Association between pre-diagnostic leukocyte mitochondrial DNA copy number and survival among colorectal cancer patients

Keming Yang[1,2], Michele R. Forman[3], Brett H. Graham[4], Patrick O. Monahan[5], Edward L. Giovannucci[1,6,7], Immaculata De Vivo[6,7], Andrew T. Chan[6,8*], Hongmei Nan[2,9,10*]

Author affiliations:
[1] Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA;
[2] Department of Epidemiology, Richard M. Fairbanks School of Public Health, Indiana University, Indianapolis, IN, USA;
[3] Department of Nutrition Science, College of Health and Human Science, Purdue Center for Cancer Research, Purdue University, West Lafayette, IN, USA;
[4] Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA;
[5] Department of Biostatistics, School of Medicine and Richard M. Fairbanks School of Public Health, Indiana University, Indianapolis, IN, USA;
[6] Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA;
[7] Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA;
[8] Clinical and Translational Epidemiology Unit and Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA;
[9] Department of Global Health, Richard M. Fairbanks School of Public Health, Indiana University, Indianapolis, IN, USA;
[10] IU Melvin and Bren Simon Cancer Center, Indiana University, Indianapolis, IN, USA

* Co-senior/corresponding authors:

Dr. Hongmei Nan: Department of Global Health, Richard M. Fairbanks School of Public Health, Indiana University, 1050 Wishard Blvd, RG 6086, IN 46202, USA
Telephone: 317-274-3120   E-mail: hnan@iu.edu

Dr. Andrew T. Chan: Clinical and Translational Epidemiology Unit and Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, 100 Cambridge Street, 15th floor, Boston, MA 02114, USA
Telephone: 617-726-2862   E-mail: ACHAN@mgh.harvard.edu

Word count: 3460

Running head: MtDNA Copy Number and Colorectal Cancer Survival

_____

ABSTRACT

Background: Mitochondrial DNA copy number (mtDNAcn) is considered a biomarker for mitochondrial function and oxidative stress. Although previous studies have suggested a potential relationship between mtDNAcn at the time of colorectal cancer (CRC) diagnosis and CRC prognosis, findings have been inconsistent, and no study has specifically investigated the association of pre-diagnostic mtDNAcn with CRC survival.

Methods: We examined the association of pre-diagnostic leukocyte mtDNAcn (measured by qPCR) with overall and CRC-specific survival among 587 patients in Nurses' Health Study and Health Professionals Follow-Up Study. Cox models were constructed to estimate hazard ratios (HRs) and 95% confidence intervals (95% CIs).

Results: During a mean follow-up of 10.5 years, 395 deaths were identified; 180 were due to CRC. Overall, we did not observe significant associations between mtDNAcn and either overall or CRC-specific survival among all cases or by cancer location, grade, or stage. In an exploratory stratified analysis, a suggestive inverse association of mtDNAcn and overall death risk appeared among current smokers [HR (95% CI) for 1 SD decrease in mtDNAcn = 1.50 (0.98, 2.32), P-trend=0.06]. Reduced mtDNAcn and lower CRC-specific death risk was observed among patients aged ≤ 70.5 at diagnosis [HR (95% CI) for 1 SD decrease of mtDNAcn = 0.71 (0.52, 0.97), P-trend = 0.03], ≤ 5 years from blood collection to diagnosis [HR (95% CI) for 1 SD decrease in mtDNAcn = 0.65 (0.44, 0.96), P-trend = 0.03] and those consuming a low-inflammatory diet [HR (95% CI) for 1 SD decrease in mtDNAcn = 0.61 (0.42, 0.88), P-trend = 0.009].

Conclusion: no significant associations between pre-diagnostic leukocyte mtDNAcn and either overall or CRC-specific survival appeared but exploratory analysis identified potential sub-group associations.

Keywords: mitochondrial DNA copy number; mitochondrial DNA; mitochondria; leukocyte; colorectal cancer; survival

Highlights

- Mitochondrial DNA copy number (MtDNAcn) is a biomarker for mitochondrial function/oxidative stress.

- No study has examined the association of pre-diagnostic mtDNAcn with colorectal cancer (CRC) survival.

- Overall, we found leukocyte mtDNAcn may not be a significant biomarker for survival among CRC patients.

- Subgroup analyses suggested an inverse association of mtDNAcn and overall death risk among current smokers.

- MtDNAcn may have a prognostic value for CRC patients consuming a low-inflammatory diet, diagnosed ≤ 70.5 or ≤ 5 years from mtDNAcn measurement.

INTRODUCTION

Colorectal cancer (CRC) is the fourth most common cancer and the second leading cause of cancer deaths in the US.[1] The 5-year survival rate for CRC patients is 64.4%;[2] according to the American Joint Committee on Cancer (AJCC) staging system,[3] the survival rate is much higher for patients with stage I (91%) or II (82%) cancers and much lower for patients with stage IV (12%) cancers.[4] Exploring molecular markers that might be of predictive value for CRC progression would benefit the development of new prevention strategies as well as the identification of novel diagnostic and therapeutic targets.

Mitochondria are cytoplasmic organelles in eukaryotic cells and contribute to many essential functions in cellular metabolism, including cellular energy production and regulating apoptosis.[5] Mitochondria contain their own DNA, i.e., mitochondrial DNA (mtDNA), which is a double-stranded, circular DNA molecule.[5] The loop structure of mtDNA can be divided into two chains: an outer heavy chain (H) and an inner light chain (L). MtDNA contains 37 genes and a noncoding region named the displacement loop (D-loop).[5] Usually, each cell contains hundreds to thousands of copies of mtDNA, varying across different tissues.[6] Also many types of cancer cells appear to have a lower mtDNA copy number (mtDNAcn) than cells in normal tissues.[7] In addition, many population studies have utilized mtDNAcn in peripheral blood leukocytes as a surrogate marker for mitochondrial function and intracellular oxidative stress.[8-12] Defects in mitochondrial function and genome maintenance (including changes in mtDNAcn) have been associated with many mitochondrial disorders, carcinogenesis, as well as aging and both cognitive and overall health among the elderly.[7, 13-15]

Previously, several nested case-control studies have reported an association between pre-diagnostic leukocyte mtDNAcn and CRC risk,[8, 11, 12] including one from our group in which we

found an inverse association between pre-diagnostic leukocyte mtDNAcn and CRC risk.[12] In terms of CRC prognosis, a few clinical and epidemiologic studies have examined the relationship between mtDNAcn at CRC diagnosis and survival among CRC patients, with inconsistent findings.[10, 16-20] For example, one study of 598 CRC patients reported that high mtDNAcn in leukocytes at diagnosis was associated with worse survival after CRC.[10] In contrast, evidence suggesting an association between higher mtDNAcn in cancer tissue and better survival was found by one hospital-based study (n= 60 CRC patients) and the population-based Netherlands Cohort Study which included 655 CRC cases.[16, 19] Mixed evidence including positive, inverse and null associations have also been reported for mtDNAcn and prognosis of many other cancers.[21-24]

To note, it is not possible to determine if differences of mtDNAcn measured at/after diagnosis is a cause or consequence of the disease progression. Research has suggested that the loss of mitochondria may occur prior to the development of dysplasia and CRC in ulcerative colitis patients;[25] it is plausible that changes in mtDNAcn may modulate tumor behaviors and characteristics that relate to survival prior to CRC diagnosis. Yet no study has investigated whether pre-diagnostic leukocyte mtDNAcn can be a prognostic marker for CRC patients. Therefore, we examined the association between pre-diagnostic leukocyte mtDNAcn and overall as well as CRC-specific survival among 587 CRC cases, including 321 female subjects from the Nurses' Health Study (NHS) and 266 male subjects from the Health Professionals Follow-Up Study (HPFS).

