mechanism is beyond the current study's possibility. As AF directly reducing a cancer risk is biologically less plausible, it possible that, as cancer cases are determined by the first cases in the previous study, an early diagnosis of AF-affected cancers might have led to under inclusion of gastric-esophageal cancer cases in the outcome data. Or, considering the significant MR-Egger intercept P value indicating presence of a directional pleiotropy, group of the instrumented genetic variants might have acted through a pleiotropic pathway, which is a violation of the InSIDE assumption and causes false-positive findings even in the pleiotropy-robust sensitivity analyses. A future study is warranted to confirm the findings with additional dissection of the cancer subtypes, as various cancer types with different histology may occur in esophagus and stomach.

There are several limitations in this study. First, this study does not confirm whether appropriate management for AF, by rhythm control or by preventive antithrombotic treatment, may reduce cancer risks. Additional trial targeting the identified causal pathway is necessary to extend the clinical application of our findings, along with a study investigating the mechanistical explanation for the results. Second, the reason for the causal effects are different according to cancer types was not explained in this study. As a

false-negative bias is possible in MR, as genetic instrument explains a portion of the exposure of interest, other cancer risks may also be affected from AF. Yet, the identified cancers that are causally affected from AF may be prioritized for future studies investigating the efficacy of management of AF on reducing risk of a type of cancer. Third, the clinical effect size of AF on cancer risks may be different from the identified genetic effects which were small.<sup>38</sup> The genetic effect sizes from MR might have been shrunken due to the availability of genetic information to explain AF and the relatively small cases of each cancer compared to the controls. Or it is possible that AF may actually have relatively small effect on cancer risk than that from other common etiologies of malignancy. Fourth, due to the limitation of the available genetic information, detailed subtypes of atrial fibrillation (persistent or paroxysmal) or cancer (histologic classification) were not considered. Lastly, the study population was limited to European ancestry individuals, thus, generalizability of our finding is not secured for other ethnic populations.

In conclusion, AF is a causal factor for certain types of cancer. Appropriate cancer screening should be suggested in clinical guidelines for AF patients. Future trial is necessary to confirm whether appropriate management of AF may reduce the risk of cancer which is a major cause of deaths in AF patients.

## **Acknowledgements**

The study was based on the data provided by the UK Biobank consortium (application No. 53799). We thank the investigators of the previous studies who provided the valuable genetic summary statistics for this study.

## **Funding**

This work was supported by the Industrial Strategic Technology Development Program - Development of bio-core technology (10077474, Development of early diagnosis technology for acute/chronic renal failure) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea) and a grant from SNU R&DB Foundation (800-20190571). The study was performed independently by the authors.

## **Conflicts of interest**

EKC: research grants from Bayer, BMS/Pfizer, Biosense Webster, Chong Kun Dang, Daiichi-Sankyo, Dreamtech Co. Ltd., Medtronic, Samjinpharm, Sanofi-Aventis, Seers Technology, Skylabs, and Yuhan. No fees are received personally. There is no other conflicts of interests.

## **Data Availability Statement**

The main data for this study are archived in public repositories by previous studies.



## References

1. Chugh SS, Havmoeller R, Narayanan K, et al. Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. *Circulation*. 2014;129:837-847.
2. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018;392:1736-1788.
3. Farmakis D, Parissis J, Filippatos G. Insights into onco-cardiology: atrial fibrillation in cancer. *J Am Coll Cardiol*. 2014;63:945-953.
4. Aboumsallem JP, Moslehi J, de Boer RA. Reverse Cardio-Oncology: Cancer Development in Patients With Cardiovascular Disease. *J Am Heart Assoc*. 2020;9:e013754.
5. Conen D, Wong JA, Sandhu RK, et al. Risk of Malignant Cancer Among Women With New-Onset Atrial Fibrillation. *JAMA Cardiol*. 2016;1:389-396.
6. Vinter N, Christesen AMS, Fenger-Grøn M, Tjønneland A, Frost L. Atrial Fibrillation and Risk of Cancer: A Danish Population-Based Cohort Study. *J Am Heart Assoc*. 2018;7:e009543.
7. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *Bmj*. 2018;362:k601.
8. Park S, Lee S, Kim Y, et al. Short or Long Sleep Duration and CKD: A Mendelian Randomization Study. *J Am Soc Nephrol*. 2020;31:2937-2947.
9. Ong JS, Law MH, An J, et al. Association between coffee consumption and overall risk of being diagnosed with or dying from cancer among >300 000 UK Biobank participants in a large-scale Mendelian randomization study. *Int J Epidemiol*. 2019;48:1447-1456.

