OXFORD

JNCI Cancer Spectrum (2020) 4(2): pkz102

doi: 10.1093/jncics/pkz102
First published online December 12, 2019
Article

White Blood Cell Count and Risk of Incident Lung Cancer in the UK Biobank

Jason Y.Y. Wong, ScD[1,*], Bryan A. Bassig, PhD[1], Erikka Loftfield, PhD[1], Wei Hu, PhD[1], Neal D. Freedman, PhD[1], Bu-tian Ji, PhD[1], Paul Elliott, PhD[2], Debra T. Silverman, ScD[1], Stephen J. Chanock, MD[1], Nathaniel Rothman, MD[1], Qing Lan, MD, PhD[1]

[1]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA and [2]Faculty of Medicine, School of Public Health, Imperial College London, London, UK

[*]**Correspondence to:** Jason Y. Y. Wong, National Cancer Institute, 9609 Medical Center Drive, Rockville, MD 20850 (e-mail: jason.wong@nih.gov).

Abstract

Background: The contribution of measurable immunological and inflammatory parameters to lung cancer development remains unclear, particularly among never smokers. We investigated the relationship between total and differential white blood cell (WBC) counts and incident lung cancer risk overall and among subgroups defined by smoking status and sex in the United Kingdom (UK).

Methods: We evaluated 424 407 adults aged 37–73 years from the UK Biobank. Questionnaires, physical measurements, and blood were administered and collected at baseline in 2006–2010. Complete blood cell counts were measured using standard methods. Lung cancer diagnoses and histological classifications were obtained from cancer registries. Multivariable Cox regression models were used to estimate the hazard ratio (HR) and 95% confidence intervals of incident lung cancer in relation to quartiles (Q) of total WBC and subtype-specific counts, with Q1 as the reference.

Results: There were 1493 incident cases diagnosed over an average 7-year follow-up. Overall, the highest quartile of total WBC count was statistically significantly associated with elevated lung cancer risk ($HR_{Q4} = 1.67$, 95% CI = 1.41 to 1.98). Among women, increased risks were found in current smokers (n_{cases} / $n = 244$ / 19 464, $HR_{Q4} = 2.15$, 95% CI = 1.46 to 3.16), former smokers (n_{cases} / $n = 280$ / 69 198, $HR_{Q4} = 1.75$, 95% CI = 1.24 to 2.47), and never smokers without environmental tobacco smoke exposure (n_{cases} / n = 108 / 111 294, $HR_{Q4} = 1.93$, 95% CI = 1.11 to 3.35). Among men, stronger associations were identified in current smokers ($n_{case\ s}$ / $n = 329$ / 22 934, $HR_{Q4} = 2.95$, 95% CI = 2.04 to 4.26) and former smokers (n_{cases} / $n = 358$/71 616, $HR_{Q4} = 2.38$, 95% CI = 1.74 to 3.27) but not in never smokers. Findings were similar for lung adenocarcinoma and squamous cell carcinoma and were driven primarily by elevated neutrophil fractions.

Conclusions: Elevated WBCs could potentially be one of many important markers for increased lung cancer risk, especially among never-smoking women and ever-smoking men.

Lung cancer is one of the most common malignancies in the Western world, with over 312 000 newly diagnosed cases every year in the European Union (EU) (1) and an estimated 234 030 in the United States (USA) in 2018 (2). Lung cancer has an estimated 5-year survival rate of only 18.6% because of diagnoses at late stages (2), and care for lung cancer patients constitutes a substantial economic burden in both the EU and USA (3,4). Identifying disease mechanism and risk factors in at-risk groups is crucial for devising population-level interventions.

Elevated white blood cell (WBC) counts and immune markers have been associated with lung cancer in prospective cohort studies, suggesting a role for immune and inflammatory processes in lung cancer pathogenesis (5–12). However, gaps in knowledge remain. In particular, the association between elevated WBC counts and lung cancer risk has not been extensively investigated among never smokers because of challenges in accruing cases in Western populations (12). Lung cancer is largely considered to be driven by smoking (13) in Western countries. However, it remains a notable malignancy among never smokers (14,15), with estimated incidence rates of 14.4–20.8 cases per 100 000 person-years for never-smoking women and 4.8–13.7 per 100 000 person-years for never-smoking men (16,17). Differences in immunologic and inflammatory parameters such as WBC counts among these subgroups could contribute to the

Received: August 4, 2019; **Revised:** November 26, 2019; **Accepted:** December 9, 2019

varying rates. Furthermore, previous epidemiologic studies reported findings for total WBC counts, but not by WBC subtypes that could differentially influence risk. Additionally, previous studies did not consider different histological subtypes of lung cancer (ie, adenocarcinoma and squamous cell carcinoma [SCC]), which have different risk factors and pathogenesis; therefore, deeper investigation could provide insight into disease etiology.

To investigate the relationships between total and differential WBC counts and incident lung cancer risk overall and among subgroups defined by smoking status and sex, we analyzed data from the United Kingdom (UK) Biobank, a prospective cohort study of over half a million adults. This investigation expands upon previous studies by including measurements of WBC subtypes and assessment of lung cancer histology, which could provide deeper understanding of the role key immunological and inflammatory processes play in lung cancer development.

Methods

Study Design

The study design and data access procedures of the UK Biobank have been described elsewhere (http://www.ukbiobank.ac.uk/) (18,19). Briefly, the target population was adults aged 40–69 years who resided within 40 km of 22 assessment centers across the UK (18). Of the 9.2 million individuals registered in the UK's National Health Service (NHS) who were mailed invitations, 503 317 (5.5%) visited the assessment centers in 2006–2010 (19). These volunteers received physical examinations, provided biological samples including 40–50 mL of whole blood, and were administered touchscreen questionnaires. Data from 502 616 participants were initially available for our analysis. After applying the exclusion criteria and data cleaning as specified in Figure 1, the final analytic sample size was 424 407 participants.

The UK Biobank study was approved by the National Information Governance Board for Health and Social Care and the NHS North West Multicenter Research Ethics Committee. Electronic informed consent was obtained from all volunteer participants (18,19).