METHODS

Study Population

Colorectal cancer cases in the current study were identified from the NHS and the HPFS, two US-based long-term prospective cohort studies.[26, 27] Begun in 1976, the NHS enrolled 121,700 female registered nurses aged 30-55 years at baseline. In 1986, the HPFS enrolled 51,529 male health professionals aged 40-75 years at baseline. During 1989-1990, a total of 32,826 NHS participants provided blood samples, and during 1993-1995, 18,225 HPFS participants provided blood samples. In both cohorts, every two years, participants complete follow-up questionnaires and report updated information about their lifestyles, disease status, and usage of medications and supplements.[26] Every 2 to 4 years, validated semi-quantitative food frequency questionnaires (FFQ) are mailed to participants.[28, 29]

In the current analysis, we included CRC cases from previous case-control studies nested within the NHS and the HPFS.[12] All CRC cases were incident cases diagnosed after blood collection. A total of 587 patients with available mtDNAcn data and complete survival information were included in the analysis, including 321 female patients from the NHS and 266 male patients from the HPFS. The study protocol was approved by the institutional review boards of the Brigham and Women's Hospital and the Harvard T.H. Chan School of Public Health, and those of participating registries as required.

Case Diagnosis and Death Confirmation

Incident CRC cases were identified by self-report on biennial follow-up questionnaires, then confirmed by physicians blinded to exposure data through reviewing medical records and pathology reports, or cancer registry data. Deaths were ascertained from next of kin or by notification from the Post Office when a questionnaire or newsletter mailed to a cohort member was returned. Further, the National Death Index was searched for any unreported cancer deaths; overall, this approach captures >98% of deaths.[30] Causes of death were confirmed via contacting

next-of-kin, as well as data from medical records, death certificates, tumor registries, and the National Death Index. Cohort investigators reviewed all information to determine the primary cause of death. In the current study of 587 CRC patients, 395 deaths occurred during follow-up, of which 180 were due to CRC.

Measurement of Leukocyte MtDNAcn

The measurement and validation of mtDNAcn in NHS and HPFS have been described in detail previously,[31] and shown in the **Supplementary Methods**. In summary, genomic DNA was extracted from the buffy-coat fractions of leukocytes in peripheral blood by the QIAmp (Qiagen) 96-spin blood protocol, and DNA concentration was measured through pico-green quantitation using a 96-well spectrophotometer. Relative mtDNAcn (the N/S ratio, or the ratio of mitochondrial *ND2* gene copy number to the number of the genomic single-copy gene *AluYb8*) was determined by a qPCR-based assay using the high-throughout 384-well format (Applied Biosystems 7900HT Real Time PCR). To adjust for inter-assay variability, a 10ng DNA standard curve point was included on every 384-well plate as a calibrator DNA. The relative mtDNAcn for each subject was calculated by dividing the exponentiated N/S ratio of each sample by the exponentiated N/S ratio of the calibrator DNA. Log-transformation of the relative mtDNAcn value was further performed to normalize its distribution.

Data Collection of Covariates

In this study, covariates adjusted in regression models included demographic information (age at blood collection and diagnosis, race, sex), tumor characteristics (location, grade of differentiation, AJCC TNM stage), CRC family history, and lifestyle habits (smoking status, body mass index [BMI], total physical activity [metabolic equivalent (MET)-hours/week], regular aspirin and non-aspirin NSAID use, and empirical dietary inflammatory pattern/EDIP

score). Tumor information was coded during the medical record and pathology report review. Demographic and lifestyle factors were collected by self-administered questionnaires during follow-up.[28, 29] The EDIP score is a weighted sum of 18 food groups that are most predictive of three plasma inflammatory biomarkers: interleukin-6 (IL-6), C-reactive protein (CRP), and tumor necrosis factor alpha - receptor 2 (TNFα-R2). Lower EDIP (more negative) scores indicate anti-inflammatory diets, and higher EDIP (more positive) scores indicate pro-inflammatory diets. Higher EDIP scores were significantly associated with increased risk of CRC in both NHS and HPFS.[32] The development and validation of the EDIP score has been previously documented.[33, 34] Covariate information for our analyses was derived from the pre-diagnostic questionnaire cycle closest to blood collection (NHS: 1989-1990, HPFS: 1993-1995).

Statistical Analysis

We tested associations between pre-diagnostic leukocyte mtDNAcn and both overall as well as CRC-specific survival among CRC patients. Overall survival represents the risk of death from all causes. CRC-specific survival represents the risk of death from CRC. Survival time was calculated from CRC diagnosis to death or last follow-up, whichever came first.

Cox proportional hazards regression models were used to estimate the Hazard Ratios (HRs) and 95% confidence interval (CIs) for overall and CRC-specific mortality by mtDNAcn quartile and then computed for overall and CRC-specific mortality by one standard deviation (SD) decrease in mtDNAcn. The proportional hazards assumption was evaluated (and determined to be satisfied) by evaluating a time-dependent variable in cox regression; the time-dependent variable was the cross-product of quartile-specific median values of mtDNAcn (continuous) and the time-to-event (in combined data: $P = 0.11$ for overall survival, $P = 0.16$ for CRC-specific survival). Trend tests were conducted by treating the quartile-specific medians of

mtDNAcn as a continuous variable in the cox models; P values from the Wald test were used to evaluate the significance of the trend. The primary analysis was performed in NHS and HPFS separately, and the heterogeneity in the association between pre-diagnostic mtDNAcn and risk of death from both all causes and CRC between the two cohorts was assessed by the Cochran's Q statistic[35]. Without evidence of between-cohort heterogeneity, we further pooled data from NHS and HPFS to enhance statistical power, and examined the associations of interest in the pooled dataset.

In the primary analysis (Table 2), we built three models: 1). Model 1: adjusted for age at blood collection and age at diagnosis (continuous), race (white, non-white), and sex/cohort (female/male), 2). Model 2: Model 1 + tumor location (proximal colon, distal colon, rectum, unspecified), grade of differentiation (well, moderate, poor, unspecified), and AJCC TNM stage (0/I, II, III, IV, unspecified), and 3). Model 3: Model 2 + smoking status (never, past, current), BMI (in tertiles), physical activity (in tertiles), CRC family history (yes/no), regular aspirin use (yes/no), regular non-aspirin NSAID use (yes/no), and empirical dietary inflammatory pattern (EDIP) score (in tertiles).

We also examined the associations between pre-diagnostic mtDNAcn and either overall or CRC-specific survival according to the location, grade of differentiation, and stage of cancer. Exploratory stratified analyses were performed by selected covariates, including age at blood collection and diagnosis, years from blood collection to diagnosis, smoking status, BMI, physical activity, dietary inflammatory potential characterized by EDIP score, and regular aspirin use in all cases. Potential effect modifications of these stratified covariates on the association of mtDNAcn with death risk were assessed by including a cross-product term of mtDNAcn (quartiles) and each stratified variable in the cox regressions; significance of the interaction term

was evaluated by likelihood ratio tests. All statistical analyses were conducted using SAS statistical software (version 9.4, SAS Institute, Cary, NC). A two-sided P value < 0.05 was considered statistically significant.

RESULTS

Baseline characteristics of CRC patients are presented in **Table 1**. Mean age at blood collection was 62.0 years; male patients were on average older than female patients (mean age 65.8 vs. 58.8 years) at blood collection. Mean age at CRC diagnosis was 70.2 years, and the vast majority (95.6%) of patients were white. Mean years from blood collection to CRC diagnosis was 8.2 years. Mean time from CRC diagnosis to death or end of follow-up was 10.5 years. Approximately 40% of all CRC cases were in the proximal colon, 30% in the distal colon, and 24% in the rectum. In terms of cancer differentiation grade, 12% of the cases were well differentiated, 57% were moderately differentiated, and 15% were poorly differentiated. Regarding TNM stages, 27% of cancers were at stage 0/I, 23% were at stage II%, 22% were at stage III, and 14% were at stage IV, which were accompanied by distant metastasis.

We examined the association between pre-diagnostic leukocyte mtDNAcn and risk of death among CRC patients **(Table 2)**. Overall, we observed no significant association between mtDNAcn and either overall or CRC-specific survival in NHS and HPFS separately, or in the pooled dataset combining all cases. Among all 587 cases, in the multivariable-adjusted model 3, the HR (95% CI) associated with 1 SD decrease of mtDNAcn was 1.05 (0.92, 1.19) for overall survival and 0.92 (0.76, 1.11) for CRC-specific survival. No statistical heterogeneity was detected between results from NHS and HPFS (P for heterogeneity > 0.05). We further studied the association of mtDNAcn with overall and CRC-specific survival among cancers by location,

grade of differentiation, and TNM stage and observed the non-significant status for the associations between mtDNAcn and risk of death did not change **(Table 3)**.

