# Soluble Receptor for Advanced Glycation End-products (sRAGE) and colorectal cancer risk: a case-control study nested within a European prospective cohort

Elom K. Aglago[1], Sabina Rinaldi[1], Heinz Freisling[1], Li Jiao[2], David J. Hughes[3], Veronika Fedirko[4], Casper G. Schalkwijk[5], Elisabete Weiderpass[6], Christina C. Dahm[7], Kim Overvad[7,8], Anne Kirstine Eriksen[9], Cecilie Kyrø[9], Marie-Christine Boutron-Ruault[10,11], Joseph A. Rothwell[10,11], Gianluca Severi[10,11, 12], Verena Katzke[13], Tilman Kühn[13], Matthias B. Schulze[14, 15], Krasimira Aleksandrova[15,16], Giovanna Masala[17], Vittorio Krogh[18], Salvatore Panico[19], Rosario Tumino[20], Alessio Naccarati[21,22], Bas Bueno-de-Mesquita[23], Carla H. van Gils[24], Torkjel M. Sandanger[25] , Inger Torhild Gram[26], Guri Skeie[26], J. Ramón Quirós[27], Paula Jakszyn[28,29], Maria-Jose Sánchez[30,31,32,33], Pilar Amiano[32, 34], José María Huerta[32, 35], Eva Ardanaz[32, 36, 37], Ingegerd Johansson[38], Sophia Harlid[38], Aurora Perez-Cornago[39], Ana-Lucia Mayén[1], Reynalda Cordova[1], Marc J. Gunter[1], Paolo Vineis[40], Amanda J. Cross[40], Elio Riboli[40], Mazda Jenab[1]

[1]Nutrition and Metabolism Section, International Agency for Research on Cancer (IARC), Lyon, France
[2]Department of Medicine, Baylor College of Medicine, Houston, TX, USA
[3]Cancer Biology and Therapeutics Group (CBT), Conway Institute, School of Biomolecular and Biomedical Science (SBBS), University College Dublin, Ireland
[4]Department of Epidemiology, Rollins School of Public Health, Winship Cancer Institute, Emory University, Atlanta, GA, USA
[5]Department of Internal Medicine, CARIM School for Cardiovascular Diseases, Maastricht University Medical Center, The Netherlands
[6]Office of the Director, International Agency for Research on Cancer (IARC), Lyon, France
[7]Department of Public Health, Aarhus University, Denmark
[8]Department of Cardiology, Aalborg University Hospital, Aalborg, Denmark
[9]Danish Cancer Society Research Center, Copenhagen, Denmark
[10]CESP, Faculté de médecine - Université Paris-Saclay, UVSQ, INSERM, 94805, Villejuif, France
[11]Gustave Roussy, Villejuif, 94805, France
[12]Department of Statistics, Computer Science and Applications (DISIA), University of Florence, Italy
[13]German Cancer Research Center (DKFZ), Foundation under Public Law, Heidelberg, Germany
[14]Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany
[15]Institute of Nutrition Science, University of Potsdam, Nuthetal, Germany
[16]Nutrition, Immunity and Metabolism Senior Scientist Group, Department of Nutrition and Gerontology, German Institute of Human Nutrition Potsdam-Rehbruecke (DIfE), Nuthetal, Germany
[17]Cancer Risk Factors and Life-Style Epidemiology Unit, Institute for Cancer Research, Prevention and Clinical Network - ISPRO, Florence, Italy