WBC Counts

WBCs and their subtypes (ie, lymphocytes, neutrophils, monocytes, eosinophils, and basophils) were measured in whole blood obtained at baseline as described online (http://www.ukbiobank.ac.uk/). Complete blood cell counts (cells/L) were measured using a Coulter LH 750 System (Beckman Coulter, Brea, CA) as per

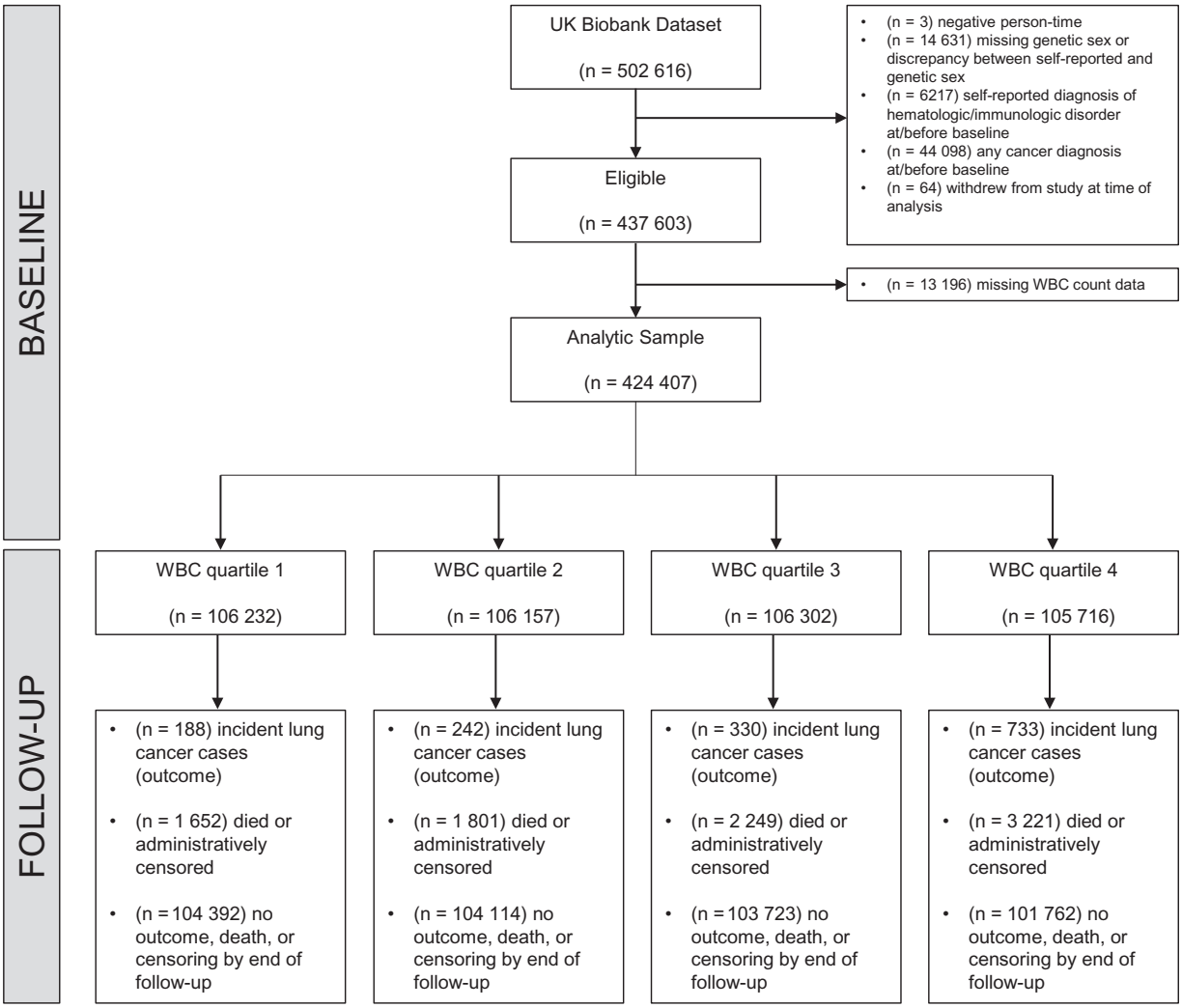

Figure 1. Flow chart of baseline exclusion criteria and the prospective follow-up. Abbreviation: WBC = white blood cell.

manufacturer's procedures. Measures of reliability/reproducibility for WBC counts are described in Appendix 1 (available online).

Lung Cancer Ascertainment

Cancer diagnoses were provided to the UK Biobank by the Health and Social Care Information Centre and the NHS Central Register (NHSCR) (http://www.ukiacr.org/). *Lung cancer diagnosis* was defined by International Classification of Diseases 10th revision (ICD-10 codes) C34.0-C34.9. ICD-O-3 code 8140 was used to define *adenocarcinoma*, whereas 8052, 8084, 8073, and 8083 were used for *SCC*.

Cohort Follow-Up

Follow-up time was counted from the date of visit to the assessment center until the date of incident lung cancer diagnosis, death, or administrative censoring (ie, January 31, 2016, for England and Wales, and November 30, 2015, for Scotland), whichever came first. The NHS Information Centre and the NHSCR Scotland provided vital status, along with date and primary underlying cause of death.

Statistical Analysis

A detailed 25-level smoking variable was created as previously described (20) (Table 1). A six-level variable for alcohol intake was utilized (*never, former, current infrequent* [less than 3 times per month], *current modest* [less than 1 drink per day], *current frequent* [1 to 3 drinks per day], and *current heavy* [greater than 3 drinks per day]) (21). Body mass index (BMI, kg/m^2) was categorized as less than 18.5; 18.5 to 25.0; 25.0 to 30.0; 30.0 to 35.0; and 35.0 or greater kg/m^2. Self-reported race/ethnicity/ancestry was categorized as European, Asian (East and South), Black (African ancestry), mixed, other, and missing/unknown/no answer. Less than 9% of the cohort lacked data on any single covariate. Indicator variables were created for missing data among categorical variables, whereas mean imputation was used to account for missing Townsend Deprivation Index (continuous).

Multivariable Cox regression models were used to estimate cause-specific hazard ratios (HR) and 95% confidence intervals (CI) of incident lung cancer diagnosis and its histological subtypes, in relation to quartiles (Q) of total WBC, lymphocyte, neutrophil, monocyte, basophil, and eosinophil counts in separate analyses. The lowest quartile (Q1) was the reference group. We also analyzed the neutrophil-to-lymphocyte ratio (NLR). Those who died or were administratively censored before the outcome across the follow-up were not counted as outcomes. We adjusted for detailed smoking (20) (reference: never smokers), race/ethnicity/ancestry (reference: European), study assessment center, age at recruitment (continuous), sex (reference: male), BMI (reference: 18.5 to 25.0 kg/m^2), Townsend Deprivation Index (continuous), and alcohol intake (reference: never drinker) (21). We also considered environmental tobacco smoke (ETS) exposure and ever having a self-reported respiratory disease as covariates. However, adjusting for these variables did not considerably change the estimates; therefore, results from the more parsimonious model are presented. Follow-up time (days) was used as the underlying timescale. Trends were tested by analyzing the WBC categories as ordinal. Supremum tests were used to assess proportional hazards assumptions.