We did not observe significant interactions between pre-selected covariates and mtDNAcn on overall or CRC-specific survival, including age at blood collection and diagnosis, years from blood collection to diagnosis, smoking status, BMI, physical activity, empirical dietary inflammatory pattern, and regular aspirin use **(Tables 4-5)**. In the exploratory analysis stratified by selected covariates, we observed a suggestive inverse association between mtDNAcn and overall survival among current smokers [Per 1 SD decrease of mtDNAcn, HR (95% CI) = 1.50 (0.98, 2.32), P for trend across mtDNAcn quartiles = 0.06] **(Table 4)**. Interestingly, the risk of CRC-specific death seemed to decrease with lower mtDNAcn among subgroups, including those aged ≤ 70.5 years (median age) at diagnosis [Per 1 SD decrease of mtDNAcn, HR (95% CI) =0.71 (0.52, 0.97), P for trend across mtDNAcn quartiles = 0.03], patients diagnosed within 5 years since blood collection[Per 1 SD decrease of mtDNAcn, HR (95% CI) = 0.65 (0.44, 0.96), P for trend across mtDNAcn quartiles = 0.03], and participants with diets of low inflammatory potential (i.e. lower EDIP score) [Per 1 SD decrease of mtDNAcn, HR (95% CI) = 0.61 (0.42, 0.88), P for trend across mtDNAcn quartiles = 0.009] **(Table 5)**.

DISCUSSION

In an analysis of 587 CRC patients from NHS and HPFS, no significant association appeared between pre-diagnostic leukocyte mtDNAcn and either overall or CRC-specific survival. To the best of our knowledge, this is the first examination of pre-diagnostic mtDNAcn, which is uninfluenced by disease status, in relation to CRC survival.

Our null associations are consistent with a few previous studies that also found no significant association between mtDNAcn in cancer tissues and CRC survival.[18, 20] In contrast, significant associations between mtDNAcn at diagnosis and CRC prognosis were reported by several other studies, though the direction of the association varied.[10, 16, 17, 19] For example, high leukocyte mtDNAcn at diagnosis was associated with poorer CRC-specific survival in 598 CRC patients [high vs. low mtDNAcn, HR (95%CI) = 1.96 (1.49, 2.59)].[10] Also, among patients with stage II/III cancers, adjuvant chemotherapy was of significant benefit to those with high, but not low mtDNAcn.[10] Conversely, among 655 CRC patients, van Osch et al. identified lower risk of death from CRC among patients with high mtDNAcn in cancer tissue [5th vs. 1st quintile, HR (95%CI) = 0.63 (0.40, 1.01)].[16] Evidence suggesting associations between lower mtDNAcn in CRC tissue and poorer survival,[19] as well as between lower mtDNAcn in peripheral blood and higher all-cause mortality among older adults in general, have also been reported.[15, 36]

Changes in mtDNAcn in tumor tissues have been correlated with clinical-pathological features of CRC. Compared to adjacent non-cancerous or normal colorectal tissue, reduced mtDNAcn in CRC tissue have been detected in many studies.[16, 19, 37] Interestingly, Feng et al. found higher mtDNAcn in CRC tissue than adjacent normal tissue only in stage I/II not in stage III/IV cancers.[38] In addition, Feng et al. reported that mtDNAcn was higher in tissues from stage I/II cancers (n=26) than in tissues from stage III/IV cancers (n=18).[38] Similarly, Lin et al. found that mtDNAcn in CRC tissue decreased with advancing stage, that is, mtDNAcn was highest in Stage I (n=22) cancer and lowest in Stage IV (n=32).[39] In our previous study of mtDNAcn and CRC risk, pre-diagnostic leukocyte mtDNAcn was lower in CRC cases than controls.[12] In our current study, mean (SD) of log-mtDNAcn in leukocytes was also lower in stage IV cancers [-

0.16 (0.35)] than in earlier-stage cancers [stage III: -0.14 (0.32), stage 0/I/II: -0.13 (0.33)]. These lines of evidence suggest that mtDNAcn may decrease during the progression of CRC.

MtDNA is susceptible to excessive oxidative damage, which may be due in part simply to its location in the mitochondria, which is in close proximity to the generation of reactive oxygen species (ROS), an intrinsic source of DNA damage.[40] Moreover, in the context of carcinogenesis, mtDNA may be a crucial target of mitochondrial ROS, and altered mtDNAcn may affect cancer growth and/or metastatic behavior.[41] Selective degradation of injured mtDNA copies is among the natural mechanisms to counteract mtDNA damage[42]. Research suggests that when mtDNA encounters oxidative damage, its copy number may initially increase in response; however, persistent damage would eventually lead to targeted mtDNA degradation and the loss of mtDNAcn.[40, 42]

In addition, both cross-sectional and longitudinal data show that mtDNAcn in peripheral blood declines during aging.[15, 36, 43, 44] Interestingly, similar age-related changes in mtDNAcn were observed in two separate studies: mtDNAcn remains steady[15] or slightly increased[43] until the fifth decade of life and significantly decreases thereafter. Lifestyle factors may also partially influence changes in leukocyte mtDNAcn. Our previous studies have shown that several lifestyle factors, including weight, waist, BMI, waist-hip ratio, trajectories of weight gain, smoking, less fruit consumption, and hyperinsulinemic lifestyle potential (based on diet, physical activity, body weight), were associated with reduced mtDNAcn in peripheral blood leukocytes.[31, 45-47]

The damage tolerance mechanism of mtDNA may help explain the above-described epidemiologic findings regarding variations with age and cancer stage. Mutated or injured copies of mtDNA and dysfunctional mitochondria may accumulate with age and the progression of cancer. In our stratified analyses, risk of CRC-specific death associated with one SD decrease of

mtDNAcn was lower among those patients who are younger than 70.5 years (median) at diagnosis [HR (95% CI) = 0.71 (0.52, 0.97)], patients diagnosed within 5 years since blood collection [HR (95% CI) = 0.65 (0.44, 0.96)], and patients who had lower dietary inflammatory potential [HR (95% CI) = 0.61 (0.42, 0.88)], although it was not statistically significant after correcting for multiple testing **(Table 5)**. Interestingly, among current smokers, a one SD decrease in mtDNAcn seemed to be associated with elevated risk of overall death [HR (95%CI) = 1.50 (0.98, 2.32)] **(Table 4)**. One of the potential mechanisms underlying such an inverse association is that poor lifestyle habits (e.g., smoking) may accelerate the exhaustion of mtDNA repair and tolerance mechanisms, lead to reduced mtDNAcn (reflecting excessive oxidative damage) which is further related to higher risk of mortality. However, this interpretation is strictly speculative, and we cannot exclude the possibility that results from these stratified subgroups could simply be the product of chance.

Interpreting the association between pre-diagnostic mtDNAcn and cancer survival could be even more complicated, because complex changes in mtDNAcn may occur prior to the development of cancer as well as during disease progression. Basic research has shown that excessive mitochondrial oxidative stress may induce cancer initiation as well as death of both normal and malignant cells.[41] Previously, one retrospective case-control study reported a positive association between leukocyte mtDNAcn measured at the time of diagnosis and risk of CRC.[48] However, prospective case-control studies nested within the NHS and the Shanghai Women's Health Study found inverse associations of pre-diagnostic leukocyte mtDNAcn with risk of CRC, while a U-shaped association was observed in the Singapore Chinese Health Study.[8, 11, 12] The leukocyte mtDNAcn was significantly inversely associated with the risk of CRC in our

previous NHS study,[12] but was not associated with CRC survival in our current study, suggesting that pre-diagnostic leukocyte mtDNAcn may be a better marker for CRC risk than for prognosis.