10. Ellervik C, Roselli C, Christophersen IE, et al. Assessment of the Relationship Between Genetic Determinants of Thyroid Function and Atrial Fibrillation: A Mendelian Randomization Study. *JAMA Cardiol.* 2019;4:144-152.
11. Roselli C, Chaffin MD, Weng LC, et al. Multi-ethnic genome-wide association study for atrial fibrillation. *Nat Genet.* 2018;50:1225-1233.
12. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature.* 2018;562:203-209.
13. Rashkin SR, Graff RE, Kachuri L, et al. Pan-cancer study detects genetic risk variants and shared genetic basis in two large cohorts. *Nat Commun.* 2020;11:4423.
14. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet.* 2017;13:e1007081.
15. Park S, Lee S, Kim Y, et al. Atrial fibrillation and kidney function: A bidirectional Mendelian randomization study. *medRxiv.* 2020:2020.2007.2031.20166207.
16. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol.* 2015;44:512-525.
17. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol.* 2016;40:304-314.
18. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *Am J Epidemiol.* 2017;186:1026-1034.
19. Banda Y, Kvale MN, Hoffmann TJ, et al. Characterizing Race/Ethnicity and Genetic Ancestry for 100,000 Subjects in the Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort. *Genetics.* 2015;200:1285-1295.

20. Burgess S, Davey Smith G, Davies NM, et al. Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res.* 2019;4:186.
21. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol.* 2017;32:377-389.
22. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet.* 2018;50:693-698.
23. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. *Genet Epidemiol.* 2016;40:597-608.
24. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife.* 2018;7.
25. Erichsen R, Christiansen CF, Mehnert F, Weiss NS, Baron JA, Sørensen HT. Colorectal cancer and risk of atrial fibrillation and flutter: a population-based case-control study. *Intern Emerg Med.* 2012;7:431-438.
26. Rahman F, Ko D, Benjamin EJ. Association of Atrial Fibrillation and Cancer. *JAMA Cardiol.* 2016;1:384-386.
27. Marijon E, Le Heuzey JY, Connolly S, et al. Causes of death and influencing factors in patients with atrial fibrillation: a competing-risk analysis from the randomized evaluation of long-term anticoagulant therapy study. *Circulation.* 2013;128:2192-2201.
28. Tagalakis V, Tamim H, Blostein M, Collet JP, Hanley JA, Kahn SR. Use of warfarin and risk of urogenital cancer: a population-based, nested case-control study. *Lancet Oncol.* 2007;8:395-402.
29. Sohara H, Amitani S, Kurose M, Miyahara K. Atrial fibrillation activates platelets and coagulation in a time-dependent manner: a study in patients with paroxysmal atrial fibrillation. *J Am Coll Cardiol.* 1997;29:106-112.

30. Hatzinikolaou-Kotsakou E, Kartasis Z, Tziakas D, et al. Atrial fibrillation and hypercoagulability: dependent on clinical factors or/and on genetic alterations? *J Thromb Thrombolysis*. 2003;16:155-161.
31. Sørensen HT, Mellekjaer L, Steffensen FH, Olsen JH, Nielsen GL. The risk of a diagnosis of cancer after primary deep venous thrombosis or pulmonary embolism. *N Engl J Med*. 1998;338:1169-1173.
32. Baron JA, Gridley G, Weiderpass E, Nyrén O, Linet M. Venous thromboembolism and cancer. *Lancet*. 1998;351:1077-1080.
33. Navi BB, Reiner AS, Kamel H, et al. Arterial thromboembolic events preceding the diagnosis of cancer in older persons. *Blood*. 2019;133:781-789.
34. Sundbøll J, Veres K, Horváth-Puhó E, Adelborg K, Sørensen HT. Risk and Prognosis of Cancer After Lower Limb Arterial Thrombosis. *Circulation*. 2018;138:669-677.
35. Eischer L, Kammer M, Traby L, Kyrle PA, Eichinger S. Risk of cancer after anticoagulation in patients with unprovoked venous thromboembolism: an observational cohort study. *J Thromb Haemost*. 2017;15:1368-1374.
36. Tavares V, Pinto R, Assis J, Pereira D, Medeiros R. Venous thromboembolism GWAS reported genetic makeup and the hallmarks of cancer: Linkage to ovarian tumour behaviour. *Biochim Biophys Acta Rev Cancer*. 2020;1873:188331.
37. Schulman S, Lindmarker P. Incidence of cancer after prophylaxis with warfarin against recurrent venous thromboembolism. *Duration of Anticoagulation Trial*. *N Engl J Med*. 2000;342:1953-1958.
38. Burgess S, Butterworth A, Malarstig A, Thompson SG. Use of Mendelian randomisation to assess potential benefit of clinical intervention. *BMJ*. 2012;345:e7325.