The analyses were further stratified by a sex-smoking combination variable (ie, never-, former-, and current-smoking women and men). The stratified analyses did not include cigar and pipe smokers. We also conducted a sensitivity analysis restricting the overall study population to European participants (94%) to assess the influence of race/ethnicity/ancestry. For the main stratified analyses, we based the quartiles of WBC count on the distribution of each subgroup because of previously established differences by sex (22,23) and smoking (24–27), as well as race/ethnicity/ancestry (22,23). We conducted sensitivity analyses of the subgroups using common WBC quartile cutoffs from the overall analyses. Multiplicative effect modification of WBC counts by a sex-smoking combination variable and age was assessed using interaction terms. The interaction analyses with the sex-smoking combination were conducted using the overall WBC quartile cutoffs. In separate analyses, never-smoking women and men were restricted to those not exposed to ETS (less than1 hour per week) to mitigate the influence of secondhand smoke.

To assess the potential influence of underlying disease bias on WBC counts, we conducted separate analyses in which cases diagnosed before 2 years of follow-up were excluded as outcomes. In additional sensitivity analyses, we restricted the overall study population to: (1) those who reported not being exposed to ETS (less than1 hour per week), 2) those who reported never having a history of respiratory diseases, and 3) those aged younger than 67 years at baseline because of potentially reduced ability to produce higher WBC counts among the oldest participants (28). All stratified, subgroup, and sensitivity analyses were defined a priori with a hypothesis-driven approach.

Two-sided *P*-values $< .05$ were considered noteworthy. We applied an extremely conservative Bonferroni corrected threshold of $\alpha = 1.7 \times 10^{-4}$ to account for multiple comparisons ($\alpha = 0.05/295$ total tests across all analyses).

Results

Study Population Characteristics

Among the participants with WBC data, we analyzed 1493 incident lung cancer cases over a maximum 10-year follow-up (7-year average). The average age at recruitment was 56.2 ± 8.1 SD years. The cumulative person-time was 2.94 million person-years (py) and the average age of lung cancer diagnosis among cases was 65.2 ± 6.0 SD years. There were 562 confirmed lung adenocarcinomas and 285 confirmed lung SCC. Differences in WBC counts between categories of various population characteristics at baseline are shown in Table 1.

Total and Differential WBC Count and Lung Cancer Risk

Overall, we observed a positive trend in lung cancer risk with increasing WBC count ($P_{trend} = 1.4 \times 10^{-12}$). Compared to the lowest quartile of total WBC, the highest quartile was associated with an estimated 1.67 (95% CI = 1.41 to 1.98, $p = 4.4 \times 10^{-9}$) times increased risk (Table 2). This relationship was primarily driven by an association with neutrophils ($HR_{Q4} = 1.57$, 95% CI = 1.33 to 1.84, $P = 7.1 \times 10^{-8}$). There was also evidence for a marginal positive association between monocytes and lung cancer risk ($HR_{Q4} = 1.17$, 95% CI = 1.00 to 1.36, $P = .04$); however, monocytes comprise only a small fraction of WBC and displayed narrow statistical variability. There was no evidence for interactions between quartiles of WBC count and age ($P_{interactions} = 0.63, 0.79, 0.88$ across quartiles).

Table 1. White blood cell counts by population characteristics in the UK Biobank ($n = 424\,407$)

	Total white blood cell count (10^9 cells/L)			
	n, subjects	%	Mean	SD
Overall	424 407	100.0	6.88	1.87
Age at recruitment, y				
37–46	68 228	16.1	6.87	1.83
47–56	130 826	30.8	6.80	1.83
57–66	183 107	43.1	6.89	1.89
67–76	42 246	10.0	7.08	1.94
Genetic sex				
Female	224 139	52.8	6.86	1.83
Male	200 268	47.2	6.90	1.91
Self-reported race/ethnicity/ancestry				
European	399 302	94.1	6.89	1.87
Black, African	6694	1.6	5.80	1.65
Asian, East and South	9955	2.3	7.20	1.78
Mixed	2522	0.6	7.06	1.85
Other	3904	0.9	6.81	2.06
Unknown, missing, no answer	2030	0.5	6.91	1.88
BMI kg/m^2				
< 18.5	1908	0.4	6.61	2.56
≥ 18.5 to < 25.0	132 902	31.3	6.56	1.87
≥ 25.0 to < 30.0	179 584	42.3	6.84	1.82
≥ 30.0 to < 35.0	73 931	17.4	7.20	1.79
≥ 35.0	28 647	6.7	7.72	1.89
Missing	7 435	1.8	7.14	2.03
Townsend Deprivation Index				
Lower ($<$ median of -2.13)	212 813	50.1	6.77	1.81
Upper ($\geq$ median of -2.13)	211 594	49.9	6.99	1.92
Smoking status and history				
Never	232 528	54.8	6.67	1.76
Current occasional smoker, smoked $<$ 100 cigarettes in lifetime	538	0.1	6.89	1.79
Current occasional smoker, smoked $\geq$ 100 cigarettes in lifetime	6656	1.6	6.92	1.74
Current occasional smoker, smoked cigars or pipes daily in past	265	0.1	7.27	1.94
Current occasional smoker, smoked cigarettes daily in past	2476	0.6	7.32	1.96
Current occasional smoker, smoked cigarettes daily in past, $\geq$ 20 cigarettes/day	1436	0.3	7.60	1.95
Current cigar or pipe smoker, former cigarette smoker	1569	0.4	8.12	2.06
Current cigar or pipe smoker, not former cigarette smoker	744	0.2	7.56	1.97
Current cigarette smoker, $<$ 10 cigarettes/day	6099	1.4	7.68	1.99
Current cigarette smoker, $\geq$ 10 to $<$ 20 cigarettes/day	12 863	3.0	8.47	2.10
Current cigarette smoker, $\geq$ 20 to $<$ 40 cigarettes/day	10 665	2.5	8.89	2.20
Current cigarette smoker, $\geq$ 40 cigarettes/day	721	0.2	9.08	2.25
Former occasional cigarette smoker, smoked $<$ 100 cigarettes in lifetime	6836	1.6	6.69	1.68
Former occasional cigarette smoker, smoked $\geq$ 100 cigarettes in lifetime	38 349	9.0	6.68	1.71
Former occasional cigarette smoker, lifetime smoking unknown	3012	0.7	6.67	1.66
Former daily cigar pipe smoker	4215	1.0	6.98	1.79
Former cigarette smoker, $<$ 20 cigarettes/day, quit $<$ 1 year ago	907	0.2	7.39	1.92
Former cigarette smoker, $\geq$ 20 cigarettes/day, quit $<$ 1 year ago	1058	0.2	7.70	1.91
Former cigarette smoker, $<$ 20 cigarettes/day, quit $\geq$ 1 to $<$ 5 year ago	4777	1.1	7.00	1.77
Former cigarette smoker, $\geq$ 20 cigarettes/day, quit $\geq$ 1 to $<$5 year ago	5886	1.4	7.34	1.91
Former cigarette smoker, $<$ 20 cigarettes/day, quit $\geq$ 5 to $<$10 year ago	5045	1.2	6.90	1.73
Former cigarette smoker, $\geq$ 20 cigarettes/day, quit $\geq$ 5 to $<$ 10 year ago	7040	1.7	7.19	1.87
Former cigarette smoker, $<$ 20 cigarettes/day, quit $\geq$10 to $<$20 year ago	8943	2.1	6.79	1.69
Former cigarette smoker, $\geq$ 20 cigarettes/day, quit $\geq$ 10 to $<$ 20 year ago	12 256	2.9	7.04	1.90
Former cigarette smoker, $<$ 20 cigarettes/day, quit $\geq$ 20 year ago	20 302	4.8	6.66	1.78
Former cigarette smoker, $<$ 20/day (less than 1/day), quit $\geq$ 20 year ago	25 329	6.0	6.87	1.79
Other/unknown/missing	3892	0.9	7.23	2.34
Self-reported ETS exposure at home or outside				
Unexposed ($<$ 1 hour/week)	343 646	81.0	6.89	1.88
Exposed ($\geq$1 hour/week)	80 761	19.0	6.84	1.81