As the first exploration of a possible relationship between pre-diagnostic mtDNAcn and survival in CRC patients, our study has several strengths, including long-term follow-up, pre-diagnostic assessment of exposure, as well as a comprehensive list of covariates. We also acknowledge some limitations, one of which is the single measurement of mtDNAcn, which is insufficient to reflect all aspects of functional alterations in mitochondria. Research has demonstrated that in addition to changes in mtDNAcn, other mtDNA alterations, such as common mtDNA genetic variants[49], mutations in the mitochondrial D-Loop,[50, 51] and mtDNA microsatellite instability,[52] are also associated with CRC risk and prognosis. In addition, increased migration of mtDNA into the nuclear genome was observed in colorectal adenocarcinomas in one study.[53] Moreover, different mtDNA alterations may be interrelated. For example, demethylation of the D-loop region was associated with elevated mtDNAcn in CRC tissues,[54] and the mtDNA mutation burden may lead to mitochondrial dysfunction.[41] Therefore, joint assessment of multiple markers of mitochondrial function (e.g., combining them into indices) in relation to the risk and survival of CRC may be more sensitive and precise than individual markers alone. Another potential limitation of our study is the lack of information on the molecular characteristics of tumors, such as the status of *KRAS, BRAF,* or *PIK3CA* mutations, though clinical-pathological information including tumor location, grade of differentiation, and AJCC TNM stages were included. Lastly, future correlation analyses of mtDNAcn in different tissues (e.g. leukocytes vs. tumor tissues) would help us better understanding the relationship of mitochondrial biology and carcinogenesis. And repeated

measurement of mtDNAcn over a longer period would further enable us to study longitudinal changes in mtDNAcn and their relationship with cancer outcomes.

In conclusion, our study did not reveal significant associations between pre-diagnostic leukocyte mtDNAcn and overall and CRC-specific survival among CRC patients. We did observe some evidence suggestive of an inverse association between mtDNAcn and overall survival among current smokers, and a positive association between mtDNAcn and CRC-specific survival among the relatively younger patients in our cohort (≤ 70.5 years at diagnosis), patients diagnosed within 5 years from blood collection, as well as patients with anti-inflammatory diets. Additional population studies are warranted, especially among patients of a more diverse age group (e.g. young adults with early-onset CRC). Additional functional studies are also needed to explore the mechanisms underlying the behavior of mtDNAcn in the context of tumor development and progression.

Authorship contribution: HN and AC were responsible for the study concept and design. HN, AC, ID were involved in mtDNAcn measurement and validation process. KY analyzed the data and drafted the manuscript. All authors (KY, MF, BG, PM, EG, ID, AC, HN) contributed to the interpretation of the results and revision of the manuscript critically for important intellectual content. All authors, external and internal, had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis.

Conflict of interest: None declared.

Data availability statement: The data of this study are available upon reasonable request. Further information including the procedures to obtain and access data from the Nurses' Health

Studies and Health Professionals Follow-up Study is described at
https://www.nurseshealthstudy.org/researchers (contact email:
nhsaccess@channing.harvard.edu) and https://sites.sph.harvard.edu/hpfs/for-collaborators/.

Supplementary file: Supplementary methods - MtDNAcn ascertainment, validation, and quality
control procedure

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA: a cancer journal for clinicians*. 2020;70:7-30
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA: a cancer journal for clinicians*. 2019;69:7-34
3. Edge sb, byrd dr, compton cc, fritz ag, greene fl, trotti a, eds. American joint committee on cancer cancer staging manual. 7th ed. New york: Springer; 2010.
4. Miller KD, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, et al. Cancer treatment and survivorship statistics, 2019. *CA: a cancer journal for clinicians*. 2019;69:363-385
5. Brandon MC, Lott MT, Nguyen KC, Spolim S, Navathe SB, Baldi P, et al. Mitomap: A human mitochondrial genome database--2004 update. *Nucleic acids research*. 2005;33:D611-613
6. Veltri KL, Espiritu M, Singh G. Distinct genomic copy number in mitochondria of different mammalian organs. *J Cell Physiol*. 1990;143:160-164
7. Reznik E, Miller ML, Senbabaoglu Y, Riaz N, Sarungbam J, Tickoo SK, et al. Mitochondrial DNA copy number variation across human cancers. *Elife*. 2016;5
8. Huang B, Gao YT, Shu XO, Wen W, Yang G, Li G, et al. Association of leukocyte mitochondrial DNA copy number with colorectal cancer risk: Results from the shanghai women's health study. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2014;23:2357-2365
9. Liu CS, Tsai CS, Kuo CL, Chen HW, Lii CK, Ma YS, et al. Oxidative stress-related alteration of the copy number of mitochondrial DNA in human leukocytes. *Free Radic Res*. 2003;37:1307-1317
10. Qu F, Chen Y, Wang X, He X, Ren T, Huang Q, et al. Leukocyte mitochondrial DNA content: A novel biomarker associated with prognosis and therapeutic outcome in colorectal cancer. *Carcinogenesis*. 2015;36:543-552
11. Thyagarajan B, Wang R, Barcelo H, Koh WP, Yuan JM. Mitochondrial copy number is associated with colorectal cancer risk. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2012;21:1574-1581
12. Yang K, Li X, Forman MR, Monahan PO, Graham BH, Joshi A, et al. Pre-diagnostic leukocyte mitochondrial DNA copy number and colorectal cancer risk. *Carcinogenesis*. 2019;40:1462-1468
13. Campisi J, Kapahi P, Lithgow GJ, Melov S, Newman JC, Verdin E. From discoveries in ageing research to therapeutics for healthy ageing. *Nature*. 2019;571:183-192

14. Parr RL, Dakubo GD, Thayer RE, McKenney K, Birch-Machin MA. Mitochondrial DNA as a potential tool for early cancer detection. *Hum Genomics*. 2006;2:252-257
15. Mengel-From J, Thinggaard M, Dalgard C, Kyvik KO, Christensen K, Christiansen L. Mitochondrial DNA copy number in peripheral blood cells declines with age and is associated with general health among elderly. *Human genetics*. 2014;133:1149-1159
16. van Osch FH, Voets AM, Schouten LJ, Gottschalk RW, Simons CC, van Engeland M, et al. Mitochondrial DNA copy number in colorectal cancer: Between tissue comparisons, clinicopathological characteristics and survival. *Carcinogenesis*. 2015;36:1502-1510
17. Wang Y, He S, Zhu X, Qiao W, Zhang J. High copy number of mitochondrial DNA predicts poor prognosis in patients with advanced stage colon cancer. *The International journal of biological markers*. 2016;31:e382-e388
18. Lee H, Cho JH, Park WJ, Jung SJ, Choi IJ, Lee JH. Loss of the association between telomere length and mitochondrial DNA copy number contribute to colorectal carcinogenesis. *Pathology oncology research : POR*. 2018;24:323-328
19. Cui H, Huang P, Wang Z, Zhang Y, Zhang Z, Xu W, et al. Association of decreased mitochondrial DNA content with the progression of colorectal cancer. *BMC cancer*. 2013;13:110
20. Mohideen AM, Dicks E, Parfrey P, Green R, Savas S. Mitochondrial DNA polymorphisms, its copy number change and outcome in colorectal cancer. *BMC research notes*. 2015;8:272
21. Xu J, Chang WS, Tsai CW, Bau DT, Davis JW, Thompson TC, et al. Mitochondrial DNA copy number in peripheral blood leukocytes is associated with biochemical recurrence in prostate cancer patients in african americans. *Carcinogenesis*. 2020;41:267-273
22. Dardaud LM, Bris C, Desquiret-Dumas V, Boisselier B, Tabouret E, Mokhtari K, et al. High mitochondrial DNA copy number is associated with longer survival in young patients with glioblastoma. *Neuro Oncol*. 2019;21:1084-1085
23. Li H, Tian Z, Zhang Y, Yang Q, Shi B, Hou P, et al. Increased copy number of mitochondrial DNA predicts poor prognosis of esophageal squamous cell carcinoma. *Oncol Lett*. 2018;15:1014-1020
24. Chen N, Wen S, Sun X, Fang Q, Huang L, Liu S, et al. Elevated mitochondrial DNA copy number in peripheral blood and tissue predict the opposite outcome of cancer: A meta-analysis. *Scientific reports*. 2016;6:37404
25. Ussakli CH, Ebaee A, Binkley J, Brentnall TA, Emond MJ, Rabinovitch PS, et al. Mitochondria and tumor progression in ulcerative colitis. *Journal of the National Cancer Institute*. 2013;105:1239-1248
26. Colditz GA, Hankinson SE. The nurses' health study: Lifestyle and health among women. *Nature reviews. Cancer*. 2005;5:388-396
27. Rimm EB, Giovannucci EL, Willett WC, Colditz GA, Ascherio A, Rosner B, et al. Prospective study of alcohol consumption and risk of coronary disease in men. *Lancet*. 1991;338:464-468
28. Feskanich D, Rimm EB, Giovannucci EL, Colditz GA, Stampfer MJ, Litin LB, et al. Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. *Journal of the American Dietetic Association*. 1993;93:790-796

29. Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. *American journal of epidemiology*. 1985;122:51-65

30. Rich-Edwards JW, Corsano KA, Stampfer MJ. Test of the national death index and equifax nationwide death search. *American journal of epidemiology*. 1994;140:1016-1019

31. Meng S, Wu S, Liang L, Liang G, Giovannucci E, De Vivo I, et al. Leukocyte mitochondrial DNA copy number, anthropometric indices, and weight change in us women. *Oncotarget*. 2016;7:60676-60686

32. Tabung FK, Liu L, Wang W, Fung TT, Wu K, Smith-Warner SA, et al. Association of dietary inflammatory potential with colorectal cancer risk in men and women. *JAMA oncology*. 2018;4:366-373

33. Tabung FK, Smith-Warner SA, Chavarro JE, Fung TT, Hu FB, Willett WC, et al. An empirical dietary inflammatory pattern score enhances prediction of circulating inflammatory biomarkers in adults. *The Journal of nutrition*. 2017;147:1567-1577

34. Tabung FK, Smith-Warner SA, Chavarro JE, Wu K, Fuchs CS, Hu FB, et al. Development and validation of an empirical dietary inflammatory index. *The Journal of nutrition*. 2016;146:1560-1570

35. Takkouche B, Cadarso-Suarez C, Spiegelman D. Evaluation of old and new tests of heterogeneity in epidemiologic meta-analysis. *American journal of epidemiology*. 1999;150:206-215

36. Moore AZ, Ding J, Tuke MA, Wood AR, Bandinelli S, Frayling TM, et al. Influence of cell distribution and diabetes status on the association between mitochondrial DNA copy number and aging phenotypes in the inchianti study. *Aging Cell*. 2018;17

37. de Araujo LF, Fonseca AS, Muys BR, Placa JR, Bueno RB, Lorenzi JC, et al. Mitochondrial genome instability in colorectal adenoma and adenocarcinoma. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*. 2015;36:8869-8879

38. Feng S, Xiong L, Ji Z, Cheng W, Yang H. Correlation between increased copy number of mitochondrial DNA and clinicopathological stage in colorectal cancer. *Oncol Lett*. 2011;2:899-903

39. Lin PC, Lin JK, Yang SH, Wang HS, Li AF, Chang SC. Expression of beta-f1-atpase and mitochondrial transcription factor a and the change in mitochondrial DNA content in colorectal cancer: Clinical data analysis and evidence from an in vitro study. *International journal of colorectal disease*. 2008;23:1223-1232

40. Shokolenko I, Venediktova N, Bochkareva A, Wilson GL, Alexeyev MF. Oxidative stress induces degradation of mitochondrial DNA. *Nucleic acids research*. 2009;37:2539-2548

41. Sabharwal SS, Schumacker PT. Mitochondrial ros in cancer: Initiators, amplifiers or an achilles' heel? *Nature reviews. Cancer*. 2014;14:709-721

42. Stein A, Sia EA. Mitochondrial DNA repair and damage tolerance. *Frontiers in bioscience (Landmark edition)*. 2017;22:920-943

43. Knez J, Winckelmans E, Plusquin M, Thijs L, Cauwenberghs N, Gu Y, et al. Correlates of peripheral blood mitochondrial DNA content in a general population. *American journal of epidemiology*. 2016;183:138-146

44.   Verhoeven JE, Revesz D, Picard M, Epel EE, Wolkowitz OM, Matthews KA, et al. Depression, telomeres and mitochondrial DNA: Between- and within-person associations from a 10-year longitudinal study. *Mol Psychiatry*. 2018;23:850-857

45.   Hang D, Nan H, Kvaerner AS, De Vivo I, Chan AT, Hu Z, et al. Longitudinal associations of lifetime adiposity with leukocyte telomere length and mitochondrial DNA copy number. *European journal of epidemiology*. 2018;33:485-495

46.   Wu S, Li X, Meng S, Fung T, Chan AT, Liang G, et al. Fruit and vegetable consumption, cigarette smoke, and leukocyte mitochondrial DNA copy number. *The American journal of clinical nutrition*. 2019;109:424-432

47.   Yang K, Forman MR, Monahan PO, Graham BH, Chan AT, Zhang X, et al. Insulinemic potential of lifestyle is inversely associated with leukocyte mitochondrial DNA copy number in us white adults. *The Journal of nutrition*. 2020

48.   Qu F, Liu X, Zhou F, Yang H, Bao G, He X, et al. Association between mitochondrial DNA content in leukocytes and colorectal cancer risk: A case-control analysis. *Cancer*. 2011;117:3148-3155

49.   Theodoratou E, Din FV, Farrington SM, Cetnarskyj R, Barnetson RA, Porteous ME, et al. Association between common mtdna variants and all-cause or colorectal cancer mortality. *Carcinogenesis*. 2010;31:296-301

50.   Guo Z, Zhao S, Fan H, Du Y, Zhao Y, Wang G. Identification of sequence polymorphisms in the d-loop region of mitochondrial DNA as a risk factor for colon cancer. *Mitochondrial DNA A DNA Mapp Seq Anal*. 2016;27:4244-4245

51.   Lievre A, Chapusot C, Bouvier AM, Zinzindohoue F, Piard F, Roignot P, et al. Clinical value of mitochondrial mutations in colorectal cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2005;23:3517-3525

52.   Kleist B, Meurer T, Poetsch M. Mitochondrial DNA alteration in primary and metastatic colorectal cancer: Different frequency and association with selected clinicopathological and molecular markers. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*. 2017;39:1010428317692246

53.   Srinivasainagendra V, Sandel MW, Singh B, Sundaresan A, Mooga VP, Bajpai P, et al. Migration of mitochondrial DNA in the nuclear genome of colorectal adenocarcinoma. *Genome medicine*. 2017;9:31

54.   Gao J, Wen S, Zhou H, Feng S. De-methylation of displacement loop of mitochondrial DNA is associated with increased mitochondrial copy number and nicotinamide adenine dinucleotide subunit 2 expression in colorectal cancer. *Molecular medicine reports*. 2015;12:7033-7038

Table 1. Baseline characteristics of patients with colorectal cancer

Characteristics	NHS	HPFS	All
Number of cases, n	321	266	587
Age at blood collection, yrs, Mean(SD)	58.8(6.7)	65.8(8.1)	62.0(8.1)
Age at CRC diagnosis, yrs, Mean(SD)	67.8(7.6)	73.1(8.6)	70.2(8.5)
Log-mtDNAcn, Mean(SD)	-0.08 (0.3)	-0.2 (0.3)	-0.1(0.3)
Sex, women%	100	0	54.7
Race, white%	98.1	92.5	95.6
Body mass index, kg/m2, Mean(SD)	25.7(5.0)	26.4(3.2)	26.0(4.3)
Physical activity, MET-hours/week, Mean(SD)	18.5(18.3)	29.6(27.6)	23.5(23.6)
Smoking status			
Never%	43.9	42.9	43.4
Past%	38.3	51.1	44.1
Current%	17.8	6.0	12.4
Empirical dietary inflammatory pattern score, Mean(SD)	-0.03(0.9)	0.04(0.9)	0.002(0.9)
Regular aspirin users%	39.3	42.5	40.7
Regular non-aspirin NSAID users%	13.4	8.3	11.1
CRC in a parent or sibling, %	17.8	18.4	18.1
Location of cancer			
Proximal colon%	46.4	32.3	40.0
Distal colon%	27.7	33.1	30.2
Rectum%	22.4	25.6	23.9
Unspecified%	3.4	9.0	6.0
Differentiation grade of cancer			
Well differentiated%	13.1	9.8	11.6
Moderately differentiated%	60.1	52.3	56.6
Poorly differentiated%	16.5	12.8	14.8
Unspecified%	10.3	25.2	17.0
AJCC TNM stage of cancer			
Stage 0/I%	23.7	32.0	27.4
Stage II%	28.0	16.5	22.8
Stage III%	22.7	20.3	21.6
Stage IV%	13.7	14.7	14.1
Unspecified%	11.8	16.5	14.0
Years from blood collection to CRC diagnosis, Mean(SD)	8.9(5.2)	7.3(4.2)	8.2(4.8)
Years from blood collection to death/end of follow-up, Mean(SD)	20.2(8.4)	16.7(7.0)	18.6(8.0)
Years from CRC diagnosis to death/end of follow-up, Mean(SD)	11.3(7.9)	9.4(6.7)	10.5(7.4)

