

# Pre-diagnosis neutrophil-to-lymphocyte ratio and mortality in individuals who develop lung cancer

Laurie Grieshober (✉ [laurie.grieshober@hci.utah.edu](mailto:laurie.grieshober@hci.utah.edu))

Huntsman Cancer Institute, University of Utah <https://orcid.org/0000-0003-2316-9143>

**Stefan Graw**

Emory University School of Public Health

**Matt J Barnett**

Fred Hutchinson Cancer Research Center

**Gary E Goodman**

Fred Hutchinson Cancer Research Center

**Chu Chen**

Fred Hutchinson Cancer Research Center, University of Washington

**Devin C Koestler**

University of Kansas Medical Center

**Carmen J Marsit**

Emory University School of Public Health

**Jennifer A Doherty**

Huntsman Cancer Institute, University of Utah

---

## Original Article

**Keywords:** Lung cancer, Small cell lung cancer, NLR, Methylation, Mortality, CARET

**DOI:** <https://doi.org/10.21203/rs.3.rs-198826/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

## Abstract

**Purpose** The neutrophil-to-lymphocyte ratio (NLR) is a marker of systemic inflammation that has been reported to be associated with survival after chronic disease diagnoses, including lung cancer. We hypothesized that the inflammatory profile reflected by pre-diagnosis NLR, rather than the well-studied pre-treatment NLR at diagnosis, may be associated with increased mortality after lung cancer is diagnosed in high risk heavy smokers.

**Methods** We examined associations between pre-diagnosis methylation derived-NLR (mdNLR) and lung cancer-specific and all-cause mortality in 372 lung cancer cases from the  $\beta$ -Carotene and Retinol Efficacy Trial (CARET). Cox proportional hazards models were adjusted for age, sex, smoking status, pack years, and time between blood draw and diagnosis, and stratified by stage of disease. Models were run separately by histotype.

**Results** Among those who developed small cell lung cancer (SCLC, N = 81), individuals with pre-diagnosis mdNLR in the highest quartile had 2.5-fold increased mortality compared to those in the lowest quartile. For each unit increase in pre-diagnosis mdNLR, we observed 22–23% increased mortality in SCLC (SCLC-specific hazard ratio [HR] = 1.23, 95% confidence interval [CI]: 1.02, 1.48; all-cause HR = 1.22, 95% CI: 1.01, 1.46). SCLC associations were strongest for current smokers at blood draw (Interaction  $P$ s = 0.03). mdNLR was not associated with mortality for individuals diagnosed with adenocarcinoma (N = 148) or squamous cell carcinoma (N = 115).

**Conclusion** Our findings suggest that increased mdNLR, representing a systemic inflammatory profile on average 4.5 years before a SCLC diagnosis, may be associated with mortality in heavy smokers who go on to develop SCLC but not other lung cancer histotypes.

## Introduction

Lung cancer is the leading cause of cancer death worldwide [1], with five-year relative survival of 24% for non-small cell lung cancer (NSCLC), which primarily includes adenocarcinoma and squamous cell histotypes, and 6% for small cell lung cancer (SCLC) [2]. More than 65% of NSCLC and 90% of SCLC patients are diagnosed at an advanced stage [3], with five-year relative survival rates of just 6% for NSCLC and 3% for SCLC among those diagnosed at an advanced stage [2]. Patient factors before diagnosis that are associated with poorer lung cancer survival include older age, male sex, weight loss, and cigarette smoke exposure [4, 5].

It is well-established that inflammatory processes are associated with risk of lung cancer [6–9], and it is plausible that an individual's systemic inflammatory profile prior to a lung cancer diagnosis may be associated with mortality. In support of this hypothesis, inflammatory conditions such as chronic bronchitis and emphysema, chronic obstructive pulmonary disease (COPD), interstitial lung disease, and diabetes are associated with increased lung cancer mortality independent of their associations with increased lung cancer risk [10–19].

Elevated neutrophil-to-lymphocyte ratio (NLR) is a marker of systemic inflammation and immune stress that has been reported to be associated with all-cause mortality in most large prospective studies of healthy individuals [20–25], as well as cancer-specific mortality in one [20] of the three aforementioned studies in which cancer-specific mortality was also examined [20, 22, 25]. A single prospective study of healthy individuals observed that higher NLR was associated with future mortality from lung cancer [26]. In cancer patients with solid tumors, higher NLR (typically assessed prior to treatment) is an independent predictor of poor prognosis regardless of treatment strategy [27–30], and recent meta-analyses confirm that these associations hold in NSCLC and SCLC [31–37].

NLR can be easily quantified using results from simple hematology testing, specifically the complete blood count (CBC) with leukocyte differentials [38]. Though traditional CBC measurement cannot be performed on archival blood samples, lineage-specific DNA methylation patterns across the genome can be leveraged to estimate blood cell proportions that can

be used to calculate methylation-derived NLR (mdNLR) [39, 40]. While NLR measured at lung cancer diagnosis most likely reflects the disease state and possibly progression [41, 42], NLR measured years prior to diagnosis provides a snapshot of the systemic inflammatory profile, which in addition to a person's health state, developed immune system, and underlying genetics, may include evidence of exposure to environmental and behavioral risk factors [43, 44]. In the present study, we examined whether pre-diagnosis mdNLR was associated with mortality in heavy smokers who later developed lung cancer, and differences by lung cancer histotype.

## Methods

Our study included 372 individuals diagnosed with lung cancer between 1994 and 2013 from the multicenter  $\beta$ -Carotene and Retinol Efficacy Trial (CARET) of heavy smokers at high risk for lung cancer [45], which we have described previously [46].

