Circulating markers of microbial translocation and host response to bacteria with risk of colorectal cancer: A prospective, nested case-control study in men

Mengyao Shi  
*Washington University School of Medicine in St. Louis*

Xiaoyu Zong  
*Washington University School of Medicine in St. Louis*

Jinhee Hur  
*Harvard University*

Brenda M Birmann  
*Harvard University*

Otoniel Martinez-Maza  
*University of California, Los Angeles*

*See next page for additional authors*

Follow this and additional works at: https://digitalcommons.wustl.edu/oa_4

Part of the *Medicine and Health Sciences Commons*

Please let us know how this document benefits you.

**Recommended Citation**

Shi, Mengyao; Zong, Xiaoyu; Hur, Jinhee; Birmann, Brenda M; Martinez-Maza, Otoniel; Epeldegui, Marta; Chan, Andrew T; Giovannucci, Edward L; and Cao, Yin, "Circulating markers of microbial translocation and host response to bacteria with risk of colorectal cancer: A prospective, nested case-control study in men." eBioMedicine. 91, 104566 (2023).

https://digitalcommons.wustl.edu/oa_4/2509
Circulating markers of microbial translocation and host response to bacteria with risk of colorectal cancer: a prospective, nested case-control study in men

Mengyao Shi,a Xiaoyu Zong,a Jinhee Hur,b,c,d Brenda M. Birmann,e Otoniel Martinez-Maza,f,g,h Marta Epeldegui,h Andrew T. Chan,e,i,j,k,l Edward L. Giovannucci,b,m and Yin Caoa,n,o,*

aDivision of Public Health Sciences, Department of Surgery, Washington University School of Medicine, St. Louis, MO, USA
bDepartment of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA
cDepartment of Food Science and Biotechnology, Sungkyunkwan University, Suwon, Gyeonggi, South Korea
dFood Clinical Research Center, Institute of Biotechnology and Bioengineering, Sungkyunkwan University, Suwon, Gyeonggi, South Korea
eChanning Division of Network Medicine, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA
fDepartment of Obstetrics and Gynecology, AIDS Institute, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA
gDepartment of Epidemiology, Fielding School of Public Health, University of California Los Angeles, Los Angeles, CA, USA
hDepartment of Microbiology, Immunology & Molecular Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA
iClinical and Translational Epidemiology Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA
jDivision of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA
kBroad Institute of MIT and Harvard, Cambridge, MA, USA
lDepartment of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, MA, USA
mDepartment of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA
nDivision of Gastroenterology, Department of Medicine, Washington University School of Medicine, St. Louis, MO, USA
oAlvin J. Siteman Cancer Center, Washington University School of Medicine, St. Louis, MO, USA

Summary

Background Gut microbial dysbiosis contributes to colorectal cancer (CRC) pathogenesis, possibly mediated in part by increased intestinal permeability to endotoxin lipopolysaccharide (LPS), microbial translocation, and subsequent endotoxemia and inflammation. However, epidemiologic evidence linking circulating markers of microbial translocation with CRC risk is limited.

Methods We conducted a prospective, nested case–control study of 261 incident CRC cases and 261 controls (matched on age and time of blood draw) among 18,159 men with pre-diagnostic blood specimens in the Health Professionals Follow-Up Study (1993–2009). We examined three complementary markers of microbial translocation and host response to bacteria, including LPS-binding protein (LBP), soluble CD14 (sCD14), and endotoxincore antibody (EndoCAb) immunoglobulin M (IgM), with subsequent risk of CRC. Unconditional logistic regressions were used to estimate odds ratios (ORs) and 95% confidence intervals (CIs).

Findings Pre-diagnostic circulating levels of sCD14 were associated with a higher risk of incident CRC. Compared to men in the lowest quartile, the multivariable OR was 1.90 (95% CI, 1.13–3.22) for men in the highest quartile (OR per standard deviation [SD] increase, 1.28; 95% CI 1.06–1.53; P_trend = 0.01). This positive association remained similar after adjusting for C-reactive protein, interleukin-6, and soluble tumor necrosis factor receptor-2, and within strata of

Abbreviations: AHEI, Alternative Healthy Eating Index; BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; CRP, C-reactive protein; CVs, coefficients of variation; ELISA, enzyme-linked immunosorbent assay; EndoCAb, endotoxin core antibody; ESD, extreme studentized deviate; FFQs, food frequency questionnaires; HPFS, Health Professionals Follow-Up Study; IgM, immunoglobulin M; IL-6, interleukin-6; IQR, interquartile range; LBP, lipopolysaccharide-binding protein; LPS, lipopolysaccharide; MD-2, myeloid differentiation factor 2; MET, metabolic equivalent of task; MV, multivariable; NSAIDs, non-steroidal anti-inflammatory drugs; OR, odds ratio; sCD14, soluble CD14; SD, standard deviation; sTNFR-2, soluble tumor necrosis factor receptor-2; TLR-4, Toll-like receptor 4; US, United States
*Corresponding author. Division of Public Health Sciences, Department of Surgery, Washington University School of Medicine, 660 S. Euclid Ave., Campus Box 8100, St. Louis, MO, 63110, USA.
E-mail address: yin.cao@wustl.edu (Y. Cao).
putative CRC risk factors. We also observed a suggestive inverse association between EndoCAb IgM and risk of CRC (OR per SD increase: 0.84; 95%CI 0.69–1.02; Ptrend = 0.09).