Notes: Abbreviation: NHS, Nurses' Health Study; HPFS, Health Professionals Follow-up Study; MET: metabolic equivalent; NSAIDs, nonsteroidal anti-inflammatory drugs; AJCC: American Joint Committee on Cancer; TNM: T-The extent (size) of the tumor, N-The spread to nearby lymph nodes, M-the spread (metastasis) to distant sites; Values are means(SD) for continuous variables and percentages for categorical variables; Empirical dietary inflammatory pattern (EDIP) scores were adjusted for energy intake using the residual method. Lower scores indicate anti-inflammatory diets whereas higher scores indicate proinflammatory diets.

Table 2. Associations of pre-diagnostic leukocyte mtDNAcn with overall and CRC-specific survival by cohort-specific quartiles of mtDNAcn

	Quartile 4	Quartile 3	Quartile 2	Quartile 1	Per 1 SD decrease	*P* for trend
Overall Survival (OS)						
NHS						
No. of cases/deaths	80/56	81/42	80/42	80/50		
HR (95%CI)- Model 1	ref	0.65 (0.44, 0.98)	0.71 (0.48, 1.07)	0.87 (0.59, 1.29)	0.95 (0.79, 1.14)	0.57
HR (95%CI)- Model 2	ref	0.74 (0.49, 1.12)	0.74 (0.49, 1.12)	0.96 (0.64, 1.43)	0.97 (0.81, 1.17)	0.76
HR (95%CI)- Model 3	ref	0.69 (0.45, 1.07)	0.61 (0.39, 0.95)	0.90 (0.59, 1.36)	0.93 (0.77, 1.13)	0.47
HPFS						
No. of cases/deaths	66/47	67/46	67/54	66/58		
HR (95%CI)- Model 1	ref	0.92 (0.61, 1.39)	0.93 (0.63, 1.39)	1.16 (0.78, 1.72)	1.06 (0.90, 1.26)	0.47
HR (95%CI)- Model 2	ref	1.03 (0.67, 1.58)	1.13 (0.75, 1.72)	1.25 (0.84, 1.86)	1.11 (0.93, 1.31)	0.24
HR (95%CI)- Model 3	ref	1.04 (0.67, 1.61)	1.18 (0.78, 1.81)	1.21 (0.80, 1.84)	1.10 (0.92, 1.31)	0.31
All						
No. of cases/deaths	147/100	147/85	147/99	146/111		
HR (95%CI)- Model 1	ref	0.79 (0.59, 1.05)	0.98 (0.74, 1.30)	0.97 (0.74, 1.29)	1.02 (0.90, 1.15)	0.79
HR (95%CI)- Model 2	ref	0.85 (0.63, 1.14)	1.00 (0.75, 1.33)	1.10 (0.83, 1.46)	1.06 (0.94, 1.21)	0.34
HR (95%CI)- Model 3	ref	0.88 (0.65, 1.19)	1.03 (0.77, 1.38)	1.07 (0.80, 1.42)	1.05 (0.92, 1.19)	0.47
CRC-specific Survival						
NHS						
No. of cases/deaths	80/34	81/23	80/22	80/26		
HR (95%CI)- Model 1	ref	0.65 (0.38, 1.12)	0.62 (0.36, 1.06)	0.82 (0.49, 1.40)	0.90 (0.71, 1.15)	0.39
HR (95%CI)- Model 2	ref	0.77 (0.45, 1.33)	0.61 (0.35, 1.06)	0.79 (0.46, 1.35)	0.87 (0.68, 1.10)	0.25
HR (95%CI)- Model 3	ref	0.67 (0.38, 1.20)	0.48 (0.26, 0.87)	0.78 (0.44, 1.39)	0.84 (0.65, 1.09)	0.19
HPFS						
No. of cases/deaths	66/19	67/19	67/17	66/20		
HR (95%CI)- Model 1	ref	0.95 (0.50, 1.80)	0.75 (0.39, 1.46)	0.87 (0.46, 1.66)	0.92 (0.70, 1.21)	0.54
HR (95%CI)- Model 2	ref	1.26 (0.64, 2.47)	1.22 (0.60, 2.50)	0.97 (0.50, 1.87)	0.98 (0.75, 1.29)	0.90
HR (95%CI)- Model 3	ref	1.31 (0.63, 2.72)	1.08 (0.51, 2.29)	0.90 (0.43, 1.89)	0.94 (0.69, 1.27)	0.67
All						
No. of cases/deaths	147/54	147/38	147/48	146/40		
HR (95%CI)- Model 1	ref	0.70 (0.46, 1.07)	0.95 (0.64, 1.40)	0.72 (0.47, 1.09)	0.90 (0.75, 1.08)	0.25
HR (95%CI)- Model 2	ref	0.84 (0.55, 1.28)	0.95 (0.63, 1.42)	0.85 (0.56, 1.31)	0.95 (0.79, 1.14)	0.56
HR (95%CI)- Model 3	ref	0.76 (0.49, 1.18)	0.95 (0.63, 1.45)	0.77 (0.49, 1.19)	0.92 (0.76, 1.11)	0.37

Notes: Model 1: adjusted for age at blood collection and age at diagnosis (continuous), race (white, non-white), and gender/cohort (women/men, for pooled dataset only); Model 2: Model 1 + tumor location (proximal colon, distal colon, rectum, unspecified), grade of differentiation (well, moderate, poor, unspecified), and AJCC TNM stage (Stage 0/I, II, III, IV, unspecified); Model 3: Model 2 + smoking status (never, past, current), BMI (in tertiles), physical activity (in tertiles), CRC family history (yes/no), regular aspirin users (yes/no), regular non-aspirin NSAID users (yes/no), Empirical dietary inflammatory pattern (EDIP) score (in tertiles); No significant statistical heterogeneity was observed between NHS and HPFS (OS, P-heterogeneity = 0.22, CRC-specific survival, P-heterogeneity=0.60, estimates for death risk per 1 SD decrease of mtDNAcn).

Table 3. Associations of pre-diagnostic leukocyte mtDNAcn with overall and CRC-specific survival, stratified by tumor location, grade, and stage

