

Original article

# Causal relationships between risk of venous thromboembolism and 18 cancers: a bidirectional Mendelian randomization analysis

Naomi Cornish <sup>1,2,\*</sup> Philip Haycock <sup>3</sup> Hermann Brenner <sup>4,5,6</sup> Jane C Figueiredo <sup>7</sup> Tessel E Galesloot,<sup>8</sup> Robert C Grant,<sup>9</sup> Mattias Johansson,<sup>10</sup> Daniela Mariosa <sup>10</sup> James McKay,<sup>10</sup> Rish Pai,<sup>11</sup> Andrew J Pellatt,<sup>12</sup> N Jewel Samadder,<sup>13</sup> Jianxin Shi,<sup>14</sup> Florian Thibord,<sup>15</sup> David-Alexandre Tréguët,<sup>16</sup> Catherine Voegelé,<sup>10</sup> Chrissie Thirlwell <sup>2,17,†</sup> Andrew Mumford<sup>1,†</sup> and Ryan Langdon<sup>3</sup>; InterLymph Consortium, INVENT-MVP Consortium

<sup>1</sup>School of Cellular and Molecular Medicine, University of Bristol, Bristol, UK, <sup>2</sup>University of Exeter Medical School, University of Exeter, Exeter, UK, <sup>3</sup>Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK, <sup>4</sup>Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany, <sup>5</sup>Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany, <sup>6</sup>German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany, <sup>7</sup>Department of Medicine, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA, <sup>8</sup>Department for Health Evidence, Radboud University Medical Center, Nijmegen, The Netherlands, <sup>9</sup>Division of Medical Oncology and Hematology, Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada, <sup>10</sup>International Agency for Research on Cancer, World Health Organization, Lyon, France, <sup>11</sup>Department of Laboratory Medicine and Pathology, Mayo Clinic, Scottsdale, AZ, USA, <sup>12</sup>Division of Cancer Medicine, MD Anderson Cancer Center, Houston, TX, USA, <sup>13</sup>Division of Gastroenterology, Mayo Clinic, Phoenix, AZ, USA, <sup>14</sup>Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA, <sup>15</sup>Population Sciences Branch, Division of Intramural Research, National Heart, Lung and Blood Institute, Framingham, MA, USA, <sup>16</sup>University of Bordeaux, Bordeaux Population Health Research Center, Bordeaux, France and <sup>17</sup>Bristol Medical School, University of Bristol, Bristol, UK

\*Corresponding author. Bristol Royal Infirmary, Research Floor 7, Queens Building, Bristol BS2 8HW, UK. E-mail: naomi.cornish@bristol.ac.uk

†Chrissie Thirlwell and Andrew Mumford contributed equally.

## Abstract

**Background:** People with cancer experience high rates of venous thromboembolism (VTE). Risk of subsequent cancer is also increased in people experiencing their first VTE. The causal mechanisms underlying this association are not completely understood, and it is unknown whether VTE is itself a risk factor for cancer.

**Methods:** We used data from large genome-wide association study meta-analyses to perform bidirectional Mendelian randomization analyses to estimate causal associations between genetic liability to VTE and risk of 18 different cancers.

**Results:** We found no conclusive evidence that genetic liability to VTE was causally associated with an increased incidence of cancer, or vice versa. We observed an association between liability to VTE and pancreatic cancer risk [odds ratio for pancreatic cancer: 1.23 (95% confidence interval: 1.08–1.40) per log-odds increase in VTE risk,  $P=0.002$ ]. However, sensitivity analyses revealed this association was predominantly driven by a variant proxying non-O blood group, with inadequate evidence to suggest a causal relationship.

**Conclusions:** These findings do not support the hypothesis that genetic liability to VTE is a cause of cancer. Existing observational epidemiological associations between VTE and cancer are therefore more likely to be driven by pathophysiological changes which occur in the setting of active cancer and anti-cancer treatments. Further work is required to explore and synthesize evidence for these mechanisms.

**Keywords:** Mendelian randomization, genetic epidemiology, deep vein thrombosis, pulmonary embolism, malignancy.

## Introduction

Venous thromboembolism (VTE), which includes deep vein thrombosis and pulmonary embolism, is the third most common cause of death from cardiovascular disease globally.<sup>1</sup> Over 20% of all VTE events occur in people with pre-existing cancer, for whom the relative risk of VTE is at least five times higher than age-matched non-cancer controls.<sup>2</sup> Evidence from in-vitro and animal models shows that many tumours directly activate platelets, produce procoagulant

proteins or alter the vascular endothelium, all of which may increase the risk of thrombosis.<sup>3</sup> Systemic anti-cancer therapy or surgery and their resultant complications, including sepsis and hospitalization, are also powerful risk factors for VTE.<sup>2</sup>

It is currently unknown whether elevated VTE risk is causally associated with cancer incidence. Over 5% of people presenting with a first VTE are subsequently diagnosed with cancer within the ensuing year,<sup>4,5</sup> and several studies have indicated that cancer risk may be elevated over the longer term

Received: 16 May 2023. Editorial Decision: 30 October 2023. Accepted: 4 December 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of the International Epidemiological Association.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

### Key Messages

- There is strong observational evidence that active cancer is associated with venous thromboembolism.
- It is currently unknown whether venous thromboembolism is a risk factor for cancer.
- We applied a bidirectional Mendelian randomization framework to appraise the causal relationships between genetic liability to venous thromboembolism and 18 different cancers.
- Overall, there was no clear evidence from Mendelian randomization that lifetime-elevated risk of venous thromboembolism is causally associated with an increased risk of cancer, or visa versa.

for people with a history of VTE.<sup>6–9</sup> Experiments in mice indicate that pro-thrombotic proteins, including tissue factor and fibrinogen, facilitate tumour growth, survival and metastasis.<sup>10,11</sup> However, observational studies examining whether treatment with antiplatelet or anticoagulant medication reduces risks of cancer have shown conflicting results.<sup>12–14</sup>

Attempts to elucidate complex causal relationships between VTE and cancer using traditional observational studies are complicated by difficulties in ascertaining direction of causality, and are susceptible to unmeasured and residual confounding from risk factors which are common to both VTE and cancer, including smoking, obesity and co-existing inflammatory conditions.<sup>2</sup> Mendelian randomization (MR) addresses some of these limitations. It employs genetic variants, typically single nucleotide polymorphisms (SNPs), as instrumental variables (IVs) to proxy the effect of an exposure on an outcome. As SNPs are randomly allocated and fixed at conception, they are unconfounded by acquired and environmental risk factors.<sup>15</sup>

Recently Chen *et al.* used MR to explore the unidirectional causal effect of liability to 14 cancers on VTE risk, and reported that genetically elevated risks of breast cancer and lymphoma may be causally associated with VTE.<sup>16</sup> However, this study was limited by relatively small sample sizes (5403 VTE cases and a median of 2321 cancer cases). Here, we use genome-wide association study (GWAS) meta-analysis data, derived from large cancer-specific consortia and >70 000

VTE cases, to perform bidirectional MR analyses of the effect of genetic liability to VTE on the risk of 18 cancers, and conversely the effect of genetic liability to cancer on the risk of VTE.

## Methods

### Data sources and genetic instruments

We obtained European-ancestry summary genetic data from GWAS meta-analyses examining risk of VTE and 18 common cancers, respectively (Table 1). Case definitions and covariate adjustment for each GWAS are shown in Supplementary Table S1 (available as Supplementary data at *IJE* online).

To examine the association between genetic liability to VTE and each cancer, we extracted risk SNPs associated with VTE at  $P < 5 \times 10^{-8}$  from a VTE GWAS conducted by Thibord *et al.*<sup>17</sup> We clumped SNPs to ensure independence,<sup>33,34</sup> then extracted summary statistics for these SNPs from each cancer risk GWAS. We harmonized exposure and outcome data to ensure that effect estimates corresponded to the same allele for each SNP. Full details of the clumping and harmonization process are described in Supplementary Methods (available as Supplementary data at *IJE* online).