**Interpretation** Microbial translocation and host response to bacteria, as reflected by sCD14, is associated with risk of incident CRC in men.

**Funding** US National Institutes of Health.

**Copyright** © 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

**Keywords:** Colorectal cancer; Microbial translocation; Gut dysbiosis

### Research in context

#### Evidence before this study
Accumulating evidence supports the role of gut microbial dysbiosis in colorectal carcinogenesis, likely mediated in part by increased permeability to endotoxin lipopolysaccharide (LPS) with subsequent endotoxemia and inflammation. However, due to challenges in measuring circulating levels of LPS, epidemiologic studies linking circulating markers of microbial translocation with colorectal cancer (CRC) are limited. LPS-binding protein (LBP), soluble CD14 (sCD14), and endotoxin core antibody (EndoCAb) immunoglobulin M (IgM) have emerged as promising markers to reflect microbial translocation, yet their associations with risk of CRC have not been comprehensively examined. We searched PubMed for articles published between Jan 1, 1990, and January 30, 2022, using the terms “LBP or sCD14 or EndoCAb” and “colorectal cancer”. One nested case-control study from the Multiethnic Cohort reported no association between pre-diagnostic plasma LBP levels and subsequent risk of CRC. However, we did not identify any epidemiologic studies that examined the association between sCD14 or EndoCAb IgM with risk of incident CRC.

#### Added value of this study
In this prospective, nested case-control study of 261 incident CRC cases and matched controls within the Health Professionals Follow-Up Study, men with plasma sCD14 levels in the highest quartile had a 90% increased risk of incident CRC compared to those in the lowest quartile. The association was independent of CRC putative risk factors and several systemic inflammatory markers. We also observed a suggestive inverse association between EndoCAb IgM and risk of incident CRC.

#### Implications of all the available evidence
Our findings lend support for the role of microbial translocation, as reflected by higher levels of plasma sCD14 and potentially lower levels of EndoCAb IgM, in colorectal carcinogenesis. Further research is needed to fully uncover the role of this emerging pathway in CRC and its promise in etiology and early detection.

### Introduction
Colorectal cancer (CRC) is the 3rd leading cancer type for both incidence and mortality in the United States (US).¹ Accumulating evidence supports the role of gut dysbiosis, a persistent imbalance of the microbial community in the gut, in colorectal carcinogenesis.² Alterations in the intestinal barrier functions resulting either from gut microbiota dysbiosis, environmental, or genetic triggers,³ promote the translocation of bacteria or their components such as the endotoxin lipopolysaccharide (LPS) into the bloodstream, causing endotoxemia, which triggers a pro-inflammatory signaling cascade and immune responses.⁴

However, human studies linking microbial translocation and host response to bacteria with risk of CRC are limited in part due to challenges in biomarker measurement. LPS is one of the potent virulence factors of Gram-negative bacterial species with a major role in both acute and chronic infections.³ However, LPS has a short half-life and can easily contaminate labware and the environment, reliable measurement of circulating LPS requires LPS-free conditions throughout the collection and measurement processes.⁵ Such conditions are challenging to achieve, particularly for archived samples collected in cohort studies.

Recently, LPS-binding protein (LBP),⁶ soluble CD14 (sCD14),⁷,⁸ and endotoxin core antibody (EndoCAb) immunoglobulin M (IgM)⁹ have emerged as promising markers to reflect microbial translocation and responses to LPS. LBP specifically binds and transfers bacterial LPS; therefore, increased production of LBP is considered an indicator of the presence of effective LPS in the body.¹⁰ sCD14, the soluble form of the LPS co-receptor and primarily produced by macrophages and hepatocytes, could also increase as part of the innate immune response to LPS.¹¹ In LPS-primed monocytes, sCD14 augments Toll-like receptor 4 (TLR-4)¹² and inflammatory biomediators. Given the long half-lives of LBP and sCD14 (24–28 h; LPS: 2 min), LBP and sCD14 have been suggested to reflect long-term exposure to
endotoxin. EndoCAb IgM, inversely associated with LPS levels, is an antibody against LPS core oligosaccharide and another indirect marker of endotoxin exposure and can be measured in healthy human plasma.\(^{9,10,15,17}\) Importantly, these three markers can be reliably measured using archived blood.\(^{9,10,15,17}\) Thus far, epidemiologic studies on these markers in relation to subsequent risk of CRC are lacking.

To address these knowledge gaps, we prospectively investigated the associations between circulating pre-diagnostic levels of plasma LBP, sCD14 and EndoCAb IgM, and risk of incident CRC, leveraging high-quality blood biospecimen and rich clinical and lifestyle data collected from a large, prospective US cohort of men, the Health Professionals Follow-Up Study (HPFS).

**Methods**

**Study population**

The HPFS is an ongoing prospective cohort study of 51,529 US male health professionals aged 40–75 years at enrolment in 1986. Briefly, participants self-reported information on lifestyle and medical history biennially through follow-up questionnaires with a response rate of approximately 90%. Detailed dietary information was collected using validated food frequency questionnaires (FFQs) every four years. The study protocol was approved by the institutional review boards of the Brigham and Women’s Hospital and Harvard T.H. Chan School of Public Health, and those of participating registries as required (IRB protocol number: 2001P001945). All participants provided informed consent upon completion of self-administered questionnaires and return of blood specimens.

