

## Article

# DNA Methylation-Based Estimates of Circulating Leukocyte Composition for Predicting Colorectal Cancer Survival: A Prospective Cohort Study

Xīn Gào <sup>1,\*</sup> , Yan Zhang <sup>1,2</sup>, Xiangwei Li <sup>1,3</sup>, Lina Jansen <sup>1</sup>, Elizabeth Alwers <sup>1</sup>, Melanie Beyer-Hudler <sup>4</sup> , Matthias Schick <sup>4</sup>, Jenny Chang-Claude <sup>5</sup>, Michael Hoffmeister <sup>1</sup>  and Hermann Brenner <sup>1,2,6</sup>

<sup>1</sup> Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 581, 69120 Heidelberg, Germany; gzhyanzi@hotmail.com (Y.Z.); xiangwei.li@dkfz.de (X.L.); ljansen@dkfz.de (L.J.); elizabeth.alwers@dkfz.de (E.A.); m.hoffmeister@dkfz.de (M.H.); h.brenner@dkfz.de (H.B.)

<sup>2</sup> German Cancer Consortium, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

<sup>3</sup> Medical Faculty, University of Heidelberg, 69120 Heidelberg, Germany

<sup>4</sup> Genomics and Proteomics Core Facility, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 580, 69120 Heidelberg, Germany; m.hudler@dkfz.de (M.B.-H.); m.schick@dkfz.de (M.S.)

<sup>5</sup> Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 581, 69120 Heidelberg, Germany; j.chang-claude@dkfz.de

<sup>6</sup> Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Im Neuenheimer Feld 460, 69120 Heidelberg, Germany

\* Correspondence: xin.gao@dkfz.de; Tel.: +49-6221421306



**Citation:** Gào, X.; Zhang, Y.; Li, X.; Jansen, L.; Alwers, E.;

Beyer-Hudler, M.; Schick, M.; Chang-Claude, J.; Hoffmeister, M.; Brenner, H. DNA Methylation-Based Estimates of Circulating Leukocyte Composition for Predicting Colorectal Cancer Survival: A Prospective Cohort Study. *Cancers* **2021**, *13*, 2948. <https://doi.org/10.3390/cancers13122948>

Academic Editors: Tomasz K. Wojdacz and Matthew Rose-Zerilli

Received: 11 May 2021

Accepted: 9 June 2021

Published: 12 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Simple Summary:** Inflammation is involved in the evolution of cancer. Leukocytes, of which the proportion can be estimated using epigenome-wide methylation data, may serve as a prognostic marker in colorectal cancer (CRC). Our aim was to investigate whether DNA methylation-based estimates of circulating leukocytes is associated with all-cause and disease-specific mortality in a prospective CRC patients' cohort. Significant associations with CRC prognosis were observed for CD4+ T cells, CD8+ T cells, B cells, NK cells, and lymphocytes, independent of age, sex, tumor stage, tumor subsite, and therapy. CD4+ T cells outperformed other leukocytes and provided added predictive value in comparison to age, sex, and tumor stage. Although cell counting is commonly used in clinical practice, DNA methylation-estimated cell proportions could be a promising tool in understanding the role of leukocytes as CRC prognostic biomarkers when using stored blood samples.

**Abstract:** Leukocytes are involved in the progression of colorectal cancer (CRC). The proportion of six major leukocyte subtypes can be estimated using epigenome-wide DNA methylation (DNAm) data from stored blood samples. Whether the composition of circulating leukocytes can be used as a prognostic factor is unclear. DNAm-based leukocyte proportions were obtained from a prospective cohort of 2206 CRC patients. Multivariate Cox regression models and survival curves were applied to assess associations between leukocyte composition and survival outcomes. A higher proportion of lymphocytes, including CD4+ T cells, CD8+ T cells, B cells, and NK cells, was associated with better survival, while a higher proportion of neutrophils was associated with poorer survival. CD4+ T cells outperformed other leukocytes in estimating the patients' prognosis. Comparing the highest quantile to the lowest quantile of CD4+ T cells, hazard ratios (95% confidence intervals) of all-cause and CRC-specific mortality were 0.59 (0.48, 0.72) and 0.59 (0.45, 0.77), respectively. Furthermore, the association of CD4+ T cells and prognosis was stronger among patients with early or intermediate CRC or patients with colon cancer. In conclusion, the composition of circulating leukocytes estimated from DNAm, particularly the proportions of CD4+ T cells, could be used as promising independent predictors of CRC survival.

**Keywords:** DNA methylation; circulating leukocytes; colorectal cancer; prognosis

## 1. Introduction

Colorectal cancer (CRC) represents the second leading cause of cancer-related deaths worldwide [1]. Although recent improvements have been made in screening strategies and treatments for CRC, the prognosis of advanced CRC is still poor [2,3]. Moreover, prognostic information provided by the anatomy-based tumor-node-metastasis staging system is incomplete. Survival rates can be significantly different among patients within the same stage [4]. Therefore, novel prognostic markers are needed to improve patient management for tertiary prevention and to elicit a better understanding of the disease processes. We previously observed that higher levels of DNA methylation (DNAm)-based mortality risk score, AgeAccelPheno and AgeAccelGrim were statistically significantly associated with poorer CRC prognosis. However, the associations were weaker for CRC-specific survival than for overall survival. Studies are needed to identify prognostic biomarkers that are more specific to CRC [5].

Immune inflammatory cells are involved in CRC progression in conflicting ways: both tumor-suppressing and tumor-promoting leukocytes are involved [6,7]. Intratumoral lymphocytes can induce tumor cell death by secreting tumor necrosis factor- $\alpha$  and interferon- $\gamma$  [8]. On the other hand, a higher infiltration of neutrophils facilitates the pathological angiogenesis that is related to tumor growth [9]. Epidemiological studies have shown that a higher neutrophil to lymphocyte ratio in peripheral blood is associated with a worse CRC survival [10–14]. These findings suggest that the leukocyte composition may represent the underlying immuno-biology to CRC progression and could serve as a promising prognostic marker for CRC [15]. However, usually, freshly drawn venous blood and strict cell processing are critical for quantifying peripheral blood leukocyte composition using traditional techniques. Therefore, leukocyte counts are often missing in large-scale clinical and epidemiological studies.

Cell lineages can be distinguished by differentially methylated DNA regions. Methods have been developed to infer the leukocyte composition using DNA methylation (DNAm) signatures that are chemically stable [16–18]. Salas and colleagues identified a library of 450 CpGs for deconvoluting CD4+ T cells, CD8+ T cells, B cells, natural killer (NK) cells, monocytes and neutrophil fraction, which were specific to DNAm signatures assayed using Illumina Infinium HumanMethylationEPIC BeadChip Kit (EPIC microarray, Illumina, Inc., San Diego, CA, USA) [18]. The method showed good accuracy of cell composition estimates, and the average coefficient of determination was 99.2% between the artificial predefined cell proportions of blood mixtures and the estimations [18]. It was further validated in actual blood samples and publicly available datasets [18].

In this work, the leukocyte composition was estimated using Salas' method based on EPIC microarray data in a large cohort of CRC patients. We then investigated the association of leukocyte composition with all-cause mortality and CRC-specific mortality of CRC patients and their capability for predicting survival.

## 2. Materials and Methods

### 2.1. Study Design and Population

The current study is based on the prospective follow-up of CRC patients of an ongoing population-based case-control study on CRC, the German Darmkrebs: Chancen der Verhütung durch Screening (DACHS) Study. Details of the DACHS study design have been described elsewhere [19–22]. Briefly, we recruited patients with a first diagnosis of CRC (ICD 10 codes: C18–C20) in 22 clinics that provide CRC surgery in the Rhine-Neckar region of Germany. In the current analysis, we included 2206 CRC patients whose information on follow-up with respect to survival outcomes and DNAm array data from blood samples taken at baseline.