To perform the analysis in the opposite direction (with genetic liability to cancer as an exposure and VTE as an outcome), we used the same process and thresholds described

**Table 1.** Source of genome-wide association study data used for bidirectional Mendelian randomization analyses

Trait	Study, date	No. of cases	No. of controls	Sample size
Venous thromboembolism	Thibord <i>et al.</i> , 2022 <sup>17</sup>	71 771	1 059 740	1 131 511
Breast cancer	Zhang <i>et al.</i> , 2020 <sup>18</sup>	133 384	113 789	247 173
Prostate cancer	Schumacher <i>et al.</i> , 2018 <sup>19</sup>	79 194	61 112	140 306
Endometrial cancer	O'Mara <i>et al.</i> , 2018 <sup>20</sup>	12 906	108 979	121 885
Colorectal cancer	Huyghe <i>et al.</i> , 2019 <sup>21</sup>	55 168	65 160	120 328
Melanoma	Landi <i>et al.</i> , 2020 <sup>22</sup>	30 134	81 415	111 549
Lung cancer	McKay <i>et al.</i> , 2017 <sup>23</sup>	29 266	56 450	85 716
Ovarian cancer	Phelan <i>et al.</i> , 2017 <sup>24</sup>	25 509	40 941	66 450
Kidney cancer	Scelo <i>et al.</i> , 2017 <sup>25</sup>	10 784	20 406	31 190
Oesophageal cancer	Gharahkhani <i>et al.</i> , 2016 <sup>26</sup>	4112	17 159	21 271
Pancreatic cancer	PanScan/PanC4, 2022 <sup>a</sup>	9055	7203	16 258
Diffuse large B cell lymphoma	Cerhan <i>et al.</i> , 2014 <sup>27</sup>	3857	7666	11 523
Chronic lymphocytic leukaemia	Berndt <i>et al.</i> , 2013 <sup>28</sup>	3100	7667	10 767
Follicular lymphoma	Skibola <i>et al.</i> , 2014 <sup>29</sup>	2728	7758	10 468
Oral cancer	Lesueur <i>et al.</i> , 2016 <sup>30</sup>	2700	5984	8684
Oropharyngeal cancer	Lesueur <i>et al.</i> , 2016 <sup>30</sup>	2433	5984	8417
Glioma	Melin <i>et al.</i> , 2017 <sup>31</sup>	4572	3286	7858
Marginal zone lymphoma	Vijai <i>et al.</i> , 2015 <sup>32</sup>	825	6221	7046
Bladder cancer	Nijmegen Bladder Cancer Study <sup>a</sup>	1799	4745	6544

<sup>a</sup> Indicates unpublished data.

above to select independent risk SNPs for each cancer from the relevant cancer GWAS (at  $P < 5 \times 10^{-8}$ ), then looked up summary statistics for the cancer risk SNPs in the VTE GWAS.

### Statistical analyses

Causal estimates from MR are underpinned by three core assumptions: (i) the genetic variants used as IVs are strongly associated with the exposure; (ii) there are no confounders of the genotype-outcome relationship; (iii) the genetic variants affect the outcome only via the exposure and not through an alternative pathway.<sup>35</sup>

We estimated the  $r^2$  (variance in phenotype explained by each IV) as described by Lee *et al.*,<sup>36</sup> using an assumed VTE prevalence of 0.2%.<sup>17</sup> Prevalence estimates for each cancer were obtained from the International Agency for Research on Cancer.<sup>37</sup> We assessed the strength of each SNP-exposure association using F statistics, estimated from summary data by dividing the square of the effect size (log-odds) by the square of the standard error. We performed Steiger filtering to exclude SNPs which explained more variance ( $r^2$ ) in the outcome than the exposure.<sup>38</sup> Assuming a causal relationship exists between an exposure and an outcome, a robust genetic IV should only explain the proportion of variance in the outcome which relates to the effect of the exposure.<sup>15</sup> Steiger filtering identifies SNPs which have a disproportionately large effect on the outcome compared with the exposure, reducing the risk of using invalid IVs which impact on the outcome via horizontal pleiotropy or which proxy a reverse causal pathway from outcome to exposure.

As recommended by published guidelines,<sup>35</sup> we used an inverse variance-weighted multiplicative random effects MR model (MR-IVW) for the primary analysis, with correction for under-dispersion when only a few SNPs were available for analysis. The MR-IVW is derived from a linear regression of the SNP-outcome and SNP-exposure associations, with each SNP weighted according to the inverse of the variance of the SNP-outcome effect. We assessed heterogeneity between the individual SNP estimates in the MR-IVW using Cochran's Q statistic. The exception to this was for marginal zone lymphoma, where only a single variant was available as a proxy; therefore the Wald ratio estimator<sup>15</sup> was used to estimate the causal effect.

Since MR-IVW assumes there is no directional pleiotropy in the MR instruments, we performed a range of sensitivity analyses, including MR-Egger, weighted-median, weighted-mode and leave-one-out analyses to test this assumption.<sup>39</sup> For phenotypes with at least three SNPs available as a proxy, we also used the MR PRESSO test to evaluate for horizontal pleiotropy and remove outlying SNPs.<sup>40</sup>

The VTE GWAS data came from a discovery cohort where some novel variants had not been replicated. Therefore, to evaluate for bias resulting from either weak instruments or 'the winner's curse',<sup>41</sup> we performed a sensitivity analysis where we limited genetic instruments for VTE to replicated loci only.<sup>17</sup>

Previous studies have reported that two powerful VTE risk variants, Factor V Leiden (rs6025, A allele) and Prothrombin G20210A (rs1799963, A allele), may be associated with cancer incidence.<sup>42–44</sup> These variants have a prevalence of ~5% and ~1%, respectively, in European populations. Notably, carriers of either of these variants have a VTE risk which is 3–5 times higher than those with wild-type alleles.<sup>45</sup> The

Prothrombin G20210A variant was excluded from the main VTE IV during the clumping process (due to absence from the reference panel). However, we performed a secondary analysis using MR Wald ratios<sup>15</sup> to examine the association between VTE risk, as proxied by these individual SNPs and risk of each cancer.

Results are presented in accordance with STROBE-MR guidelines<sup>46</sup> as the odds ratio (OR) and 95% confidence interval (CI) for each outcome per log-odds increase in the risk of exposure. *P*-values (*P*) have been adjusted for multiple testing using a false-discovery rate correction (*FDR-P*). All analyses were performed in R version 4.0.3 using the 'TwoSampleMR'<sup>33</sup> and MR-PRESSO<sup>40</sup> packages.

## Results

### Mendelian randomization analyses of the association between genetic liability to venous thromboembolism and cancer

After selecting independent VTE-risk SNPs ( $P < 5 \times 10^{-8}$ ,  $r^2 \leq 0.001$ ), there were 73 SNPs available as genetic instruments for VTE. These variants explained approximately 3% of the variance in VTE risk in the VTE GWAS cohort.<sup>17,33,36</sup> The number of instrumental variables varied for each VTE-cancer analysis (Table 2), as some VTE SNPs were either unavailable for assessment in the cancer GWAS studies, were excluded by Steiger filtering or could not be harmonized between the datasets due to coding-strand ambiguities. Summary data for the SNPs used in each analysis are shown in Supplementary Table S2 (available as Supplementary data at *IJE* online).

We estimated the OR for each cancer per log-odds increase in genetic liability to VTE using MR-IVW analysis (Figure 1). Increased risk of VTE was associated with an increased risk of pancreatic cancer [OR 1.23 (95% CI, 1.08–1.40),  $P = 0.002$ , *FDR-P* = 0.05]. A much weaker association in the same direction was seen for ovarian cancer (OR 1.05 (95% CI, 1.00–1.11),  $P = 0.04$ , *FDR-P* = 0.29) and endometrial cancer [OR 1.06 (95% CI, 1.00–1.12),  $P = 0.05$ , *FDR-P* = 0.31]. Sensitivity analyses showed inconsistent estimates of effect between the MR-IVW, MR-Egger, weighted-median and weighted-mode estimates for pancreatic cancer (Figure 2A). There was significant heterogeneity in the VTE IV estimates for pancreatic, ovarian and endometrial cancer as assessed by Cochran's Q statistic, and the MR-PRESSO global test showed evidence for horizontal pleiotropy (Supplementary Table S3, available as Supplementary data at *IJE* online). Graphical assessment of the leave-one-out plots, single SNP plots and funnel plots identified an outlying SNP (rs687289) which was confirmed by the MR-PRESSO outlier test (Supplementary Figures S1–S3, available as Supplementary data at *IJE* online). Removal of this SNP from the analysis virtually abolished the association between VTE and pancreatic, ovarian and endometrial cancer (Figure 2B).