We assayed DNA methylation in the archival whole blood samples using the Illumina HumanMethylationEPIC BeadArray, followed by standard normalization and preprocessing procedures [46, 47]. We estimated proportions of six blood cell types (B cell, CD4T, CD8T, natural killer (NK), neutrophil, monocyte) for each case in our normalized methylation dataset using constrained projection of the EPIC IDOL-optimized cell mixture deconvolution matrix with the "projectCellType\_CP" function from the FlowSorted.Blood.EPIC package in R [40]. Continuous mdNLR was calculated as the ratio of predicted neutrophil and lymphocyte (sum of B cell, CD4T, CD8T, and NK proportions) proportions, and we discretized mdNLR into quartiles based on the distribution from all 372 cases (Q1 0.39–1.42, Q2 1.43–1.89, Q3 1.90–2.46, Q4 2.47–16.90), representing increasing levels of systemic inflammation.

We evaluated associations between pre-diagnosis mdNLR and lung cancer-specific and all-cause mortality using multivariable-adjusted Cox proportional hazards models fit to all cases, and separately for adenocarcinoma, squamous cell carcinoma, and SCLC histotypes. We defined time to event as years from lung cancer diagnosis to death or December 31, 2013, whichever occurred first. Stage data were not available for cases ascertained between 2005 and 2013 due to passive follow-up procedures implemented after 2005 and for those for which medical records could not be otherwise obtained. Therefore, our models included a strata variable to allow for differing baseline hazards by early (stage I/II), late (stage III/IV), or unknown stage. Models were *a priori* adjusted for variables assessed at the time the blood samples for methylation assays were drawn, based on biologic plausibility, including age, sex, smoking status, pack years, and time between blood draw and diagnosis. We assessed study covariates, such as body mass index (BMI), enrollment year, intervention arm, occupational asbestos exposure, race, years since quit smoking, and cigarettes smoked per day for potential confounding of mortality models based on a  $\geq 10\%$  change in continuous mdNLR hazard ratio estimates for all cases combined or by histotype in the *a priori* adjusted models. No additional covariates were included in our final models based on this threshold. In SCLC models, we explored effect modification by age (dichotomized at the mean in SCLC, 64.1 years), intervention arm, sex, smoking history (dichotomized at the mean in SCLC, 59.3 pack years), smoking status, and time between blood draw and diagnosis (dichotomized at the mean in SCLC, 4.5 years) by performing stratified analyses of the final, adjusted models. We evaluated statistical interaction between the dichotomous stratification variables and continuous mdNLR using product term *P*-values. Interaction models for intervention arm were also adjusted for the respective first level variable when assessing interactions since those variables were not included in the final, adjusted models. We performed a sensitivity analysis of our main models, overall and by histotype, excluding individuals diagnosed within two years of blood draw. Analytical modeling was performed in SAS 9.4 (Cary, NC). Statistical significance was defined as  $P < 0.05$  in two-sided tests.

## Results

Participant characteristics at blood draw are summarized overall and by histotype in Table 1. Whole blood was collected on average 4.9 (range 0.1 to 19.3) years prior to diagnosis for the 372 cases. Lung cancer cases were on average 64 years

old at blood draw, mostly white, and had mean smoking histories ranging from 57 to 62 pack years. The histotype distribution for the 372 cases was: adenocarcinoma (N = 148), squamous cell carcinoma (N = 115), SCLC (N = 81), NSCLC not otherwise specified (NOS; N = 16), and unknown (N = 12). Approximately 40% of adenocarcinoma and SCLC cases were female compared to 23% of squamous cell cases. Average time from blood draw to diagnosis was about 4.5 years in each histotype, and slightly longer in all cases combined due to the additional cases with missing histotype data, diagnosed after 2005 (mean 4.9 years). More than half of the cases, and 73% of the SCLC cases, were diagnosed at late stage (III/IV), though stage was missing for 13–23% of cases. Median time from diagnosis to death was shortest for SCLC (8.4 months), and longest for squamous cell carcinoma (12 months).

Table 1  
 Characteristics of lung cancer cases overall and by histotype

	All lung cancer cases <sup>a</sup>	Adenocarcinoma	Squamous cell	Small cell
	(N = 372)	(N = 148)	(N = 115)	(N = 81)
<b>Age at blood draw</b> , years; mean (SD)	64.3 (5.6)	64.2 (5.6)	64.5 (5.6)	64.1 (5.9)
45 to < 55; N (%)	23 (6)	10 (7)	8 (7)	5 (6)
55 to < 60; N (%)	57 (15)	22 (15)	18 (16)	14 (17)
60 to < 65; N (%)	120 (32)	52 (35)	31 (27)	24 (30)
65 to < 70; N (%)	109 (29)	37 (25)	38 (33)	25 (31)
≥ 70; N (%)	63 (17)	27 (18)	20 (17)	13 (16)
<b>Age at diagnosis</b> , years; mean (SD)	69.2 (6.2)	68.6 (5.8)	69.1 (5.9)	68.6 (6.0)
<b>BMI<sup>c</sup></b> ; mean (SD)	27.5 (4.8)	27.5 (4.8)	27.1 (4.9)	28.0 (5.1)
Normal (≥ 18.5 and < 25); N (%)	110 (30)	46 (31)	37 (32)	21 (26)
Overweight (≥ 25 and < 30); N (%)	161 (43)	58 (39)	47 (41)	39 (48)
Obese (≥ 30); N (%)	97 (26)	43 (29)	29 (25)	20 (25)
<b>Enrollment year</b> , N (%)				
1985–1986	21 (6)	12 (8)	4 (3)	3 (4)
1987–1988	20 (5)	6 (4)	5 (4)	5 (6)
1989–1990	90 (24)	38 (26)	23 (20)	19 (23)
1991–1992	170 (46)	60 (41)	59 (51)	40 (49)
1993–1994	71 (19)	32 (22)	24 (21)	14 (17)
<b>Race White</b> ; N (%)	356 (96)	143 (97)	107 (93)	78 (96)
<b>Sex</b> , female; N (%)	132 (35)	63 (43)	26 (23)	33 (41)
<b>Current smoker at blood draw</b> ; N (%)	245 (66)	89 (60)	87 (76)	51 (63)
<b>Pack-years at blood draw</b> ; mean (SD)	58.8 (21.8)	57.3 (20.0)	62.0 (24.5)	59.3 (21.7)