**Ascertainment of cases and selection of controls**

A blood specimen was returned by 18,159 men in the HPFS on ice by overnight courier between 1993 and 1995. More than 95% of samples arrived within 24 h of collection. Procedures for blood collection, handling and storage have been summarized previously.\(^{18}\) Participants who reported a diagnosis of CRC in biennial questionnaires (FFQs) every four years. The study protocol was approved by the institutional review boards of the Brigham and Women’s Hospital and Harvard T.H. Chan School of Public Health, and those of participating registries as required (IRB protocol number: 2001P001945). All participants provided informed consent upon completion of self-administered questionnaires and return of blood specimens.

Potential confounders, effect modifiers or mediators were identified from the literature\(^{19,20}\) and incorporated into a directed acyclic graph that was used to guide the modeling strategy (Supplementary Fig. S1). Detailed information on putative CRC risk factors was obtained from self-reported questionnaires beginning in 1986 through the time of blood draw (1993–95). This included adult height, current weight, family history of CRC in one or more first-degree relatives, smoking status (never or ever smoked) and pack-years of smoking among ever smokers, physical activity in metabolic equivalent of task (MET)-hours, and regular use (≥2 tablets/week) of aspirin or non-steroidal anti-inflammatory drugs (NSAIDs). We used adult height and current weight information to calculate the body mass index (BMI, kg/m\(^2\)). Using FFQs, intake of alcohol, red and processed meat, dietary fiber, total folate (from food and supplements), total calcium and total vitamin D was updated every 4 years. Diet quality was assessed by the Alternative Healthy Eating Index (AHEI)-2010 excluding the alcohol component. We used cumulative average BMI, physical activity, aspirin/NSAIDs use, and dietary intake variables by averaging the values from 1986 to the time of blood draw. Biennially, participants reported history of colonoscopy or sigmoidoscopy during the previous 2 years.
We also used extant data on three systemic inflammatory markers measured previously in this nested case-control samples, including C-reactive protein (CRP), interleukin-6 (IL-6), and soluble tumor necrosis factor receptor-2 (sTNFR-2). In brief, CRP was measured using a highly sensitive immunoturbidimetric assay (Denka Seiken Co), IL-6 was measured with an ultra-sensitive ELISA (R&D Systems), and sTNFR-2 was measured by ELISA (R&D Systems). The CVs were 7.8% for CRP, 12.1% for IL-6, and 10.1% for sTNFR-2.

Statistical analyses
We first corrected the biomarker concentrations of LBP, sCD14, and EndoCAb IgM using the average batch correction method to minimize the potential influence of laboratory batch-related variability. Generalized extreme studentized deviate (ESD) many-outlier detection was performed to detect outliers and set them as missing in further analyses. Spearman correlation coefficients were calculated between plasma markers of microbial translocation and plasma markers of inflammation among control participants. We also performed factor analysis on these biomarkers to derive three biomarker patterns. We used the iterated principal factor method to extract factors, followed by orthogonal varimax rotation. We retained three factors based on eigenvalues and inspection of scree plots.

To examine the association between LBP, sCD14, EndoCAb IgM and subsequent risk of CRC, we categorized each plasma marker into quartiles based on their distributions among controls. Unconditional logistic regressions were used to estimate the quartile-specific odds ratios (ORs) and 95% confidence intervals (CIs) for CRC, using the lowest quartiles as the respective reference categories. Tests for trend were performed using the median value for each quartile as a continuous variable. We estimated OR (95% CI) per 1 standard deviation (SD) increase. To allow for comparison across different samples or populations, we also estimated unstandardized regression coefficient and OR per 1000 ng/ml increase in LBP, per 100 ng/ml increase in sCD14, and per 10 MMU/ml increase in EndoCAb IgM. We assessed linearity using Box–Tidwell test and conducted restricted cubic splines to non-parametrically examine the possible non-linear relation between biomarkers and risk of CRC. We tested for non-linearity using the likelihood ratio test, comparing the model with only the linear term to the model with the linear and the cubic spline terms. As the planned stratified analysis disassembled the matching, unconditional instead of conditional logistic regression was used to estimate OR and 95% CIs to retain samples and allow for direct comparison of effect estimates throughout the analyses.

In the multivariable analyses, we adjusted for the matching factors (age and time of blood draw) and for race (white, non-white), BMI (continuous), family history of CRC (yes, no), prior history of sigmoidoscopy/colonoscopy (yes, no), regular use of aspirin (yes, no), regular use of NSAIDs (yes, no), pack-years of smoking (continuous), physical activity (continuous), intake of alcohol, red and processed meat, dietary fiber, total folate (from foods and supplements), total calcium and total vitamin D (all continuous), and AHEI-2010 without alcohol (continuous). To evaluate whether the identified associations were mediated by major systemic inflammatory markers associated with CRC, we further adjusted for CRP, IL-6, and sTNFR-2 (all continuous) to estimate the controlled direct effect of our exposures. We also conducted sensitivity analyses after excluding cases diagnosed within 2 years of blood draw and their matched controls to diminish the potential influence of subclinical neoplasia on circulating marker levels.