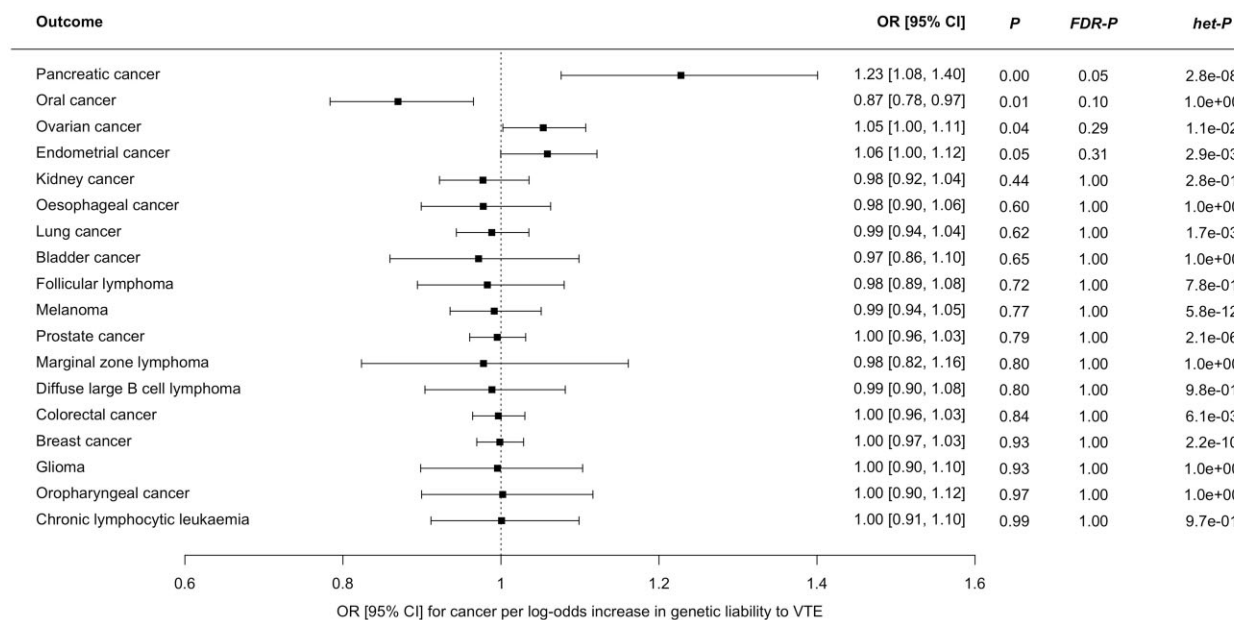
There was weak evidence from the MR-IVW analysis for a small inverse association between genetic risk of VTE and risk of oral cancer [OR 0.87 (95% CI, 0.78–0.97),  $P = 0.01$ , *FDR-P* = 0.10]. Sensitivity analyses showed consistent estimates of effect (Figure 2) with no indication of directional pleiotropy identified from the MR-Egger or MR-PRESSO analysis (Supplementary Table S3, available as Supplementary data at *IJE* online). There were no obvious outliers on inspection of the funnel plots and leave-one-out plots (Supplementary Figures S1 and S2, available as Supplementary data at *IJE* online).

**Table 2.** Number of genetic instruments for venous thromboembolism used for Mendelian randomization analyses, associated  $r^2$  and mean F statistic

Outcome GWAS	VTE SNPs available <sup>a</sup>	SNPs excluded <sup>a</sup>	VTE SNPs used <sup>a</sup>	$r^2$ for VTE	Mean F statistic
Breast cancer	68	4	64	0.031	208
Prostate cancer	66	2	64	0.031	210
Endometrial cancer	73	7	66	0.032	205
Colorectal cancer	68	2	66	0.031	204
Melanoma	68	4	64	0.031	209
Lung cancer	64	4	60	0.031	219
Ovarian cancer	68	6	62	0.031	213
Kidney cancer	70	9	61	0.031	218
Oesophageal cancer	70	6	64	0.032	210
Pancreatic cancer	45	5	40	0.023	254
Diffuse large B cell lymphoma	71	16	55	0.030	237
Chronic lymphocytic leukaemia	71	23	48	0.030	266
Follicular lymphoma	71	23	48	0.029	264
Oral cancer	68	15	53	0.030	242
Oropharyngeal cancer	68	18	50	0.029	255
Glioma	69	26	43	0.028	283
Marginal zone lymphoma	71	29	42	0.028	292
Bladder cancer	66	26	40	0.028	307

GWAS, genome-wide association study;  $r^2$ , phenotypic variance explained; SNP, single nucleotide polymorphism; VTE, venous thromboembolism.

<sup>a</sup> For each cancer outcome GWAS, 'VTE SNPs available' refers to the number of VTE-risk SNPs for which a direct correlate or proxy could be identified in the cancer GWAS study; 'SNPs excluded' refers to the number of VTE SNPs which could not be harmonized due to coding-strand ambiguities or which were excluded after Steiger filtering. 'VTE SNPs used' refers to the final number of genetic instruments for VTE used in each analysis.



**Figure 1.** Forest plot showing estimates from Mendelian randomization inverse variance-weighted estimates of the effect of genetic liability to venous thromboembolism as an exposure on 18 cancers as outcomes. CI, confidence interval; FDR-P, false-discovery corrected  $P$ -value; het-P, heterogeneity  $P$ -value for Cochran's  $Q$  statistic; OR, odds ratio; VTE, venous thromboembolism

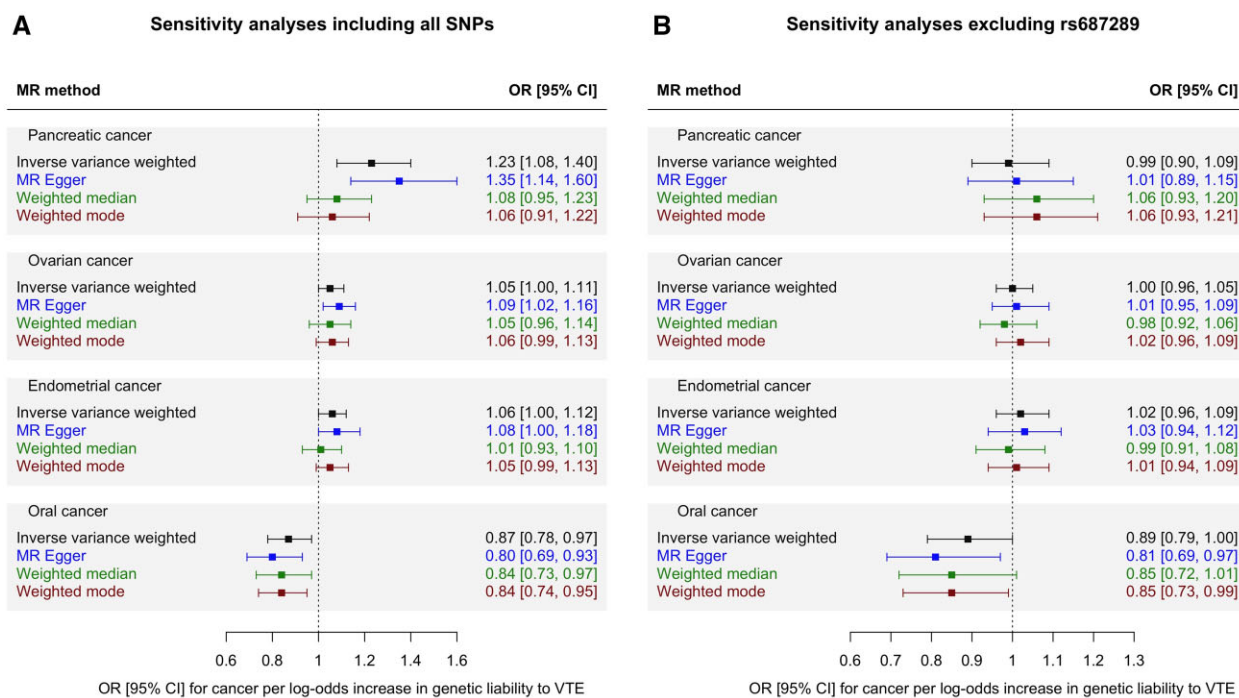
We performed two additional sensitivity analyses: first using only 31 replicated VTE SNPs as instrumental variables; secondly using all available VTE SNPs with no Steiger filtering applied. These results were similar to the primary analysis for all 18 cancers (Supplementary Tables S4 and S5, Supplementary Figures S4 and S5, available as Supplementary data at *IJE* online).