	All lung cancer cases <sup>a</sup>	Adenocarcinoma	Squamous cell	Small cell
<b>Years since quit smoking at blood draw<sup>b</sup>; mean (SD)</b>	6.4 (4.9)	6.2 (5.2)	7.3 (4.3)	6.0 (4.8)
<b>Active intervention arm; N (%)</b>	193 (52)	77 (52)	59 (51)	43 (53)
<b>Asbestos exposure; N (%)</b>	62 (17)	24 (16)	23 (20)	11 (14)
<b>Stage; N (%)</b>				
Early stage (I/II)	78 (21)	37 (25)	38 (33)	3 (4)
Late stage (III/IV)	215 (58)	85 (57)	62 (54)	59 (73)
Unknown	79 (21)	26 (18)	15 (13)	19 (23)
<b>Months from diagnosis to death<sup>d</sup> or end of follow-up; median (IQR)</b>	9.6 (26.4)	10.8 (45.6)	12.0 (39.6)	8.4 (9.6)
<b>Years between blood draw and diagnosis; mean (SD)</b>	4.9 (3.3)	4.5 (2.9)	4.6 (2.4)	4.5 (2.7)
<b>mdNLR; mean (SD)</b>	2.18 (1.46)	2.11 (1.22)	2.38 (1.91)	2.08 (1.32)
Abbreviations: BMI = Body Mass Index; SD = Standard Deviation.				
<sup>a</sup> All lung cancer cases" includes adenocarcinoma, squamous cell, and small cell, as well as 16 cases for whom histotype was NSCLC, NOS and 12 cases with unknown histotype.				
<sup>b</sup> Former smokers only (n = 127).				
<sup>c</sup> BMI is missing for two participants (N = 1 Squamous cell and N = 1 Small cell) and two participants were underweight (N = 1 Adeno, N = 1 Squamous cell). BMI cut-points per National Heart, Lung, and Blood Institute definition: underweight (< 18.5), normal (≥ 18.5 and < 25), overweight (≥ 25 and < 30), and obese (≥ 30).				
<sup>d</sup> Death events collected through end of follow-up in 2015.				

Among SCLC cases, we observed a statistically significant 23% increased lung cancer-specific mortality (hazard ratio [HR] = 1.23, 95% confidence interval [CI]: 1.02, 1.48) and 22% increased all-cause mortality (HR = 1.22, 95% CI: 1.01, 1.46) for each unit increase in pre-diagnosis mdNLR (Table 2). We observed similar results for quartiled mdNLR in SCLC, with Q4 vs Q1 mdNLR HRs of 2.49 (95% CI: 1.15, 5.40) for SCLC-specific mortality and 2.44 (95% CI: 1.13, 5.26) for all-cause mortality. We observed a linear trend across increasing mdNLR quartiles that was statistically significant for increased SCLC-specific and all-cause mortality in SCLC (*P*-trends = 0.04). Mortality was not associated with continuous or quartiled mdNLR in all cases combined, adenocarcinoma cases, or squamous cell carcinoma cases. Our sensitivity analysis restricting to individuals diagnosed two or more years after blood draw produced similar results overall and by histotype (**Supplementary Table 1**). SCLC mortality estimates for Q4 vs Q1 were strengthened (SCLC-specific HR = 3.54, CI: 1.37, 9.14; all-cause HR = 3.37, CI: 1.33, 8.57), with similar estimates of linear trend (*P*-trends = 0.03); however, the continuous unit-change models lost statistical significance after excluding the 23% of SCLC cases diagnosed within two years of blood draw (SCLC-specific HR = 1.21, CI: 0.96, 1.53; all-cause HR = 1.19, CI: 0.95, 1.50).

Table 2  
mdNLR and mortality<sup>a</sup> for all lung cancer cases and by histotype.

mdNLR	All lung cancer cases <sup>b</sup>			Adenocarcinoma			Squamous cell carcinoma			Small cell		
	Death	Case	HR (95% CI)	Death	Case	HR (95% CI)	Death	Case	HR (95% CI)	Death	Case	HR (95% CI)
	N	N	N	N	N	N	N	N	N	N	N	N
<b>Lung cancer-specific mortality</b>												
<b>Continuous</b>	313	372	0.98 (0.90, 1.06)	117	148	1.02 (0.86, 1.20)	94	115	0.92 (0.81, 1.04)	77	81	1.23 (1.02, 1.48)
<b>Q1</b> (lowest inflammation)	77	93	Ref	32	44	Ref	17	20	Ref	24	24	Ref
<b>Q2</b>	75	93	0.96 (0.69, 1.33)	24	32	0.85 (0.48, 1.52)	26	34	0.71 (0.36, 1.41)	20	21	1.60 (0.84, 3.02)
<b>Q3</b>	80	93	1.05 (0.76, 1.45)	28	32	1.06 (0.62, 1.81)	24	29	0.85 (0.42, 1.72)	18	21	1.41 (0.70, 2.82)
<b>Q4</b> (highest inflammation)	81	93	1.12 (0.80, 1.55)	33	40	1.20 (0.71, 2.04)	27	32	0.71 (0.37, 1.34)	15	15	2.49 (1.15, 5.40)
	<b>P-trend</b>		0.44	<b>P-trend</b>		0.36	<b>P-trend</b>		0.45	<b>P-trend</b>		0.04
<b>All-cause mortality</b>												
<b>Continuous</b>	357	372	0.97 (0.90, 1.04)	137	148	0.97 (0.84, 1.12)	113	115	0.92 (0.83, 1.03)	80	81	1.22 (1.01, 1.46)
<b>Q1</b> (lowest inflammation)	88	93	Ref	39	44	Ref	20	20	Ref	24	24	Ref
<b>Q2</b>	89	93	0.97 (0.72, 1.31)	29	32	0.80 (0.48, 1.36)	34	34	0.77 (0.41, 1.44)	21	21	1.67 (0.89, 3.15)
<b>Q3</b>	90	93	1.05 (0.78, 1.43)	31	32	1.03 (0.63, 1.70)	28	29	0.87 (0.45, 1.66)	20	21	1.46 (0.74, 2.89)
<b>Q4</b> (highest inflammation)	90	93	1.06 (0.78, 1.44)	38	40	1.10 (0.67, 1.81)	31	32	0.70 (0.38, 1.26)	15	15	2.44 (1.13, 5.26)
	<b>P-trend</b>		0.62	<b>P-trend</b>		0.49	<b>P-trend</b>		0.33	<b>P-trend</b>		0.04
Abbreviations: HR = Hazard Ratio, mdNLR = methylation-derived neutrophil-to-lymphocyte ratio												
<sup>a</sup> Mortality is estimated using Cox proportional hazards models adjusted for age, sex, smoking status, pack years at blood draw, and time between blood draw and diagnosis. Stage (early (I/II), late (III/IV), unknown) is included as a strata variable.												
<sup>b</sup> "All lung cancer cases" includes adenocarcinoma, squamous cell, and small cell, as well as 16 cases for whom histotype was NSCLC, NOS and 12 cases with unknown histotype.												

