RESEARCH ARTICLE

Cancer Epidemiology

Circulating white blood cell traits and colorectal cancer risk: A Mendelian randomisation study

Andrei-Emil Constantinescu[1,2,3] | Caroline J. Bull[1,2,3] | Nicholas Jones[4] | Ruth Mitchell[1,2] | Kimberley Burrows[1,2] | Niki Dimou[5] | Stéphane Bézieau[6] | Hermann Brenner[7,8,9] | Daniel D. Buchanan[10,11,12] | Mauro D'Amato[13,14,15] | Mark A. Jenkins[16] | Victor Moreno[17,18,19,20] | Rish K. Pai[21] | Caroline Y. Um[22] | Emily White[23,24] | Neil Murphy[5] | Marc Gunter[5] | Nicholas J. Timpson[1,2] | Jeroen R. Huyghe[24] | Emma E. Vincent[1,2,3]

[1]MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

[2]Bristol Medical School, Population Health Sciences, University of Bristol, Bristol, UK

[3]School of Translational Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

[4]Institute of Life Science, Swansea University Medical School, Swansea, UK

[5]Nutrition and Metabolism Branch, International Agency for Research on Cancer, World Health Organization, Lyon, France

[6]Service de Génétique Médicale, Centre Hospitalier Universitaire (CHU) Nantes, Nantes, France

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

[8]Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany

[9]German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany

[10]Colorectal Oncogenomics Group, Department of Clinical Pathology, The University of Melbourne, Parkville, Victoria, Australia

[11]University of Melbourne Centre for Cancer Research, Victorian Comprehensive Cancer Centre, Parkville, Victoria, Australia

[12]Genetic Medicine and Family Cancer Clinic, The Royal Melbourne Hospital, Parkville, Victoria, Australia

[13]Department of Medicine and Surgery, LUM University, Casamassima, Italy

[14]Gastrointestinal Genetics Lab, CIC bioGUNE – BRTA, Derio, Spain

[15]Ikerbasque, Basque Foundation for Science, Bilbao, Spain

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

[17]Oncology Data Analytics Program, Catalan Institute of Oncology-IDIBELL, L'Hospitalet de Llobregat, Barcelona, Spain

[18]CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain

[19]Department of Clinical Sciences, Faculty of Medicine, University of Barcelona, Barcelona, Spain

[20]ONCOBEL Program, Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet de Llobregat, Barcelona, Spain

[21]Department of Laboratory Medicine and Pathology, Mayo Clinic Arizona, Scottsdale, Arizona, USA

[22]Department of Population Science, American Cancer Society, Atlanta, Georgia, USA

[23]Department of Epidemiology, University of Washington School of Public Health, Seattle, Washington, USA

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

Correspondence
Andrei-Emil Constantinescu, Oakfield House, Oakfield Grove, Bristol, UK.
Email: andrei.constantinescu@bristol.ac.uk

Abstract

Observational studies have suggested a protective role for eosinophils in colorectal cancer (CRC) development and implicated neutrophils, but the causal relationships remain unclear. Here, we aimed to estimate the causal effect of circulating white blood cell (WBC) counts (N = ~550 000) for basophils, eosinophils, monocytes, lymphocytes and neutrophils on CRC risk (N = 52 775 cases and 45 940 controls) using Mendelian randomisation (MR). For comparison, we also examined this relationship using individual-level data from UK Biobank (4043 incident CRC cases and 332 773 controls) in a longitudinal cohort analysis. The inverse-variance weighted (IVW) MR analysis suggested a protective effect of increased basophil count and eosinophil count on CRC risk [OR per 1-SD increase: 0.88, 95% CI: 0.78-0.99, P = .04; OR: 0.93, 95% CI: 0.88-0.98, P = .01]. The protective effect of eosinophils remained [OR per 1-SD increase: 0.88, 95% CI: 0.80-0.97, P = .01] following adjustments for all other WBC subtypes, to account for genetic correlation between the traits, using multivariable MR. A protective effect of increased lymphocyte count on CRC risk was also found [OR: 0.84, 95% CI: 0.76-0.93, P = 6.70e-4] following adjustment. Consistent with MR results, a protective effect for eosinophils in the cohort analysis in the fully adjusted model [RR per 1-SD increase: 0.96, 95% CI: 0.93-0.99, P = .02] and following adjustment for the other WBC subtypes [RR: 0.96, 95% CI: 0.93-0.99, P = .001] was observed. Our study implicates peripheral blood immune cells, in particular eosinophils and lymphocytes, in CRC development, highlighting a need for mechanistic studies to interrogate these relationships.

KEYWORDS
colorectal cancer, eosinophils, Mendelian randomisation, UK biobank, White blood cell count

What's new?
Alterations to circulating white blood cell counts have been associated with colorectal cancer, but whether these changes have causal effects remains unexplored. Using multivariable Mendelian randomisation, the authors assessed the independent causal effects of white blood cell counts by cell subtype by adjusting for shared genetic architecture. In parallel, they ran the largest longitudinal cohort study to explore causal effects between white blood cell counts and colorectal cancer using UK Biobank data. The evidence suggests that elevated eosinophil and lymphocyte counts may have a protective effect against colorectal cancer, adding new insights into colorectal cancer pathogenesis.

1 | INTRODUCTION

Colorectal cancer (CRC) accounts for over 10% of all worldwide cancer cases and is the second leading cause of cancer-related deaths globally.[1-3] Overall, the number of CRC cases is rising[4] and is alarmingly on the rise in younger people (aged <50 years).[5-7] Given current challenges, and the estimation that 50% of CRC cases may be preventable,[8] a focus on identifying novel risk factors, and subsequent prophylactic and treatment options are warranted to limit the future healthcare burden.