Furthermore, we performed stratified analyses for the association of per SD increase in sCD14 with risk of CRC according to age at blood draw (<, ≥65 years), time since blood draw (<, ≥6 years), BMI (<25, ≥25 kg/m²), alcohol intake (<15, ≥15 g/day), smoking status (never, ever), prior history of sigmoidoscopy/colonoscopy (yes, no), family history of CRC (yes, no), regular use of aspirin (yes, no) and AHEI-2010 without alcohol (<, ≥ median). Tests for interaction were performed using the Wald test based on a cross-product term of each potential effect modifier and per SD increase in biomarker.

We also assessed whether the association between sCD14 and risk of CRC differed according to age of CRC diagnosis and tumor location, using polytomous logistic regressions (age of CRC diagnosis < median age vs. ≥ median age vs. controls; colon cancer vs. rectal cancer vs. controls). P-values for heterogeneity were estimated based on the Wald test. We used SAS version 9.4 (SAS Institute, Inc) for all statistical analyses. All statistical tests were two-sided, and a P-value less than .05 was considered statistically significant.

Role of the funders
The funders had no role in study design; the collection, analysis, and interpretation of data; writing of the report; or the decision to submit for publication.

Results
Over up to 16 years of follow-up after blood collection, a total of 261 incident CRC cases and 261 matched controls were included in the analysis (Table 1). At the time of blood draw, participants who subsequently developed CRC were more likely than controls to have family history of CRC and less likely to have prior history of sigmoidoscopy/colonoscopy or to use NSAIDs regularly. Also, cases had lower intake of vitamin D and higher circulating levels of CRP and IL-6 than controls.
A total of 259 cases and 260 controls for LBP (2 missing measurement, 1 outlier), 253 cases and 259 controls for sCD14 (3 outliers), and 253 cases and 259 controls for EndoCAb IgM (10 outliers) were included in the final analyses. LBP and sCD14 levels were higher among cases than controls, while EndoCAb IgM levels were lower. Among controls, most baseline characteristics did not differ systematically by quartile of the three biomarkers (Supplementary Tables S1–S3). sCD14 and LBP were modestly but significantly correlated (Spearman r = 0.28, P < 0.001), while no significant correlation was found between EndoCAb IgM and sCD14 or LBP (Supplementary Table S4). In addition, LBP and sCD14 were modestly correlated with CRP, IL-6, and sTNFR-2 (Spearman r ranged from 0.25 to 0.50), whereas EndoCAb IgM was weakly negatively correlated with sTNFR-2 (r = −0.16) but not correlated with CRP (r = −0.10) and IL-6 (r = −0.02). Using factor analysis, three biomarker patterns were identified and named based on which biomarkers loaded most highly on the pattern, including CRP/IL-6 pattern, sTNFR-2 dominant pattern, and EndoCAb IgM dominant pattern. LBP contributed to CRP/IL-6 pattern, while sCD14 contributed to both the CRP/IL-6 and sTNFR-2 dominant patterns. In contrast, EndoCAb IgM led its own pattern (Supplementary Table S5).

Higher levels of plasma sCD14 were associated with an increased risk of CRC after adjusting for known or putative CRC risk factors (Table 2 and Supplementary Fig. S2B). Compared to men with sCD14 levels in the lowest quartile (Q1), men with levels in the highest quartile (Q4) had a 90% increased risk of CRC (OR, 1.90; 95% CI, 1.13–3.22; P trend = 0.01), with a 28% higher risk per SD increase in sCD14 (OR, 1.28; 95%CI 1.06–1.53). Additional adjustment for markers of inflammation, including CRP, IL-6, and sTNFR-2, only slightly attenuated the magnitude of this association. These associations also remained robust after excluding CRC cases diagnosed within 2 years of blood draw (Supplementary Table S6) or using conditional logistic regression (Supplementary Table S7).