### 2.2. Data Collection

CRC patients were diagnosed between 2003 and 2010 and were invited by their treating physicians during the first hospital stay due to CRC and notified to the study

center at the German Cancer Research Center, after receiving informed consent. Trained interviewers collected patients' sociodemographic, medical and lifestyle information using a standardized questionnaire at the earliest possible convenience, either during hospital stay or shortly thereafter at patients' homes. Moreover, data on tumor characteristics and treatment were extracted from medical records. Information on newly diagnosed diseases and recurrences was provided by physicians at 3- and 5-year follow-up. Data on date and cause of death were obtained from the local population registers and public health authorities. The study was approved by the Ethics Committees of the Medical Faculty of the University of Heidelberg (ID: 310/2001) and the Medical Chambers of Baden-Württemberg (ID: M-198-02) and Rhineland-Palatinate (ID: 837.419.02 (3637)).

### 2.3. DNA Methylation and Leukocyte Composition Estimation

Blood samples were collected after the interview and were shipped to the study center and stored at  $-80\text{ }^{\circ}\text{C}$  until whole blood DNA extraction. Genome-wide DNAm signatures were assayed using the EPIC microarray (Illumina, Inc., San Diego, CA, USA) according to the manufacturer's protocol at the Genomics and Proteomics Core Facility of the German Cancer Research Center. Preprocessing and normalization of DNAm data was conducted following the pipeline proposed by Lehne and colleagues [23]. Probes with detection  $p$  value  $> 0.01$ , with missing value  $> 10\%$ , probes targeting the sex chromosomes, cross-reactive probes and polymorphic CpGs were excluded, leaving 787,231 CpGs for analyses. The proportions of CD4+ T cells, CD8+ T cells, B cells, NK cells, monocytes and neutrophils were estimated using a deconvolution method, which was implemented with the FlowSorted.Blood.EPIC package of R [24].

### 2.4. Statistical Methods

The proportion of categorical covariates was calculated to describe the characteristics of the study population. The distribution of baseline characteristics of patients at different stages were compared using Chi square test. Wilcoxon–Mann–Whitney test was performed to compare medians of leukocyte composition across categorical variables. Jonckheere–Terpstra test was used to determine the significance of the trend of ordinal variables. Correlations between leukocyte components were assessed using Pearson correlation coefficients and scatter plots.

Associations between leukocyte composition and survival were assessed using Cox regression, from which hazard ratios (HRs) and 95% confidence intervals (95% CIs) were derived. We ran a “clinical model” adjusted for covariates that can be easily obtained in clinical settings, including age, sex, stage, CRC subsite, batch, and timing of blood sampling. A “comprehensive model” additionally adjusted for body mass index (BMI,  $\text{kg}/\text{m}^2$ ), smoking status (never, former and current smokers), and Charlson comorbidity index (CCI) score was applied to evaluate potential confounding by these covariates. The characteristics of all covariates except for the timing of blood sampling are shown in Table 1. Complete case analysis was applied to address missing data ( $<0.5\%$ ) when adjusting for these variables. The timing of blood sampling was grouped into four categories including: (1) prior to surgery or within 1 month after surgery ( $n = 1006$ ); (2) more than 1 month after surgery, but not receiving chemotherapy or radiotherapy ( $n = 569$ ); (3) during chemo- or radiotherapy, or within 5 months after the first treatment ( $n = 278$ ); (4) more than 5 months after chemo- or radiotherapy ( $n = 353$ ) (Figure S1). The proportional hazards assumption was tested using the Schoenfeld Residuals. Furthermore, we accounted for delayed entry time by incorporating the time elapsed between diagnosis and the starting of patients' enrolment in the standard Cox models.

**Table 1.** Baseline characteristics of the patient cohort.

Characteristics	n (%) *
Sex	
Women	910 (41.2)
Men	1296 (58.8)
Age at diagnosis	
[33, 65) years	717 (32.5)
[65, 75) years	776 (35.2)
[75, 96) years	713 (32.3)
Tumor stage	
I	400 (18.2)
II	759 (34.6)
III	726 (33.1)
IV	309 (14.1)
Tumor subsite	
Distal colon †	738 (33.5)
Proximal colon ‡	796 (36.1)
Rectum	669 (30.4)
Body mass index at diagnosis	
<25 kg/m <sup>2</sup>	834 (38.0)
25–30 kg/m <sup>2</sup>	932 (42.4)
>30 kg/m <sup>2</sup>	430 (19.6)
Smoking status	
Never	907 (41.1)
Former	947 (43.0)
Current	350 (15.9)
Charlson comorbidity index	
0 (no comorbidity)	1281 (58.1)
1 (mild comorbidity)	479 (21.7)
2+ (moderate comorbidity)	446 (20.2)

Note: \* 12 missing values for tumor stage, 10 missing values for body mass index, 2 missing values for smoking status and 3 missing values for tumor subsite. Complete case analysis was applied when adjusting for the variables with missing values. † The distal colon includes the descending colon and the sigmoid colon. ‡ The proximal colon includes the cecum, the ascending colon, the right flexure, and the transverse colon and the left flexure.

Adjusted survival curves were plotted to assess whether the association differed depending on tumor stage, tumor subsite and the timing of blood sampling. The difference between survival curves across quartiles of leukocyte composition was evaluated using the G-rho family of tests. Furthermore, predictive accuracy and discriminating ability of leukocyte composition were evaluated using Harrell's concordance statistics (C-statistics) and were compared with age, sex and stage. Cubic spline models were also performed to assess the dose–response relationship between CD4+ T cell proportion and CRC mortality in various therapy groups.

Correlation matrix, survival curves and cubic spline curves were produced using the R 3.6.0 with the packages corrplot, survminer and survival, respectively [25]. Hazard ratios and Harrell's C-statistics were derived using the PROC PHREG procedure in SAS version 9.4 (SAS Institute, Cary, NC, USA). Statistical significance was defined by  $p < 0.05$  in two-sided testing.

### 3. Results

#### 3.1. Characteristics of the Study Population

Table 1 summarizes the demographic and clinical characteristics of the study population at baseline. Among the 2206 CRC patients, there were more men (58.8%) than women (41.2%). More than two-thirds of the study population (67.5%) were older than 65 years, 62.0% of patients were overweight or obese, 15.9% of patients were current smoker at baseline, and more than 40% had a relevant comorbidity (CCI > 0). Of the patients, 18.2%, 34.6%, 33.1% and 14.1% were diagnosed with stage I, II, III and IV cancer, respectively. Colon cancer and rectal cancer accounted for 69.6% and 30.4% of CRC patients, respectively. The distribution of age, BMI, smoking status, CCI and tumor sub-site were statistically significantly different across the tumor stages (Supplementary Materials Table S1). A lower lymphocyte proportion and a higher neutrophil proportion were observed in males, older patients, and patients diagnosed at advanced stages, with rectal cancer and with higher CCI (Table S2). Moreover, CD4+ T cells, CD8+ T cells, B cells, NK cells, and monocytes were statistically positively correlated with each other, whereas neutrophils were statistically inversely correlated with other leukocyte subtypes (Figure S2).