We examined the MR Wald ratios for the association between VTE risk proxied by either Factor V Leiden (rs6025) or Prothrombin G20210A (rs1799963), and cancer (Figure 3; Supplementary Table S6, available as Supplementary data at *IJE* online). Summary data for Factor V Leiden were available for all cancers. Summary data for Prothrombin G20210A were unavailable for six of 18 cancers (endometrial cancer, kidney

cancer, lung cancer, marginal zone lymphoma, pancreatic cancer and prostate cancer). There was a very weak inverse association between VTE risk, proxied by Factor V Leiden, and colorectal cancer [OR 0.95 (95% CI, 0.90–1.00),  $P=0.04$ ,  $FDR-P=0.59$ ]. There were no associations between Factor V Leiden or Prothrombin G20210A and any other cancer.

### Mendelian randomization analyses of the association between genetic liability to cancer and venous thromboembolism

We performed the MR analyses in the reverse direction, using genetic liability to cancer as an exposure and VTE as an outcome. The number of instrumental variables used for each cancer is shown in Table 3. Summary data for the SNPs used



**Figure 2.** Mendelian randomization sensitivity analyses of genetic liability to venous thromboembolism as an exposure and risk of four cancers (pancreatic, ovarian, endometrial and oral cancer) which showed an association ( $P \leq 0.05$ ) in the MR-IVW analysis. (A) shows sensitivity analyses including all SNPs. (B) shows sensitivity analyses with rs687289 removed. CI, confidence interval; FDR- $P$ , false-discovery corrected  $P$ -value; het- $P$ , heterogeneity  $P$ -value for Cochran's Q statistic; MR-IVW, Mendelian randomization inverse variance-weighted estimates; OR, odds ratio; SNP, single nucleotide polymorphism; VTE, venous thromboembolism

in each analysis are shown in [Supplementary Table S7](#) (available as [Supplementary data](#) at *IJE* online).

The MR-IVW analysis showed very weak evidence for a small inverse association between genetic liability to oropharyngeal cancer and risk of VTE [OR 0.93 (95% CI, 0.86–1.00),  $P = 0.05$ ,  $FDR-P = 0.48$ ; [Figure 4](#)]. Only three SNPs were available as instrumental variables, which precluded MR-PRESSO analysis. However, these SNPs displayed significant heterogeneity as measured by Cochran's Q statistic ( $P = 0.004$ ). This was reflected by the wide estimate of effect in the MR Egger analysis [OR 0.97 (95% CI, 0.61–1.52),  $P = 0.90$ , MR Egger intercept  $-0.02$ , intercept standard error 0.08]. Full results are shown in [Supplementary Table S8](#) (available as [Supplementary data](#) at *IJE* online). There were no other associations between genetically proxied risk of cancer and VTE in the primary analysis.

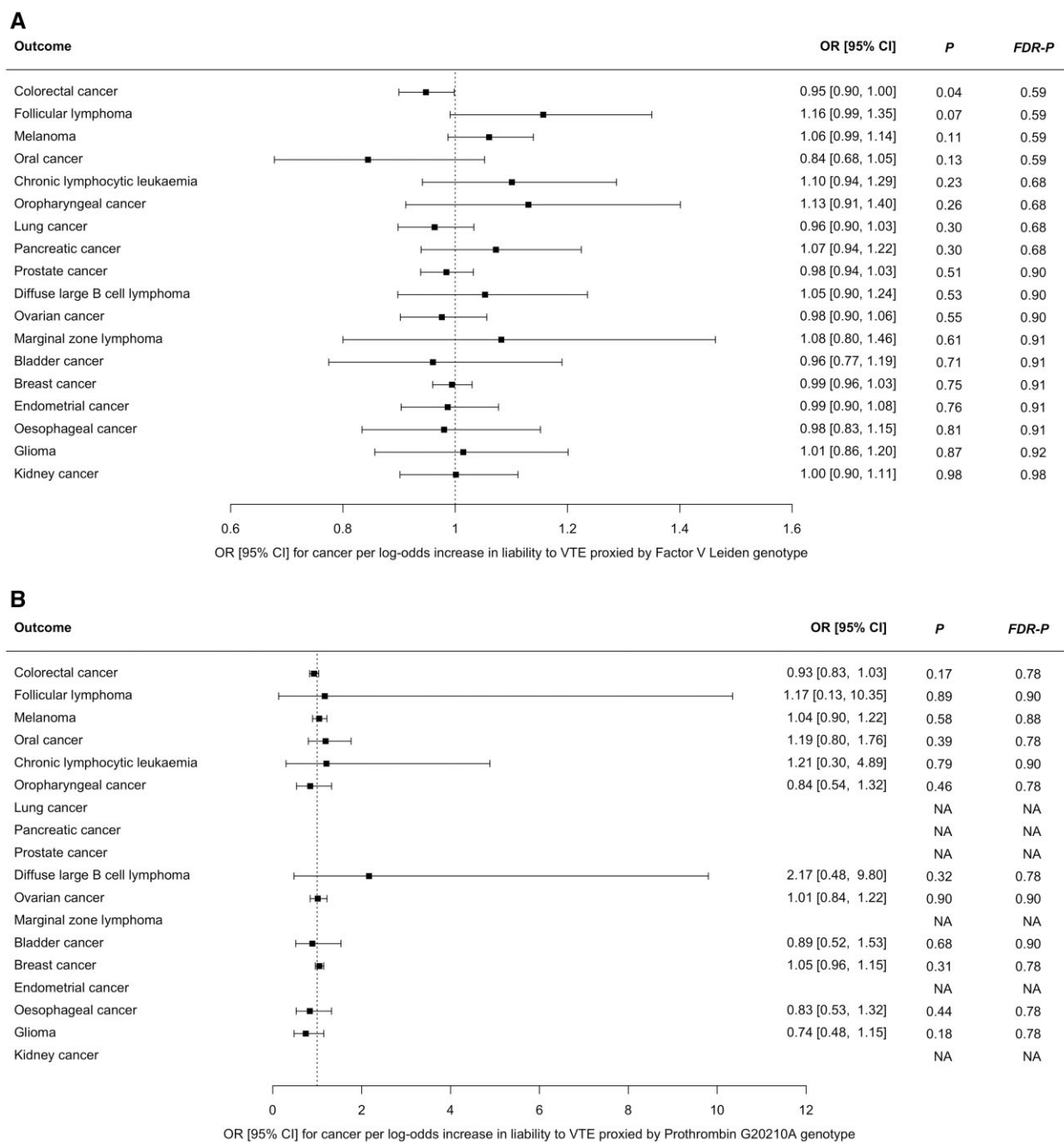
## Discussion

We observed an association between genetic liability to VTE and increased risk of pancreatic cancer. Pancreatic cancer has consistently been linked with very high rates of VTE in conventional epidemiological studies.<sup>47–49</sup> However, our MR sensitivity analyses indicated that the association between genetic liability to VTE and pancreatic cancer was largely driven by a single outlying SNP (rs687289). We observed weaker associations between genetic liability to VTE and risks of ovarian and endometrial cancer. Evidence for these associations was minimal after correction for multiple testing, and attenuated further towards the null in sensitivity analyses with rs687289 removed.