Table 1: Characteristics of CRC cases and controls at blood draw, HPFS 1993-1995.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Cases (n = 261)</th>
<th>Controls (n = 261)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at blood draw, years</td>
<td>65.7 (8.4)</td>
<td>65.6 (8.4)</td>
</tr>
<tr>
<td>Blood draw to diagnosis, years</td>
<td>6.9 (4.1)</td>
<td>NA</td>
</tr>
<tr>
<td>White, n (%)</td>
<td>245 (93.9)</td>
<td>253 (96.9)</td>
</tr>
<tr>
<td>Body mass index, kg/m²</td>
<td>26.2 (3.0)</td>
<td>25.2 (2.6)</td>
</tr>
<tr>
<td>Family history of CRC, n (%)</td>
<td>51 (19.5)</td>
<td>39 (14.9)</td>
</tr>
<tr>
<td>Prior history of sigmoidoscopy/colonoscopy, n (%)</td>
<td>149 (57.1)</td>
<td>167 (64.0)</td>
</tr>
<tr>
<td>Regular use of aspirin (≥2 tablets/week), n (%)</td>
<td>112 (42.9)</td>
<td>120 (46.0)</td>
</tr>
<tr>
<td>Regular use of NSAIDs (≥2 tablets/week), n (%)</td>
<td>28 (10.7)</td>
<td>34 (13.0)</td>
</tr>
<tr>
<td>Ever smokers, n (%)</td>
<td>144 (55.2)</td>
<td>144 (55.2)</td>
</tr>
<tr>
<td>Pack-years among ever smokers, median (IQR)</td>
<td>15.0 (8.0–27.0)</td>
<td>15.0 (5.0–26.0)</td>
</tr>
<tr>
<td>Physical activity, MET-hours/week, median (IQR)</td>
<td>25.6 (12.1–40.2)</td>
<td>25.5 (12.8–37.9)</td>
</tr>
<tr>
<td>Dietary intake</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol, g/day, median (IQR)</td>
<td>7.3 (0.9–18.3)</td>
<td>6.9 (1.9–16.9)</td>
</tr>
<tr>
<td>Red and processed meats, servings/week, median (IQR)</td>
<td>5.9 (3.8–9.0)</td>
<td>6.1 (3.5–8.9)</td>
</tr>
<tr>
<td>Dietary fiber, g/day</td>
<td>22.2 (6.3)</td>
<td>22.7 (6.6)</td>
</tr>
<tr>
<td>Total folate, μg/day</td>
<td>496 (201)</td>
<td>537 (240)</td>
</tr>
<tr>
<td>Total calcium, mg/day</td>
<td>909 (328)</td>
<td>940 (370)</td>
</tr>
<tr>
<td>Total vitamin D, IU/day, median (IQR)</td>
<td>403 (243-546)</td>
<td>410 (262–636)</td>
</tr>
<tr>
<td>AHEI-2010 without alcohol</td>
<td>48.1 (10.6)</td>
<td>48.3 (10.4)</td>
</tr>
<tr>
<td>Plasma markers of inflammation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRP, mg/L, median (IQR)</td>
<td>1.43 (0.72–2.83)</td>
<td>1.06 (0.57–2.17)</td>
</tr>
<tr>
<td>IL-6, pg/ml, median (IQR)</td>
<td>1.61 (1.00–2.68)</td>
<td>1.39 (0.93–2.37)</td>
</tr>
<tr>
<td>sTNFR-2, pg/ml, median (IQR)</td>
<td>2751 (2356–3273)</td>
<td>2684 (2323–3273)</td>
</tr>
<tr>
<td>Plasma markers of microbial translocation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LBP, ng/ml, median (IQR)</td>
<td>4340 (3650–5009)</td>
<td>4274 (3694–4839)</td>
</tr>
<tr>
<td>sCD14, ng/ml, median (IQR)</td>
<td>1212 (1031–1365)</td>
<td>1191 (1075–1320)</td>
</tr>
<tr>
<td>EndoCAb IgM, MMU/ml, median (IQR)</td>
<td>47.5 (32.6–64.4)</td>
<td>51.8 (35.9–73.2)</td>
</tr>
</tbody>
</table>

AHEI, Alternative Healthy Eating Index; CRC, colorectal cancer; CRP, C-reactive protein; EndoCAb, endotoxin core antibody HPFS, Health Professionals Follow-up Study (men only); IgM, immunoglobulin M; IL-6, interleukin-6; IQR, interquartile range; LBP, lipopolysaccharide-binding protein; MET, metabolic equivalent of task; NA, not applicable; NSAID, non-steroidal anti-inflammatory drug; sCD14, soluble CD14; SD, standard deviation; sTNFR-2, soluble tumor necrosis factor receptor-2. Data were presented as mean (SD) unless otherwise indicated. Cumulative average value was used by averaging the values from baseline to the time of blood draw. Batch corrected concentrations were used.
In contrast, plasma levels of LBP (OR per SD increase, 1.02; 95%CI, 0.86–1.21; \( P_{\text{trend}} = 0.88 \)) was not associated with risk of CRC. For EndoCAb IgM, we also identified a suggestive inverse association with subsequent risk of CRC (OR per SD increase, 0.84; 95%CI, 0.69–1.02; \( P_{\text{trend}} = 0.09 \)). ORs (95%CI) per 1000 ng/ml increase in LBP, per 100 ng/ml increase in sCD14, and per 10 MMU/ml increase in EndoCAb IgM were reported in Supplementary Table S8. The restricted cubic splines showed a pattern similar to the quartile analysis (Supplementary Fig. S2).

Stratified analyses for sCD14 and CRC according to age at blood draw, time since blood draw, BMI, alcohol intake, smoking status, prior history of sigmoidoscopy/colonoscopy, regular use of aspirin and AHEI-2010 score without alcohol did not reveal any statistically significant interactions (all \( P_{\text{interaction}} \geq 0.32 \); Fig. 1 and Supplementary Table S9). Nonetheless, somewhat stronger associations were observed among men who were younger at blood draw (OR per SD increase, 1.43; 95%CI, 1.05–1.96; \( P_{\text{trend}} = 0.03 \)) than among those of older age at blood draw (OR per SD increase, 1.20; 95%CI, 0.96–1.51; \( P_{\text{trend}} = 0.27 \), and among individuals with higher AHEI-2010 score without alcohol (above-median [48.2]: OR per SD increase, 1.40; 95%CI, 1.08–1.80; \( P_{\text{trend}} = 0.02 \)) compared to those with lower AHEI-2010 score without alcohol (below-median: OR per SD increase, 1.17; 95%CI, 0.90–1.51; \( P_{\text{trend}} = 0.19 \)).