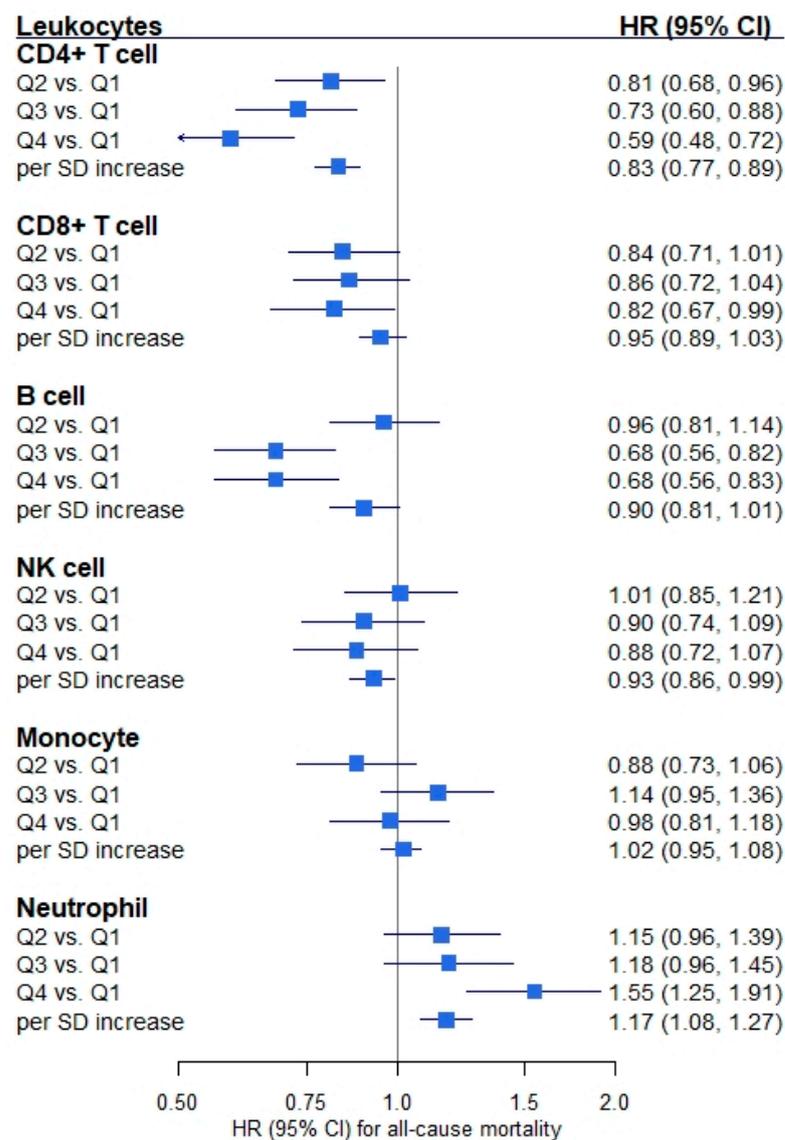
#### 3.2. Association of Leukocyte Composition with CRC Prognosis

Figure 1 shows that proportions of all leukocyte subtypes except monocytes were statistically significantly associated with all-cause mortality of CRC patients. HRs (95% CIs) for the association of all-cause mortality with highest (vs. lowest) quartiles of the proportions were: CD4+ T cells, 0.59 (0.48, 0.72); CD8+ T cells, 0.82 (0.67, 0.99); B cells, 0.68 (0.56, 0.83); NK cells, 0.88 (0.72, 1.07); monocytes, 0.98 (0.81, 1.18); and neutrophils, 1.55 (1.25, 1.91), respectively. However, only CD4+ T cell, B cell and neutrophil proportions were statistically significantly associated with CRC-specific mortality (Figure 2). HRs (95% CIs) for the comparison of the highest with the lowest quartile of the proportions were: CD4+ T cells, 0.59 (0.45, 0.77); CD8+ T cells, 0.88 (0.69, 1.14); B cells, 0.70 (0.54, 0.90); NK cells, 1.03 (0.79, 1.36); monocytes, 1.06 (0.83, 1.35); and neutrophils, 1.32 (1.00, 1.75), respectively. Furthermore, sensitivity analyses showed that some of the associations between leukocyte proportions and CRC prognosis were attenuated after additional adjustment for BMI, smoking status and CCI. However, statistically significant associations with CRC prognosis remained for CD4+ T cells, B cells and neutrophils (Table S2).

#### 3.3. Subgroup Analyses

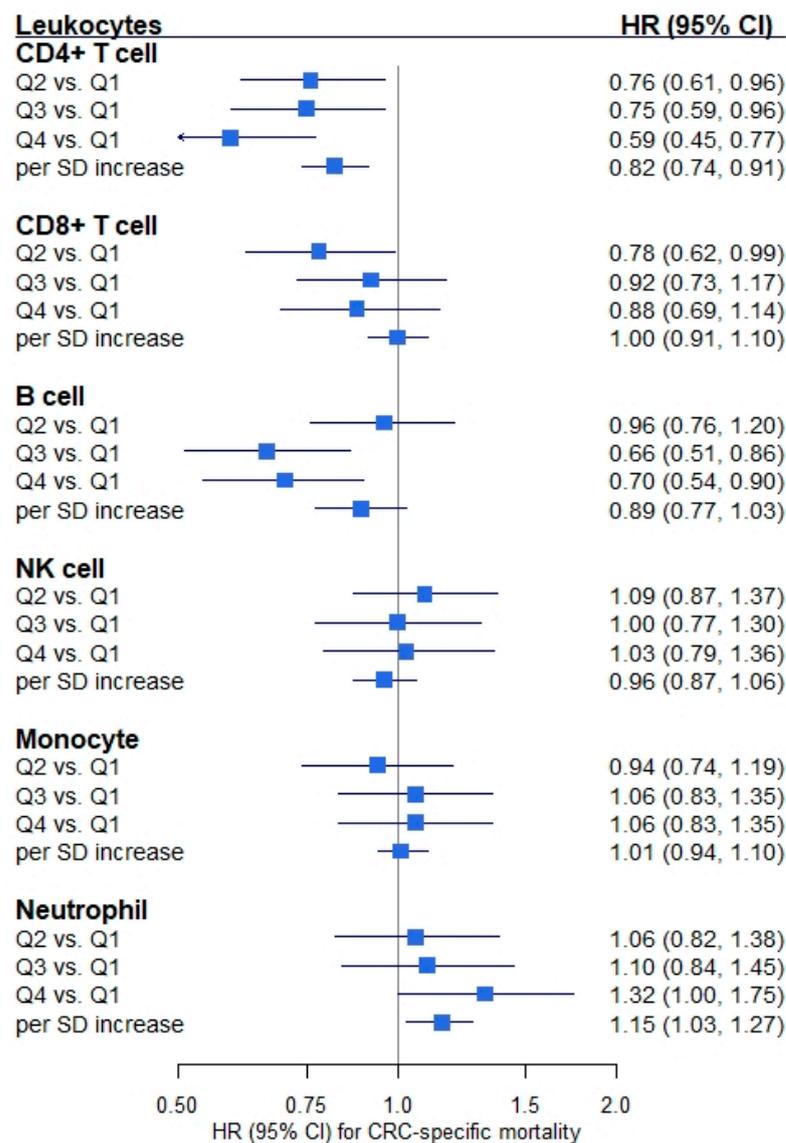
Figure 3 presents the overall and CRC-specific survival curves across quartiles of the proportion of CD4+ T cells by tumor stage. We observed statistically significant associations between a higher proportion of CD4+ T cells and improved overall survival in the early and intermediate stages (I–III) but not in the advanced stage (IV). Moreover, the proportion of CD4+ T cells was statistically significantly positively associated with CRC-specific survival in patients with stage III cancers.