The variant rs687289 is found in intron 2 of the *ABO* blood group gene. The VTE-risk allele at this SNP is in high

linkage disequilibrium with an allele which determines non-O blood group (rs8176719).<sup>50</sup> Non-O blood group is associated with increased risk of a range of phenotypes including cardiovascular disease and several cancers, including pancreatic and ovarian cancer.<sup>51,52</sup> One possibility is that the MR association between VTE and pancreatic cancer results from horizontal pleiotropy (i.e. rs687289 exerting an effect on both VTE and pancreatic cancer through independent biological pathways). People with non-O blood group have higher levels of von-Willebrand factor and LDL-cholesterol, both of which may potentially contribute to VTE.<sup>50</sup> The mechanism by which blood group affects cancer risk is unknown, although it is hypothesized that ABO antibodies interact with aberrant glycoproteins expressed on pancreatic tumour cells.<sup>53</sup> It is also plausible that the association between VTE and pancreatic cancer is mediated by ABO blood group. Further multivariable MR analyses were beyond the scope of this study but would be helpful in evaluating this hypothesis. Last, since the prevalence of different ABO blood groups varies geographically,<sup>54</sup> the associations driven by this SNP could indicate confounding by population stratification. Although all GWAS data were drawn from genetically inferred European-ancestry participants, this encompasses a heterogeneous population, defined by different GWAS using varying principal component clustering models. Therefore, there may be genetic drift between the cohorts included in the VTE and cancer studies.

There was weak evidence from both the MR-IVW and MR sensitivity analyses that genetic liability to VTE was associated with slightly reduced risk of oral cancer [MR-IVW OR 0.87 (95% CI, 0.78–0.97),  $P = 0.01$ ,  $FDR-P = 0.10$ ]. This is surprising, given that procoagulant proteins are frequently over-expressed by oral cancer cells, which suggests growth of



**Figure 3.** Mendelian randomization Wald ratios for association between (A) liability to VTE, proxied by Factor V Leiden only, and risk of 18 cancers; and (B) liability to VTE, proxied by Prothrombin G20210A only, and risk of 12 cancers. NA indicates cancers for which the Prothrombin G20210A variant was not available in the genome-wide association study summary data. CI, confidence interval; FDR-P, false-discovery corrected *P*-value; OR, odds ratio; VTE, venous thromboembolism

this tumour occurs in a prothrombotic environment.<sup>55</sup> Paradoxically, epidemiological studies have shown that for people presenting with VTE, the risk of subsequent oral cancer diagnosis is relatively low compared with other cancers,<sup>56</sup> an observation which adds credibility to our results. However, since most people who experience VTE start treatment with anticoagulant therapy,<sup>4</sup> this could theoretically confound any associations between VTE and oral cancer.<sup>12</sup> A few small studies have previously described that SNPs in genes encoding two coagulation proteins, plasminogen activator inhibitor 1 (PAI-1) (rs1799889) and Factor XIII (rs5985), respectively, are associated with risk of oral squamous cell carcinoma.<sup>57,58</sup> Neither of these SNPs were

identified as VTE-risk variants in the VTE GWAS,<sup>17</sup> and therefore these were not represented in our analysis. Their specific role in oral cancer carcinogenesis is unclear. Given the lack of a consistent biological mechanism to explain our MR finding, our result should be interpreted with caution.

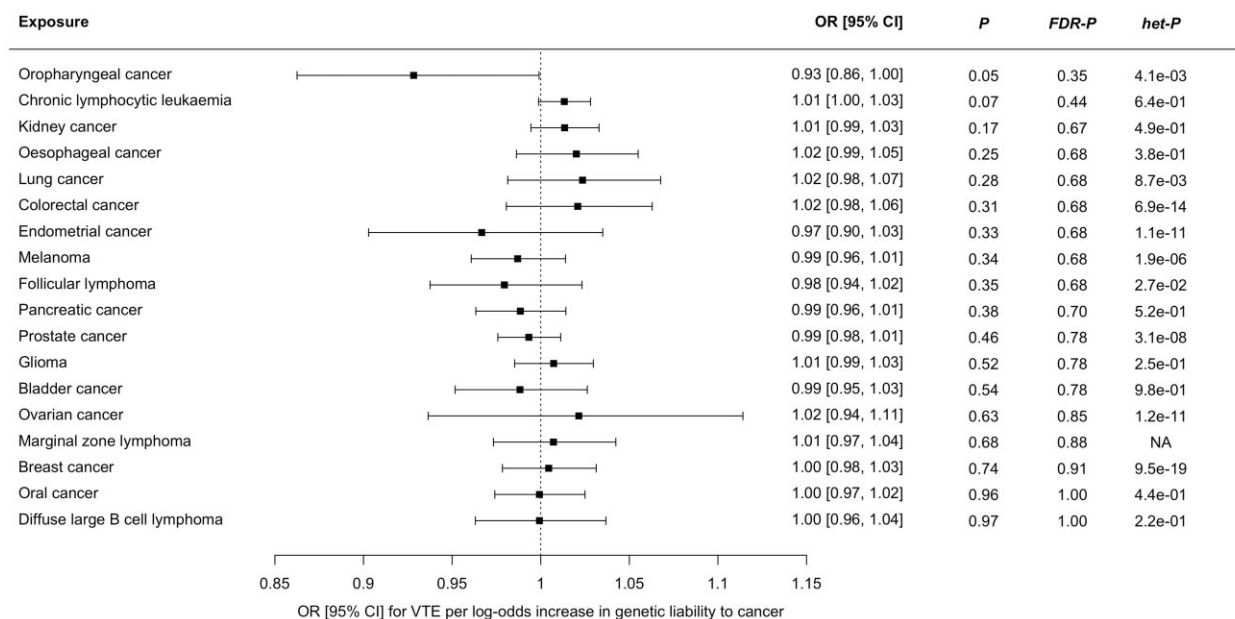
For the MR analyses in the cancer-VTE direction, which examined genetic liability to 18 cancers as exposures and VTE as an outcome, we found no clear evidence that genetic predisposition for any cancer was associated with increased risk of VTE, after correction for multiple testing. There are several caveats which should be considered in the interpretation of this result. This analysis estimates the impact of lifetime elevated genetic risk of cancer on risk of VTE, and will

**Table 3.** Number of genetic instruments for cancer used for Mendelian randomization analyses, associated  $r^2$  and mean F statistic

Exposure GWAS	Cancer-risk SNPs <sup>a</sup>	SNPs unavailable <sup>a</sup>	SNPs excluded <sup>a</sup>	Cancer SNPs used <sup>a</sup>	$r^2$ for cancer	Mean F statistic
Breast cancer	156	1	5	150	0.037	90
Prostate cancer	137	6	8	123	0.063	108
Endometrial cancer	16	0	0	16	0.013	44
Colorectal cancer	56	0	3	53	0.021	64
Melanoma	38	0	1	37	0.052	211
Lung cancer	15	1	0	14	0.013	83
Ovarian cancer	12	0	1	11	0.010	71
Kidney cancer	18	0	0	18	0.034	70
Oesophageal cancer	5	0	0	5	0.011	32
Pancreatic cancer	9	0	1	8	0.015	54
Diffuse large B cell lymphoma	4	0	0	4	0.012	27
Chronic lymphocytic leukaemia	8	0	0	8	0.048	70
Follicular lymphoma	2	0	0	2	0.032	48
Oral cancer	6	0	0	6	0.018	35
Oropharyngeal cancer	3	0	0	3	0.014	38
Glioma	5	0	0	5	0.028	80
Marginal zone lymphoma	1	0	0	1	0.009	10
Bladder cancer	2	0	0	2	0.006	31

GWAS, genome-wide association study; IV, instrumental variable;  $r^2$ , phenotypic variance explained; SNP, single-nucleotide polymorphism; VTE, venous thromboembolism.

<sup>a</sup> For each cancer exposure GWAS, ‘Cancer-risk SNPs’ refers to the total number of SNPs which could potentially be used as an IV for each cancer; ‘SNPs unavailable’ refers to the number of cancer SNPs for which no direct correlate could be identified in the VTE GWAS study; ‘SNPs excluded’ refers to the number of SNPs which could not be harmonized due to coding-strand ambiguities or which were excluded after Steiger filtering. ‘Cancer SNPs used’ refers to the final number of genetic instruments for cancer used in each analysis.