Stratified model results for continuous mdNLR and SCLC mortality are presented in Table 3. We observed stronger associations between mdNLR and SCLC-specific mortality in current smokers (HR = 2.00, 95% CI: 1.32, 3.03) versus former smokers (HR = 1.14, 95% CI: 0.83, 1.56), with interaction  $P = 0.03$ . We also observed stronger associations among those assigned to the placebo arm (HR = 1.86, 95% CI: 1.29, 2.69) versus the active intervention (HR = 1.17, 95% CI: 0.85, 1.60), and males (HR = 1.46, 95% CI: 1.07, 1.98) versus females (HR = 1.09, 95% CI: 0.85, 1.41), though these stratified results did not show evidence of statistical interaction (interaction  $P$ s  $\geq 0.44$ ). HRs were similar in magnitude for strata defined by mean age at diagnosis, mean pack years, and mean time between blood draw and diagnosis. SCLC stratified all-cause mortality results were similar to those for SCLC-specific mortality.



Table 3  
mdNLR and mortality in small cell lung cancer cases by subgroup.

Variable	Strata definition	Lung cancer-specific mortality			All-cause mortality		
		Death	Case	HR (95% CI)	Death	Case	HR (95% CI)
		N	N		N	N	
Age at blood draw <sup>b</sup>	< 64.1 years	38	39	1.21 (0.96, 1.52)	38	39	1.21 (0.96, 1.52)
	≥ 64.1 years	39	42	1.22 (0.77, 1.95)	42	42	1.21 (0.77, 1.91)
	Interaction <i>P</i> <sup>c</sup>	0.88			0.98		
Intervention arm	Active	41	43	1.17 (0.85, 1.60)	42	43	1.17 (0.85, 1.60)
	Placebo	36	38	<b>1.86 (1.29, 2.69)</b>	38	38	<b>1.74 (1.24, 2.42)</b>
	Interaction <i>P</i> <sup>c</sup>	0.64			0.71		
Sex	Female	31	33	1.09 (0.85, 1.41)	33	33	1.10 (0.86, 1.42)
	Male	46	48	<b>1.46 (1.07, 1.98)</b>	47	48	<b>1.49 (1.10, 2.02)</b>
	Interaction <i>P</i> <sup>c</sup>	0.55			0.44		
Smoking history at blood draw <sup>b</sup>	< 59.3 pack years	41	44	1.12 (0.86, 1.46)	44	44	1.14 (0.88, 1.48)
	≥ 59.3 pack years	36	37	1.28 (0.97, 1.70)	36	37	1.28 (0.97, 1.70)
	Interaction <i>P</i> <sup>c</sup>	0.60			0.45		
Smoking status at blood draw	Former	28	30	1.14 (0.83, 1.56)	29	30	1.12 (0.83, 1.51)
	Current	49	51	<b>2.00 (1.32, 3.03)</b>	51	51	<b>1.96 (1.30, 2.96)</b>
	Interaction <i>P</i> <sup>c</sup>	<b>0.03</b>			<b>0.03</b>		
Time between blood draw and diagnosis <sup>b</sup>	< 4.5 years	39	41	1.35 (0.95, 1.91)	40	41	1.34 (0.95, 1.89)

	Lung cancer-specific mortality			All-cause mortality		
<b>≥ 4.5 years</b>	38	40	1.20 (0.93, 1.54)	40	40	1.16 (0.91, 1.49)
Interaction <i>P</i> <sup>c</sup>	0.76			0.68		
Abbreviations: HR = Hazard Ratio, mdNLR = methylation-derived neutrophil-to-lymphocyte ratio						
<sup>a</sup> Mortality is estimated using Cox proportional hazards models adjusted for age, sex, smoking status, pack years at blood draw, and time between blood draw and diagnosis. Stage (early (I/II), late (III/IV), unknown) is included as a strata variable.						
<sup>b</sup> Mean values among SCLC cases used to define strata.						
<sup>c</sup> Interaction <i>P</i> -value for the product term between continuous mdNLR and the dichotomized covariate. For intervention arm, both the main-effect and interaction terms were added to calculate the interaction <i>P</i> since intervention arm was not included in the fully-adjusted models.						