Figure 4 shows the tumor subsite-specific association of the proportion of CD4+ T cells with overall and CRC-specific survival. Among patients with colon cancer, a higher proportion of CD4+ T cells was statistically significantly associated with better survival, and the association was weaker for CRC-specific survival than for overall survival. However, the association between the proportion of CD4+ T cells and both outcomes was not statistically significant among patients with rectal cancer.



**Figure 1.** Forest plot of the associations between leukocyte composition and all-cause mortality. Note: Hazard ratio (HR) and 95% confidence interval (CI) were derived from Cox regression model adjusting for age, sex, tumor stage, tumor subsite, timing of blood sampling and batch ( $n_{\text{patients}} = 2194$ ,  $n_{\text{deaths}} = 1071$ ). An HR of one means that there is no difference in survival between the two groups. Abbreviation: Q1–Q4, quartile 1 (lowest)–quartile 4 (highest); SD, standard deviation.

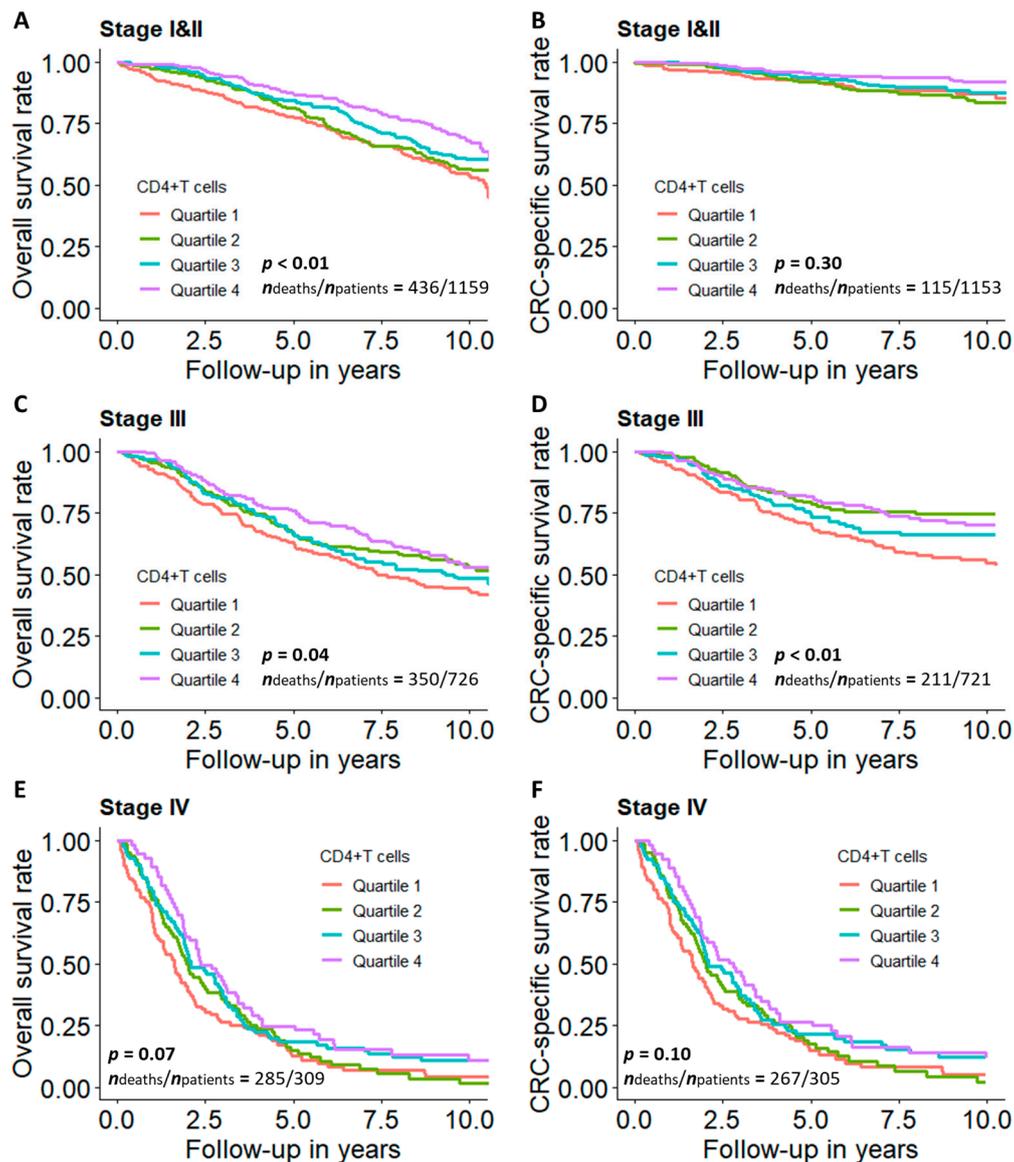
Additionally, patients in the second-highest quartile of CD4+ T cell proportion had the poorest survival among the subgroup whose blood was sampled during or within 5 months after initiation of chemo- or radiotherapy treatment, which is different from the other subgroups (Figure S3). Moreover, as shown in the dose–response curves, a “U-shaped” dose–response relationship was observed among patients who received surgery, chemotherapy and radiotherapy (or only received surgery and radiotherapy) (Figure S4).



**Figure 2.** Forest plot of the associations between leukocyte composition and colorectal cancer (CRC)-specific mortality. Note: Hazard ratio (HR) and 95% confidence interval (CI) were derived from Cox regression model adjusting for age, sex, tumor stage, tumor subsite, timing of blood sampling and batch ( $n_{\text{patients}} = 2179$ ,  $n_{\text{deaths}} = 593$ ). An HR of one means that there is no difference in survival between the two groups. Abbreviation: Q1–Q4, quartile 1 (lowest)–quartile 4 (highest); SD, standard deviation.

### 3.4. Predictive Utility of CD4+ T Cell Proportion and All Leukocyte Proportions Combined

Table 2 presents the discriminative ability of various combinations of predictors, including age, sex, tumor stage, and leukocyte compositions. The prediction of CRC prognosis was modestly improved after adding the proportion of CD4+ T cells into the model, overall and by tumor stage. In all stages, as well as stages I–II and stage III, simultaneous consideration of all leukocyte subtypes combined slightly outperformed exclusive consideration of CD4+ T cell proportions for prediction of CRC prognosis. However, in stage IV, the models including CD4+ T cell proportions only, had an even better performance than the models including all leukocyte subtypes.