**Figure 4.** Forest plot showing estimates from Mendelian randomization analyses of the effect of genetic liability to 18 cancers as exposures on risk of venous thromboembolism as an outcome. The Mendelian randomization inverse variance-weighted estimates are shown for all cancers except marginal zone lymphoma, where the Wald ratio is shown, as only a single instrumental variable was available. CI, confidence interval; FDR-P, false-discovery corrected  $P$ -value; het-P, heterogeneity  $P$ -value for Cochran’s Q statistic; OR, odds ratio; VTE, venous thromboembolism

not capture time-dependent causal associations which occur due to acute changes in the context of active or progressive malignancy. Biological mechanisms leading to VTE include interruption to venous blood flow, vascular endothelial dysfunction and/or blood hypercoagulability.<sup>3</sup> The pathophysiology of VTE in people with cancer may be different from that in people who develop unprovoked VTE (in the absence of cancer): for example, tumour-induced endothelial hypoxia may contribute to VTE in the former group but not the latter.<sup>3</sup> As our outcome VTE GWAS cohort was derived from a heterogeneous case group (rather than a cancer-specific

cohort), this may have reduced our power to detect causal associations between genetic liability to cancer and VTE.<sup>17</sup> Future MR studies, using IVs which proxy time-dependent or intermediate exposure phenotypes, may be helpful to explore the association between cancer and VTE.

### Study limitations

In addition to the aforementioned limitations, we acknowledge that power to detect causal associations between cancer and VTE may have been limited for cancers with smaller GWAS case numbers. Since VTE in the context of cancer has

a high mortality rate,<sup>2</sup> it is also likely that cancer patients are under-represented in the VTE GWAS meta-analysis due to survival bias and selection bias of some of the contributing sub-studies.<sup>17</sup> This may distort the ability to detect causal associations with an MR approach.

Given the long latency period of many cancers<sup>41</sup> and polygenic nature of both cancer and VTE, the possibility of mis-attributing the direction of causality remains a concern. To address this, we performed a bidirectional analysis and used Steiger filtering<sup>38</sup> to select instrumental variables for each analysis (Supplementary Tables S2 and S7). Interestingly, in the MR of liability to cancer (exposure) on VTE (outcome), this process excluded a single SNP in the ABO blood group gene from the IV of both pancreatic and ovarian cancer (rs687289 and rs115478735, respectively). Steiger filtering may be unreliable if there is a significant difference in measurement error between the exposure and outcome GWAS.<sup>38</sup> We performed a sensitivity analysis in which no Steiger filtering was used, and found broadly similar results to the primary analysis (Supplementary Figure S6) but with much wider estimates for the effect of liability to pancreatic and ovarian cancer on VTE. This may be worth re-exploring in future when larger GWAS for these cancers become available.

### Comparison with wider literature

Although associations between VTE and cancer have been rigorously investigated with conventional epidemiological approaches,<sup>2</sup> to our knowledge there are no published MR analyses dedicated to examining the causal effect of genetic liability to VTE on cancer risk. One previous MR study has examined the causal effect of genetic liability to multiple cancers on VTE risk.<sup>16</sup> The authors of this study reported a trend towards reduced VTE risk in the context of genetic predisposition to melanoma, and increased VTE risk in the context of genetic predisposition to non-Hodgkin lymphoma and breast cancer, although evidence for these associations diminished after correction for multiple testing. In contrast our study, which used data from GWAS with much larger case numbers for both VTE and each cancer, did not replicate these associations. Qin *et al.* have also recently reported a unidirectional MR analysis examining the effect of genetic liability to breast cancer on VTE risk, and similarly did not find any association.<sup>59</sup>

Several small case-control studies have applied regression analyses to examine whether carriers of single thrombophilia gene polymorphisms, including Factor V Leiden (rs6025) and Prothrombin G20210A (rs1799963), are at increased risk of cancer.<sup>60</sup> Two groups previously reported that Prothrombin G20210A was associated with an increased risk of gastrointestinal and colorectal cancer, respectively.<sup>42,44</sup> In contrast, Vossen *et al.*<sup>43</sup> found that heterozygous carriers of either Prothrombin G20210A or Factor V Leiden had a reduced risk of colorectal cancer. Using an MR Wald ratio analysis, we found very weak evidence that the Factor V Leiden allele was associated with a slightly reduced risk of colorectal cancer [OR 0.94 (95% CI 0.89–1.00),  $P = 0.04$ ,  $FDR-P = 0.59$ ]. Asymptomatic carriers of Factor V Leiden have been shown to have accelerated formation of activated protein C.<sup>61</sup> This enzyme has effects on endothelial barrier integrity and inflammation which appear to be independent of coagulation pathways.<sup>62</sup> Therefore, the inverse association between Factor V Leiden and colorectal cancer risk may result from a biological interaction which is independent of thrombosis.

Alternatively, the result could reflect confounding by population stratification. Prothrombin G20210A genotype data were only available for 12 of the 18 cancers; however, we did not find any associations between this variant and cancer as assessed by MR Wald ratios.

### Conclusions

We present a bidirectional MR analysis examining associations between genetic liability to VTE and 18 different cancers, using summary data from large GWAS meta-analyses. Our findings do not support the hypothesis that genetic liability to VTE is a cause of cancer. Genetic liability to VTE was associated with increased risk of pancreatic cancer and slightly reduced risk of oral cancer, but there was inadequate evidence to suggest a causal relationship. Further work is required to establish whether and how biological pathways involving ABO blood group contribute to epidemiological associations between VTE and pancreatic cancer. Additional mechanistic studies may elucidate causal relationships between active cancer and VTE, as well as the role of VTE in cancer progression.

### Ethics approval

This analysis used GWAS summary data only, and therefore no additional ethics approval was required. Approval was obtained by the original GWAS studies by their appropriate ethics review boards: see details in the source GWAS publications.

### Data availability

R scripts used for the analysis are available via GitHub [<https://github.com/NaomiC-0/Mendelian-randomization-analysis-of-VTE-and-Cancer>]. Harmonised summary data for all SNPs included in this analysis are available in the Supplementary data. Full summary statistics are publicly available via: the Open GWAS database [<https://gwas.mrcieu.ac.uk>] for ovarian cancer (accession number: ieu-a-1120) and prostate cancer (accession number: ieu-b-85); the European Bioinformatics Institute GWAS Catalogue [<https://www.ebi.ac.uk/gwas>] for endometrial cancer (accession number: GCST006464), lung cancer (accession number: GCST004748) and oesophageal cancer (accession number: GCST003739); and the Breast Cancer Association Consortium [<https://bcac.ccge.medschl.cam.ac.uk>] for breast cancer. PanScan and PanC4 individual-level data are available through dbGaP (<https://www.ncbi.nlm.nih.gov/gap/>, accession numbers phs000206.v5.p3 and phs000648.v1.p1, respectively). Application to the relevant GWAS consortium is required for full summary statistics for the remaining phenotypes.

### Supplementary data

Supplementary data are available at *IJE* online.

### Author contributions

N.C., R.L., A.M. and C.T. designed the study. H.B., J.F., T.G., R.G., M.J., D.M., J.M., R.P., A.P., N.S., J.S., F.T., D-A.T. and C.V. contributed to the acquisition and analysis of



primary data used for the analyses. N.C. performed the final analyses and drafted the manuscript with R.L. and P.H. All authors critically revised and approved the final manuscript.

## Funding

This work was supported by the Wellcome Trust [grant number 225541/Z/22/Z (GW4 Clinical Academic Training Programme for Health Professionals) to N.C.] and Cancer Research UK (grant number C18281/A29019 to R.L. and P.H.). J.M. was supported by the US National Cancer Institute (grant number 5U01CA257679-02). D-A.T. was supported by the EPIDEMIO-M-VT Senior Chair from the University of Bordeaux initiative of excellence (IdEX). For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. Full funding acknowledgements for all contributing GWAS can be found in the source publications.

For the colorectal cancer GWAS: Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO) was funded by: National Cancer Institute, National Institutes of Health, U.S. Department of Health and Human Services (R01 CA059045, U01 CA164930, R01 201407, R01 CA244588) and in part through the NIH/NCI Cancer Center Support Grant P30 CA015704. Scientific Computing Infrastructure at Fred Hutch was funded by ORIP grant S10OD028685. Colon CFR was funded by NCI/NIH award U01 CA167551. For acknowledgements and funding of specific contributing studies, please see [<https://docs.google.com/document/d/1CW3cEYMzAabVdASrvC27p0g8vljQ6DJt/edit>].