Further, plasma sCD14 levels were not differentially associated with risk of CRC according to age of CRC diagnosis or tumor location (all \( P_{\text{heterogeneity}} \geq 0.51 \) (Table 3), although a somewhat stronger association was observed for cases younger than the median age (73.8 years) (OR per SD increase, 1.42; 95%CI, 1.12–1.82; \( P_{\text{trend}} = 0.03 \)), compared to those diagnosed at older ages (OR per SD increase, 1.10; 95%CI, 0.87–1.40; \( P_{\text{trend}} = 0.38 \)).

**Discussion**

In this prospective, nested case–control study of US men, higher levels of plasma sCD14 were associated with a higher risk of CRC, independent of putative risk factors and systemic inflammatory markers including CRP, IL-6, and sTNFR-2. We also observed a suggestive inverse association between EndoCAb IgM and risk of CRC. However, we found no association between LBP and subsequent risk of CRC. Our findings provide

### Table 2: Association of circulating markers of microbial translocation with risk of CRC, HPFS 1993-2009.

<table>
<thead>
<tr>
<th>Biomarkers</th>
<th>Quartile</th>
<th>Per SD increase</th>
<th>( P_{\text{trend}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LBP</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases/controls</td>
<td>73/65</td>
<td>52/65</td>
<td>52/65</td>
</tr>
<tr>
<td>Median, ng/ml</td>
<td>3103</td>
<td>3904</td>
<td>4499</td>
</tr>
<tr>
<td>Cut points, ng/ml</td>
<td>1250–3691</td>
<td>3692–4268</td>
<td>4269–4832</td>
</tr>
<tr>
<td>Age-adjusted OR (95% CI)</td>
<td>1 (referent)</td>
<td>0.75 (0.46–1.24)</td>
<td>0.71 (0.43–1.17)</td>
</tr>
<tr>
<td>MV-adjusted OR (95% CI)</td>
<td>1 (referent)</td>
<td>0.76 (0.46–1.28)</td>
<td>0.83 (0.37–1.67)</td>
</tr>
<tr>
<td>MV-adjusted OR (95% CI)</td>
<td>1 (referent)</td>
<td>0.84 (0.48–1.44)</td>
<td>0.59 (0.34–1.03)</td>
</tr>
<tr>
<td><strong>sCD14</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases/controls</td>
<td>55/65</td>
<td>56/65</td>
<td>67/65</td>
</tr>
<tr>
<td>Median, ng/ml</td>
<td>649–1074</td>
<td>1075–1189</td>
<td>1190–1317</td>
</tr>
<tr>
<td>Age-adjusted OR (95% CI)</td>
<td>1 (referent)</td>
<td>1.04 (0.62–1.73)</td>
<td>1.23 (0.74–2.03)</td>
</tr>
<tr>
<td>MV-adjusted OR (95% CI)</td>
<td>1 (referent)</td>
<td>1.29 (0.75–2.20)</td>
<td>1.45 (0.85–2.49)</td>
</tr>
<tr>
<td>MV-adjusted OR (95% CI)</td>
<td>1 (referent)</td>
<td>1.30 (0.73–2.30)</td>
<td>1.54 (0.86–2.74)</td>
</tr>
<tr>
<td><strong>EndoCAb IgM</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases/controls</td>
<td>77/65</td>
<td>69/65</td>
<td>60/64</td>
</tr>
<tr>
<td>Median, MMU/ml</td>
<td>27.2</td>
<td>43.6</td>
<td>61.3</td>
</tr>
<tr>
<td>Cut points, MMU/ml</td>
<td>10.1–35.8</td>
<td>35.9–51.7</td>
<td>51.8–73.0</td>
</tr>
<tr>
<td>Age-adjusted OR (95% CI)</td>
<td>1 (referent)</td>
<td>0.89 (0.55–1.44)</td>
<td>0.77 (0.47–1.25)</td>
</tr>
<tr>
<td>MV-adjusted OR (95% CI)</td>
<td>1 (referent)</td>
<td>0.88 (0.53–1.44)</td>
<td>0.83 (0.50–1.38)</td>
</tr>
<tr>
<td>MV-adjusted OR (95% CI)</td>
<td>1 (referent)</td>
<td>0.96 (0.57–1.62)</td>
<td>0.89 (0.52–1.53)</td>
</tr>
</tbody>
</table>