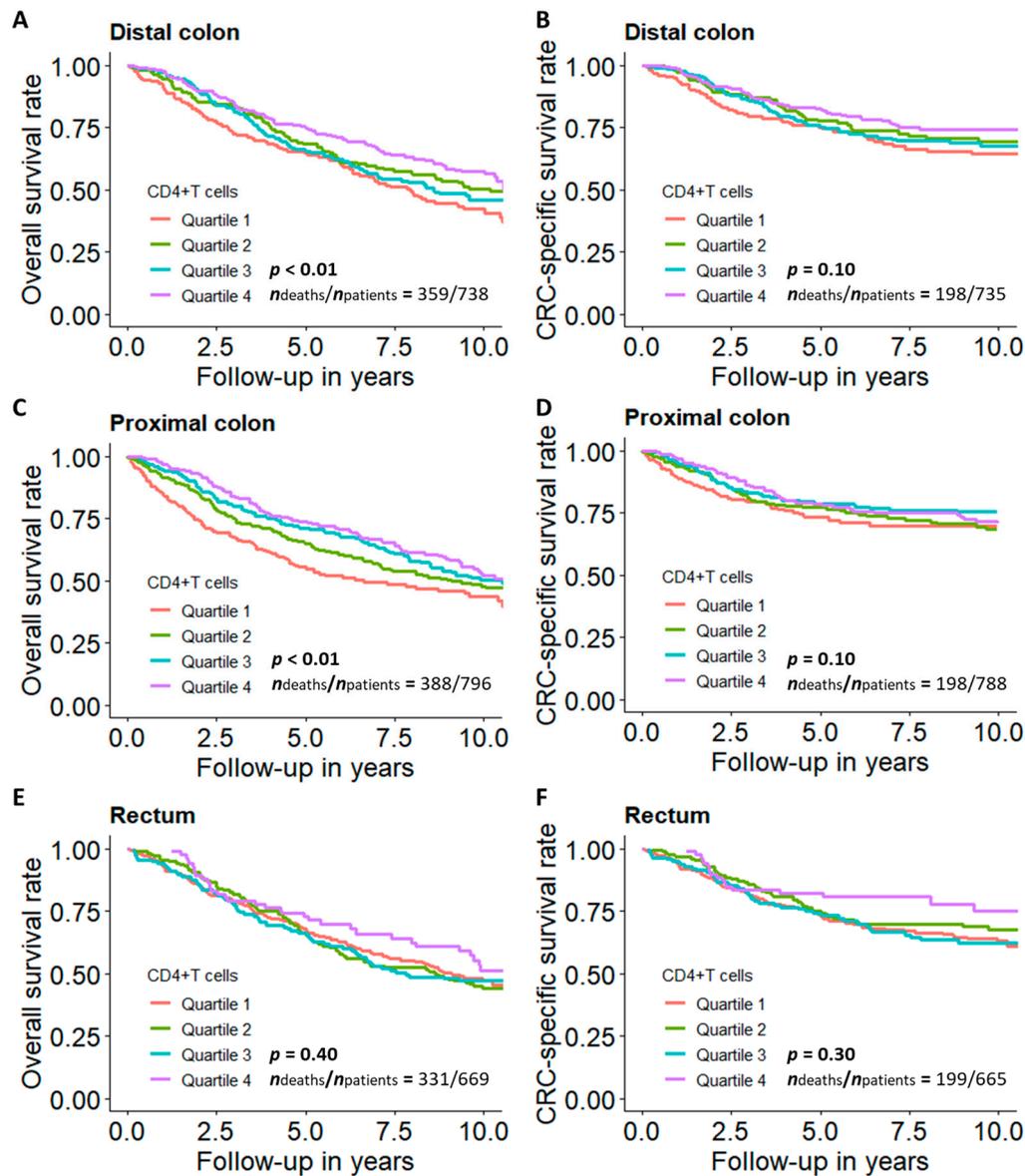
**Figure 3.** Survival curves for CRC patients across quartiles of CD4+ T cell proportions by tumor stage. (A) Overall and (B) CRC-specific survival curves among stage I and II patients; (C) overall and (D) CRC-specific survival curves among stage III patients; (E) overall and (F) CRC-specific survival curves among stage IV patients. Note: Survival curves were adjusted for age, sex, tumor subsite, timing of blood sampling and batch. *p* values were derived from G-rho family of tests. Quartile 1: lowest quartile; Quartile 4: highest quartile.

**Table 2.** Harrell’s C-statistics (95% confidence interval) for all-cause mortality and colorectal cancer (CRC)-specific mortality prediction.

Combination of Predictors		All-Cause Mortality	CRC-Specific Mortality
All stages	Age + sex + stage	0.739 (0.723, 0.754)	0.809 (0.792, 0.825)
	Age + sex + stage + CD4(+)T cell	0.743 (0.728, 0.758)	0.813 (0.797, 0.829)
	Age + sex + stage + all leukocyte subtypes	0.744 (0.729, 0.759)	0.814 (0.798, 0.830)
Stage I and II	Age + sex	0.693 (0.667, 0.718)	0.612 (0.559, 0.666)
	Age + sex + CD4(+)T cell	0.703 (0.678, 0.729)	0.630 (0.578, 0.682)
	Age + sex + all leukocyte subtypes	0.706 (0.680, 0.731)	0.643 (0.593, 0.693)

Table 2. Cont.

Combination of Predictors		All-Cause Mortality	CRC-Specific Mortality
Stage III	Age + sex	0.653 (0.622, 0.683)	0.608 (0.568, 0.648)
	Age + sex + CD4(+)T cell	0.659 (0.629, 0.689)	0.622 (0.583, 0.660)
	Age + sex + all leukocyte subtypes	0.662 (0.632, 0.691)	0.626 (0.587, 0.665)
Stage IV	Age + sex	0.557 (0.520, 0.594)	0.557 (0.519, 0.595)
	Age + sex + CD4(+)T cell	0.597 (0.562, 0.631)	0.600 (0.564, 0.635)
	Age + sex + all leukocyte subtypes	0.593 (0.557, 0.629)	0.599 (0.562, 0.636)



**Figure 4.** Survival curves for CRC patients across quartiles of CD4+ T cell proportions by tumor subsite. (A) Overall and (B) CRC-specific survival curve among patients with distal colon cancer; (C) overall and (D) CRC-specific survival curve among patients with proximal colon cancer; (E) overall and (F) CRC-specific survival curve among patients with rectal cancer. Note: Survival curves were adjusted for age, sex, tumor stage, timing of blood sampling and batch. *p* values were derived from G-rho family of tests. Quartile 1: lowest quartile; Quartile 4: highest quartile.

#### 4. Discussion

This analysis is the first prospective study to investigate the association of DNAm-based estimates of leukocyte composition with CRC prognosis. Patients with higher estimated proportions of CD4+ T cells, CD8+ T cells and B cells had lower all-cause and CRC-specific mortality, and patients with a higher proportion of neutrophils had higher all-cause and CRC-specific mortality. Among all leukocyte subtypes, CD4+ T cells presented the strongest association with CRC prognosis. A one-SD increment of the proportion of CD4+ T cells (7%) was associated with 17% and 18% lower all-cause and CRC-specific mortality, respectively. Subgroup analyses showed that the association of CD4+ T cells with CRC prognosis was statistically significant in stage I–III and colon cancer patients.

CD4+ T cells, including T helper (Th) cells and regulatory T (Treg) cells, function to instigate and shape adaptive immune responses. Th cells can target tumor cells in various ways, either directly by eliminating tumor cells through cytolytic mechanisms or indirectly by activating and directing innate immune cells, B cells and CD8+ T cells [26]. Studies have shown that intratumoral CD4+ T cell infiltration can suppress tumor progression [27,28]. On the other hand, Treg cells, which are essential to maintain self-tolerance and immune cell homeostasis, are involved in the tumor progression by establishing an immunosuppressive tumor microenvironment [29]. Suppression of tumor-specific CD4+ T cells by Treg was associated with a worse CRC prognosis in a cohort of 62 patients [30]. CD8+ T cells, also known as cytotoxic T lymphocytes, can infiltrate into tumor tissue and kill malignant cells by releasing cytotoxic granules or by recognizing Fas ligand on the surface of target cells. A higher density of intratumoral CD8+ T cells was associated with improved CRC prognosis, which is in line with our findings [31,32].

B lymphocytes are involved in cell-mediated adaptive immunity and can inhibit tumor progression by secreting tumor-reactive antibodies and priming T cells [33]. Various studies reported that the tumor-infiltrating B cell density was positively associated with CRC prognosis [34,35]. However, regulatory B cells can suppress Th1 and CD8+ T cell responses and subsequently promote tumor progression, similar to Treg cells [33]. Nevertheless, evidence on the association of circulating CD4+ and CD8+ T cells and B cells with CRC prognosis is scarce. Further studies are needed to confirm our findings.