White blood cells (WBCs) are commonly measured in routine blood tests and are divided into five subtypes: basophils, eosinophils, lymphocytes, monocytes and neutrophils.[9] Alterations to circulating WBC counts have been found to play a role in disease risk, severity and progression, CRC development and mortality.[10-13] For example, higher circulating counts of basophils and eosinophils, cells with a role in IgE-mediated immunity, have been associated with reduced CRC risk and increased survival.[14-17] Higher absolute lymphocyte count [such as T, B and natural killer (NK) cells] has also been associated with better overall survival of CRC.[18,19] By contrast, a high absolute monocyte count was found to be associated with worse CRC survival,[18,19] consistent with their potential role in tumour progression and metastasis.[20] Finally, neutrophils, a critical part of the innate immune system,[21] have also been associated with poor CRC overall survival.[18,19]

Traditional (ie, cross-sectional, case-control or cohort) studies account for the majority of epidemiologic analyses undertaken on CRC, which can suffer from certain limitations, such as confounding and reverse causation which can bias effect estimates.[22] Mendelian randomisation (MR) is a method in genetic epidemiology which could overcome these issues by using single-nucleotide polymorphisms (SNPs) to proxy for an exposure of interest to estimate the effect of an exposure on an outcome.[23] It operates akin to a randomised control trial (RCT), as alleles are randomly assigned at birth.[23] We have used MR to estimate the total effect of each WBC trait on CRC,[24] and, to account for correlation between WBC traits, multivariable (MV) MR to estimate the direct effect of these exposures on CRC by adjusting for their shared genetic architecture.[24]

In our study, we aimed to investigate the relationship between circulating WBC subtypes and CRC incidence. First, an MR analysis was undertaken using the most comprehensive genome-wide association study (GWAS) for both WBC subtype counts and incident CRC available. Second, for comparison, we ran the largest longitudinal cohort study between WBCs and CRC to date using UK Biobank. Together these analyses allow us to compare between genetically proxied and observational estimates to thoroughly explore the relationship between circulating WBC counts and CRC development.

2 | METHODS

2.1 | Study design

We aimed to investigate the relationship between circulating WBCs and CRC using a genetic epidemiologic and complementary observational approach. First, MR analyses were undertaken to estimate the effect of

WBC subtype counts on CRC. We then performed a multivariable MR (MVMR) analysis where the direct effect of each WBC subtype count was estimated by adding all five WBC subtypes into the model (Figure 1A). STROBE-MR guidelines were followed (STROBE-MR Supplement).[25] Here, units were interpreted as odds ratio (OR) for CRC per a normalised SD (1-SD) increase in WBC count. Next, a prospective longitudinal cohort study was undertaken using UK Biobank individual-level data. Here, subtype-specific WBC counts were studied individually, and then each was adjusted for each of the other traits (Figure 1B). STROBE guidelines were followed (STROBE Cohort Supplement). Here, units were interpreted as risk ratio (RR) for CRC per a normalised SD (1-SD) increase in WBC count.

2.2 | WBC count GWAS data

Summary statistics for WBC subtype counts were obtained from the 'Blood Cell Consortium' (BCX) meta-analysis, the largest study being UKBB (N = ~562 243).[26] Genetic sex, age, age^2, study-specific covariates and PCs 1 to 10 were used as covariates. A brief description of each study included in the meta-analysis is available in Table S1. Only variants which did not display heterogenous effects across studies were selected. Specific details on QC steps and association testing are available in the source manuscripts. Summary statistics for WBC counts were downloaded from: http://www.mhi-humangenetics.org/en/resources/.

2.3 | CRC GWAS data

GWAS summary statistics for overall CRC risk and anatomical subsite were taken from the most comprehensive meta-analyses to date[27]:
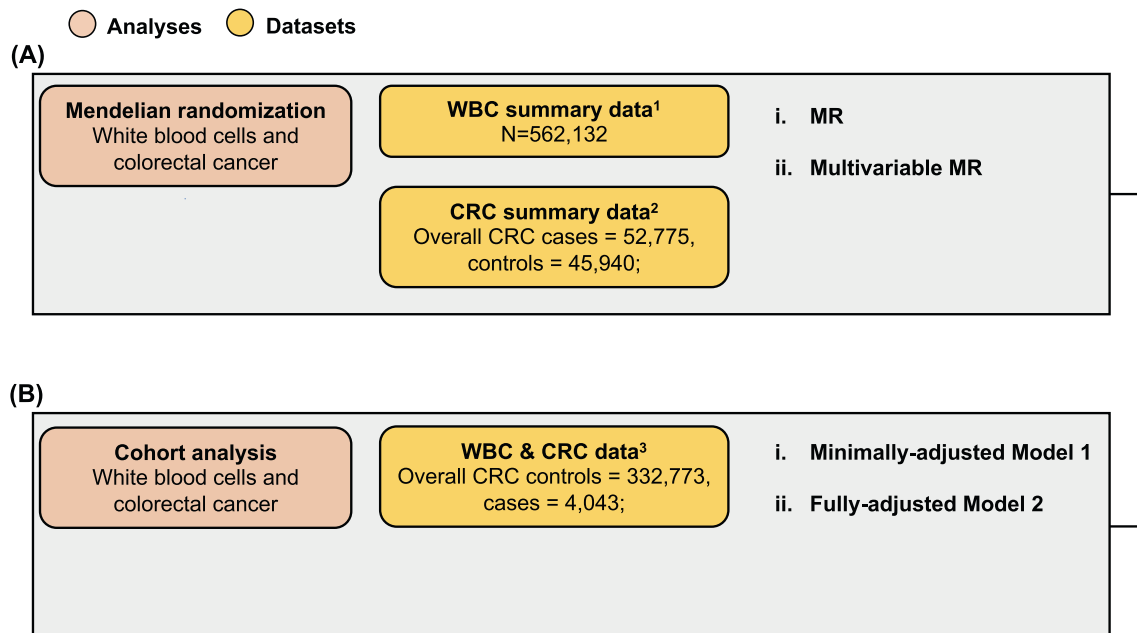

FIGURE 1 Study design. We triangulated findings from two study designs: a Mendelian randomisation analysis (A) and a longitudinal cohort analysis (B) to estimate the causal effect of WBC on CRC risk. [Color figure can be viewed at wileyonlinelibrary.com]

colon cancer (split into proximal and distal colon cancer) and rectal cancer. Sixty-six studies, which include those of the Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO), Colorectal Cancer Transdisciplinary Study (CORECT) and Colon Cancer Family Registry (CCFR) consortia were meta-analysed.[27] The final sample was predominantly of European ancestry with 5.36% of East Asian ancestry. These were included due to their similar genetic architecture with regard to CRC risk.[28] Genetic sex, age, study-specific variables and PCs were used as covariates. An overview of all consortia included in the CRC meta-analysis is available in Table S2, and a breakdown of the sample sizes for each set of CRC summary statistics is presented in Table S3.