[18]Epidemiology and Prevention Unit, Department of Research, Fondazione IRCCS Istituto Nazionale dei Tumori Via Venezian, Milano, Italy
[19]Dipartimento di Medicina Clinica e Chirurgia Federico II University, Naples, Italy
[20]Cancer Registry and Histopathology Department, Provincial Health Authority (ASP), Ragusa, Italy
[21] Italian Institute for Genomic Medicine (IIGM), c/o IRCCS Candiolo, Candiolo, Torino, Italy
[22]Candiolo Cancer Institute, FPO-IRCCS, Candiolo, Torino, Italy
[23]Former senior scientist, Dept. for Determinants of Chronic Diseases (DCD), National Institute for Public Health and the Environment (RIVM), BA Bilthoven, The Netherlands
[24]University Medical Center Utrecht, Utrecht, the Netherlands
[25]Department of Community Medicine, Faculty of Health Sciences, UiT -The Arctic University of Norway, Tromsø, Norway
[26]Faculty of Health Sciences, Department of Community Medicine, University of Tromsø, The Arctic University of Norway
[27]Public Health Directorate, Asturias Ciriaco Miguel Vigil St, 9 33006 Oviedo, Spain
[28] Unit of Nutrition and Cancer, Cancer Epidemiology Research Programme, Catalan Institute of Oncology (ICO-IDIBELL), Barcelona, Spain
[29] Blanquerna School of Health Sciences, Ramon Llull University, Barcelona, Spain
[30]Escuela Andaluza de Salud Pública (EASP), Granada, Spain
[31]Instituto de Investigación Biosanitaria ibs.GRANADA, Granada, Spain
[32]Centro de Investigación Biomédica en Red de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain
[33]Department of Preventive Medicine and Public Health, University of Granada, Granada, Spain
[34]Public Health Division of Gipuzkoa; BioDonostia Research Institute, Donostia-San Sebastian, Spain
[35]Department of Epidemiology, Murcia Regional Health Council, IMIB-Arrixaca, Murcia, Spain.
[36]Navarra Public Health Institute, Pamplona, Spain
[37]IdiSNA, Navarra Institute for Health Research, Pamplona, Spain
[38]Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden
[39]Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford
[40]School of Public Health, Imperial College London, London, UK

**Running title:** sRAGE concentrations and colorectal cancer

**Keywords:** sRAGE, colorectal cancer, advanced glycation end-product, inflammation, polymorphism
**Financial support:** Funding [WCRF 2015/1391, Principal Investigator, Mazda Jenab] was obtained from Wereld Kanker Onderzoek Fonds (WKOF), as part of the World Cancer Research Fund International grant programme.

**Conflict of interest statement:** The authors declare no potential conflicts of interest.

**Abbreviations used:** ADAM10, A Disintegrin And Metalloproteinase Domain 10; AGE, advanced glycation end-products; AGER, Advanced Glycosylation End-Product Specific Receptor ; BMI, body mass index; CRC, colorectal cancer; CRP, C-reactive protein; CV, coefficients of variation; GLO1, Glyoxalase I; EPIC, European Prospective Investigation into Cancer and Nutrition; IARC, International Agency for Research on Cancer; mRNA, messenger ribonucleic acid ; NF-κB, nuclear factor kappa B; OR, odds ratio; RAGE, receptor for AGE; RNF5, Ring Finger Protein 5 ; SD, standard deviation; SNP, single nucleotide polymorphism; sRAGE, soluble receptor for AGE; TNFα, tumor necrosis factor alpha; WC, wait circumference; WHR, waist-to-hip ratio

**Corresponding author:** Dr. Mazda Jenab, Address: 150 Cours Albert Thomas, 69372 Lyon Cedex 08, Email: jenabm@iarc.fr, Tel: +33 472 73 80 82

**Data sharing statement:** For information on how to submit an application for gaining access to EPIC data and/or biospecimens, please follow the instructions at http://epic.iarc.fr/access/index.php

**Disclaimer:** Where authors are identified as personnel of the International Agency for Research on Cancer / World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer / World Health Organization.

Conflict of interest statement: The authors have declared no conflicts of interest.

Word count: 4,000

Number of tables and figures: 4 tables and 2 figures

**Abstract**

**Background:** Overexpression of the Receptor for Advanced Glycation End-product (RAGE) has been associated with chronic inflammation, which in turn has been associated with increased colorectal cancer (CRC) risk. Soluble RAGE (sRAGE) competes with RAGE to bind its ligands, thus potentially preventing RAGE-induced inflammation.

**Methods:** To investigate whether sRAGE and related genetic variants are associated with CRC risk, we conducted a nested case-control study in the European Prospective Investigation into Cancer and Nutrition (EPIC). Plasma sRAGE concentrations were measured by ELISA in 1,361 CRC matched case-control sets. Twenty-four single nucleotide polymorphisms (SNPs) encoded in the genes associated with sRAGE concentrations were available for 1,985 CRC cases and 2,220 controls. Multivariable-adjusted odds ratios (ORs) and 95% confidence intervals (CIs) were computed using conditional and unconditional logistic regression for CRC risk and circulating sRAGE and SNPs, respectively.

**Results:** Higher sRAGE concentrations were inversely associated with CRC ($OR_{Q5vs.Q1}$=0.77, 95%CI=0.59-1.00). Sex-specific analyses revealed that the observed inverse risk association was restricted to men ($OR_{Q5vs.Q1}$=0.63, 95%CI=0.42-0.94) whereas no association was observed in women ($OR_{Q5vs.Q1}$=1.00, 95%CI=0.68-1.48, $P_{heterogeneity}$ for sex=*0.006*). Participants carrying minor allele of rs653765 (promoter region of *ADAM10*) had lower CRC risk (C vs. T, OR=0.90; 95%CI=0.82-0.99).