NK cells can kill tumor cells without any priming or prior activation. Tang and colleagues observed in 447 CRC patients that patients with an increased percentage of peripheral blood NKs had a higher 3-year survival rate [36]. It was consistent with our results that a marginal inverse association between NK proportion and all-cause mortality was observed among CRC patients whose circulating leukocyte composition was not affected by the therapies. Monocytes are a major innate immune component of the mononuclear phagocyte system. A meta-analysis of 16 prospective studies comprising 3826 CRC patients showed that increased absolute blood monocyte counts were significantly associated with worse overall survival [37]. However, no significant association between the proportion of blood monocytes and CRC prognosis was observed in our study. The difference between the results of the two studies could be explained by the different measures of peripheral blood monocyte (cell count vs. cell proportion).

Neutrophils are the most abundant leukocytes in blood and an essential part of the innate immune system. In line with our findings, a higher neutrophil to lymphocyte ratio in peripheral blood is associated with higher mortality among CRC patients, overall or by tumor stage, which suggests that lymphocytes and neutrophils may play opposite roles in CRC progression [10–14]. Evidence also shows that neutrophils may play both pro- and antitumor roles, depending on the heterogeneous subsets [38,39]. The association between neutrophil levels and CRC prognosis is controversial. Rao et al. observed that elevated intratumoral neutrophils in CRC were significantly associated with malignant phenotype and adverse prognosis in 229 patients [40]. However, Wikberg et al. found in a study of 448 patients that low infiltration of neutrophils in the tumor front was indicative of a worse prognosis [41]. Thus, research with larger CRC patient cohorts is needed to investigate the relationship of neutrophils with survival.

DNAm-base markers have shown prognostic value for CRC patients in previous studies [5,42]. Our study further revealed that blood DNAm-based leukocyte composition, especially the proportion of CD4+ T cells, could be used as an independent marker to enhance clinical judgment of prognosis in patients with early or intermediate stage CRC. Leukocyte composition can also be used to explore potential mechanisms underlying immune responses to tumor growth. The positive association of each subtype of circulating lymphocytes with CRC overall survival suggests that lymphocytes, at the interface of innate and adaptive immunity, suppress CRC progression. Our data suggest that specific lymphocytes and neutrophils may be potentially utilized as pathogenic effectors and therapeutic targets.

CD4+ T cells were the strongest predictor for CRC prognosis among leukocyte subtypes in our study. However, the associations between CD4+ T cell proportion and CRC survival outcomes were weak and statistically nonsignificant among patients with advanced CRC. A possible explanation is that the survival for advanced CRC is generally extremely poor and the case numbers in this group were the smallest, which limited the statistical power to detect possible associations. Among early- and intermediate-stage patients, the predictive performance of circulating CD4+ cell proportion was better for all-cause mortality than for CRC-specific mortality, even after controlling for comorbidity score. It suggested that circulating CD4+ cell may not serve as a specific predictor for CRC prognosis. In subsite specific analyses, the association of a higher proportion of CD4+ T cells with better survival was statistically significant among colon cancer patients but not rectal cancer patients, of whom the sample size was smaller.

This study has several strengths including the prospective study design, large sample size, long-term follow-up, the well-recorded causes of death, and detailed information on the study population. Moreover, the composition of the leukocyte population was estimated using DNAm array data that is stable and can be obtained from stored whole blood samples.

There are also limitations. First, blood samples were obtained at various points of time shortly before or after surgery and potentially other therapy. However, leukocyte composition can be affected by surgery, chemo- and radiotherapy administration [43]. The percentage of neutrophils was higher, and the percentage of lymphocytes was lower shortly after surgery and chemotherapy or radiotherapy (Figure S1). We therefore adjusted for the timing of blood sampling relative to treatment in all Cox models. Survival curves stratified by the timing were performed. Significant and consistent associations of CD4+ T cell proportions with CRC prognosis were observed among all groups of patients except those whose blood sample was taken during, or within 5 months of the initiation of, chemo- or radiotherapy. These patterns suggest that one should avoid collecting blood samples within the course of chemo- or radiotherapy when using leukocyte estimates to predict CRC prognosis in research and clinical practice. Second, the absolute leukocyte counting information is missing in our study. Even though the DNAm-based estimation of leukocyte composition has been validated in various studies, comparison of the DNAm-based method with the traditional cell counting method is still needed in further studies. However, DNA methylation estimated cell proportions could help scientists to further understand the role of cell counts as prognostic biomarkers in CRC by using archived blood samples. Third, CRC is recognized as a heterogeneous disease, and its behavior differs between molecular subtypes [44], information on which was only available for a proportion of our patients and did not allow sufficiently powered analyses in our study. Further studies with comprehensive information on CRC molecular subtypes are needed to investigate the association across molecular subtypes. Lastly, the current high cost of DNAm microarray makes this technique difficult to be promoted in clinical settings. However, this study provides a potential tool to improve the prediction of CRC prognosis and to assist the physician's decision-making. It also provides a foundation for further study of inspiration. The DNAm-based markers may gain wider use if a low-price DNAm microarray technique is developed and commercialized.

## 5. Conclusions

Our findings suggest that blood DNAm-based estimates of leukocyte distribution, in particular the proportion of CD4+ T cells, have the potential to improve the accuracy of prognostic judgment for patients with early and intermediate CRC. However, validation by other studies, in particular studies with information on molecular subtypes of CRC and leukocyte count, is warranted to corroborate and extend the clinical importance of this observation.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/cancers13122948/s1>, Table S1: Baseline characteristics of the patient cohort according to tumor stages. Table S2: Median and interquartile range of leukocyte proportions by characteristics of the study population at baseline. Table S3: Associations of leukocyte composition with all-cause and colorectal cancer (CRC)-specific mortality in total study population (sensitivity analyses). Table S4: Associations of leukocyte composition with all-cause and colorectal cancer (CRC)-specific mortality among patients whose blood sample was taken prior to surgery, more than 1 month after surgery but not receiving chemo- or radiotherapy or more than 5 months after the first chemo- or radiotherapy. Figure S1: Leukocytes proportion according to the time of blood collection relative to treatments. Figure S2: Correlations of leukocyte subtype proportions. Figure S3: Survival curves for CD4+ T cell proportion by the timing of blood sampling. Figure S4: Adjusted restricted cubic splines for CRC patients across quartiles of CD4+ T cells proportion by therapies.

**Author Contributions:** Conceptualization, X.G. and H.B.; formal analysis, X.G. and Y.Z.; funding acquisition, J.C.-C. and M.H.; investigation, M.B.-H. and M.S.; methodology, X.G.; project administration, X.G.; supervision, H.B.; visualization, X.G.; writing—original draft, X.G.; writing—review and editing, Y.Z., X.L., L.J., E.A., J.C.-C., M.H. and H.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the German Research Council (BR 1704/6-1, BR 1704/6-3, BR 1704/6-4, CH 117/1-1, HO 5117/2-1, HE 5998/2-1, KL 2354/3-1, RO 2270/8-1 and BR 1704/17-1), the Interdisciplinary Research Program of the National Center for Tumor Diseases (NCT), Germany, and the German Federal Ministry of Education and Research (01KH0404, 01ER0814, 01ER0815, 01ER1505A and 01ER1505B). The sponsors had no role in the study design, in the collection, analysis and interpretation of data and preparation, review or approval of the manuscript.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committees of the Medical Faculty of the University of Heidelberg (ID: 310/2001) and the Medical Chambers of Baden-Württemberg (ID: M-198-02) and Rhineland-Palatinate (ID: 837.419.02 (3637)).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to ethical and data security requirements.