(continued)

Table 1. (continued)

	Total white blood cell count (10^9 cells/L)			
	n, subjects	%	Mean	SD
History of respiratory disease (ie, asthma, chronic obstructive pulmonary disease, emphysema, bronchitis, bronchiectasis, interstitial lung disease, asbestosis, pulmonary fibrosis, or other respiratory problems)				
No	371 250	87.5	6.83	1.85
Yes	53 157	12.5	7.20	1.99
Alcohol consumption				
Never	18 460	4.3	7.00	2.01
Former	14 862	3.5	7.22	2.08
Current infrequent, < 3 times/month	94 731	22.3	7.06	1.98
Current modest, < 1 drink/day	105 131	24.8	6.78	1.84
Current frequent, ≥ 1 to ≤ 3 drinks/day	153 623	36.2	6.76	1.73
Current heavy, > 3 drinks/day	36 404	8.6	6.96	1.98
Missing	1196	0.3	7.09	1.95

Abbreviations: BMI = body mass index; ETS = environmental tobacco smoke.

WBC Count and Lung Cancer Risk among Subgroups Defined by Smoking Status and Sex

Among the 232 528 never smokers, the highest quartile of WBC count, associated with increased risk of lung cancer, was compared to the lowest quartile (HR_{Q4} = 1.52, 95% CI = 1.03 to 2.25, P = .03; HR_{Q3} = 1.00, 95% CI = 0.64 to 1.50, P = .99; HR_{Q2} = 1.10, 95% CI = 0.75 to 1.62, P = .63).

Among women, risk estimates were elevated among never smokers (HR_{Q4} = 1.82, 95% CI = 1.10 to 3.00, P = .02), former smokers (HR_{Q4} = 1.75, 95% CI = 1.24 to 2.47, $P = 1.4 \times 10^{-3}$), and current smokers (HR_{Q4} = 2.15, 95% CI = 1.46–3.16, $P = 1.0 \times 10^{-4}$) (Table 3). Among never-smoking women, excluding those who were exposed to ETS nominally increased the risk estimate (HR_{Q4} = 1.93, 95% CI = 1.11 to 3.35, P = .02) (Table 3).

Among men, the associations were moderately stronger for former smokers (HR_{Q4} = 2.38, 95% CI = 1.74 to 3.27, $P = 6.3 \times 10^{-8}$) and especially for current smokers (HR_{Q4} = 2.95, 95% CI = 2.04 to 4.26, $P = 8.5 \times 10^{-9}$). However, noteworthy associations were not observed for never-smoking men, with or without excluding those exposed to ETS (Table 3).

Based on the apparent differences among the stratified risk estimates, we tested for multiplicative interactions between WBC counts and the smoking-sex combination variable. Compared to never-smoking men, there was no evidence of effect modification for never-smoking women ($P_{interaction}$ = 0.22); however, there was evidence among former-smoking women ($P_{interaction}$ = 0.04), current-smoking women ($P_{interaction}$ = 0.02), former-smoking men ($P_{interaction} = 2.4 \times 10^{-3}$) and current-smoking men $P_{interaction} = 7.0 \times 10^{-4}$).

We also analyzed current-smoking women and men stratified by smoking intensity. The associations between WBC count and lung cancer risk were slightly stronger among those who currently smoke less than 20 cigarettes per day, compared with 20 or greater cigarettes per day (Women: HR_{Q4} = 2.55, 95% CI = 1.17 to 5.57, P = .02 vs. HR_{Q4} = 0.60, 95% CI = 0.26 to 1.40, P = .24; Men: HR_{Q4} = 5.23, 95% CI = 2.12 to 12.90, $P = 3.2 \times 10^{-4}$ vs. HR_{Q4} = 1.83, 95% CI = 0.67 to 4.98, P = .24). There was evidence for effect modification among women ($P_{interaction}$ = 0.03), but not among men ($P_{interaction}$ = 0.11).

WBC Count, Lung Adenocarcinoma, and SCC

We also analyzed histological subtypes of lung cancer and found similar trends for lung adenocarcinoma and SCC (Table 2) relative to the overall analyses. Notably, there were similar risk estimates for the analyses of WBC count and lung adenocarcinoma (HR_{Q4} = 1.68, 95% CI = 1.28 to 2.22, $P = 2.0 \times 10^{-4}$) and SCC (HR_{Q4} = 1.55, 95% CI = 1.04 to 2.29, P = .03), compared with overall lung cancer. Additionally, there was evidence for an association between lymphocytes and SCC (HR_{Q4} = 1.46, 95% CI = 1.03 to 2.06, P = .03). Although the number of cases was small, we also stratified the analyses of lung adenocarcinoma by smoking status and sex. Similar patterns were found as the analyses for overall lung cancer (Supplementary Table 1 available online).