## Discussion

To our knowledge, our study is the first to assess whether NLR estimated years before diagnosis is associated with mortality among individuals who go on to develop lung cancer. In this study of heavy smokers from CARET, we observed that pre-diagnosis mdNLR was associated with increased mortality for SCLC cases, but not for adenocarcinoma cases or squamous cell carcinoma cases.

Approximately 15% of lung cancer diagnoses are SCLC [2]. SCLC is the most aggressive lung cancer histotype with distinctive tumor behavior characterized by rapid growth, early and widespread metastases, genomic instability, and acquired chemoresistance [48]. Median survival in SCLC patients is just seven months [49]; we observed a median survival of 8.4 months in our 81 SCLC patient subset from the CARET study. SCLC is not amenable to early detection by screening due to its short preclinical phase [50], so smoking cessation and improved treatments are the main targets for reducing mortality from this highly lethal and primarily smoking-related cancer [48, 50]. There are currently over 200 ongoing and recruiting clinical trials for SCLC [51], yet biomarkers for targeted therapy selection and immunotherapy in SCLC remain scarce [52].

NLR is an index of systemic inflammation that estimates the balance between the innate and adaptive immune systems [27]. Immune homeostasis is a complex and dynamic process that includes maintaining relatively constant component leukocyte proportions within physiologic ranges [43, 53]. Therefore, elevated NLR may indicate immune dysregulation that is evident from abnormal CBC components, such as high neutrophil or low lymphocyte counts, or the ratio measure may indicate low-grade immune dysregulation despite within-range CBCs. When measured prior to treatment at lung cancer diagnosis, higher NLR is thought to reflect the disease state and likelihood of progression since higher neutrophil counts have been shown to promote metastasis [54–56], and lower lymphocyte counts have been observed to be associated with loss of tumor suppressor activities [57].

We previously reported that higher pre-diagnosis mdNLR was not associated with an increased risk of developing SCLC (Odds Ratio = 1.06, 95% CI: 0.77, 1.47) in 68 matched case-control pairs of high-risk heavy smokers from CARET [47]. Here, we observed that among 81 SCLC cases (including the 68 SCLC patients from the aforementioned case-control study), higher pre-diagnosis mdNLR was associated with increased mortality. Individuals in the highest quartile of mdNLR had 2.5-fold increased mortality compared to those in the lowest quartile. Higher mdNLR was most strongly associated with

increased SCLC-specific and all-cause mortality in current smokers, those assigned to the placebo arm, and males compared to their counterpart stratum. The systemic inflammatory profile indicated by higher NLR could indicate a lesser ability to mount a robust immune response to a developing cancer and/or a favorable environment for the pathogenesis of more aggressive SCLC molecular histotypes [58, 59]. Given the short preclinical period of SCLC and the lack of association between mdNLR and SCLC risk in our previous work, we hypothesize that higher NLR measured years before a clinical SCLC diagnosis may reflect a systemic low-grade inflammatory profile that enables poorer post-diagnosis survival rather than occult carcinogenesis. Our sensitivity analysis excluding 23% SCLC cases who were diagnosed within two years of blood draw supports this hypothesis since results were similar, and even stronger for comparisons of the top to the bottom mdNLR quartile (2.5-fold increased SCLC mortality among all SCLC and 3.5-fold increased SCLC mortality in individuals diagnosed more than two years after their blood draw).

In the extensive literature on NLR and mortality in lung cancer patients, pre-treatment NLR is typically measured at diagnosis or up to 30 days prior to treatment [37, 60], and it has been reported to be associated with mortality in meta-analyses of both NSCLC and SCLC [31–37]. However, since blood was drawn on average 4.9 years (median 4.7 years) prior to lung cancer diagnosis in our study, these studies are not directly comparable to ours. One other study currently available in preprint is similar to our work in that respect—a study of 205 lung cancer cases from the “Give Us a Clue to Cancer and Heart Disease” cohorts (CLUE I/II) [61], with mdNLR measured a median of 14 years prior to diagnosis. They found that each standard deviation increase in pre-diagnosis mdNLR was associated with increased lung cancer-specific mortality (HR = 1.27, 95% CI: 1.08, 1.50). This association was strongest in adenocarcinoma cases (N = 67, HR = 2.12, 95% CI: 1.41, 3.19), but no results were presented for SCLC due to limited sample size (N = 29). In contrast to the CLUE I/II study which included never smokers [61], our study only includes heavy smokers, and our participants were older and had shorter times from blood draw to diagnosis. In addition, their mdNLR mean was lower, and standard deviations smaller, than those observed in the present study for all lung cancer cases (CLUE I/II mean 1.48 and SD 0.82, CARET mdNLR mean 2.18 and SD 1.46). Meta-analyses of pre-treatment NLR and mortality in lung cancer patients report NLR cut-offs for mortality associations between 2.2 and 5.9 [31–33], with a median NLR cut-off of 3.7 identified across 20 SCLC studies [34]. In our study, blood was collected on average 4.5 years prior to diagnosis for SCLC cases, so mdNLR was more consistent with adult population-level estimates of NLR (from populations with respective mean ages 52 and 48 years) [62, 63]. We did not examine associations using the pre-treatment NLR literature-based cut-offs, as just 3.7% of the SCLC cases in our study had pre-diagnosis mdNLR > 5, and 4.9% had mdNLR > 3.7.

We estimated methylation-based blood cell type proportions from archival samples based on CpGs that were recently identified using deconvolution algorithms applied to EPIC 850K CpG array data (i.e. EPIC-optimized) by Salas et al [40], in contrast to our prior work in which we used CpGs that were identified using the 450K CpG array data [46, 47]. Cell type estimates, and therefore mdNLR, obtained from the two arrays are highly correlated in our study (mdNLR Spearman  $r = 0.99$ ,  $P = 7.0E-301$ ) and in the literature [40]. We opted to use the now available EPIC-optimized method for cell type estimation in this publication since 69% of the EPIC-optimized CpGs are unique to the EPIC array [40].