**Acknowledgments:** We thank the study participants and the interviewers who collected the data. We also thank the following hospitals and cooperating institutions that recruited patients for this study: Chirurgische Universitätsklinik Heidelberg, Klinik am Gesundbrunnen Heilbronn, St Vincentiuskrankenhaus Speyer, St Josefskrankenhaus Heidelberg, Chirurgische Universitätsklinik Mannheim, Diakonissenkrankenhaus Speyer, Krankenhaus Salem Heidelberg, Kreiskrankenhaus Schwetzingen, St Marienkrankenhaus Ludwigshafen, Klinikum Ludwigshafen, Stadtklinik Frankenthal, Diakoniekrankenhaus Mannheim, Kreiskrankenhaus Sinsheim, Klinikum am Plattenwald Bad Friedrichshall, Kreiskrankenhaus Weinheim, Kreiskrankenhaus Eberbach, Kreiskrankenhaus Buchen, Kreiskrankenhaus Mosbach, Enddarmzentrum Mannheim, Kreiskrankenhaus Brackenheim and Cancer Registry of Rhineland-Palatinate, Mainz.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; 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. [[CrossRef](#)]
2. Brenner, H.; Kloor, M.; Pox, C.P. Colorectal cancer. *Lancet* **2014**, *383*, 1490–1502. [[CrossRef](#)]

3. Brenner, H.; Chen, C. The colorectal cancer epidemic: Challenges and opportunities for primary, secondary and tertiary prevention. *Br. J. Cancer* **2018**, *119*, 785–792. [CrossRef] [PubMed]
4. Galon, J.; Mlecnik, B.; Bindea, G.; Angell, H.K.; Berger, A.; Lagorce, C.; Lugli, A.; Zlobec, I.; Hartmann, A.; Bifulco, C.; et al. Towards the introduction of the ‘Immunoscore’ in the classification of malignant tumours. *J. Pathol.* **2014**, *232*, 199–209. [CrossRef]
5. Gào, X.; Zhang, Y.; Boakye, D.; Li, X.; Chang-Claude, J.; Hoffmeister, M.; Brenner, H. Whole blood DNA methylation aging markers predict colorectal cancer survival: A prospective cohort study. *Clin. Epigenetics* **2020**, *12*, 184. [CrossRef]
6. Kraus, S.; Arber, N. Inflammation and colorectal cancer. *Curr. Opin. Pharmacol.* **2009**, *9*, 405–410. [CrossRef]
7. Terzić, J.; Grivennikov, S.; Karin, E.; Karin, M. Inflammation and colon cancer. *Gastroenterology* **2010**, *138*, 2101–2114. [CrossRef]
8. Lança, T.; Silva-Santos, B. The split nature of tumor-infiltrating leukocytes: Implications for cancer surveillance and immunotherapy. *Oncoimmunology* **2012**, *1*, 717–725. [CrossRef]
9. Zumsteg, A.; Christofori, G. Corrupt policemen: Inflammatory cells promote tumor angiogenesis. *Curr. Opin. Oncol.* **2009**, *21*, 60–70. [CrossRef]
10. Li, M.X.; Liu, X.M.; Zhang, X.F.; Zhang, J.F.; Wang, W.L.; Zhu, Y.; Dong, J.; Cheng, J.W.; Liu, Z.W.; Ma, L.; et al. Prognostic role of neutrophil-to-lymphocyte ratio in colorectal cancer: A systematic review and meta-analysis. *Int. J. Cancer* **2014**, *134*, 2403–2413. [CrossRef]
11. Ozdemir, Y.; Akin, M.L.; Sucullu, I.; Balta, A.Z.; Yucel, E. Pretreatment neutrophil/lymphocyte ratio as a prognostic aid in colorectal cancer. *Asian Pac. J. Cancer Prev.* **2014**, *15*, 2647–2650. [CrossRef]
12. Chen, Z.Y.; Raghav, K.; Lieu, C.H.; Jiang, Z.Q.; Eng, C.; Vauthey, J.N.; Chang, G.J.; Qiao, W.; Morris, J.; Hong, D.; et al. Cytokine profile and prognostic significance of high neutrophil-lymphocyte ratio in colorectal cancer. *Br. J. Cancer* **2015**, *112*, 1088–1097. [CrossRef]
13. Ikeguchi, M.; Ashida, K.; Saito, H. New Prognostic Indicator is Useful for Predicting the Survival of Patients with Unresectable Advanced Colorectal Cancer. *Hepatogastroenterology* **2015**, *62*, 859–862.
14. Dimitriou, N.; Felekouras, E.; Karavokyros, I.; Alexandrou, A.; Pikoulis, E.; Griniatsos, J. Neutrophils to lymphocytes ratio as a useful prognosticator for stage II colorectal cancer patients. *BMC Cancer* **2018**, *18*, 1202. [CrossRef]
15. Baron, U.; Türbachova, I.; Hellwag, A.; Eckhardt, F.; Berlin, K.; Hoffmuller, U.; Gardina, P.; Olek, S. DNA methylation analysis as a tool for cell typing. *Epigenetics* **2006**, *1*, 55–60. [CrossRef]
16. Houseman, E.A.; Accomando, W.P.; Koestler, D.C.; Christensen, B.C.; Marsit, C.J.; Nelson, H.H.; Wiencke, J.K.; Kelsey, K.T. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinform.* **2012**, *13*, 86. [CrossRef] [PubMed]
17. Heiss, J.A.; Breitling, L.P.; Lehne, B.; Kooner, J.S.; Chambers, J.C.; Brenner, H. Training a model for estimating leukocyte composition using whole-blood DNA methylation and cell counts as reference. *Epigenomics* **2017**, *9*, 13–20. [CrossRef] [PubMed]
18. Salas, L.A.; Koestler, D.C.; Butler, R.A.; Hansen, H.M.; Wiencke, J.K.; Kelsey, K.T.; Christensen, B.C. An optimized library for reference-based deconvolution of whole-blood biospecimens assayed using the Illumina HumanMethylationEPIC BeadArray. *Genome Biol.* **2018**, *19*, 64. [CrossRef]
19. Brenner, H.; Chang-Claude, J.; Seiler, C.M.; Rickert, A.; Hoffmeister, M. Protection from colorectal cancer after colonoscopy: A population-based, case-control study. *Ann. Intern. Med.* **2011**, *154*, 22–30. [CrossRef]
20. Brenner, H.; Chang-Claude, J.; Jansen, L.; Knebel, P.; Stock, C.; Hoffmeister, M. Reduced risk of colorectal cancer up to 10 years after screening, surveillance, or diagnostic colonoscopy. *Gastroenterology* **2014**, *146*, 709–717. [CrossRef]
21. Hoffmeister, M.; Jansen, L.; Rudolph, A.; Toth, C.; Kloor, M.; Roth, W.; Bläker, H.; Chang-Claude, J.; Brenner, H. Statin use and survival after colorectal cancer: The importance of comprehensive confounder adjustment. *J. Natl. Cancer Inst.* **2015**, *107*, djv045. [CrossRef]
22. Weigl, K.; Chang-Claude, J.; Knebel, P.; Hsu, L.; Hoffmeister, M.; Brenner, H. Strongly enhanced colorectal cancer risk stratification by combining family history and genetic risk score. *Clin. Epidemiol.* **2018**, *10*, 143–152. [CrossRef]
23. Lehne, B.; Drong, A.W.; Loh, M.; Zhang, W.; Scott, W.R.; Tan, S.T.; Afzal, U.; Scott, J.; Jarvelin, M.R.; Elliott, P.; et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol.* **2015**, *16*, 37. [CrossRef]
24. Salas, L.; Koestler, D. FlowSorted.Blood.EPIC: Illumina EPIC Data on Immunomagnetic Sorted Peripheral Adult Blood Cells. R Package Version 161. 2020. Available online: <https://github.com/immunomethylomics/FlowSorted.Blood.EPIC> (accessed on 1 April 2020).
25. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019. Available online: <http://www.r-project.org/index.html> (accessed on 1 April 2020).
26. Borst, J.; Ahrends, T.; Bąbała, N.; Melief, C.J.M.; Kastenmüller, W. CD4<sup>+</sup> T cell help in cancer immunology and immunotherapy. *Nat. Rev. Immunol.* **2018**, *18*, 635–647. [CrossRef]
27. Kuwahara, T.; Hazama, S.; Suzuki, N.; Yoshida, S.; Tomochika, S.; Nakagami, Y.; Matsui, H.; Shindo, Y.; Kanekiyo, S.; Tokumitsu, Y.; et al. Intratumoural-infiltrating CD4<sup>+</sup> and FOXP3<sup>+</sup> T cells as strong positive predictive markers for the prognosis of resectable colorectal cancer. *Br. J. Cancer* **2019**, *121*, 659–665. [CrossRef]
28. Toor, S.M.; Murshed, K.; Al-Dhaheri, M.; Khawar, M.; Abu Nada, M.; Elkord, E. Immune Checkpoints in Circulating and Tumor-Infiltrating CD4<sup>+</sup> T Cell Subsets in Colorectal Cancer Patients. *Front. Immunol.* **2019**, *10*, 2936. [CrossRef]