Additional Analyses

In the 2-year lagged analyses, increased total WBC counts remained associated with increased lung cancer risk (Supplementary Table 2 available online). Additionally, comparable results were found when restricting to Europeans (Supplementary Table 3 available online). Furthermore, excluding those who self-reported ever having a respiratory disease or ETS exposure, as well as excluding participants who were aged 67 years or older at recruitment from the analyses did not influence the results (data not shown). In the subgroup analyses, using subgroup-specific or overall quartile cut points did not affect the findings (Supplementary Tables 4 and 5 available online). Lastly, we found that increasing NLR was associated with increased lung cancer risk overall ($HR_{per\ 1.0\ increase}$ = 1.14, 95% CI = 1.09 to 1.19, $P = 8.3 \times 10^{-9}$), among never- and current-smoking women, and ever-smoking men (Supplementary Table 6 available online).

Discussion

In a prospective cohort study of nearly half a million adults from the UK, we observed that elevated total WBC count was associated with increased risk of developing lung cancer. The risk estimates among current-, former-, and never-smoking women were comparable to the overall analyses, while the associations were stronger among former- and current-smoking men. Similar elevated risk estimates were found for the two major subtypes of lung cancer, adenocarcinoma and SCC. The associations between WBC counts and lung cancer were primarily driven by elevated neutrophil counts, a key component of innate immunity which constitutes 40%–60% of WBCs. Similar to

Table 2. Associations between total and differential white blood cell counts and lung cancer risk in the UK Biobank ($n = 424\,407$)

		I) Overall lung cancer ($n = 1493$)					II) Lung adenocarcinoma ($n = 562$)					III) Lung squamous cell carcinoma ($n = 285$)				
Quartile	Cell counts, 10^9 cells/L	No. of incident cases	HR	95% CI lower	95% CI upper	*P*	No. of incident cases	HR	95% CI lower	95% CI upper	*P*	No. of incident cases	HR	95% CI lower	95% CI upper	*P*
Total WBC																
1	≤5.64	188	1.00				76	1.00				34	1.00			
2	>5.64-≤6.66	242	1.04	0.86	1.26	.68	116	1.30	0.97	1.73	.08	34	0.74	0.46	1.19	.21
3	>6.66-≤7.86	330	1.13	0.94	1.35	.20	121	1.13	0.85	1.52	.40	56	0.89	0.58	1.38	.61
4	>7.86	733	1.67	1.41	1.98	$4.4\text{x}10^{-9}$**	249	1.68	1.28	2.22	$2.0\text{x}10^{-4}$*	161	1.55	1.04	2.29	.03*
					P_{trend}	$1.4\text{x}10^{-12}$**				P_{trend}	$2.8\text{x}10^{-4}$*				P_{trend}	$4.7\text{x}10^{-4}$*
Lymphocytes																
1	≤1.52	297	1.00				127	1.00				51	1.00			
2	>1.52-≤1.89	280	0.91	0.77	1.07	.26	109	0.83	0.64	1.07	.14	54	1.02	0.69	1.50	.92
3	>1.89-≤2.30	389	1.00	0.86	1.17	.98	143	0.89	0.69	1.13	.32	58	0.86	0.58	1.25	.42
4	>2.30	524	1.13	0.97	1.31	.12	181	0.95	0.75	1.21	.68	121	1.46	1.03	2.06	.03*
					P_{trend}	.04*				P_{trend}	.91				P_{trend}	.03*
Neutrophils																
1	≤3.27	210	1.00				95	1.00				36	1.00			
2	>3.27-≤4.02	232	0.87	0.72	1.05	.15	82	0.72	0.54	0.97	.03*	30	0.60	0.37	0.97	.04*
3	>4.02-≤4.97	337	1.06	0.89	1.26	.52	143	1.10	0.85	1.44	.47	66	1.02	0.68	1.55	.91
4	>4.97	711	1.57	1.33	1.84	$7.1\text{x}10^{-8}$**	240	1.40	1.09	1.81	.01*	152	1.51	1.03	2.21	.04*
					P_{trend}	$2.9\text{x}10^{-13}$**				P_{trend}	$6.3\text{x}10^{-5}$**				P_{trend}	$1.4\text{x}10^{-4}$**
Basophils																
1	0	487	1.00				178	1.00				89	1.00			
2	>0-≤0.02	232	0.84	0.69	1.01	.06	111	0.90	0.67	1.21	.48	26	0.54	0.33	0.89	.01*
3	>0.02-<0.04	268	0.86	0.71	1.03	.10	112	0.83	0.62	1.12	.23	63	1.07	0.71	1.61	.74
4	≥0.04	503	1.12	0.98	1.30	.11	159	0.91	0.71	1.17	.47	106	1.21	0.88	1.68	.24
					P_{trend}	.02*				P_{trend}	.52				P_{trend}	.02*
Monocytes																
1	≤0.37	277	1.00				131	1.00				35	1.00			
2	>0.37-≤0.45	285	0.95	0.81	1.12	.56	106	0.79	0.61	1.02	.08	56	1.38	0.90	2.11	.14
3	>0.45-≤0.57	335	0.91	0.77	1.07	.25	135	0.84	0.66	1.08	.17	65	1.20	0.80	1.82	.38
4	>0.57	592	1.17	1.00	1.36	.04*	188	0.92	0.73	1.17	.50	128	1.51	1.03	2.22	.04*
					P_{trend}	.02*				P_{trend}	.73				P_{trend}	.06
Eosinophils																
1	≤0.10	511	1.00				206	1.00				83	1.00			
2	>0.10-≤0.14	150	1.01	0.84	1.22	.94	69	1.08	0.81	1.43	.61	29	1.14	0.74	1.77	.56
3	>0.14-≤0.21	378	0.93	0.81	1.06	.26	124	0.79	0.64	1.00	.04*	79	1.08	0.79	1.47	.64
4	>0.21	408	0.91	0.79	1.04	.15	143	0.84	0.68	1.05	.12	83	0.97	0.71	1.33	.85
					P_{trend}	.11				P_{trend}	.04*				P_{trend}	.86

*P-values < .05.

**P-values below a Bonferroni corrected alpha threshold of $1.7\text{x}10^{-4}$.

Quartiles of WBC counts were based on the distribution of the overall study population. Cox regression models were adjusted for age, sex, BMI, race/ethnicity/ancestry, assessment center, detailed smoking status/intensity/history, alcohol consumption, and Townsend Deprivation Index as surrogate for socioeconomic status.

Abbreviations: HR = hazard ratio; CI = confidence interval; WBC = white blood cell.

Discrepancy in counts is due to missing blood cell data.