CI, confidence interval; CRC, colorectal cancer; CRP, C-reactive protein; EndoCAb, endotoxin core antibody; HPFS, Health Professionals Follow-up Study (men only); IgM, immunoglobulin M; IL-6, interleukin-6; LBP, lipopolysaccharide-binding protein; MV, multivariable; OR, odds ratio; sCD14, soluble CD14; SD, standard deviation; sTNFR-2, soluble tumor necrosis factor receptor-2. aQuartiles of the three biomarkers were defined based on the distribution in the controls. bSD for LBP: 942 ng/ml; SD for sCD14: 2.30; SD for sTNFR-2: 1.56. MMU/mI. \( P \) for trend was calculated using the median of each quartile as a continuous variable. cAdjusted for matching factors (age and time of blood draw [continuous], and race [white, non-white]). dAdditionally adjusted for body mass index (continuous), family history of CRC (yes, no), prior history of sigmoidoscopy/colonoscopy (yes, no), regular use of aspirin (yes, no), regular use of non-steroidal anti-inflammatory drugs (yes, no), pack-years of smoking (continuous), physical activity (continuous), alcohol intake, red and processed meat, dietary fiber, total folate (from foods and supplements), calcium, and vitamin D intake (all continuous), and Alternative Healthy Eating Index-2010 score without alcohol (continuous). eAdditionally adjusted for CRP, IL-6, and sTNFR-2 (all continuous). fAdditionally adjusted for CRP, IL-6, and sTNFR-2 (all continuous).
epidemiologic evidence that microbial translocation, as reflected by plasma sCD14 and potentially EndoCAb IgM, may play important roles in colorectal carcinogenesis, and if validated, may have the promise to serve as a pre-diagnostic marker.

Sparse but growing evidence support the link between sCD14 and CRC. For instance, miRNA levels of CD14 were found to be downregulated in 188 microsatellite stable CRC cases compared to normal mucosa.\(^{28}\) In addition, polymorphisms of CD14 such as rs4914 were associated with CRC risk in 479 cases and 486 controls in the Chinese Han population.\(^{29}\) We are not aware of other reports of a positive association between pre-diagnostic plasma sCD14 and subsequent risk of CRC. Notably, the association we observed for sCD14 with CRC was largely independent of systemic inflammatory markers for CRC, including CRP, IL-6, and sTNFR-2,\(^{25-27}\) indicating that sCD14 or biologic processes captured by sCD14 may independently contribute to colorectal carcinogenesis or operate through inflammatory pathways not captured by the three inflammatory markers. Comprehensive profiling of inflammatory markers is needed to elucidate the independent contribution of sCD14 and related pathways to colorectal carcinogenesis. While validations are warranted, the suggestive stronger associations for participants who are younger or have a healthier diet at blood draw, and for cases diagnosed at younger ages, if confirmed, support the promise of using sCD14 for early CRC detection among individuals not considered at high risk based on other known risk factors. Of particular interest is whether sCD14 is a pre-diagnostic marker for early-onset CRC diagnosed under age 50 or 55, which has been increasing substantially in the US\(^{30}\) and gut dysbiosis has been hypothesized to be the major driver with limited empirical evidence.\(^{31}\)

---

**Fig. 1:** Associations of sCD14 (per SD increase) with risk of CRC according to age at blood draw, time since blood draw, and putative CRC risk factors, HPFS 1993–2009. AHEI, Alternative Healthy Eating Index; BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; HPFS, Health Professionals Follow-up Study (men only); OR, odds ratio; sCD14, soluble CD14; SD, standard deviation.

**Table 3:** Association of sCD14 with risk of CRC according to age at diagnosis and tumor location, HPFS 1993–2009.
Findings on LBP and CRC have been mixed. Germline polymorphisms in LBP, such as rs2232596\(^{16}\) and rs5743533,\(^{17}\) were associated with increased risk of CRC. However, similar to our findings, a recent nested case–control study of 819 CRC from the Multiethnic Cohort (MEC) reported no association between pre-diagnostic plasma LBP levels and risk of CRC.\(^{26}\) It is hypothesized that delayed processing or long-term storage may affect LBP measurements and in part have led to a spurious outcome. Indeed, the range of LBP levels observed in the MEC samples (cases: 2–108 μg/ml; controls: 4–100 μg/ml) stretched notably higher than that typically observed in normal healthy populations (1–24 μg/ml).\(^{33}\) While LBP levels in our study sample had much narrower distributions (cases: 1.8–7.2 μg/ml; controls: 1.7–8.0 μg/ml), we similarly did not identify an association between pre-diagnostic LBP and CRC risk. Additional studies are warranted.

To our knowledge, microbial translocation reflected by suppressed levels of EndoCAb IgM was largely reported among HIV-infected population,\(^{10,13,27}\) and no prior studies have examined the association between EndoCAb IgM and CRC risk. Although not statistically significant after multivariable adjustment, we observed a suggestive inverse association between EndoCAb IgM and CRC risk. While future studies with greater power are needed, the totality of our findings supports the need to further evaluate the role of circulating markers of microbial translocation and host response to bacteria in CRC development. As indirect markers of gut barrier hyperpermeability, plasma levels of LBP, sCD14, and EndoCAb IgM predominately capture responses to LPS, a consequence of increased gut permeability and/or mucosal damage.\(^{3}\) As we increasingly recognized the role of non-bacterial species and the importance of trans kingdom interactions in CRC, studies on markers such as (1 → 3)-β-D-Glucan,\(^{34,37}\) a fungal translocation marker recently found to be correlated with LPS and sCD14, may be promising. Studies of other recently emerged surrogate markers of gut barrier dysfunction (e.g., zonulin or fatty acid binding protein-2\(^{38}\)) , markers that are crucial for the regulation of host-microbiota interplay (e.g., regenerating islet-derived protein 3 A\(^{39}\)), and direct measurement of intestinal permeability (e.g., oral probes of \(^{13}\)C-mannitol or rhamnose\(^{40}\)), along with deep characterization of the gut microbiome, may contribute further insights on the role of microbial translocation and gut barrier dysfunction in the etiology and prevention of CRC.