2.4 | Genetic data processing

To select valid MR instruments, summary statistics for exposures were processed using the 'TwoSampleMR' R package.[29,30] The presence of correlated instruments, that is, those in linkage disequilibrium (LD), can bias MR estimates.[29,30] Therefore, the exposure SNPs were clumped ($r^2 = 0.001$, window $= 10$ Mb, P-value threshold $= 5e-8$) using the 1000 Genomes European dataset[31] as a reference panel. Following this step, the exposure and outcome datasets were 'harmonised', that is, had their effect alleles placed on the same reference strand.[32] SNPs with incorrect but unambiguous strand references were corrected, while those with ambiguous strand references were removed.

2.5 | MR analysis of WBC counts on CRC risk

Primary MR analyses were undertaken using the inverse-variance weighted (IVW) method, which is the multiplicative random-effects meta-analysis of the estimated effect of all exposure SNPs on the CRC outcome.[33] Conditional F-statistics were calculated to detect weak instrument bias for each exposure SNP as previously described.[34] Several sensitivity analyses were undertaken to compare with the main IVW estimates. The presence of vertical pleiotropy, that is, when a trait is downstream of the genetic variant but on the same biological pathway as the exposure,[35] was measured using Cochran's Q heterogeneity test. Horizontal pleiotropy, when some or all instruments for a trait act through a different pathway to the exposure,[35] can violate one of the main MR assumptions. A number of sensitivity MR analyses were undertaken to identify horizontal pleiotropy: MR-Egger (where the regression intercept is not constrained to zero),[36] Weighted median (the median of all SNP ratio estimates, where each ratio is weighted by the inverse of the variance),[37] Weighted modes (assumes that the most frequent estimate in a set of instruments is zero)[38] and MR-PRESSO (detects individual SNPs contributing to horizontal pleiotropy).[39] The direction of the causal relationship between WBC traits and CRC risk was tested using the MR Steiger method, which uses Steiger's test to test the difference between the Pearson correlations of genetic variants with both the exposure and outcome.[30]

2.6 | Multivariable MR analysis of WBC counts on CRC risk

The IVW method was used for the MVMR analysis. First, a pair-wise analysis between all five WBC subtype counts was undertaken, where the proportion of variance explained (PVE) for SNPs used to instrument a WBC trait was estimated in the other four WBC subtypes using the previously described methodology.[34] The direct effect of each WBC subtype was estimated by adding all five WBC subtypes into the MVMR model. Bias arising from weak instruments was also determined here. This was undertaken using the methodology described by Sanderson et al, where a generalised version of Cochran's Q was employed to evaluate instrument strength.[40] Standard Cochran's Q statistic was calculated to detect the presence of heterogeneity. For those traits with an F-statistic <10, a follow-up MVMR analysis was done accounting for the presence of weak instruments.

2.7 | UK Biobank phenotypic data

Between the years 2007 and 2010, UKBB participants visited assessment centres (N = 22) throughout the UK.[41] The individuals had their health records linked, their genomes sequenced and underwent multiple evaluations, such as self-report questionnaires and medical examinations.[41] The latter includes the analysis of blood cell samples using Beckman Coulter LH750 instruments designed for high throughput screening.[42] Total WBC count and WBC subtype percentage (%) were measured, with absolute WBC subtype count derived as 'WBC subtype %/100 × total WBC' and expressed as 10^9 cells/L.[42] The blood sampling date variable was split into year, month, day and minutes (passed since the start of the day of the appointment visit). Additional variables were gathered including recruitment centre, sampling device ID, age, genetic sex, principal components 1 to 10, body mass index (BMI), Townsend deprivation index, smoking and alcohol drinker status (self-report questionnaire—UKBB codes 20116 and 20117). CRC cases were identified through hospital inpatient records coded to the 10th version of the International Classification of Disease (ICD-10). UK Biobank data was granted under application code 81499.

2.8 | Filtering and selection criteria

The UKBB dataset underwent a series of steps prior to further analyses. Withdrawn participants and those of non-European ancestry were excluded. Viable controls and incident CRC cases were defined using methodology previously described by Burrows et al[43] (Table S4). Here, we defined incident CRC cases as those diagnosed at least 1 year after blood sampling. Participants with no WBC measurement data or sampling date were removed, as were those who were known to be pregnant, have chronic conditions (eg, HIV, blood cancers, thalassaemia) or undergoing erythropoietin treatment, as in Astle et al[44] and Chen et al,[26] given the effects of these traits on WBC

measurements. Those with acute conditions (eg, upper respiratory infections) diagnosed less than 3 months prior to blood sampling were excluded. Finally, missing values in 'Townsend Deprivation Index', 'Body mass index', 'Smoking status' and 'Alcohol drinker status' variables were removed.

2.9 | Cohort study between WBC count and CRC

We conducted a cohort analysis between circulating WBCs and incident CRC. WBC count values were log-transformed, after which they were adjusted for the following covariates: assessment centre, sex, age, age^2, PCs 1 to 10, as in Chen et al's GWAS.[26] The resulting residuals were rank-inverse normal transformed and then used in a logistic regression on CRC incidence. This main observational analysis was termed 'Model 1', a minimally adjusted model. An additional fully adjusted analysis was undertaken, termed 'Model 2', where BMI, Townsend DI, smoker status and alcohol drinker status were added as additional covariates. Following this, another pair of analyses was run, where all five WBC subtype counts were added together into the model to reduce potential bias due to their correlated values. Analyses, where each WBC trait was studied individually, were termed as 'univariable', while those where they were added together were termed as 'multivariable'.