**Conclusion**: Pre-diagnostic sRAGE concentrations were inversely associated with CRC risk in men but not in women. A SNP located within *ADAM10* gene pertaining to RAGE shedding, was associated with CRC risk.

**Impact:** Further studies are needed to confirm our observed sex difference in the association and better explore the potential involvement of genetic variants of sRAGE in CRC development.

## Introduction

Advanced glycation end-products (AGEs) are a heterogeneous group of molecules formed by non-enzymatic reactions between reducing sugars and proteins, lipids or nucleic acids (1). AGEs are produced endogenously, but diet and lifestyle are likely the largest contributors to the overall AGEs pool particularly from high-temperature processed food products which contain high amounts of AGEs and/or their precursors (2-4). Glycated proteins tend to become dysfunctional and agglutinate with other reacting molecules to create cross-links and aggregates which can accumulate within diverse tissues in the body (5). The accumulation of AGEs throughout the life course is thought to contribute to intracellular signalling alterations, chronic low-level inflammation and a decrease in tissue functionality (6).

AGEs are recognized by a multi-ligand cell-surface protein receptor, known as the Receptor for Advanced Glycation End-products (RAGE). RAGE consists of an extracellular N-terminal, a transmembrane helix, and an intracellular C-terminal tail (7). RAGE is expressed at low levels in most tissue types except the lung in which the expression is generally high (8). Overexpression of RAGE and its high activity have been demonstrated in various cancers including in the colon, breast, brain, prostate and in the ovaries (9). Binding of AGEs to their receptor triggers a signalling cascade leading to intracellular inflammation with activation of nuclear factor kappa B (NF-κB), increased secretion of cytokines and chemokines, and elevated production of reactive oxygen and nitrogen species (10).

Soluble RAGE (sRAGE) is a free circulating isoform of RAGE that also binds AGEs and acts as a decoy for RAGE. In contrast to RAGE, binding of AGEs to sRAGE does not induce inflammation and oxidative stress (8). Although the concentration of sRAGE is likely insufficient to bind all circulating AGEs (11), higher sRAGE levels had been associated with low inflammation and lower risk of several chronic diseases, including cancers (12). The variability in sRAGE concentrations is considerably affected by a combination of genetic and environmental factors (13). sRAGE levels have been reported to be elevated in women *vs.* men, younger *vs.* older individuals, and individuals with normal weight *vs.* with overweight and obesity (14-17). Furthermore, genetic determinants of sRAGE expression have also been identified and include single nucleotide polymorphisms (SNPs) located within Advanced Glycosylation End-Product Specific Receptor (*AGER*), A Disintegrin And Metalloproteinase Domain 10 (*ADAM10*), Glyoxalase I (*GLO1*), and Ring Finger Protein 5 (*RNF5*) genes (17-21).

We hypothesised that higher circulating sRAGE levels are inversely associated with colorectal cancer (CRC) development. Previously, only two prospective studies have investigated the association, and showed an inverse association of high sRAGE concentrations with CRC risk among Finnish male smokers (22) and women with overweight and obesity (23). However, there is sparse data from other prospective studies, and there is a need to carefully investigate possible differences in the association by sex or lifestyle factors. To address these gaps, we studied the association between pre-diagnostic levels of circulating sRAGE and risk of CRC in a large, multinational European prospective cohort. We

also investigated whether SNPs, reported to be related to sRAGE levels or RAGE function, are associated with CRC risk.