Characteristics of tumor	No. of cases/deaths	log-mtDNAcn mean (SD)	Multivariable HR (95% CI) for mtDNAcn quartiles					*P* for trend
			Quartile 4	Quartile 3	Quartile 2	Quartile 1	Per 1 SD decrease	
			Overall Survival (OS)					
Location								
Proximal colon	235/168	-0.14 (0.33)	ref	0.93 (0.59, 1.49)	1.13 (0.70, 1.84)	1.11 (0.68, 1.80)	1.07 (0.87, 1.32)	0.53
Distal colon	177/113	-0.12 (0.31)	ref	1.35 (0.73, 2.50)	1.22 (0.65, 2.28)	1.72 (0.89, 3.34)	1.24 (0.93, 1.66)	0.15
Rectum	140/87	-0.14 (0.35)	ref	0.59 (0.29, 1.23)	0.77 (0.38, 1.53)	0.80 (0.42, 1.56)	0.93 (0.68, 1.26)	0.62
Grade of differentiation								
Well - Moderate	400/254	-0.13 (0.32)	ref	1.11 (0.76, 1.63)	1.14 (0.78, 1.66)	1.28 (0.88, 1.86)	1.11 (0.94, 1.31)	0.20
Poor	87/67	-0.11 (0.38)	ref	1.52 (0.58, 3.96)	0.71 (0.28, 1.75)	0.95 (0.41, 2.16)	0.93 (0.65, 1.34)	0.69
AJCC Stage								
Stage 0/I/II	295/175	-0.13 (0.33)	ref	0.70 (0.44, 1.12)	0.86 (0.54, 1.36)	0.86 (0.55, 1.34)	0.97 (0.80, 1.19)	0.78
Stage III	127/88	-0.14 (0.32)	ref	1.22 (0.60, 2.48)	0.73 (0.35, 1.54)	1.24 (0.65, 2.39)	1.06 (0.79, 1.43)	0.68
Stage IV	83/77	-0.16 (0.35)	ref	1.02 (0.40, 2.63)	1.26 (0.52, 3.04)	1.68 (0.70, 4.02)	1.27 (0.86, 1.87)	0.23
			CRC-specific Survival					
Location								
Proximal colon	235/77	-0.14 (0.33)	ref	0.74 (0.38, 1.45)	0.68 (0.34, 1.35)	0.66 (0.31, 1.41)	0.82 (0.60, 1.13)	0.22
Distal colon	177/50	-0.12 (0.31)	ref	2.17 (0.78, 6.00)	2.60 (0.90, 7.48)	2.29 (0.68, 7.73)	1.44 (0.86, 2.41)	0.16
Rectum	140/37	-0.14 (0.35)	ref	0.29 (0.09, 0.99)	0.67 (0.23, 1.94)	0.60 (0.20, 1.80)	0.80 (0.49, 1.33)	0.39
Grade of differentiation								
Well - Moderate	400/103	-0.13 (0.32)	ref	0.95 (0.54, 1.66)	1.06 (0.59, 1.90)	0.91 (0.49, 1.70)	0.98 (0.75, 1.28)	0.88
Poor	87/44	-0.10 (0.38)	ref	0.78 (0.23, 2.65)	1.20 (0.41, 3.51)	0.61 (0.19, 2.00)	0.87 (0.53, 1.44)	0.58
AJCC Stage								
Stage 0/I/II	295/32	-0.13 (0.33)	ref	0.49 (0.18, 1.35)	0.56 (0.21, 1.49)	0.44 (0.14, 1.37)	0.70 (0.43, 1.14)	0.15
Stage III	127/44	-0.14 (0.32)	ref	1.38 (0.56, 3.40)	0.71 (0.26, 1.94)	0.80 (0.30, 2.10)	0.85 (0.56, 1.30)	0.46
Stage IV	83/74	-0.16 (0.35)	ref	1.01 (0.39, 2.59)	1.15 (0.47, 2.81)	1.50 (0.62, 3.62)	1.20 (0.81, 1.77)	0.36

Note: HRs (95% CIs) were adjusted for the following variables (except for the stratified covariate): age at blood collection and age at diagnosis (continuous), race (white, non-white), and gender/cohort (women/men), tumor location (proximal colon, distal colon, rectum, unspecified), grade of differentiation (well,

moderate, poor, unspecified), and AJCC TNM stage (Stage 0/I, II, III, IV, unspecified), smoking status (never, past, current), BMI (in tertiles), physical activity (in tertiles), CRC family history (yes/no), regular aspirin users (yes/no), regular non-aspirin NSAID users (yes/no), empirical dietary inflammatory pattern (EDIP) score (in tertiles).

Table 4. Association of pre-diagnostic leukocyte mtDNAcn with overall survival, stratified by selected covariates

Covariates	No. of cases/deaths	log-mtDNAcn (mean (SD))	Multivariable HR (95% CI) for mtDNAcn quartiles					P for trend	P for interaction
			Quartile 4	Quartile 3	Quartile 2	Quartile 1	Per 1 SD decrease		
Age at blood collection*									0.97
≤ median (62.3 years)	293/140	-0.10 (0.33)	ref	0.74 (0.44, 1.25)	0.80 (0.49, 1.31)	1.32 (0.79, 2.19)	1.10 (0.87, 1.38)	0.43	
> median (62.3 years)	294/255	-0.17 (0.32)	ref	1.03 (0.69, 1.55)	1.24 (0.82, 1.87)	1.07 (0.74, 1.55)	1.04 (0.89, 1.22)	0.64	
Age at diagnosis*									0.80
≤ median (70.5 years)	295/161	-0.06 (0.32)	ref	0.78 (0.49, 1.24)	0.57 (0.36, 0.90)	1.06 (0.64, 1.75)	0.92 (0.73, 1.14)	0.43	
> median (70.5 years)	292/234	-0.21 (0.32)	ref	1.09 (0.70, 1.68)	1.46 (0.96, 2.21)	1.17 (0.80, 1.72)	1.08 (0.92, 1.27)	0.34	
Years from blood collection to diagnosis									0.70
≤ 5 years	168/138	-0.13 (0.34)	ref	0.92 (0.51, 1.66)	0.53 (0.29, 0.95)	0.98 (0.54, 1.77)	0.92 (0.71, 1.20)	168/138	
≤ 10 years	386/280	-0.11 (0.31)	ref	0.92 (0.64, 1.33)	1.00 (0.70, 1.43)	1.12 (0.78, 1.59)	1.06 (0.90, 1.24)	386/280	
> 10 years	201/115	-0.18 (0.35)	ref	0.82 (0.45, 1.52)	0.99 (0.54, 1.81)	1.18 (0.68, 2.04)	1.10 (0.86, 1.41)	201/115	
Smoking status									0.45
Never	255/164	-0.14 (0.32)	ref	1.02 (0.64, 1.62)	0.90 (0.57, 1.42)	1.25 (0.80, 1.95)	1.08 (0.88, 1.32)	0.45	
Past	259/180	-0.12 (0.33)	ref	0.64 (0.40, 1.01)	1.08 (0.70, 1.68)	0.73 (0.46, 1.16)	0.94 (0.77, 1.15)	0.54	
Current	73/51	-0.15 (0.34)	ref	1.12 (0.38, 3.26)	2.22 (0.75, 6.57)	2.27 (0.84, 6.14)	**1.50 (0.98, 2.32)**	**0.06**	
BMI*									0.54
15.6 ~ 24.0 kg/m2	195/124	-0.09 (0.36)	ref	0.83 (0.47, 1.46)	0.83 (0.46, 1.50)	0.87 (0.51, 1.48)	0.94 (0.74, 1.19)	0.60	
24.0 ~ 26.9 kg/m2	196/139	-0.16 (0.30)	ref	0.78 (0.44, 1.37)	1.14 (0.69, 1.90)	1.07 (0.69, 1.89)	1.07 (0.85, 1.36)	0.57	
26.9 ~ 49.1 kg/m2	196/132	-0.16 (0.31)	ref	0.78 (0.43, 1.40)	0.90 (0.49, 1.66)	0.96 (0.54, 1.70)	1.02 (0.80, 1.31)	0.86	
Physical activity*									0.35
0 ~ 9.2 METs-hours/wk	195/130	-0.14 (0.34)	ref	1.23 (0.70, 2.16)	1.76 (1.01, 3.08)	1.27 (0.75, 2.15)	1.15 (0.92, 1.43)	0.21	
9.2 ~ 26.9 METs-hours/wk	196/127	-0.11 (0.32)	ref	0.97 (0.57, 1.66)	0.89 (0.53, 1.49)	1.01 (0.58, 1.77)	0.99 (0.77, 1.26)	0.92	
26.9 ~ 177.3 METs-hours/wk	196/138	-0.16 (0.32)	ref	0.51 (0.29, 0.92)	0.69 (0.41, 1.17)	0.86 (0.52, 1.42)	1.00 (0.79, 1.25)	0.97	
Dietary inflammatory potential*									0.14
Low (-4.1 ~ -0.3)	195/125	-0.11 (0.36)	ref	1.04 (0.61, 1.77)	1.08 (0.64, 1.83)	0.62 (0.34, 1.11)	0.84 (0.66, 1.08)	0.17	
Median (-0.3 ~ 0.3)	196/132	-0.14 (0.30)	ref	0.68 (0.37, 1.24)	0.63 (0.37, 1.10)	1.41 (0.82, 2.43)	1.13 (0.88, 1.46)	0.33	
High (0.3 ~ 3.9)	196/138	-0.16 (0.31)	ref	0.92 (0.53, 1.60)	1.74 (1.02, 2.98)	1.11 (0.65, 1.87)	1.12 (0.90, 1.40)	0.31	
Regular aspirin users									0.70
No	348/221	-0.12 (0.31)	ref	0.62 (0.42, 0.92)	0.98 (0.66, 1.47)	0.85 (0.57, 1.27)	1.00 (0.83, 1.19)	0.96	
Yes	239/174	-0.16 (0.35)	ref	1.75 (1.06, 2.90)	1.05 (0.65, 1.69)	1.23 (0.78, 1.93)	1.04 (0.86, 1.26)	0.70	