2.10 | Working environment

All analyses were performed with R version 4.1.2 (Bird Hippie) in a Linux environment supported by the University of Bristol's Advanced Computing Research Centre (ACRC). Genetic data preparation, as well as the MR analyses, were undertaken with the 'TwoSampleMR' R package.[29,30] The MVMR analyses were undertaken with the 'MVMR' R package.[45]

2.11 | Patient and public involvement

Patients or the public were not involved in the design, conduct, reporting or dissemination plans of our research.

3 | RESULTS

3.1 | Effect of WBC count on CRC

Before performing the MR analysis, the average F-statistic for each WBC trait was estimated to detect the presence of weak instrument bias, which is generally indicated by an average F-statistic <10.[46] For overall CRC, these were 64.48 (basophil count, 171 SNPs), 124.72 (eosinophil count, 396 SNPs), 105.85 (lymphocyte count, 444 SNPs), 147.44 (monocyte count, 477 SNPs)

and 98.84 (neutrophil count, 387 SNPs), indicating strong MR instruments (Table S5).

The IVW method showed evidence of a protective effect of basophil count (1-SD increase in WBC count OR: 0.88, 95% CI: 0.78-0.99, P-value: .037; Figure 2, Figure S1) and eosinophil count for CRC risk (OR: 0.93, 95% CI: 0.88-0.98, P-value: .012; Figure 2, Figure S1). All results, including those for sex- and site-specific CRC, are available in Table S6.

Next, the presence of vertical and horizontal pleiotropy in the MR analyses was analysed. Cochran's heterogeneity test indicated the presence of heterogeneity in all but one WBC trait-CRC pair (basophil count-male CRC, $P_{HET} = .104$; Table S7). Following this, the MR-Egger test for horizontal pleiotropy was performed. Here, evidence for this type of pleiotropy was identified for eosinophil count and overall ($P_{PLT} = .015$) CRC risk, suggesting a possible bias of MR estimates (Table S7). Although the MR-PRESSO method identified the presence of SNP horizontal-pleiotropic outliers, there was little evidence that the removal of these outliers contributed to a notable shift in the point estimates (Table S8).

3.2 | Multivariable MR of WBC count on CRC

A pair-wise analysis of the proportion of variance explained (PVE) for SNPs instrumenting each WBC trait indicated a low PVE for each of the other WBC traits, with the exception of basophil count (2.44% vs 2.39% when instrumenting neutrophil count; Table S9). The overall results indicated that statistical power should not suffer to a large degree by adding all five WBC subtype counts into the MVMR analysis. Here, the MVMR IVW method estimated a protective effect of eosinophil count (OR: 0.88, 95% CI: 0.80-0.97, P-value: .011) and lymphocyte count (OR: 0.84, 95% CI: 0.76-0.93, P-value: .0007) on CRC risk (Figure 2). Site-specific CRC estimates are available in Table S10.

Sensitivity analyses were undertaken to assess heterogeneity and the presence of weak instruments in the MVMR analysis. There was evidence of heterogeneity in all WBC trait-CRC pairs (Table S11). Conditional F-statistics showed evidence of weak instruments ($F < 10$) for basophil count (Table S11). Based on these results, an additional MVMR analysis was run adjusting for weak instruments for basophil count (Table S11).

3.3 | Cohort observational analysis between WBC count and CRC

In total, 336 816 UKBB participants remained after passing the filtering and selection criteria (Figure S2). Of these, 332 773 were controls and 4043 were incident CRC cases. When split by genetic sex, there were 154 629 male and 178 144 female controls and 2316 male and 1727 female cases. Those with CRC were more likely to be male (57% vs 46%), had a higher average age (60.7 vs 55.8 years), had slightly higher BMI (28.0 vs 27.4 kg/m^2) and were more likely to have
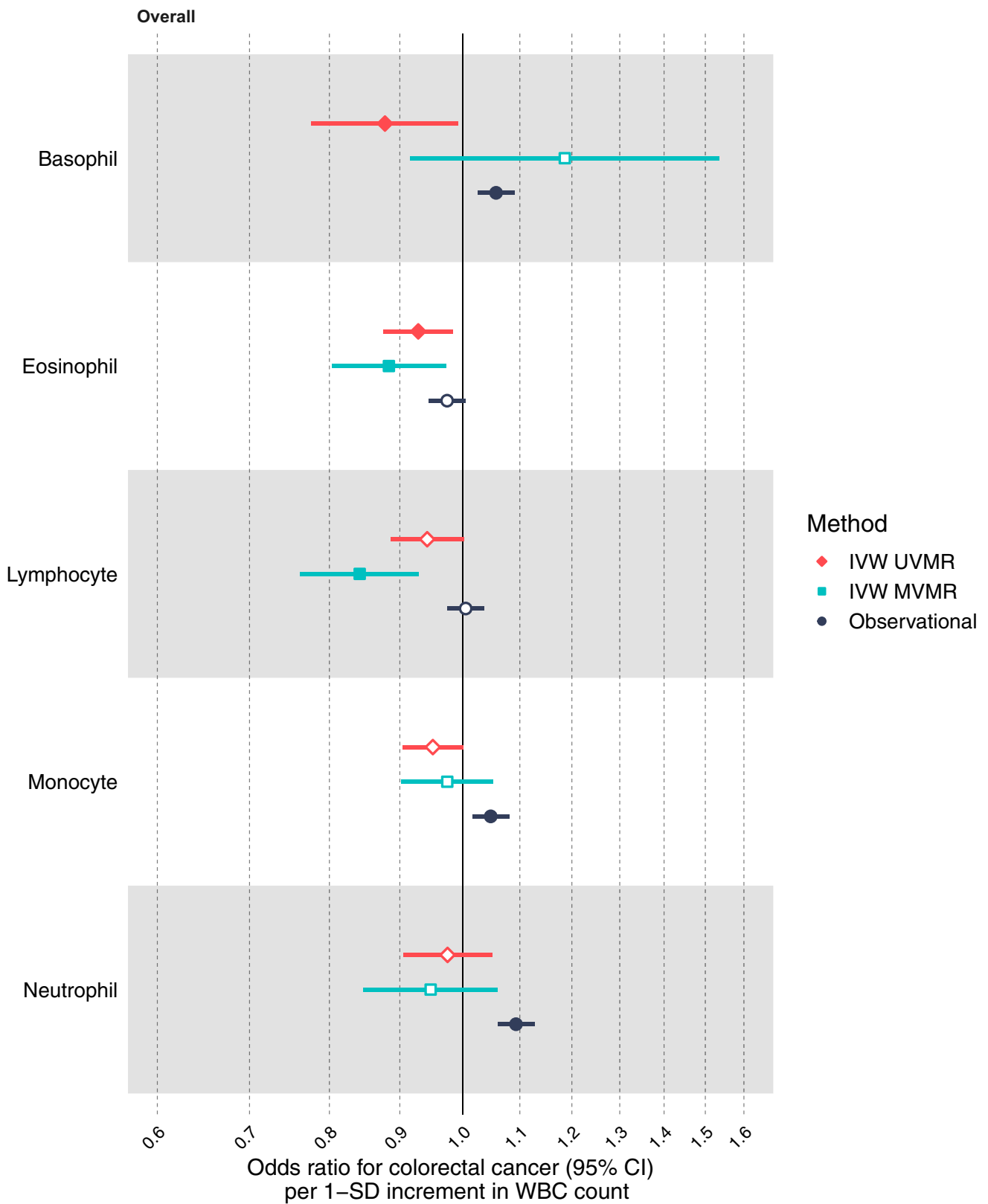
FIGURE 2 The relationship between WBC count and CRC risk based on UV, MV two-sample MR and cohort observational analyses. Each WBC trait is presented on the X-axis. The estimated effect is presented on the Y-axis. Point estimates were filled where the P-value was less than .05. Results are interpreted as ORs (95% CI) for CRC risk per 1-SD normalised increment in WBC count. [1]Chen et al[26]; [2]Huyghe et al[27] (https://doi.org/10.1136/gutjnl-2020-321534); [3]UK Biobank.[47] [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Baseline characteristics of UK Biobank study sample.