## Materials and methods

### Study population and data collection

We used a case-control design nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. EPIC is an ongoing multicentre prospective cohort with 521,324 participants (70% women) recruited from 23 study centres located in 10 European countries (Denmark, France, Germany, Greece, Italy, the Netherlands, Norway, Spain, Sweden, and the United Kingdom). The rationale and methods of the EPIC study, including information on the recruitment of the participants as well as data collection have been described previously (24). Participants gave written informed consent before joining the EPIC study. Participant's health history, anthropometry, socio-demographic and standardised lifestyle variables including education, smoking, and physical activity were collected by questionnaire at baseline, prior to disease onset or diagnosis. Physical activity was based on the Cambridge physical activity index: inactive (sedentary job and no recreational activity), moderately inactive (sedentary job with <0.5 h recreational activity per day/or standing job with no recreational activity), moderately active (sedentary job with 0.5 to 1 h recreational activity per day/ or standing job with 0.5 h recreational activity per day/ or physical job with no recreational activity) or active (sedentary job with >1 h recreational activity per day/or standing job with >0.5 h recreational activity per day/or physical job with at least some recreational activity/or heavy manual job) (25). Dietary intake was assessed at recruitment by validated centre-specific questionnaires. In each of the study centres, blood samples were drawn at recruitment (≈80% of participants provided blood samples) and stored in liquid nitrogen (−196°C, liquid nitrogen) at the International Agency for Research on Cancer (IARC) biobank, or in local biobanks (at -150°C in nitrogen vapour in Denmark; -80°C freezers at Malmö and Umeå centres in Sweden) (24).

### Follow-up for cancer incidence and vital status

Vital status follow-up (98.4% complete) is collected by record linkage with regional and/or national mortality registries in all countries except Germany and Greece, and the Italian centre of Naples, where data are collected actively. Incident cancer cases were identified through record linkage with regional cancer registries or using a combination of methods, including health insurance records, cancer and pathology registries, and active follow-up through participants and their relatives. CRC cases were eligible if they were first incident and histologically-confirmed. Cases were defined using the International Classification of Diseases for Oncology (ICD-O). Colon cancers were defined as tumours that occurred in the cecum, appendix, ascending colon, hepatic flexure, transverse colon, splenic flexure, descending and sigmoid colon (C18.0-C18.7), and overlapping and or unspecified origin tumours (C18.8 and C18.9). Rectal cancers were defined as tumours that occurred at the recto-sigmoid junction (C19) or rectum (C20). Cancers of the anal canal were excluded.

205 Case-control design

206 From baseline onwards, 1,413 first incident CRC cases with available blood samples were identified
207 (until June 2003 as endpoint) among all the total 2,476 CRC cases ascertained **(Figure 1)**. For each
208 identified case, one control was matched by incidence density sampling from all cohort members alive
209 and cancer-free at the time of diagnosis of the index case. Cases and controls were matched by age (±1
210 year), sex, centre, and blood collection details including time (±3 hours), fasting pre-venepuncture (<3,
211 3-6, and >6 hours); and additionally among women only, by menopausal status (pre-, peri-, and
212 postmenopausal), and hormone replacement therapy (HRT) use at the time of blood collection (yes/no).
213 After exclusion of participants with incomplete matched case sets (n=16), those with extreme sRAGE
214 levels (n=3 controls and 1 case with sRAGE concentrations unusually high i.e. >mean+4 standard
215 deviation), and 32 cases and matched controls from Greece due to unforeseen data restriction issues,
216 1,361 cases and 1,361 matched controls were included in the sRAGE analysis. Among EPIC
217 participants, 4,487 participants (until December 2012 as endpoint, 2,148 CRC cases and matched 2,339
218 controls) have been previously genotyped. After exclusion of 100 CRC cases and 100 matched controls
219 from Greece, and 82 participants with missing lifestyle variable, 1,985 CRC cases and 2,220 matched
220 controls were included in the genetic analysis. Among the participants who have been genotyped, 972
221 CRC cases and 767 non-cases overlap with case-control sets in whom sRAGE measurements were
222 conducted.
223
224 Laboratory analyses

225 Circulating sRAGE concentrations were measured in citrated plasma samples by ELISA (Quantikine,
226 R&D Systems, MN, USA), following the manufacturer's instructions. Previous studies have reported
227 that sRAGE is stable in plasma over a long period of time (26). Analyses were run with case-control
228 sets randomized across batches (n=40 batches, with an average of 35 case-control pairs analysed per
229 batch). Intra- and inter-batch coefficients of variation (CV) were assessed by measuring 3 different
230 samples used as quality controls in duplicate in each. Mean intra- and inter-batch CVs were 1.25% and
231 6.0%, respectively. C-reactive protein (CRP) concentrations were determined using a high-sensitivity
232 assay (Beckman-Coulter, Woerden, The Netherlands).
233
234 DNA genotyping and genetic variants selection