Our findings on the association between circulating markers of microbial translocation and risk of CRC are biologically plausible. Some putative CRC risk factors, such as high fat diet and alcohol can damage the intestinal barrier, resulting in either direct induction of epithelial cell damage and intracellular signaling responses or indirect interaction with the intestinal microbiota.\(^{3}\) The impaired gut barrier is partially characterized by increased gut permeability, which permits translocation of LPS and other pathogens into the circulation, activates TLR-4, and promotes endotoxemia and inflammation. More specifically, in the circulation, LPS is first recognized by LBP, which mediates interaction of LPS with CD14 on the cell surface; CD14 then splits LPS aggregates into monomeric molecules and presents them to the TLR4-myeloid differentiation factor 2 (MD-2) complex.\(^{31}\) Once LPS binds the TLR4-MD-2 complex and is recognized, TLR4 is activated, leading to the production of pro-inflammatory cytokines which are necessary to activate immune responses.\(^{32}\) In addition, disruption of the gut barrier contributes to a shift in the composition of the microbiota, with enrichment of deleterious bacterial populations (e.g. \(Bacteroides\)\(^{33}\)) and depletion or reduction of commensal bacterial species (e.g. \(Clostridium\)\(^{34}\)). Microbial translocation can further lead to the formation of pro-inflammatory cytokines and release of reactive oxygen species.\(^{35}\) During these processes, chronic inflammation, uncontrolled epithelial cell proliferation, and colorectal neoplasia formation can be potentiated, contributing to the initiation and/or progression of CRC.

The major strength of our study is the nested case–control design embedded with a well-characterized prospective cohort with both high-quality pre-diagnostic blood collection and long-term follow up of disease outcomes. This unparalleled setting allowed us to examine the association between three complementary biomarkers of microbial translocation with subsequent risk of incident CRC. We also implemented a rigorous study design during biomarker measurement. Notably, the CVs for each of the marker were <5% after batch correction, which are considered ideal in epidemiologic studies.\(^{40}\) Furthermore, leveraging previous measurement of well-known systemic inflammatory markers, we reported minimal mediation of the positive association between sCD14 and CRC by CRP, IL-6, and sTNFR-2, highlighting the need of in-depth investigations of both sCD14 related and inflammation related pathways. Finally, the collection of detailed information on lifestyle and medical history also allowed us to adjust for a list of potential covariates to minimize confounding. We have also rigorously adjusted for age throughout the analyses. Nonetheless, several limitations also warrant consideration. First, a single measurement of plasma markers may not represent long-term concentrations. While prior studies reported reasonable stability of sCD14 over a 2-year period,\(^{13}\) studies with repeated blood collection will allow us to evaluate the role of sCD14 trajectories in relation to CRC risk. Measurement error due to categorization of the exposure variables may also exist. However, both continuous and restricted cubic spline analyses support the positive association between sCD14 and risk of CRC. Second, residual confounding due to measurement error in the confounders (such as smoking, physical activity, and dietary intake) or...
unmeasured confounding cannot be ruled out completely. Third, when testing interaction, additive interaction is preferred in terms of both causality and public health but cannot be assessed in matched case-control studies. 27 Forth, the vast majority of our study population comprised white male health professionals, thus limiting the generalizability of the findings to women or other racial/ethnic groups. Studies are also warranted to evaluate the clinical utility of sCD14, in early detection and risk prediction.

In conclusion, in this prospective, nested case-control study of US men, higher levels of plasma sCD14 were associated with an increased risk of incident CRC and we also observed a suggestive inverse association between EndoCaB IgM and subsequent risk of CRC. Our findings support for the role of microbial translocation in colorectal carcinogenesis and point to the value of further investigations of biomarkers of these processes for CRC etiology and prevention.

Declaration of interests

AC has served as an investigator on studies supported by Pfizer, Freerome, and Zoe Ltd. outside the submitted work. AC also serves as a consultant for Pfizer, Boehringer Ingelheim, Bayer Pharma AG outside the submitted work. YC has served as a consultant for GeneoScopy outside the submitted work. 

Acknowledgments

The authors would like to acknowledge the contribution to this study from central cancer registries supported through the Centers for Disease Control and Prevention’s National Program of Cancer Registries (NPCR) and/or the National Cancer Institute’s Surveillance, Epidemiology, and End Results (SEER) Program. Central registries may also be supported by state agencies, universities, and cancer centers. Participating central cancer registries include the following: Alabama, Alaska, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, Florida, Georgia, Hawaii, Idaho, Indiana, Iowa, Kentucky, Louisiana, Massachusetts, Maine, Maryland, Michigan, Mississippi, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Puerto Rico, Rhode Island, Seattle SEER Registry, South Carolina, Tennessee, Texas, Utah, Virginia, West Virginia, Wyoming.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.ebiom.2023.104566.

References


34 Ellis RJ, Iudicello JE, Heaton RK, et al. Markers of gut barrier function and microbial translocation associate with lower gut microbial diversity in people with HIV. *Virology.* 2021;13(10).


40 Tworoger SS, Hankinson SE. Use of biomarkers in epidemiologic studies: minimizing the influence of measurement error in the study design and analysis. *Cancer Causes Control.* 2006;17(7):889–899.