29. Chaudhary, B.; Elkord, E. Regulatory T Cells in the Tumor Microenvironment and Cancer Progression: Role and Therapeutic Targeting. *Vaccines* **2016**, *4*, 28. [[CrossRef](#)]
30. Betts, G.; Jones, E.; Junaid, S.; El-Shanawany, T.; Scurr, M.; Mizen, P.; Kumar, M.; Jones, S.; Rees, B.; Williams, G.; et al. Suppression of tumour-specific CD4<sup>+</sup> T cells by regulatory T cells is associated with progression of human colorectal cancer. *Gut* **2012**, *61*, 1163–1171. [[CrossRef](#)]
31. Reissfelder, C.; Stamova, S.; Gossmann, C.; Braun, M.; Bonertz, A.; Walliczek, U.; Grimm, M.; Rahbari, N.N.; Koch, M.; Saadati, M.; et al. Tumor-specific cytotoxic T lymphocyte activity determines colorectal cancer patient prognosis. *J. Clin. Investig.* **2015**, *125*, 739–751. [[CrossRef](#)]
32. Glaire, M.A.; Domingo, E.; Sveen, A.; Bruun, J.; Nesbakken, A.; Nicholson, G.; Novelli, M.; Lawson, K.; Oukrif, D.; Kildal, W.; et al. Tumour-infiltrating CD8<sup>+</sup> lymphocytes and colorectal cancer recurrence by tumour and nodal stage. *Br. J. Cancer* **2019**, *121*, 474–482. [[CrossRef](#)] [[PubMed](#)]
33. Yuen, G.J.; Demissie, E.; Pillai, S. B lymphocytes and cancer: A love-hate relationship. *Trends Cancer* **2016**, *2*, 747–757. [[CrossRef](#)]
34. Berntsson, J.; Svensson, M.C.; Leandersson, K.; Nodin, B.; Micke, P.; Larsson, A.H.; Eberhard, J.; Jirstrom, K. The clinical impact of tumour-infiltrating lymphocytes in colorectal cancer differs by anatomical subsite: A cohort study. *Int. J. Cancer* **2017**, *141*, 1654–1666. [[CrossRef](#)]
35. Edin, S.; Kaprio, T.; Hagström, J.; Larsson, P.; Mustonen, H.; Böckelman, C.; Strigård, K.; Gunnarsson, U.; Haglund, C.; Palmqvist, R. The Prognostic Importance of CD20<sup>+</sup> B lymphocytes in Colorectal Cancer and the Relation to Other Immune Cell subsets. *Sci. Rep.* **2019**, *9*, 19997. [[CrossRef](#)]
36. Tang, Y.P.; Xie, M.Z.; Li, K.Z.; Li, J.L.; Cai, Z.M.; Hu, B.L. Prognostic value of peripheral blood natural killer cells in colorectal cancer. *BMC Gastroenterol.* **2020**, *20*, 31. [[CrossRef](#)]
37. Wen, S.; Chen, N.; Peng, J.; Ling, W.; Fang, Q.; Yin, S.F.; He, X.; Qiu, M.; Hu, Y. Peripheral monocyte counts predict the clinical outcome for patients with colorectal cancer: A systematic review and meta-analysis. *Eur. J. Gastroenterol. Hepatol.* **2019**, *31*, 1313–1321. [[CrossRef](#)]
38. Grecian, R.; Whyte, M.K.B.; Walmsley, S.R. The role of neutrophils in cancer. *Br. Med. Bull.* **2018**, *128*, 5–14. [[CrossRef](#)]
39. Mizuno, R.; Kawada, K.; Itatani, Y.; Ogawa, R.; Kiyasu, Y.; Sakai, Y. The Role of Tumor-Associated Neutrophils in Colorectal Cancer. *Int. J. Mol. Sci.* **2019**, *20*, 529. [[CrossRef](#)] [[PubMed](#)]
40. Rao, H.L.; Chen, J.W.; Li, M.; Xiao, Y.B.; Fu, J.; Zeng, Y.X.; Cai, M.Y.; Xie, D. Increased intratumoral neutrophil in colorectal carcinomas correlates closely with malignant phenotype and predicts patients' adverse prognosis. *PLoS ONE* **2012**, *7*, e30806. [[CrossRef](#)]
41. Wikberg, M.L.; Ling, A.; Li, X.; Öberg, Å.; Edin, S.; Palmqvist, R. Neutrophil infiltration is a favorable prognostic factor in early stages of colon cancer. *Hum. Pathol.* **2017**, *68*, 193–202. [[CrossRef](#)]
42. Dugué, P.A.; Bassett, J.K.; Joo, J.E.; Jung, C.H.; Ming Wong, E.; Moreno-Betancur, M.; Schmidt, D.; Makalic, E.; Li, S.; Severi, G.; et al. DNA methylation-based biological aging and cancer risk and survival: Pooled analysis of seven prospective studies. *Int. J. Cancer* **2018**, *142*, 1611–1619. [[CrossRef](#)]
43. van Meir, H.; Nout, R.A.; Welters, M.J.; Loof, N.M.; de Kam, M.L.; van Ham, J.J.; Samuels, S.; Kenter, G.G.; Cohen, A.F.; Melief, C.J.; et al. Impact of (chemo)radiotherapy on immune cell composition and function in cervical cancer patients. *Oncoimmunology* **2017**, *6*, e1267095. [[CrossRef](#)] [[PubMed](#)]
44. Wang, W.; Kandimalla, R.; Huang, H.; Zhu, L.; Li, Y.; Gao, F.; Goel, A.; Wang, X. Molecular subtyping of colorectal cancer: Recent progress, new challenges and emerging opportunities. *Semin. Cancer Biol.* **2019**, *55*, 37–52. [[CrossRef](#)] [[PubMed](#)]