235 DNA was extracted from buffy coats from citrated blood samples at the Center for Inherited Disease
236 Research (CIDR, Johns Hopkins University) using the HumanOmniExpressExome-8v1–2 array as
237 described elsewhere (27). All SNPs met criteria for quality control for genotyping call rate (above 95%).
238 Candidate SNPs selected for our study were those previously associated with sRAGE levels. Most of
239 these SNPs appear to be located within the *AGER* gene, with rs2070600 being the most important and
240 explaining 22% of the variability in sRAGE concentrations in Caucasians (17). In addition to *AGER*,
241 four additional genes contain SNPs associated with sRAGE: *RNF5*, a neighbouring gene which encodes

for RAGE (28), *ADAM10* encodes for metalloproteinases involved in the shedding of RAGE ectodomain to form sRAGE (29), and *GLO1* encodes for glyoxalase enzyme responsible to metabolise methylglyoxal and prevent aberrant AGEs formation (30). The main SNPs are from *AGER* (rs2070600, rs1800625, rs1800624, rs184003, rs2854050), *ADAM10* (rs653765) and *RNF5* (rs9469089) (17-21,31-38). We additionally considered less-studied SNPs located within *AGER* (rs1035798, rs1800684, rs3131300, rs3134940, rs2269422, rs2853807, rs9391855, rs17846798), *ADAM10* (rs514049), *RNF5* (rs57409105, rs41268928, rs17493811), and *GLO1* (rs4746, rs1130534, rs1049346, rs6932648, rs10484854). The choice of this supplementary group of SNPs was based on the potential influence and interactions they may have in modulating sRAGE levels directly or through AGEs (13,17,21,31,39-41).

Genotype distributions were in Hardy-Weinberg equilibrium (cutoff of *P*-value=$1x10^{-3}$) for all the SNPs considered, with the exception of rs6932648 which was consequently excluded from the analysis. The selected SNPs and their characteristics are detailed in **Supplementary Table 1**. To select the independent variants, Linkage Disequilibrium (LD) pruning (LD≤1%) was performed using NCI LDlink tools (https://ldlink.nci.nih.gov). We found the following independent variants (highly correlated variants are in brackets): rs2070600 (rs41268928, rs9391855, rs2854050), rs1800625 (rs3131300, rs3134940), rs1800624 (rs17846798), rs4746 (rs1130534, rs10484854), rs17846798 (rs57409105), rs9469089, rs1800684, rs2269422, rs2853807, rs1049346, rs17493811, and rs653765 (rs514049). A flowchart outlining the selection of the independent SNPs is detailed in **Supplementary Figure 1**.

Among the 767 control subjects who had both sRAGE and genetic data, we assessed the association between the independent genetic variants and log-transformed sRAGE levels using linear regression models (**Supplementary Table 2**). The SNPs in the following genes were significantly associated with sRAGE levels: *AGER* (rs2070600, rs1800625), *RNF5* (rs9469089), and *GLO1* (rs4746). Although rs653765 (*ADAM10*) was not associated with sRAGE levels, we decided to conserve it in our analysis for two main reasons: first, as a major variant of metalloproteinases which are involved in the shedding of the ectodomain of RAGE to produce sRAGE; second, this variant was previously associated with sRAGE levels in other populations (21). Overall, five SNPs (rs2070600, rs1800625, rs9469089, rs4746, rs653765) were examined for the association with CRC risk.

Statistical analysis

Case-control differences in baseline characteristics were evaluated using Student's paired t-test and Wilcoxon's signed-rank test for continuous variables and Kruskal–Wallis test for categorical variables. Spearman rank correlation was used to correlate sRAGE levels to anthropometry, dietary intakes and other biomarkers. We divided sRAGE concentrations into quintiles based on the distribution in the control group. Conditional logistic regression was used to compute odds ratios (ORs) and 95% confidence intervals (CIs) for the associations between circulating levels of sRAGE and CRC risk. We ran two different models by including for each successive model additional adjustment variables

incrementally. Model 1 (crude) was conditioned on the matching factors. Model 2 was additionally adjusted for body mass index (BMI), height, education (none, primary, technical and professional, secondary, higher), physical activity (inactive, moderately inactive, moderately active, active), smoking status, duration, and intensity (never; cigarettes/day 1-<=15, 16-<=25, >26; former smokers <=10, 11-<=20, >20 years, occasional), dietary energy, and intakes of alcohol, red and processed meat, dietary fibre, and dairy products. Dietary factors included as adjustment factors have been previously associated with CRC and/or sRAGE levels (42). P-values for the linear trend (*P* for trend) were obtained by including the median value of each quintile as a continuous variable in the model. We also examined sRAGE levels as a continuous variable, per standard deviation (SD) increment.