Like most NLR studies, our study is limited by a single timepoint of estimated mdNLR. Given that NLR is dynamic in the presence of acute physiologic stress such as infections and disease development, any regression dilution bias in our prospective assessment would be expected to attenuate mortality associations [64]. This may have impacted our ability to observe associations between mdNLR and mortality in adenocarcinoma and squamous cell histotypes. Though we were able to examine mortality within each histotype, histotype data was missing for 7.5% of cases and stage data were missing for 21% of cases. Our study is observational in design and included a limited number of SCLC cases, so our results must be replicated to ensure that they did not arise due to chance or due to confounding by missing or inadequately captured covariates. Since CARET was a phase III chemoprevention trial, a major strength of our study is detailed participant and outcome data. Trial eligibility required that all participants have heavy smoking histories, so all

cases had similar exposure to high-risk smoking behavior, making our study robust to confounding of the mdNLR and mortality associations by smoking.

Our results suggest that higher pre-diagnosis mdNLR, which may indicate a low-grade systemic inflammatory profile, is associated with poorer post-diagnosis survival following the most aggressive form of lung cancer, SCLC. Our study provides preliminary evidence suggesting that pre-diagnosis CBCs in heavy smokers at high risk of lung cancer could possibly be leveraged to provide patient-level prognostic information that ultimately may have applications in risk stratification as well as aiding clinical treatment choice and monitoring [41, 55].

## **Declarations**

### **Funding**

The research reported in this publication was supported by the National Center for Advancing Translational Sciences (NCATS) of the NIH under Award Number TL1 TR002540 and the National Cancer Institute (NCI) of the NIH R01 CA151989 (to J.A. Doherty), the Munck-Pfefferkorn Fund at Dartmouth College (to J.A. Doherty and C.J. Marsit), the Huntsman Cancer Foundation (to J.A. Doherty), and the Kansas IDeA Network of Biomedical Research Excellence Bioinformatics Core (to D.C. Koestler), and supported in part by the National Institute of General Medical Science (NIGMS) award P20 GM103418 (to D.E. Wright), and the NCI under award numbers P30 CA042014 (to M.E. Beckerle), P30 CA168524 (to R.A. Jensen), and R01 CA111703 (to C. Chen). Support for CARET is from NCI grants UM1 CA167462 and U01 CA63673 (to G.E. Goodman) and U01 CA167462 (to C. Chen). The funding bodies had no roles in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

### **Conflicts of interest/Competing interests**

The authors declare that they have no competing interests.

### **Ethics approval**

All procedures performed in studies involving human participants were in accordance with the ethical standards of the Institutional Review Boards for each participating CARET institution (full list by study site, including Federalwide Assurance Numbers, are included in Electronic Supplementary Information; Supplementary Table 2), overseen by the CARET Coordinating Center (Fred Hutchinson Cancer Research Center, Seattle, WA), and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

### **Consent to participate**

Written informed consent was obtained from all CARET participants.

### **Consent for publication**

Not applicable

### **Availability of data and material**

The data that support the findings of this study are available from CARET but restrictions apply to the availability of these data, which were used in agreement with CARET for the current study, and so are not publicly available. Data are available from the authors upon request and with permission of CARET (<http://www.compass.fhcrc.org/caretWeb/requests/requesting.aspx>).

## Code availability

Not applicable

## References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68: 394-424.
2. Howlader N NA, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). (1975-2017) SEER Cancer Statistics Review. Bethesda, MD: National Cancer Institute.
3. Howlader N, Forjaz G, Mooradian MJ, et al (2020) The Effect of Advances in Lung-Cancer Treatment on Population Mortality. *N. Engl. J. Med.* 383: 640-649.
4. Islami F, Goding Sauer A, Miller KD, et al (2018) Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States. *CA Cancer J. Clin.* 68: 31-54.
5. Alberg AJ, Brock MV, Ford JG, Samet JM, Spivack SD (2013) Epidemiology of lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest.* 143: e1S-e29S.
6. Shiels MS, Pfeiffer RM, Hildesheim A, et al (2013) Circulating inflammation markers and prospective risk for lung cancer. *J. Natl. Cancer Inst.* 105: 1871-1880.
7. Pine SR, Mechanic LE, Enewold L, et al (2011) Increased levels of circulating interleukin 6, interleukin 8, C-reactive protein, and risk of lung cancer. *J. Natl. Cancer Inst.* 103: 1112-1122.
8. Huang JY, Larose TL, Luu HN, et al (2020) Circulating markers of cellular immune activation in prediagnostic blood sample and lung cancer risk in the Lung Cancer Cohort Consortium (LC3). *Int. J. Cancer.* 146: 2394-2405.
9. Olvera Alvarez HA, Kubzansky LD, Campen MJ, Slavich GM (2018) Early life stress, air pollution, inflammation, and disease: An integrative review and immunologic model of social-environmental adversity and lifespan health. *Neurosci. Biobehav. Rev.* 92: 226-242.
10. Leduc C, Antoni D, Charloux A, Falcoz P-E, Quoix E (2017) Comorbidities in the management of patients with lung cancer. *Eur. Respir. J.* 49: 1601721.
11. Gullón JA, Suárez I, Medina A, Rubinos G, Fernández R, González I (2011) Role of emphysema and airway obstruction in prognosis of lung cancer. *Lung Cancer.* 71: 182-185.
12. Gao YH, Guan WJ, Liu Q, et al (2016) Impact of COPD and emphysema on survival of patients with lung cancer: A meta-analysis of observational studies. *Respirology.* 21: 269-279.
13. Gibiot Q, Monnet I, Levy P, et al (2020) Interstitial Lung Disease Associated with Lung Cancer: A Case-Control Study. *Journal of clinical medicine.* 9: 700.
14. Lin H, Lu Y, Lin L, Meng K, Fan J (2019) Does chronic obstructive pulmonary disease relate to poor prognosis in patients with lung cancer?: A meta-analysis. *Medicine (Baltimore).* 98: e14837.
15. Wang P, Zhu M, Zhang D, et al (2019) The relationship between chronic obstructive pulmonary disease and non-small cell lung cancer in the elderly. *Cancer medicine.* 8: 4124-4134.
16. Mina N, Soubani AO, Cote ML, et al (2012) The relationship between chronic obstructive pulmonary disease and lung cancer in African American patients. *Clin. Lung Cancer.* 13: 149-156.
17. Han S, Lee YJ, Park JS, et al (2019) Prognosis of non-small-cell lung cancer in patients with idiopathic pulmonary fibrosis. *Sci. Rep.* 9: 12561.
18. Goto T, Maeshima A, Oyamada Y, Kato R (2014) Idiopathic pulmonary fibrosis as a prognostic factor in non-small cell lung cancer. *Int. J. Clin. Oncol.* 19: 266-273.