## Figure Legend

**Figure 1. Study flow diagram.** SNPs = single nucleotide polymorphism, GWAS = genome-wide association study, GERA = Genetic Epidemiology Research on Adult Health and Aging.

**Figure 2. Causal estimates from kidney function on site-specific cancer risk.** The x-axes indicate the odds ratios and confidence intervals, from one standard deviation increase in the genetic predisposition for atrial fibrillation. The vertical broken line indicates 1 OR. The black round causal estimates are from the multiplicative random effect inverse variance weighted method. The white round causal estimates are from the MR-Egger regression with bootstrapped standard error. The white square causal estimates are from the penalized weighted median method.

**Table 1.** Summary of the study population of the previous genome-wide association study meta-analysis.

Component of this study	Phenotype	Included cohort	Study population which derived the summary statistics	Remarks
Genetic instrument	Paroxysmal/persistent AF	Summary statistics from AFGen consortium Broad AF study, UK Biobank, and Biobank Japan	537,409 European ancestry individuals with 55,144 AF cases	Identified independent 84 sentinel SNPs with a genome-wide significant level $P < 1 \times 10^{-8}$ ) association with AF risk
Outcome summary statistics <sup>a</sup>	First identified site-specific cancer	UK Biobank	408,786 European ancestry individuals with 48,961 cancer cases	Cancer types were grouped to 17 site-specific cancer types to secure statistical power. Testis cancer data was obtained only in the UK Biobank due to data availability.
		Genetic Epidemiology Research on Adult Health and Aging (GERA)	66,526 European ancestry individuals with 16,001 cancer cases	

AF = atrial fibrillation, SNPs = single nucleotide polymorphisms

<sup>a</sup>The patient age and cancer grade/stage was various at the time of cancer diagnosis in the cohorts which derived the outcome summary statistics. The detailed information according to the types of site-specific cancer is available in the supplemental material attached to the previously published study by Sara R. Rashkin et al. (URL: [https://static-content.springer.com/esm/art%3A10.1038%2Fs41467-020-18246-6/MediaObjects/41467\\_2020\\_18246\\_MOESM1\\_ESM.pdf](https://static-content.springer.com/esm/art%3A10.1038%2Fs41467-020-18246-6/MediaObjects/41467_2020_18246_MOESM1_ESM.pdf)).