Table 3. Associations between total white blood cell counts and overall lung cancer risk in the UK Biobank by cigarette smoking status and sex

I) women							II) men						
Smoking status	Quartile, WBC count, 10^9 cells/L	No. of incident cases	HR	95% CI lower	95% CI upper	P	Smoking status	Quartile, WBC count, 10^9 cells/L	No. of incident cases	HR	95% CI lower	95% CI upper	P
Never smokers	1, ≤5.55	26	1.00				Never smokers,	1, ≤5.50	22	1.00			
$n = 134\,209$	2, >5.55-≤6.51	33	1.27	0.76	2.13	.37	$n = 98\,319$	2, >5.50-≤6.41	17	0.77	0.41	1.46	.42
	3, >6.51-≤7.67	27	1.09	0.64	1.88	.75		3, >6.41-≤7.50	15	0.63	0.32	1.22	.17
	4, >7.67	41	1.82	1.10	3.00	.02*		4, >7.50	24	0.96	0.53	1.75	.90
					P_{trend}	.04*						P_{trend}	.81
Never smokers, no ETS exposure	1, ≤5.55	21	1.00				Never smokers, no ETS exposure,	1, ≤5.50	15	1.00			
$n = 111\,294$	2, >5.55-≤6.51	30	1.43	0.82	2.49	.21	$n = 75\,918$	2, >5.50-≤6.41	14	0.89	0.43	1.85	.75
	3, >6.51-≤7.67	22	1.11	0.61	2.03	.74		3, >6.41–≤7.50	13	0.74	0.35	1.58	.44
	4, >7.67	35	1.93	1.11	3.35	.02*		4, >7.50	20	1.09	0.55	2.16	.81
					P_{trend}	.05*						P_{trend}	.87
Former smokers	1, ≤5.60	52	1.00				Former smokers,	1, ≤5.70	55	1.00			
$n = 69\,198$	2, >5.60–≤6.60	58	1.05	0.72	1.53	.79	$n = 71\,616$	2, >5.70–≤6.70	65	1.11	0.77	1.59	.59
	3, >6.60–≤7.73	71	1.31	0.91	1.88	.14		3, >6.70–≤7.82	82	1.33	0.94	1.88	.11
	4, >7.73	99	1.75	1.24	2.47	$1.4x10^{-3*}$		4, >7.82	156	2.38	1.74	3.27	$6.3x10^{-8**}$
					P_{trend}	$3.6x10^{-4*}$						P_{trend}	$7.7x10^{-10**}$
Current smokers	1, ≤6.60	38	1.00				Current smokers,	1, ≤6.58	37	1.00			
$n = 19\,464$	2, >6.60–≤7.85	51	1.25	0.82	1.90	.30	$n = 22\,934$	2, >6.58–≤7.90	62	1.37	0.91	2.06	.13
	3, >7.85–≤9.30	63	1.45	0.96	2.17	.08		3, >7.90–≤9.39	93	2.04	1.39	3.00	$2.6x10^{-4*}$
	4, >9.30	92	2.15	1.46	3.16	$1.0x10^{-4**}$		4, >9.39	137	2.95	2.04	4.26	$8.5x10^{-9**}$
					P_{trend}	$3.3x10^{-5**}$						P_{trend}	$1.6x10^{-11**}$

*P-values $< .05$.

**P-values below a Bonferroni corrected alpha threshold of 1.7×10^{-4}.

The analyses did not include cigar and pipe smokers. Cox regression models were adjusted for age, BMI, race/ethnicity/ancestry, assessment center, alcohol consumption, and Townsend Deprivation Index as a surrogate for socio-economic status. Discrepancy in counts is due to missing blood cell data. WBC quartiles were based on the distribution of each sex-smoking subgroup.

Abbreviations: HR = hazard ratio; CI = confidence interval; ETS = environmental tobacco smoke; WBC = white blood cell.

We tested multiplicative effect modification of WBC counts by the smoking-sex combination variable using interaction terms. Compared with never-smoking men, there was no evidence of effect modification for never-smoking women ($P_{interaction} = 0.22$); however, there was evidence among formerly smoking women ($P_{interaction} = 0.04$), currently smoking women ($P_{interaction} = 0.02$), formerly smoking men ($P_{interaction} = 2.4 \times 10^{-3}$), and currently smoking men ($P_{interaction} = 7.0 \times 10^{-4}$).

previous studies (29–31), we found that increasing NLR was associated with increased risk. Notably, our study is one of the first of sufficient power to investigate the relationship between WBC count and lung cancer risk among never-smoking women, an under-studied group. Our findings are consistent and robust across sensitivity analyses.

In the stratified analyses, there was evidence for a stronger relationship between WBC count and lung cancer risk among smokers. Additionally, the associations were generally greater among men, with risk estimates being highest for current-smoking men. Among current cigarette smokers, most of the moderate-to-heavy smokers who smoked ≥20 cigarettes per day were men (61%), which could partially explain the findings. However, the stronger associations among men remained after performing the analyses among current smokers adjusting for and stratified by smoking intensity. Notably, we found positive associations between WBC counts and lung cancer risk among never-smoking women but for not never-smoking men, which could be due to the smaller sample size for never-smoking men. Although WBC counts could contribute to lung cancer development, lung carcinogenesis among never-smoking women is strongly linked to underlying driver mutations and pathways (eg, Epidermal growth factor receptor mutations [32], echinoderm microtubule associated protein-like 4 - anaplastic lymphoma kinase gene fusions [33], proto-oncogene tyrosine-protein kinase rearrangements [34]).

Our findings are concordant with previous epidemiologic studies of WBC count and lung cancer development (9–12,35) (Supplementary Table 7 available online). An investigation in the Women's Health Initiative (WHI) found positive associations between WBC count and lung cancer among postmenopausal women (12). However, the WHI study did not detect an association among never-smoking women (estimates were not reported) (12). Our study found a noteworthy association among never-smoking women that was marginally stronger when further restricting to those without ETS exposure. Several factors could have contributed to the discrepancies including statistical power and differences in analyses and the source populations. Notably, the age-range of the WHI participants (50–79 years) was higher than for the UK Biobank (37–73 years). Additionally, the UK Biobank was a population-based cohort study, whereas the WHI combined randomized trials with an observational study, which could have introduced heterogeneity.

Similar to our investigation, an analysis of three male cohort studies from the USA and UK found increased lung cancer incidence and mortality with elevated WBC (9). Additionally, a cohort study of 4831 men and women from Wisconsin found that those with elevated WBC counts over 6.4×10^9 cells/L had nearly triple the risk compared to those under this threshold (10). Further, the Blue Mountains Eye Study and the second National Health and Nutrition Examination Survey (NHANES II) found non-statistically significantly positive associations between WBC count and lung cancer mortality (11,35).