Characteristic	Control, N = 332773[a]	Case, N = 4043[a]
Sex		
Female	178 144/332 773 (54%)	1727/4043 (43%)
Male	154 629/332 773 (46%)	2316/4043 (57%)
Age (years)	55.8 (8.1)	60.7 (6.6)
Body mass index	27.4 (4.7)	28.0 (4.7)
Smoking status		
Never	183 987/332 773 (55%)	1861/4043 (46%)
Previous	114 349/332 773 (34%)	1790/4043 (44%)
Current	34 437/332 773 (10%)	392/4043 (9.7%)
Alcohol drinker status		
Never	10 655/332 773 (3.2%)	122/4043 (3.0%)
Previous	10 982/332 773 (3.3%)	140/4043 (3.5%)
Current	311 136/332 773 (93%)	3781/4043 (94%)
Townsend deprivation index	−1.4 (3.0)	−1.6 (3.0)
Basophil count	0.0 (0.1)	0.0 (0.1)
Eosinophil count	0.2 (0.1)	0.2 (0.1)
Lymphocyte count	1.9 (0.6)	1.9 (0.6)
Monocyte count	0.5 (0.2)	0.5 (0.2)
Neutrophil count	4.2 (1.4)	4.3 (1.4)
Overall WBC count	6.9 (1.7)	7.0 (1.8)

[a]n/N (%); Mean (SD).

been cigarette smokers in the past (46% vs 55% never smokers and 44% vs 34% previous smokers; Table 1).

As a percentage of the total WBC count based on the median values, basophils accounted for 0.3%, eosinophils 2.11%, lymphocytes 28.16%, monocytes 6.78% and neutrophils 60.39% (Table S12). A pair-wise correlation matrix between each WBC subtype demonstrated correlation coefficients equal to or below 0.3 (Figure S3). Batch variables (eg, blood sample device and sampling date), Townsend DI, and alcohol drinker status explained some of the variances in WBC count (0%-0.66%). Depending on the WBC subtype, genetic sex explained between 0.23% and 2.93% of the variance, BMI explained between 0.14% and 2.64%, and smoking status explained between 0.44% and 3.85% (Figure S4).

In the main analysis ('Model 1'—the minimally adjusted model), basophil count was positively associated with CRC risk (RR: 1.06, 95% CI: 1.02-1.09, P-value: .0005), as was monocyte (RR: 1.05, 95% CI: 1.02-1.08, P-value: .003) and neutrophil count (RR: 1.09, 95% CI: 1.06-1.13, P-value: 1.94E−08; Figure 2, Table S13). In 'Model 2', eosinophil count was negatively associated with CRC risk (overall CRC RR: 0.96, 95% CI: 0.93-0.99, P-value: .022; Table S13).

Observational associations were re-computed by adding all five WBC subtype counts together, largely showing agreement with the main analysis (Table S14).

4 | DISCUSSION

In our study, we aimed to estimate the effects of five circulating WBC subtypes on CRC risk by using a combined genetic epidemiologic and longitudinal cohort framework. Through the aid of MVMR, we were able to assess the independent causal effect of WBC counts by adjusting for their shared genetic architecture. Taken together, the evidence across analyses suggests a potential protective effect of increased circulating eosinophil and lymphocyte count on CRC risk.