19. Zhu L, Cao H, Zhang T, et al (2016) The Effect of Diabetes Mellitus on Lung Cancer Prognosis: A PRISMA-compliant Meta-analysis of Cohort Studies. *Medicine (Baltimore)*. 95: e3528.
20. Song M, Graubard BI, Rabkin CS, Engels EA (2021) Neutrophil-to-lymphocyte ratio and mortality in the United States general population. *Sci. Rep.* 11: 464.
21. Colicino E, Marioni R, Ward-Caviness C, et al (2020) Blood DNA methylation sites predict death risk in a longitudinal study of 12, 300 individuals. *Aging*. 12: 14092-14124.
22. Fest J, Ruiter TR, Groot Koerkamp B, et al (2019) The neutrophil-to-lymphocyte ratio is associated with mortality in the general population: The Rotterdam Study. *Eur. J. Epidemiol.* 34: 463-470.
23. Kim S, Eliot M, Koestler DC, Wu W-C, Kelsey KT (2018) Association of Neutrophil-to-Lymphocyte Ratio With Mortality and Cardiovascular Disease in the Jackson Heart Study and Modification by the Duffy Antigen Variant. *JAMA cardiology*. 3: 455-462.
24. Shah N, Parikh V, Patel N, et al (2014) Neutrophil lymphocyte ratio significantly improves the Framingham risk score in prediction of coronary heart disease mortality: Insights from the National Health and Nutrition Examination Survey-III. *Int. J. Cardiol.* 171: 390-397.
25. Abete I, Lu Y, Lassale C, Verschuren M, van der Schouw Y, Bueno-de-Mesquita B (2019) White cell counts in relation to mortality in a general population of cohort study in the Netherlands: a mediating effect or not? *BMJ Open*. 9: e030949.
26. Kang J, Chang Y, Ahn J, et al (2019) Neutrophil-to-lymphocyte ratio and risk of lung cancer mortality in a low-risk population: A cohort study. *Int. J. Cancer*. 145: 3267-3275.
27. Templeton AJ, McNamara MG, Seruga B, et al (2014) Prognostic role of neutrophil-to-lymphocyte ratio in solid tumors: a systematic review and meta-analysis. *J. Natl. Cancer Inst.* 106: dju124.
28. Dolan RD, Lim J, McSorley ST, Horgan PG, McMillan DC (2017) The role of the systemic inflammatory response in predicting outcomes in patients with operable cancer: Systematic review and meta-analysis. *Sci. Rep.* 7: 16717.
29. Li X, Dai D, Chen B, Tang H, Xie X, Wei W (2018) The value of neutrophil-to-lymphocyte ratio for response and prognostic effect of neoadjuvant chemotherapy in solid tumors: A systematic review and meta-analysis. *J. Cancer*. 9: 861-871.
30. Choi N, Kim JH, Chie EK, Gim J, Kang H-C (2019) A meta-analysis of the impact of neutrophil-to-lymphocyte ratio on treatment outcomes after radiotherapy for solid tumors. *Medicine*. 98: e15369-e15369.
31. Yu Y, Qian L, Cui J (2017) Value of neutrophil-to-lymphocyte ratio for predicting lung cancer prognosis: A meta-analysis of 7,219 patients. *Mol Clin Oncol*. 7: 498-506.
32. Gu XB, Tian T, Tian XJ, Zhang XJ (2015) Prognostic significance of neutrophil-to-lymphocyte ratio in non-small cell lung cancer: a meta-analysis. *Sci. Rep.* 5: 12493.
33. Peng B, Wang YH, Liu YM, Ma LX (2015) Prognostic significance of the neutrophil to lymphocyte ratio in patients with non-small cell lung cancer: a systemic review and meta-analysis. *Int. J. Clin. Exp. Med.* 8: 3098-3106.
34. Lu Y, Jiang J, Ren C (2020) The clinicopathological and prognostic value of the pretreatment neutrophil-to-lymphocyte ratio in small cell lung cancer: A meta-analysis. *PLoS One*. 15: e0230979.
35. Lee SF, Luk H, Wong A, Ng CK, Wong FCS, Luque-Fernandez MA (2020) Prediction model for short-term mortality after palliative radiotherapy for patients having advanced cancer: a cohort study from routine electronic medical data. *Sci. Rep.* 10: 5779.
36. Wang Z, Zhan P, Lv Y, et al (2019) Prognostic role of pretreatment neutrophil-to-lymphocyte ratio in non-small cell lung cancer patients treated with systemic therapy: a meta-analysis. *Translational lung cancer research*. 8: 214-226.
37. Lohinai Z, Bonanno L, Aksarin A, et al (2019) Neutrophil-lymphocyte ratio is prognostic in early stage resected small-cell lung cancer. *PeerJ*. 7: e7232-e7232.