The interrelationship between inflammation and lung cancer development is widely accepted (36); however, underlying mechanisms remain unclear. Chronic inflammation of lung tissue caused by exposure to hazardous particulates could promote apoptosis and necrosis, which induces the release of reactive oxygen species. These genotoxic chemicals can cause deleterious changes to tumor suppressors or proto-oncogenes with prolonged exposure. Indeed, various circulating immunological and inflammatory markers were statistically significantly associated with lung cancer risk (6,7).

The role of lymphocytes is well-documented with respect to lung cancer prognosis, but their relationship with future risk is unclear. We did not detect associations between lymphocyte count and future risk of developing lung cancer. Rather, we found that the WBC-lung cancer association was driven by neutrophils. Taken together, these findings suggest that early changes to innate response could be more pertinent to future disease risk and etiology among cancer-free individuals, whereas cell-mediated immune response may be more reflective of cancer progression. Indeed, neutrophils are considered inflammatory "first responders" that have been shown in animal models to mediate clearance of early tumor cells (37) but can promote drive tumor progression (38). In the few studies that found positive associations between NLR and future risk (29,31), the exposure-response could have been related to increased neutrophils more so than changes in lymphocyte counts.

This study has several strengths. First, the large sample size provided adequate statistical power to detect modest effects overall and stratified by smoking status and sex. Second, we had data on histological subtypes of lung cancer, providing greater insight into disease etiology. Third, lung cancer diagnoses were obtained from comprehensive government cancer registries linked to multiple sources. Fourth, we could establish temporality between WBC counts and lung cancer development with the prospective cohort study design. Fifth, we had extensive information on smoking, which allowed tight control of residual confounding from tobacco use.

This study has limitations. We only had sufficient WBC data from a single baseline blood draw and therefore could not assess the influence of trajectories in WBC counts on lung cancer risk. We assumed that WBC count at baseline reflected levels in the etiologic window for lung carcinogenesis. WBC count varies in the short-term in response to infection and environmental exposures. Further, longitudinal studies have found U-shaped patterns in WBC counts over the life course (39). However, resting-state WBC counts at single timepoints have been shown in many studies to be associated with future risk of numerous chronic diseases and mortality (40–45).

In summary, we found that increased total WBC count was associated with increased lung cancer risk. This relationship was primarily driven by neutrophil fractions, which are crucial components of innate immune response. The associations among current-, former-, and never-smoking women were comparable with the overall analyses, whereas stronger associations were found among former- and current-smoking men. Notably, this was one of the few studies to report associations among never-smoking women. Our study provides further evidence that immunological and inflammatory processes contribute to lung carcinogenesis, particularly among never-smoking women and ever-smoking men.

Funding

This work was supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute. This research was conducted using resources from the UK Biobank (Application 28072). The UK Biobank was established by the Wellcome Trust, the Medical Research Council, the UK Department of Health, and the Scottish Government. The UK Biobank has also received funding from the Welsh Assembly Government, the British Heart Foundation, and Diabetes United Kingdom.

Notes

Conflict of interest statement: The authors declare no conflict of interest.

Acknowledgments: We thank Information Management Services for technical and programming support. SAS version 9.4 (SAS Institute Inc) and the computational resources of the National Institutes of Health's High-Performance Computing Biowulf cluster were used to conduct the analyses.

Authors' contributions: Study concept and design: JYYW, NR, QL. Acquisition, analysis, or interpretation of data: JYYW, BAB, EL, WH, NF, B-TJ, PE, DTS, SJC, NR, QL. Manuscript composition: JYYW, NR, QL, BAB. Critical revision of the manuscript for important intellectual content: All authors. Statistical analysis: JYYW, EL, BAB. Administrative, technical, or material support: EL, NF, BAB, WH, B-TJ. Study supervision: QL, NR, SJC, DTS.

Ethics approval and consent to participate: The UK Biobank study was approved by the National Information Governance Board for Health and Social Care and the NHS North West Multicenter Research Ethics Committee. Electronic informed consent was obtained from all volunteer participants. The study was performed in accordance with the Declaration of Helsinki.

Consent for publication: Not applicable.

Data availability: Data are publicly available at: https://www.ukbiobank.ac.uk/register-apply/.