The breast cancer genome-wide association analyses for BCAC and CIMBA were supported by Cancer Research UK (PPRPGM-Nov20100002, C1287/A10118, C1287/A16563, C1287/A10710, C12292/A20861, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692, C8197/A16565) and the Gray Foundation, 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).

The International Lung Cancer Consortium (ILCCO) was supported by grant U19 CA203654

Funding for the kidney cancer GWAS was provided by the US National Institutes of Health (NIH), National Cancer Institute (U01CA155309) for those studies coordinated by IARC, and by the intramural research programme of the National Cancer Institute, US NIH, for those studies coordinated by the NCI.

For the melanoma GWAS: the MelaNostrum Consortium was funded by the intramural Research Program of the National Institutes of Health (NIH), National Cancer Institute (NCI), Division of Cancer Epidemiology and Genetics, and by the participating institutes in Italy, Spain and Greece. The GenoMEL study was funded by the European Commission under the 6th Framework Programme (contract no. LSHC-CT-2006-018702), by Cancer Research UK Programme Awards (C588/A4994 and C588/A10589), by a Cancer Research UK Project Grant (C8216/A6129) and by a grant from the US National Institutes of Health (NIH; CA83115). This research was also supported by the intramural Research Program of the NIH, National Cancer Institute (NCI), Division of Cancer Epidemiology and Genetics, and by grants from the National Health and Medical Research Council (NHMRC) of Australia. For acknowledgements and funding of specific contributing studies, please see the source publication Landi *et al.* (*Nat Genet* 2020).

## 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. Where authors are identified as affiliated with the National Heart, Lung and Blood Institute or the National Institute of Health, the views expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Heart, Lung and Blood Institute or the National Institute of Health.

## Acknowledgements

This work has been conducted using GWAS summary data (including publicly available data) from separate GWAS consortia. We are grateful to all the consortia and their participants. These include the: Breast Cancer Association Consortium (BCAC), Barrett's and Esophageal Adenocarcinoma Consortium (BEACON), Esophageal Adenocarcinoma Genetics Consortium (EAGLE), Endometrial Cancer Association Consortium (ECAC), Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO), Genetic Epidemiology of Glioma International Consortium, International Lung Cancer Consortium (ILCCO), International Lymphoma Epidemiology Consortium (InterLymph), INVENT-VTE and MVP consortia, MelaNostrum Consortium, GenoMEL Consortium, Oncoarray Consortium, Ovarian Cancer Association Consortium (OCAC), Pancreatic Cancer Cohort Consortium (PanScan) and Pancreatic Cancer Case-Control Consortium (PanC4), Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome Consortium (PRACTICAL), the Nijmegen Bladder Cancer Study (NBCS) and Nijmegen Biomedical Study. This work used the computational facilities of the Advanced Computing Research Centre, University of Bristol [<http://www.bristol.ac.uk/acrc>].

## Conflict of interest

R.C.G. is an advisory or honorary for Astrazeneca, Eisai and Knight Therapeutics and is in receipt of a Graduate

Scholarship from Pfizer. No other conflicts of interest declared.

## References

- Wendelboe AM, Raskob GE. Global burden of thrombosis: epidemiologic aspects. *Circ Res* 2016;118:1340–47.
- Timp JF, Braekkan SK, Versteeg HH *et al.* Epidemiology of cancer-associated venous thrombosis. *Blood* 2013;122:1712–23.
- Abdol Razak NB, Jones G, Bhandari M *et al.* Cancer-associated thrombosis: an overview of mechanisms, risk factors, and treatment. *Cancers (Basel)* 2018;10:380.
- Cordeanu E-M, Jambert L, Tusch J *et al.* The conundrum of occult cancer screening in venous thromboembolism: lessons from the REMOTEV registry. *Med Kaunas Lith* 2022;58:913.
- Sørensen HT, Mellemejkær L, Steffensen FH *et al.* The risk of a diagnosis of cancer after primary deep venous thrombosis or pulmonary embolism. *N Engl J Med* 1998;338:1169–73.
- Douketis JD, Gu C, Piccioli A *et al.* The long-term risk of cancer in patients with a first episode of venous thromboembolism. *J Thromb Haemost* 2009;7:546–51.
- Rønnow Sand J, Troelsen FS, Nagy D *et al.* Increased cancer risk in patients with kidney disease and venous thromboembolism: a population-based cohort study. *Thromb Haemost* 2022;123:1165–76.
- Murchison JT, Wylie L, Stockton DL. Excess risk of cancer in patients with primary venous thromboembolism: a national, population-based cohort study. *Br J Cancer* 2004;91:92–95.
- Sørensen HT, Sværke C, Farkas DK *et al.* Superficial and deep venous thrombosis, pulmonary embolism and subsequent risk of cancer. *Eur J Cancer* 2012;48:586–93.
- Palumbo JS, Talmage KE, Massari JV *et al.* Tumor cell-associated tissue factor and circulating hemostatic factors cooperate to increase metastatic potential through natural killer cell-dependent and-independent mechanisms. *Blood* 2007;110:133–41.
- Palumbo JS, Kombrinck KW, Drew AF *et al.* Fibrinogen is an important determinant of the metastatic potential of circulating tumor cells. *Blood* 2000;96:3302–309.
- Haaland GS, Falk RS, Straume O *et al.* Association of warfarin use with lower overall cancer incidence among patients older than 50 years. *JAMA Intern Med* 2017;177:1774–80.
- Elwood P, Protty M, Morgan G *et al.* Aspirin and cancer: biological mechanisms and clinical outcomes. *Open Biol* 12:220124.
- Ifimi AA, Rodríguez-Bernal CL, Peiró S *et al.* Association of vitamin K and non-vitamin K oral anticoagulant use and cancer incidence in atrial fibrillation patients. *Clin Pharmacol Ther* 2022;111:200–208.
- Zheng J, Baird D, Borges M-C *et al.* Recent developments in Mendelian randomization studies. *Curr Epidemiol Rep* 2017;4:330–45.
- Chen X, Hong X, Luo S *et al.* Causal relationship between 14 site-specific cancers and venous thromboembolism. *Cancer Innov* 2022;1:316–27.
- Thibord F, Klarin D, Brody JA *et al.*; Global Biobank Meta-Analysis Initiative; Estonian Biobank Research Team; 23andMe Research Team; Biobank Japan; CHARGE Hemostasis Working Group. Cross-ancestry investigation of venous thromboembolism genomic predictors. *Circulation* 146:1225–42.
- Zhang H, Ahearn TU, Lecarpentier J *et al.*; GEMO Study Collaborators. Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nat Genet* 2020;52:572–81.
- Schumacher FR, Al Olama AA, Berndt SI *et al.*; Genetic Associations and Mechanisms in Oncology (GAME-ON)/Elucidating Loci Involved in Prostate Cancer Susceptibility (ELLIPSE) Consortium. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat Genet* 2018;50:928–36.
- O'Mara TA, Glubb DM, Amant F *et al.* Identification of nine new susceptibility loci for endometrial cancer. *Nat Commun* 2018;9:3166.
- Huyghe JR, Bien SA, Harrison TA *et al.* Discovery of common and rare genetic risk variants for colorectal cancer. *Nat Genet* 2019;51:76–87.
- Landi MT, Bishop DT, MacGregor S *et al.*; MelaNostrum Consortium. Genome-wide association meta-analyses combining multiple risk phenotypes provide insights into the genetic architecture of cutaneous melanoma susceptibility. *Nat Genet* 2020;52:494–504.
- McKay JD, Hung RJ, Han Y *et al.*; SpiroMeta Consortium. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nat Genet* 2017;49:1126–32.
- Phelan CM, Kuchenbaecker KB, Tyrer JP *et al.*; OPAL Study Group. Identification of 12 new susceptibility loci for different histotypes of epithelial ovarian cancer. *Nat Genet* 2017;49:680–91.
- Scelo G, Purdue MP, Brown KM *et al.* Genome-wide association study identifies multiple risk loci for renal cell carcinoma. *Nat Commun* 2017;8:15724.
- Gharahkhani P, Fitzgerald RC, Vaughan TL *et al.*; Wellcome Trust Case Control Consortium 2 (WTCCC2). Genome-wide association studies in oesophageal adenocarcinoma and Barrett's oesophagus: a large-scale meta-analysis. *Lancet Oncol* 2016;17:1363–73.
- Cerhan JR, Berndt SI, Vijai J *et al.* Genome-wide association study identifies multiple susceptibility loci for diffuse large B cell lymphoma. *Nat Genet* 2014;46:1233–28.
- Berndt SI, Skibola CF, Joseph V *et al.* Genome-wide association study identifies multiple risk loci for chronic lymphocytic leukemia. *Nat Genet* 2013;45:868–76.
- Skibola CF, Berndt SI, Vijai J *et al.* Genome-wide association study identifies five susceptibility loci for follicular lymphoma outside the HLA region. *Am J Hum Genet* 2014;95:462–71.
- Lesueur C, Diergaard B, Olshan AF *et al.* Genome-wide association analyses identify new susceptibility loci for oral cavity and pharyngeal cancer. *Nat Genet* 2016;48:1544–50.
- Melin BS, Barnholtz-Sloan JS, Wrensch MR *et al.*; GliomaScan Consortium. Genome-wide association study of glioma subtypes identifies specific differences in genetic susceptibility to glioblastoma and non-glioblastoma tumors. *Nat Genet* 2017;49:789–94.
- Vijai J, Wang Z, Berndt SI *et al.* A genome-wide association study of marginal zone lymphoma shows association to the HLA region. *Nat Commun* 2015;6:5751.
- Hemani G, Zheng J, Elsworth B *et al.* The MR-Base platform supports systematic causal inference across the human phenome. Loos R (ed.). *Life* 2018;7:e34408.
- Auton A, Brooks LD, Durbin RM *et al.*; 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* 2015;526:68–74.
- Burgess S, Davey Smith G, Davies NM *et al.* Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res* 2020;4:186.
- Lee SH, Goddard ME, Wray NR *et al.* A Better Coefficient of Determination for Genetic Profile Analysis. *Genet Epidemiol* 2012;36:214–24.
- International Agency for Research on Cancer, World Health Organization. *Cancer Today Data and Methods*. 2020. [https://gco.iarc.fr/today/online-analysis-table?v=2020&mode=cancer&mode\\_population=continents&population=900&populations=900\\_908&key=asr&sex=0&cancer=39&type=0&statistic=5&prevalence=0&population\\_group=0&ages\\_group%5B%5D=0&ages\\_group%5B%5D=17&group\\_cancer=1&include\\_nmssc=0&include\\_nmssc\\_other=1#collapse-group-0-4](https://gco.iarc.fr/today/online-analysis-table?v=2020&mode=cancer&mode_population=continents&population=900&populations=900_908&key=asr&sex=0&cancer=39&type=0&statistic=5&prevalence=0&population_group=0&ages_group%5B%5D=0&ages_group%5B%5D=17&group_cancer=1&include_nmssc=0&include_nmssc_other=1#collapse-group-0-4) (9 November 2022, date last accessed).
- Hemani G, Tilling K, Smith GD. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet* 2017;13:e1007081.
- Burgess S, Bowden J, Fall T *et al.* Sensitivity analyses for robust causal inference from mendelian randomization analyses with multiple genetic variants. *Epidemiology* 2017;28:30–42.
- Verbanck M, Chen C-Y, Neale B *et al.* Detection of widespread horizontal pleiotropy in causal relationships inferred from