38. Michaud DS, Houseman EA, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT (2015) Understanding the Role of the Immune System in the Development of Cancer: New Opportunities for Population-Based Research. *Cancer Epidemiol. Biomarkers Prev.* 24: 1811-1819.
39. Koestler DC, Usset J, Christensen BC, et al (2017) DNA Methylation-Derived Neutrophil-to-Lymphocyte Ratio: An Epigenetic Tool to Explore Cancer Inflammation and Outcomes. *Cancer Epidemiol. Biomarkers Prev.* 26: 328-338.
40. Salas LA, Koestler DC, Butler RA, et al (2018) An optimized library for reference-based deconvolution of whole-blood biospecimens assayed using the Illumina HumanMethylationEPIC BeadArray. *Genome Biol.* 19: 64.
41. Derman BA, Macklis JN, Azeem MS, et al (2017) Relationships between longitudinal neutrophil to lymphocyte ratios, body weight changes, and overall survival in patients with non-small cell lung cancer. *BMC Cancer.* 17: 141.
42. Kitamura T, Qian BZ, Pollard JW (2015) Immune cell promotion of metastasis. *Nat. Rev. Immunol.* 15: 73-86.
43. Kotas ME, Medzhitov R (2015) Homeostasis, inflammation, and disease susceptibility. *Cell.* 160: 816-827.
44. Schneider G, Schmidt-Supprian M, Rad R, Saur D (2017) Tissue-specific tumorigenesis: context matters. *Nature reviews. Cancer.* 17: 239-253.
45. Omenn GS, Goodman GE, Thornquist MD, et al (1996) Risk factors for lung cancer and for intervention effects in CARET, the Beta-Carotene and Retinol Efficacy Trial. *J. Natl. Cancer Inst.* 88: 1550-1559.
46. Grieshober L, Graw S, Barnett MJ, et al (2020) AHRR methylation in heavy smokers: associations with smoking, lung cancer risk, and lung cancer mortality. *BMC Cancer.* 20: 905.
47. Grieshober L, Graw S, Barnett MJ, et al (2018) Methylation-derived Neutrophil-to-Lymphocyte Ratio and Lung Cancer Risk in Heavy Smokers. *Cancer Prev. Res.* 11: 727-734.
48. Gazdar AF, Bunn PA, Minna JD (2017) Small-cell lung cancer: what we know, what we need to know and the path forward. *Nat. Rev. Cancer.* 17: 725-737.
49. Wang S, Tang J, Sun T, et al (2017) Survival changes in patients with small cell lung cancer and disparities between different sexes, socioeconomic statuses and ages. *Sci. Rep.* 7: 1339-1339.
50. Cuffe S, Moua T, Summerfield R, Roberts H, Jett J, Shepherd FA (2011) Characteristics and Outcomes of Small Cell Lung Cancer Patients Diagnosed During Two Lung Cancer Computed Tomographic Screening Programs in Heavy Smokers. *J. Thorac. Oncol.* 6: 818-822.
51. Wang S, Zimmermann S, Parikh K, Mansfield AS, Adjei AA (2019) Current Diagnosis and Management of Small-Cell Lung Cancer. *Mayo Clin. Proc.* 94: 1599-1622.
52. Taniguchi H, Sen T, Rudin CM (2020) Targeted Therapies and Biomarkers in Small Cell Lung Cancer. *Front. Oncol.* 10: 741-741.
53. Scheiermann C, Frenette PS, Hidalgo A (2015) Regulation of leucocyte homeostasis in the circulation. *Cardiovasc. Res.* 107: 340-351.
54. Grecian R, Whyte MKB, Walmsley SR (2018) The role of neutrophils in cancer. *Br. Med. Bull.* 128: 5-14.
55. Coffelt SB, Wellenstein MD, de Visser KE (2016) Neutrophils in cancer: neutral no more. *Nature Reviews Cancer.* 16: 431-446.
56. Kolaczkowska E, Kubes P (2013) Neutrophil recruitment and function in health and inflammation. *Nature Reviews Immunology.* 13: 159-175.
57. Engblom C, Pfirschke C, Pittet MJ (2016) The role of myeloid cells in cancer therapies. *Nature Reviews Cancer.* 16: 447-462.
58. Rudin CM, Poirier JT, Byers LA, et al (2019) Molecular subtypes of small cell lung cancer: a synthesis of human and mouse model data. *Nature reviews. Cancer.* 19: 289-297.
59. Poirier JT, Gardner EE, Connis N, et al (2015) DNA methylation in small cell lung cancer defines distinct disease subtypes and correlates with high expression of EZH2. *Oncogene.* 34: 5869-5878.

60. Vano Y-A, Oudard S, By M-A, et al (2018) Optimal cut-off for neutrophil-to-lymphocyte ratio: Fact or Fantasy? A prospective cohort study in metastatic cancer patients. PLoS One. 13: e0195042.
61. Zhao NR, Mengyuan; Koestler, Devin C.; Lu, Jiayun; Marsit, Carmen J.; Kelsey, Karl T.; Platz, Elizabeth A.; Michaud, Dominique S. (2021) Methylation-derived Inflammatory Measures and Lung Cancer Risk and Survival. medRxiv.
62. Suh B, Shin DW, Kwon HM, et al (2017) Elevated neutrophil to lymphocyte ratio and ischemic stroke risk in generally healthy adults. PLoS One. 12: e0183706.
63. Azab B, Camacho-Rivera M, Taioli E (2014) Average values and racial differences of neutrophil lymphocyte ratio among a nationally representative sample of United States subjects. PLoS One. 9: e112361.
64. Clarke R, Shipley M, Lewington S, et al (1999) Underestimation of Risk Associations Due to Regression Dilution in Long-term Follow-up of Prospective Studies. Am. J. Epidemiol. 150: 341-353.