Consistent with our study, Prizment et al found that eosinophil count (tertiles Q3 and Q2 vs Q1) was negatively associated with odds of developing colon, but not of rectal cancer.[16] Similar results have been reported for other cancers; Wong et al reported a negative trend between increasing eosinophil count quartiles and lung adenocarcinoma odds in a UKBB study,[47] while a similar study looking at prostate cancer showed a negative association between eosinophil count quintiles Q3-5, as well as a per 1-SD increase in the trait (HR 0.96 vs OR 0.93 for CRC in our analysis).[48]

Eosinophils have a well-established role in allergic diseases, including asthma and allergic rhinitis.[49] Indeed, MR analyses have also reported that eosinophil count levels affect the risk of developing allergic disease[44,50] and a recent systematic review investigating the relationship between allergies and cancer reported evidence for a reduced risk of CRC in those with allergic diseases.[51]

Here, our results suggest that the immune response through eosinophils provides protection against tumour development. Indeed, in several neoplasias, including CRC, eosinophils have been found to play an anti-tumourigenic role and are a source of anti-tumourigenic molecules, such as eosinophil-derived neurotoxin (EDN).[52,53] Experimental studies have also found a tumour-protective effect of IgE.[54] Increased eosinophil recruitment to the CRC tumour site has also been associated with better survival, even when adjusting for the effects of CD8+ T-cells,[11] and eosinophil-specific granule secretion of granzyme A has been linked with the killing of CRC cells.[55]

In addition to eosinophil count, we also found a protective effect of lymphocyte count on CRC risk. While this was not apparent in the MR analysis, the MVMR estimates indicated lower ORs for CRC across all anatomical subsites (proximal colon cancer trended towards protective) with increased lymphocyte counts. The multivariable fully adjusted 'Model 2' in the cohort analysis also indicated that there could be a negative association with CRC risk. It is not surprising that we found a protective effect of higher circulating levels of lymphocytes with CRC odds given their established role in combatting tumour development.[56] Tumour-infiltrating lymphocytes (TILs) like CD8+ T-cells help antagonise tumour growth through direct action and recruitment of other immune cells.[56] High levels of TILs were previously associated with better CRC overall survival and disease-free survival.[19,56] In support of our findings, two observational studies found higher lymphocyte counts compared to cases vs controls a year to 6 months prior to CRC diagnosis.[57,58] However, these results could indicate production and recruitment of lymphocytes to the site of precancerous or undetected tumours rather than a causal effect.

4.1 | Limitations

There are several limitations to our study. With regard to the cohort analysis, only baseline blood measurements were available. This assumes that WBC counts were constant and did not allow us to establish a relationship between a trend in WBC count and its relationship with CRC odds. Nevertheless, baseline WBC count measurements have previously been shown to be associated with disease risk,[47,48,59,60] making their study in relation to disease development a worthwhile endeavour. Also, in the cohort analysis, incident CRC cases were defined as those diagnosed at least 1 year after blood sampling, in order to not diminish the number of cases to a large degree. However, as CRC develops over a long period, our cohort analysis may have not excluded all participants with undiagnosed CRC.

With regards to the MR analysis, the genetic instruments used here proxied for lifetime variation of WBC count. Therefore, the MR analysis cannot be used to infer how large changes over a short timespan might affect CRC development. Regarding the MVMR method, caution should be applied when investigating traits with very weak instruments, as it cannot reliably adjust for those traits.[45] This was the case for basophil count, as the F-statistic was estimated to be between 4.7 and 4.8 (Table S12). Therefore, despite pointing to an increased detrimental effect compared to the main MVMR analysis, ORs derived from the weak-MVMR analysis should be interpreted with this in mind.

5 | CONCLUSION

In summary, the results generated here provide evidence for a protective causal effect of elevated levels of circulating eosinophil and lymphocyte counts on CRC risk. Going forward, additional research is needed to disentangle the biological mechanisms and pinpoint-specific pathways through which eosinophils and lymphocytes might protect against CRC development.

AUTHOR CONTRIBUTIONS

Andrei-Emil Constantinescu, Caroline J. Bull, Jeroen R. Huyghe and Emma E. Vincent conceived the idea for the paper. Andrei-Emil Constantinescu conducted the analysis. All authors contributed to the interpretation of the findings. Andrei-Emil Constantinescu, Caroline J. Bull and Emma E. Vincent wrote the manuscript. All authors critically revised the paper for intellectual content and approved the final version of the manuscript. The work reported in the paper has been performed by the authors, unless clearly specified in the text.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Summary statistics for WBC counts were downloaded from the following website: http://www.mhi-humangenetics.org/en/resources/. The summary-level GWAS data on outcomes used in our study were made available following an application to the Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO): https://www.fredhutch.org/en/research/divisions/public-health-sciences-division/research/cancer-prevention/genetics-epidemiology-colorectal-cancer-consortium-gecco.html. Scripts associated with our study are available on GitHub: https://github.com/andrewcon/wbc-crc. Further information is available from the corresponding author upon request.

ETHICS STATEMENT

All participants provided written informed consent, and each study was approved by the relevant research ethics committee or institutional review board. UK Biobank received ethical approval from the NHS National Research Ethics Service North West (11/NW/0382; 16/NW/0274) and was conducted in accordance with the Declaration of Helsinki. All participants provided written informed consent before enrolment in the study.

ORCID

Andrei-Emil Constantinescu https://orcid.org/0000-0002-7974-5245

Caroline J. Bull https://orcid.org/0000-0002-2176-5120

Daniel D. Buchanan https://orcid.org/0000-0003-2225-6675

Victor Moreno https://orcid.org/0000-0002-2818-5487

Neil Murphy https://orcid.org/0000-0003-3347-8249

TWITTER

Andrei-Emil Constantinescu 🐦 *_andreicon*

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68:394-424. doi:10.3322/CAAC.21492

2. Sung H, Siegel RL, Rosenberg PS, Jemal A. Emerging cancer trends among young adults in the USA: analysis of a population-based cancer registry. *Lancet Public Health*. 2019;4:e137-e147. doi:10.1016/S2468-2667(18)30267-6

3. Mauri G, Sartore-Bianchi A, Russo AG, Marsoni S, Bardelli A, Siena S. Early-onset colorectal cancer in young individuals. *Mol Oncol*. 2019; 13:109. doi:10.1002/1878-0261.12417

4. Xi Y, Xu P. Global colorectal cancer burden in 2020 and projections to 2040. *Transl Oncol*. 2021;14:101174. doi:10.1016/J.TRANON.2021.101174

5. Chambers AC, Dixon SW, White P, Williams AC, Thomas MG, Messenger DE. Demographic trends in the incidence of young-onset colorectal cancer: a population-based study. *Br J Surg*. 2020;107:595. doi:10.1002/BJS.11486

6. SEER. *Surveillance, Epidemiology, and End Results (SEER) Dental Program*. Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health; 2019.