- Mendelian randomization between complex traits and diseases. *Nat Genet* 2018;50:693–98.
41. Yarmolinsky J, Wade KH, Richmond RC *et al.* Causal inference in cancer epidemiology: what is the role of mendelian randomization? *Cancer Epidemiol Biomarkers Prev* 2018;27:995–1010.
  42. Baghdad I, Erguibi D, Chehab F *et al.* Risk of colorectal cancer and clotting factor gene polymorphisms in Moroccan Population. *Int J Adv Res (Indore)* 2017;5:1141–46.
  43. Vossen CY, Hoffmeister M, Chang-Claude JC *et al.* Clotting factor gene polymorphisms and colorectal cancer risk. *J Clin Oncol* 2011;29:1722–27.
  44. Pihusch R, Danzl G, Scholz M *et al.* Impact of thrombophilic gene mutations on thrombosis risk in patients with gastrointestinal carcinoma. *Cancer* 2002;94:3120–26.
  45. Arachchilage DJ, Mackillop L, Chandratheva A *et al.* Thrombophilia testing: a British society for haematology guideline. *Br J Haematol* 2022;198:443–58.
  46. Skrivankova VW, Richmond RC, Woolf BAR *et al.* Strengthening the reporting of observational studies in epidemiology using mendelian randomization: the STROBE-MR statement. *JAMA* 2021;326:1614–21.
  47. Strongman H, Gadd S, Matthews A *et al.* Medium and long-term risks of specific cardiovascular diseases in survivors of 20 adult cancers: a population-based cohort study using multiple linked UK electronic health records databases. *Lancet* 2019;394:1041–54.
  48. Horsted F, West J, Grainge MJ. Risk of venous thromboembolism in patients with cancer: a systematic review and meta-analysis. *PLoS Med* 2012;9:e1001275.
  49. Khorana AA, Francis CW. Risk prediction of cancer-associated thrombosis: appraising the first decade and developing the future. *Thromb Res* 2018;164(Suppl 1):S70–76.
  50. Li S, Schooling CM. A phenome-wide association study of ABO blood groups. *BMC Med* 2020;18:334.
  51. Staley JR, Blackshaw J, Kamat MA *et al.* PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics* 2016;32:3207–209.
  52. Kamat MA, Blackshaw JA, Young R *et al.* PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics* 2019;35:4851–53.
  53. Hofmann BT, Stehr A, Dohrmann T *et al.* ABO blood group IgM isoagglutinins interact with tumor-associated o-glycan structures in pancreatic cancer. *Clin Cancer Res* 2014;20:6117–26.
  54. Anstee D. The relationship between blood groups and disease. *Blood* 2010;115:4635–43.
  55. Racine F, Soudet S, Sevestre M-A *et al.* The coagulome of oral squamous cell carcinoma: examining the role and regulation of coagulation in oral cancers using a systems approach. *Curr Opin Otolaryngol Head Neck Surg* 2023;31:73–77.
  56. Haen P, Mege D, Crescence L *et al.* Thrombosis risk associated with head and neck cancer: a review. *Int J Mol Sci* 2019;20:2838.
  57. Vylliotis A, Yapijakis C, Nkenke E *et al.* Effect of thrombosis-related gene polymorphisms upon oral cancer: a regression analysis. *Anticancer Res* 2013;33:4033–39.
  58. Vairaktaris E, Serefoglou Z, Avgoustidis D *et al.* Gene polymorphisms related to angiogenesis, inflammation and thrombosis that influence risk for oral cancer. *Oral Oncol* 2009;45:247–53.
  59. Qin Q, Pan Y, Wang L. No evidence for a genetic causal effect of breast cancer on venous thromboembolism: a Mendelian randomization study. *J Thromb Thrombolysis* 2023. doi: [10.1007/s11239-023-02871-1](https://doi.org/10.1007/s11239-023-02871-1).
  60. Tavares V, Neto BV, Vilas-Boas MI *et al.* Impact of hereditary thrombophilia on cancer-associated thrombosis, tumour susceptibility and progression: a review of existing evidence. *Biochim Biophys Acta Rev Cancer* 2022;1877:188778.
  61. Rühl H, Berens C, Winterhagen FI *et al.* Increased activated protein c response rates reduce the thrombotic risk of factor V Leiden carriers but not of prothrombin 20210G>A carriers. *Circ Res* 2019;125:523–34.
  62. Feistritz C, Riewald M. Endothelial barrier protection by activated protein C through PAR1-dependent sphingosine 1-phosphate receptor-1 crossactivation. *Blood* 2005;105:3178–84.