**Table 2.** Summary-level MR results.

Site-specific cancer	Number of cases (UK Biobank /GERA)	Cochran's Q statistics P value	MR-Egger test P value for directional	MR method	OR (95% CI)	P
Oropharynx	930/293	0.545	0.069	Inverse variance weighted	0.996 (0.984, 1.007)	0.475
				MR-Egger	1.003 (0.988, 1.017)	0.334
				Penalized weighted	0.976 (0.958, 0.994)	0.008
Thyroid	527/235	0.541	< 0.001	Inverse variance weighted	0.994 (0.986, 1.002)	0.161
				MR-Egger	0.99 (0.98, 0.999)	0.018
				Penalized weighted	0.994 (0.984, 1.004)	0.234
Lung	1728/757	0.840	0.205	Inverse variance weighted	1.012 (1.004, 1.021)	0.005
				MR-Egger	1.013 (1.002, 1.024)	0.012
				Penalized weighted	1.014 (1.003, 1.026)	0.017
Breast	13903/3978	0.240	0.306	Inverse variance weighted	1.003 (1.0005, 1.005)	0.017
				MR-Egger	1.003 (1.0005, 1.005)	0.007
				Penalized weighted	1.003 (1.0005, 1.005)	0.018
Pancreas	471/192	0.982	0.556	Inverse variance weighted	1.028 (1.01, 1.046)	0.002
				MR-Egger	1.031 (1.003, 1.06)	0.016
				Penalized weighted	1.031 (1.005, 1.057)	0.019
Esophagus or stomach	929/162	0.691	0.039	Inverse variance weighted	0.999 (0.998, 0.9996)	0.001
				MR-Egger	0.999 (0.998, 0.9997)	0.002
				Penalized weighted	0.999 (0.998, 0.9996)	0.003
Colon	2897/896	0.043	0.142	Inverse variance weighted	1 (0.998, 1.002)	0.909
				MR-Egger	0.998 (0.996, 1)	0.037
				Penalized weighted	1 (0.997, 1.003)	0.897
Rectum	1808/283	0.129	< 0.001	Inverse variance weighted	1.001 (0.994, 1.008)	0.794
				MR-Egger	1.009 (1.001, 1.017)	0.007
				Penalized weighted	0.999 (0.988, 1.009)	0.786
Prostate	7441/3351	0.034	0.039	Inverse variance weighted	0.998 (0.996, 1.001)	0.16
				MR-Egger	0.998 (0.996, 1.001)	0.074
				Penalized weighted	0.997 (0.994, 0.999)	0.01
Bladder	1550/692	0.914	0.230	Inverse variance weighted	0.991 (0.971, 1.011)	0.37
				MR-Egger	0.979 (0.95, 1.009)	0.091
				Penalized weighted	0.991 (0.953, 1.03)	0.647
Kidney	1021/317	0.370	0.054	Inverse variance weighted	0.994 (0.983, 1.005)	0.255
				MR-Egger	0.986 (0.971, 1.001)	0.033
				Penalized weighted	0.978 (0.96, 0.996)	0.018
Cervix	5998/565	0.893	0.119	Inverse variance weighted	1.006 (1.001, 1.01)	0.017
				MR-Egger	1.006 (1.001, 1.012)	0.013
				Penalized weighted	1.007 (1.001, 1.013)	0.015
Ovary	1006/253	0.755	0.885	Inverse variance weighted	1.013 (0.996, 1.031)	0.143
				MR-Egger	1.022 (0.998, 1.047)	0.033
				Penalized weighted	1.017 (0.995, 1.04)	0.137
Endometrium	1414/623	0.814	0.722	Inverse variance weighted	1.016 (1.009, 1.023)	< 0.001
				MR-Egger	1.013 (1.001, 1.025)	0.009
				Penalized weighted	1.015 (1.005, 1.025)	0.004
Lymphocytic leukemia	594/258	0.672	0.244	Inverse variance weighted	0.984 (0.967, 1.002)	0.081
				MR-Egger	0.975 (0.948, 1.002)	0.038
				Penalized weighted	0.963 (0.936, 0.99)	0.007
Melanoma	4271/2506	0.883	0.830	Inverse variance weighted	1.021 (1.01, 1.032)	< 0.001
				MR-Egger	1.02 (1.003, 1.036)	0.009
				Penalized weighted	1.024 (1.008, 1.04)	0.004
Non-Hodgkin lymphoma	1760/640	0.610	0.862	Inverse variance weighted	0.995 (0.992, 0.999)	0.005
				MR-Egger	0.995 (0.991, 0.999)	0.005
				Penalized weighted	0.994 (0.99, 0.998)	0.007

OR = odds ratio, CI = confidence interval, GERA = Genetic Epidemiology Research on Adult Health and Aging

The effect sizes of the causal estimates were from one standard deviation increase in the genetic predisposition for atrial fibrillation.



84 SNPs associated with atrial fibrillation identified from previous GWAS meta-analysis of European ancestry individuals (N = 482,265/55,144 case/controls)

4 SNPs strongly associated with potential confounders (obesity, hypertension, diabetes, dyslipidemia, thyroid disease, smoking)

13 non-overlapping or palindromic SNPs

Genetically predicted atrial fibrillation by 67 SNPs

GWAS meta-analysis for 17 site-specific cancer in European ancestry individuals [UK Biobank (48,961/359,825 case/controls) and GERA cohort (16,001/50,525 case/controls)]

**Summary-level  
Mendelian randomization**

● Inverse variance weighted ○ MR-Egger (bootstrap) □ Penalized weighted median