7. Araghi M, Soerjomataram I, Bardot A, et al. Changes in colorectal cancer incidence in seven high-income countries: a population-based study. *Lancet Gastroenterol Hepatol*. 2019;4:511-518. doi:10.1016/S2468-1253(19)30147-5

8. Hull MA. Nutritional prevention of colorectal cancer. *Proc Nutr Soc*. 2021;80:59-64. doi:10.1017/S0029665120000051

9. Nicholson LB. The immune system. *Essays Biochem*. 2016;60:275-301. doi:10.1042/EBC20160017

10. Wu J, Ge X, Zhu W, et al. Values of applying white blood cell counts in the prognostic evaluation of resectable colorectal cancer. *Mol Med Rep*. 2019;19:2330-2340. doi:10.3892/MMR.2019.9844

11. Prizment AE, Vierkant RA, Smyrk TC, et al. Tumor eosinophil infiltration and improved survival of colorectal cancer patients: Iowa Women's health study. *Mod Pathol*. 2016;29:516-527. doi:10.1038/modpathol.2016.42

12. Rosman Y, Hornik-Lurie T, Meir-Shafrir K, et al. Changes in peripheral blood eosinophils may predict colorectal cancer—a retrospective study. *World Allergy Org J*. 2022;15:100696. doi:10.1016/J.WAOJOU.2022.100696

13. Iijima K, Fujibayashi K, Okumura M, Sasabe N, Gunji T. Number of eosinophils and incidence of cancer in a Japanese population: a single institution study. *Ningen Dock Int*. 2019;6:56-61.

14. Liu Q, Luo D, Cai S, Li Q, Li X. Circulating basophil count as a prognostic marker of tumor aggressiveness and survival outcomes in colorectal cancer. *Clin Transl Med*. 2020;9:1-12. doi:10.1186/S40169-019-0255-4

15. Wei Y, Zhang X, Wang G, et al. The impacts of pretreatment circulating eosinophils and basophils on prognosis of stage I–III colorectal cancer. *Asia Pac J Clin Oncol*. 2018;14:e243-e251. doi:10.1111/AJCO.12871

16. Prizment AE, Anderson KE, Visvanathan K, Folsom AR. Inverse association of eosinophil count with colorectal cancer incidence: atherosclerosis risk in communities study. *Cancer Epidemiol Biomarkers Prev*. 2011;20:1861-1864. doi:10.1158/1055-9965.EPI-11-0360

17. Ma W, Yang J, Li P, Lu X, Cai J. Association between allergic conditions and colorectal cancer risk/mortality: a meta-analysis of prospective studies. *Sci Rep*. 2017;7:1-7. doi:10.1038/s41598-017-04772-9

18. Yamamoto T, Kawada K, Obama K. Inflammation-related biomarkers for the prediction of prognosis in colorectal cancer patients. *Int J Mol Sci*. 2021;22:8002. doi:10.3390/IJMS22158002

19. Tanio A, Saito H, Uejima C, et al. A prognostic index for colorectal cancer based on preoperative absolute lymphocyte, monocyte, and neutrophil counts. *Surg Today*. 2019;49:245-253. doi:10.1007/S00595-018-1728-6/TABLES/2

20. Olingy CE, Dinh HQ, Hedrick CC. Monocyte heterogeneity and functions in cancer. *J Leukoc Biol*. 2019;106:309-322. doi:10.1002/JLB.4RI0818-311R

21. Quail DF, Amulic B, Aziz M, et al. Neutrophil phenotypes and functions in cancer: a consensus statement. *J Exp Med*. 2022;219:39. doi:10.1084/JEM.20220011/213202

22. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet*. 2014;23:R89-R98. doi:10.1093/hmg/ddu328

23. Davey Smith G, Ebrahim S, Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003;32:1-22. doi:10.1093/ije/dyg070

24. Carter AR, Sanderson E, Hammerton G, et al. Mendelian randomisation for mediation analysis: current methods and challenges for implementation. *Eur J Epidemiol*. 2021;36:465-478. doi:10.1007/S10654-021-00757-1

25. Skrivankova VW, Richmond RC, Woolf BAR, et al. Strengthening the reporting of observational studies in epidemiology using Mendelian randomisation (STROBE-MR): explanation and elaboration. *BMJ*. 2021;375. doi:10.1136/BMJ.N2233

26. Chen MH, Raffield LM, Mousas A, et al. Trans-ethnic and ancestry-specific blood-cell genetics in 746,667 individuals from 5 global populations. *Cell*. 2020;182:1198-1213.e14. doi:10.1016/j.cell.2020.06.045

27. Huyghe JR, Bien SA, Harrison TA, et al. Discovery of common and rare genetic risk variants for colorectal cancer. *Nat Genet*. 2019;51: 76. doi:10.1038/S41588-018-0286-6

28. Zhang B, Jia WH, Matsuda K, et al. Large-scale genetic study in East Asians identifies six new loci associated with colorectal cancer risk. *Nat Genet*. 2014;46:533. doi:10.1038/NG.2985

29. Hemani G, Zheng J, Elsworth B, et al. The MR-based platform supports systematic causal inference across the human phenome. *Elife*. 2018;7:e34408.