Note: Subgroups by BMI, physical activity, and dietary inflammatory potential were defined by the tertiles of BMI, total MET-hours/week, and empirical dietary inflammatory pattern (EDIP) score, respectively. The EDIP scores were adjusted for energy intake using the residual method; lower scores indicate anti-inflammatory diets whereas higher scores indicate proinflammatory diets. HRs (95% CIs) were adjusted for the following variables (except for the stratified covariate): age at blood collection and age at diagnosis (continuous), race (white, non-white), and gender/cohort (women/men), tumor location (proximal colon, distal colon, rectum, unspecified), grade of differentiation (well, moderate, poor, unspecified), and AJCC TNM stage (Stage 0/I, II, III, IV), smoking status (never, past, current), BMI (in tertiles), physical activity (in tertiles), CRC family history (yes/no), regular aspirin users (yes/no), regular non-aspirin NSAIDs users (yes/no), Empirical dietary inflammatory pattern (EDIP) score (in tertiles). * The covariate was further adjusted in each tertile as a continuous variable.

Table 5. Association of pre-diagnostic leukocyte mtDNAcn with CRC-specific survival, stratified by selected covariates

Covariates	No. of cases/deaths	log-mtDNAcn (mean (SD))	Quartile 4	Quartile 3	Quartile 2	Quartile 1	Per 1 SD decrease	P for trend	P for interaction
			Multivariable HR (95% CI) for mtDNAcn quartiles						
Age at blood collection*									0.64
≤ median (62.3 years)	293/80	-0.10 (0.33)	ref	0.85 (0.44, 1.67)	0.71 (0.37, 1.35)	1.00 (0.48, 2.07)	0.94 (0.68, 1.28)	0.68	
> median (62.3 years)	294/100	-0.17 (0.32)	ref	0.85 (0.45, 1.62)	1.10 (0.58, 2.10)	0.76 (0.41, 1.41)	0.91 (0.70, 1.19)	0.49	
Age at diagnosis*									0.33
≤ median (70.5 years)	295/87	-0.06 (0.32)	ref	0.73 (0.40, 1.33)	0.45 (0.23, 0.85)	0.63 (0.30, 1.35)	**0.71 (0.52, 0.97)**	**0.03**	
> median (70.5 years)	292/93	-0.21 (0.32)	ref	1.20 (0.59, 2.48)	1.65 (0.86, 3.16)	0.96 (0.51, 1.83)	1.00 (0.76, 1.31)	0.99	
Years from blood collection to diagnosis									0.37
≤ 5 years	168/67	-0.13 (0.34)	ref	0.81 (0.38, 1.73)	0.46 (0.21, 1.03)	0.43 (0.18, 1.06)	**0.65 (0.44, 0.96)**	**0.03**	
≤ 10 years	386/129	-0.11 (0.31)	ref	0.87 (0.53, 1.43)	0.90 (0.54, 1.48)	0.70 (0.40, 1.25)	0.87 (0.69, 1.11)	0.27	
> 10 years	201/51	-0.18 (0.35)	ref	0.34 (0.10, 1.10)	0.99 (0.37, 2.63)	0.93 (0.35, 2.43)	1.17 (0.76, 1.80)	0.47	
Smoking status									0.91
Never	255/73	-0.14 (0.32)	ref	0.88 (0.43, 1.84)	1.00 (0.53, 1.90)	1.21 (0.62, 2.37)	1.09 (0.81, 1.47)	0.58	
Past	259/84	-0.12 (0.33)	ref	0.69 (0.37, 1.27)	0.89 (0.48, 1.66)	0.47 (0.23, 0.98)	0.77 (0.57, 1.04)	0.09	
Current	73/23	-0.15 (0.34)	ref	0.78 (0.15, 4.04)	2.42 (0.44, 13.42)	2.52 (0.47, 13.50)	1.71 (0.82, 3.55)	0.15	
BMI*									0.89
15.6 ~ 24.0 kg/m2	195/68	-0.09 (0.36)	ref	0.66 (0.30, 1.46)	0.63 (0.29, 1.34)	0.65 (0.31, 1.36)	0.81 (0.58, 1.13)	0.20	
24.0 ~ 26.9 kg/m2	196/58	-0.16 (0.30)	ref	0.99 (0.40, 2.47)	1.62 (0.70, 3.77)	0.92 (0.33, 2.59)	1.06 (0.69, 1.63)	0.79	
26.9 ~ 49.1 kg/m2	196/54	-0.16 (0.31)	ref	1.03 (0.43, 2.46)	1.01 (0.41, 2.52)	0.65 (0.26, 1.68)	0.83 (0.56, 1.23)	0.35	
Physical activity*									0.68
0 ~ 9.2 METs-hours/wk	195/56	-0.14 (0.34)	ref	1.01 (0.41, 2.48)	1.76 (0.80, 3.89)	0.76 (0.33, 1.74)	0.99 (0.71, 1.37)	0.93	
9.2 ~ 26.9 METs-hours/wk	196/67	-0.11 (0.32)	ref	0.95 (0.45, 2.00)	0.93 (0.45, 1.91)	0.67 (0.29, 1.56)	0.86 (0.61, 1.22)	0.40	
26.9 ~ 177.3 METs-hours/wk	196/57	-0.16 (0.32)	ref	0.71 (0.28, 1.79)	0.90 (0.37, 2.16)	1.00 (0.43, 2.32)	1.04 (0.72, 1.51)	0.82	
Dietary inflammatory potential*									0.45
Low (-4.1 ~ -0.3)	195/62	-0.11 (0.36)	ref	0.77 (0.37, 1.62)	0.61 (0.27, 1.39)	0.29 (0.12, 0.73)	**0.61 (0.42, 0.88)**	**0.009**	
Median (-0.3 ~ 0.3)	196/62	-0.14 (0.30)	ref	0.47 (0.18, 1.22)	0.82 (0.37, 1.82)	1.55 (0.68, 3.54)	1.26 (0.86, 1.85)	0.24	
High (0.3 ~ 3.9)	196/56	-0.16 (0.31)	ref	1.37 (0.56, 3.31)	2.55 (1.11, 5.88)	0.41 (0.14, 1.19)	0.89 (0.62, 1.29)	0.55	

Regular aspirin users								0.83
No	348/106	-0.12 (0.31)	ref	0.51 (0.28, 0.91)	0.78 (0.44, 1.39)	0.58 (0.31, 1.08)	0.83 (0.63, 1.10)	0.20
Yes	239/74	-0.16 (0.35)	ref	1.75 (0.84, 3.60)	1.14 (0.54, 2.40)	0.80 (0.38, 1.67)	0.89 (0.66, 1.21)	0.46

Notes: Subgroups by BMI, physical activity, and dietary inflammatory potential were defined by the tertiles of BMI, total MET-hours/week, and empirical dietary inflammatory pattern (EDIP) score, respectively. The EDIP scores were adjusted for energy intake using the residual method; lower scores indicate anti-inflammatory diets whereas higher scores indicate proinflammatory diets. HRs (95% CIs) were adjusted for the following variables (except for the stratified covariate): age at blood collection and age at diagnosis (continuous), race (white, non-white), and gender/cohort (women/men), tumor location (proximal colon, distal colon, rectum, unspecified), grade of differentiation (well, moderate, poor, unspecified), and AJCC TNM stage (Stage 0/I, II, III, IV), smoking status (never, past, current), BMI (in tertiles), physical activity (in tertiles), CRC family history (yes/no), regular aspirin users (yes/no), regular non-aspirin NSAIDs users (yes/no), Empirical dietary inflammatory pattern (EDIP) score (in tertiles). * The covariate was further adjusted in each tertile as a continuous variable.