References

1. LuCE Report on Lung Cancer: Challenges in Lung Cancer in Europe. https://www.lungcancereurope.eu/wp-content/uploads/2017/10/LuCE-Report-final.pdf; 2017.
2. SEER. National Cancer Institute, Stat Facts: Lung and Bronchus Cancer; 2018. https://seer.cancer.gov/statfacts/html/lungb.html.
3. Luengo-Fernandez R, Leal J, Gray A, et al. Economic burden of cancer across the European Union: a population-based cost analysis. *Lancet Oncol*. 2013;14(12):1165–1174.
4. Mariotto AB, Yabroff KR, Shao Y, et al. Projections of the cost of cancer care in the United States: 2010-2020. *J Natl Cancer Inst*. 2011;103(2):117–128.
5. O'Callaghan DS, O'Donnell D, O'Connell F, et al. The role of inflammation in the pathogenesis of non-small cell lung cancer. *J Thorac Oncol*. 2010;5(12):2024–2036.
6. Shiels MS, Shu XO, Chaturvedi AK, et al. A prospective study of immune and inflammation markers and risk of lung cancer among female never smokers in Shanghai. *Carcinogenesis*. 2017;38(10):1004–1010.
7. Shiels MS, Pfeiffer RM, Hildesheim A, et al. Circulating inflammation markers and prospective risk for lung cancer. *J Natl Cancer Inst*. 2013;105(24):1871–1880.
8. Pine SR, Mechanic LE, Enewold L, et al. Increased levels of circulating interleukin 6, interleukin 8, C-reactive protein, and risk of lung cancer. *J Natl Cancer Inst*. 2011;103(14):1112–1122.
9. Phillips AN, Neaton JD, Cook DG, et al. The leukocyte count and risk of lung cancer. *Cancer*. 1992;69(3):680–684.
10. Sprague BL, Trentham-Dietz A, Klein BE, et al. Physical activity, white blood cell count, and lung cancer risk in a prospective cohort study. *Cancer Epidemiol Biomarkers Prev*. 2008;17(10):2714–2722.
11. Shankar A, Wang JJ, Rochtchina E, et al. Association between circulating white blood cell count and cancer mortality: a population-based cohort study. *Arch Intern Med*. 2006;166(2):188–194.
12. Margolis KL, Rodabough RJ, Thomson CA, et al. Prospective study of leukocyte count as a predictor of incident breast, colorectal, endometrial, and lung cancer and mortality in postmenopausal women. *Arch Intern Med*. 2007;167(17):1837–1844.
13. Friedman GD, Fireman BH. The leukocyte count and cancer mortality. *Am J Epidemiol*. 1991;133(4):376–380.
14. Thun MJ, Henley SJ, Burns D, et al. Lung cancer death rates in lifelong nonsmokers. *J Natl Cancer Inst*. 2006;98(10):691–699.
15. Thun MJ, Hannan LM, Adams-Campbell LL, et al. Lung cancer occurrence in never-smokers: an analysis of 13 cohorts and 22 cancer registry studies. *PLoS Med*. 2008;5(9):e185.
16. Wakelee HA, Chang ET, Gomez SL, et al. Lung cancer incidence in never smokers. *J Clin Oncol*. 2007;25(5):472–478.
17. Rivera GA, Wakelee H. Lung cancer in never smokers. *Adv Exp Med Biol*. 2016;893:43–57.
18. Sudlow C, Gallacher J, Allen N, et al. UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015;12(3):e1001779.
19. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of sociodemographic and health-related characteristics of UK biobank participants with those of the general population. *Am J Epidemiol*. 2017;186(9):1026–1034.
20. Loftfield E, Cornelis MC, Caporaso N, et al. Association of coffee drinking with mortality by genetic variation in caffeine metabolism: findings from the UK biobank. *JAMA Intern Med*. 2018;178(8):1086–1097.
21. Loftfield E, Zhou W, Graubard BI, et al. Predictors of mosaic chromosome Y loss and associations with mortality in the UK Biobank. *Sci Rep*. 2018;8(1):12316.
22. Bain BJ. Ethnic and sex differences in the total and differential white cell count and platelet count. *J Clin Pathol*. 1996;49(8):664–666.
23. Lim E-M, Cembrowski G, Cembrowski M, et al. Race-specific WBC and neutrophil count reference intervals. *Int J Lab Hematol*. 2010;32(6p2):590–597.
24. Higuchi T, Omata F, Tsuchihashi K, et al. Current cigarette smoking is a reversible cause of elevated white blood cell count: cross-sectional and longitudinal studies. *Prev Med Rep*. 2016;4:417–422.
25. Abel GA, Hays JT, Decker PA, et al. Effects of biochemically confirmed smoking cessation on white blood cell count. *Mayo Clin Proc*. 2005;80(8):1022–1028.
26. Sunyer J, Mu Oz A, Peng Y, et al. Longitudinal relation between smoking and white blood cells. *Am J Epidemiol*. 1996;144(8):734–741.
27. Jensen EJ, Pedersen B, Frederiksen R, et al. Prospective study on the effect of smoking and nicotine substitution on leucocyte blood counts and relation between blood leucocytes and lung function. *Thorax*. 1998;53(9):784–789.
28. Aminzadeh Z, Parsa E. Relationship between age and peripheral white blood cell count in patients with sepsis. *Int J Prev Med*. 2011;2(4):238–242.
29. Kang J, Chang Y, Ahn J, et al. Neutrophil-to-lymphocyte ratio and risk of lung cancer mortality in a low-risk population: a cohort study. *Int J Cancer*. 2019;145(12):3267–3275.
30. Sanchez-Salcedo P, de-Torres JP, Martinez-Urbistondo D, et al. The neutrophil to lymphocyte and platelet to lymphocyte ratios as biomarkers for lung cancer development. *Lung Cancer*. 2016;97:28–34.
31. Grieshober L, Graw S, Barnett MJ, et al. Methylation-derived neutrophil-to-lymphocyte ratio and lung cancer risk in heavy smokers. *Cancer Prev Res*. 2018;11(11):727–734.
32. Seow WJ, Matsuo K, Hsiung CA, et al. Association between GWAS-identified lung adenocarcinoma susceptibility loci and EGFR mutations in never-smoking Asian women, and comparison with findings from Western populations. *Hum Mol Genet*. 2017;26(2):454–465.
33. Martelli MP, Sozzi G, Hernandez L, et al. EML4-ALK rearrangement in non-small cell lung cancer and non-tumor lung tissues. *Am J Pathol*. 2009;174(2):661–670.
34. Sehgal K, Patell R, Rangachari D, et al. Targeting ROS1 rearrangements in non-small cell lung cancer with crizotinib and other kinase inhibitors. *Transl Cancer Res*. 2018;7(Suppl 7): S779–S786.
35. Erlinger TP, Muntner P, Helzlsouer KJ. WBC count and the risk of cancer mortality in a national sample of U.S. adults: results from the Second National Health and Nutrition Examination Survey mortality study. *Cancer Epidemiol Biomarkers Prev*. 2004;13(6):1052–1056.
36. Coussens LM, Werb Z. Inflammation and cancer. *Nature*. 2002;420(6917):860–867.
37. Xue W, Zender L, Miething C, et al. Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. *Nature*. 2007;445(7128):656–660.
38. Singel KL, Segal BH. Neutrophils in the tumor microenvironment: trying to heal the wound that cannot heal. *Immunol Rev*. 2016;273(1):329–343.
39. Chmielewski PP, Borysławski K, Chmielowiec K, et al. The association between total leukocyte count and longevity: evidence from longitudinal and cross-sectional data. *Ann Anat*. 2016;204:1–10.
40. Chao TT, Hsieh CH, Lin JD, et al. Use of white blood cell counts to predict metabolic syndrome in the elderly: a 4 year longitudinal study. *Aging Male*. 2014;17(4):230–237.
41. Pei C, Chang JB, Hsieh CH, et al. Using white blood cell counts to predict metabolic syndrome in the elderly: a combined cross-sectional and longitudinal study. *Eur J Intern Med*. 2015;26(5):324–329.
42. Huh JY, Ross GW, Chen R, et al. Total and differential white blood cell counts in late life predict 8-year incident stroke: The Honolulu Heart Program. *J Am Geriatr Soc*. 2015;63(3):439–446.
43. Nilsson G, Hedberg P, Ohrvik J. White blood cell count in elderly is clinically useful in predicting long-term survival. *J Aging Res*. 2014;2014:1–6.
44. Brown DW, Ford ES, Giles WH, et al. Associations between white blood cell count and risk for cerebrovascular disease mortality: NHANES II Mortality Study, 1976-1992. *Ann Epidemiol*. 2004;14(6):425–430.
45. Ahmadi-Abhari S, Luben RN, Wareham NJ, et al. Seventeen year risk of all-cause and cause-specific mortality associated with C-reactive protein, fibrinogen and leukocyte count in men and women: the EPIC-Norfolk study. *Eur J Epidemiol*. 2013;28(7):541–550.