30. Hemani G, Tilling K, Davey SG. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet*. 2017;13:e1007081. doi:10.1371/JOURNAL.PGEN.1007081

31. Auton A, Abecasis GR, Altshuler DM, et al. A global reference for human genetic variation. *Nature*. 2015;526:68-74. doi:10.1038/nature15393

32. Wootton RE, Sallis HM. Let's call it the effect allele: a suggestion for GWAS naming conventions. *Int J Epidemiol*. 2020;49:1734-1735. doi:10.1093/IJE/DYAA149

33. Burgess S, Dudbridge F, Thompson SG. Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods. *Stat Med*. 2016;35:1880-1906. doi:10.1002/sim.6835

34. Shim H, Chasman DI, Smith JD, et al. A multivariate genome-wide association analysis of 10 LDL subfractions, and their response to statin treatment, in 1868 Caucasians. *PLoS One*. 2015;10:120758. doi:10.1371/JOURNAL.PONE.0120758

35. Burgess S, Davey Smith G, Davies NM, et al. Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res*. 2020;4:186. doi:10.12688/wellcomeopenres.15555.2

36. 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. doi:10.1093/ije/dyv080

37. 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. doi:10.1002/gepi.21965

38. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol*. 2017;46:1985-1998. doi:10.1093/ije/dyx102

39. 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. doi:10.1038/s41588-018-0099-7

40. Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable mendelian randomization in the single-sample and two-sample summary data settings. *Int J Epidemiol*. 2019;48:713-727. doi:10.1093/ije/dyy262

41. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562:203-209. doi:10.1038/s41586-018-0579-z

42. Sheard S, Nicholls R, Froggatt J. *UK Biobank Haematology Data Companion Document*. Stockport, UK: Oxford University Press; 2016.

43. Burrows K, Bull CJ, Dudding T, et al. *Genome-Wide Association Study of Cancer Risk in UK Biobank*. Bristol, UK: University of Bristol; 2021. doi:10.5523/bris.aed0u12w0ede20olb0m77p4b9

44. Astle WJ, Elding H, Jiang T, et al. The allelic landscape of human blood cell trait variation and links to common complex disease. *Cell*. 2016;167:1415-1429.e19. doi:10.1016/j.cell.2016.10.042

45. Sanderson E, Spiller W, Bowden J. Testing and correcting for weak and pleiotropic instruments in two-sample multivariable Mendelian randomization. *Stat Med*. 2021;40:5434-5452. doi:10.1002/SIM.9133

46. Burgess S, Thompson SG. Bias in causal estimates from Mendelian randomization studies with weak instruments. *Stat Med*. 2011;30(11):1312-1323.

47. Wong JYY, Bassig BA, Loftfield E, et al. White blood cell count and risk of incident lung cancer in the UK Biobank. *JNCI Cancer Spectr*. 2020;4:pkz102. doi:10.1093/JNCICS/PKZ102

48. Watts EL, Perez-Cornago A, Kothari J, Allen NE, Travis RC, Key TJ. Hematologic markers and prostate cancer risk: a prospective analysis in UK Biobank. *Cancer Epidemiol Biomarkers Prev*. 2020;29:1615-1626. doi:10.1158/1055-9965.EPI-19-1525

49. Benson VS, Hartl S, Barnes N, Galwey N, van Dyke MK, Kwon N. Blood eosinophil counts in the general population and airways disease: a comprehensive review and meta-analysis. *Eur Respir J*. 2022;59:2004590. doi:10.1183/13993003.04590-2020

50. Morrison J, Knoblauch N, Marcus JH, Stephens M, He X. Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics. *Nat Genet*. 2020;52:740-747. doi:10.1038/s41588-020-0631-4

51. Karim AF, Westenberg LEH, Eurelings LEM, Otten R, Gerth Van Wijk R. The association between allergic diseases and cancer: a systematic review of the literature. *Neth J Med*. 2019;77(2):42-66.

52. Varricchi G, Galdiero MR, Loffredo S, et al. Eosinophils: the unsung heroes in cancer? *Oncoimmunology*. 2018;7(2):e1393134. doi:10.1080/2162402X.2017.1393134

53. Arnold IC, Artola-Boran M, Gurtner A, et al. The GM-CSF-IRF5 signaling axis in eosinophils promotes antitumor immunity through activation of type 1 T cell responses. *J Exp Med*. 2020;217:e20190706. doi:10.1084/JEM.20190706/152117

54. Platzer B, Elpek KG, Cremasco V, et al. IgE/FcεRI-mediated antigen cross-presentation by dendritic cells enhances anti-tumor immune responses. *Cell Rep*. 2015;10:1487. doi:10.1016/J.CELREP.2015.02.015

55. Legrand F, Driss V, Delbeke M, et al. Human eosinophils exert TNF-α and granzyme A-mediated tumoricidal activity toward colon carcinoma cells. *J Immunol*. 2010;185:7443-7451. doi:10.4049/JIMMUNOL.1000446

56. Idos GE, Kwok J, Bonthala N, Kysh L, Gruber SB, Qu C. The prognostic implications of tumor infiltrating lymphocytes in colorectal cancer: a systematic review and meta-analysis. *Sci Rep*. 2020;10:1-14. doi:10.1038/s41598-020-60255-4

57. Boursi B, Mamtani R, Hwang WT, Haynes K, Yang YX. A risk prediction model for sporadic CRC based on routine lab results. *Dig Dis Sci*. 2016;61:2076-2086. doi:10.1007/S10620-016-4081-X/TABLES/4

58. Goshen R, Mizrahi B, Akiva P, et al. Predicting the presence of colon cancer in members of a health maintenance organisation by evaluating analytes from standard laboratory records. *Br J Cancer*. 2017;116:944-950. doi:10.1038/bjc.2017.53

59. Siedlinski M, Jozefczuk E, Xu X, et al. White blood cells and blood pressure indices. *Circulation*. 2020;141:1307. doi:10.1161/CIRCULATIONAHA.119.045102

60. Welsh C, Welsh P, Mark PB, et al. Association of total and differential leukocyte counts with cardiovascular disease and mortality in the UK Biobank. *Arterioscler Thromb Vasc Biol*. 2018;38:1415-1423. doi:10.1161/ATVBAHA.118.310945

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Constantinescu A-E, Bull CJ, Jones N, et al. Circulating white blood cell traits and colorectal cancer risk: A Mendelian randomisation study. *Int J Cancer*. 2023; 1-10. doi:10.1002/ijc.34691