Stratified analyses were performed by anatomical sub-sites (colon *vs*. rectal cancers, proximal colon *vs.* distal colon cancers), sex (men *vs.* women), age groups (<50, >=50-<55, >=55-<60, >=60-<65, >=65), smoking (never, former, ever), alcohol intake (tertiles), physical activity (inactive, moderately inactive, moderately active, active), BMI (<25, >=25-<30, >=30 kg/m$^2$); and below or above sex-specific recommended cut-offs for waist circumference (WC, men, 94 cm, women, 80 cm) and waist-to-hip ratio (WHR, men, 0.90, women, 0.85), and in women by menopausal status (pre-, post and perimenopause). The cut-offs for WC and WHR were based on the WHO's definitions of central adiposity in European men and women (43). Additional stratified analyses were conducted for CRP (tertiles) as a marker of inflammation. P-values for heterogeneity were calculated using the Wald test. For sub-group analyses by anthropometric measures, individual models were run for BMI, WC and WHR in men and women separately (model 2 without BMI). In sensitivity analyses, we excluded cases diagnosed during the first 2 years of follow-up and rerun the analyses.

We assessed the association between the genetic variants and CRC risk using data of all participants genotyped in EPIC to increase the statistical power of the analysis. The associations between the five independent genetic variants and CRC risk were assessed by unconditional logistic regression models. Two models were run, an unadjusted model and a multivariable-adjusted model, adjusted for sex, age, BMI, smoking status, alcohol, and country. Additive (major allele=0, heterozygotous=1, minor allele=2), dominant (major allele=0, heterozygotous+minor allele=1) and recessive models (major allele+ heterozygotous=0, minor allele=1) were run for the genetic variants. In sensitivity analyses, we analysed the participants with overlapping genetic and sRAGE concentrations data. All the statistical analyses were performed using Stata 14.0 (StataCorp, College Station, TX, USA). *P*-values <0.05 was considered statistically significant.

## Results

Baseline characteristics and sRAGE levels in cases and controls are presented in **Table 1**. Compared to controls, CRC cases have higher BMI, WC, WHR and CRP concentrations, and consume more alcohol and less dairy products and fruit and vegetables. sRAGE concentrations were slightly lower in CRC cases than controls (1086 *versus* 1130 pg/mL) but this was mainly observed among men (982 versus 1066 pg/mL in male cases versus controls, respectively); whereas among women sRAGE was 1185 pg/mL in cases and 1191 pg/mL in controls. BMI, WC, WHR, and alcohol intake were all negatively correlated with sRAGE levels whereas sugar and confectionaries, fruit and vegetable, and cereals intakes showed positive correlations (**Supplementary Table 3).** Women with higher sRAGE levels have lower CRP concentrations (Spearman rho=-0.156, p=0.004).

sRAGE and CRC risk

sRAGE concentrations were inversely associated with CRC risk in multivariable-adjusted analyses (OR comparing the highest to the lowest quintile $OR_{Q5vs.Q1}$=0.75, 95%CI=0.58-0.98, $P_{trend}$=*0.035*, **Table 2**). Sub-group analyses by sex showed an inverse risk association for men ($OR_{Q5vs.Q1}$=0.63, 95%CI=0.42-0.94, $P_{trend}$=*0.001*) but not in women ($OR_{Q5vs.Q1}$=0.94, 95%CI=0.63-1.38, $P_{trend}$=*0.754*; $P_{heterogeneity}$=*0.006*). In men, sRAGE was associated with a lower risk of both colon cancer (OR per SD increment, OR =0.84, 95%CI=0.70-0.99) and rectal cancer (OR=0.80, 95%CI=0.64-0.99) with no heterogeneity across anatomical subsites ($P_{heterogeneity}$=*0.607*) (**Table 3**). The magnitude of the inverse association appeared stronger for distal colon cancer (OR=0.61, 95%CI=0.44-0.84) compared to proximal cancer (OR=0.94, 95%CI=0.69-1.29) but no heterogeneity was observed ($P_{heterogeneity}$=*0.671*). In women, no association was found between sRAGE and colon (OR=0.99, 95%CI=0.85-1.15) or rectal cancer (OR=1.06, 95%CI=0.86-1.32). Stratified analyses by age groups, BMI categories, WC and WHR cut-offs, and smoking status showed no significant differences across strata (**Figure 2**). Women in higher CRP tertiles tended to have higher CRC risk associated with sRAGE ($P_{heterogeneity}$ across=*0.011*) (**Figure 2**).

Analyses of genetic variants

**Table 4** presents the association of the genetic variants with CRC risk. While comparing minor allele *vs.* major allele, rs1800625 (*AGER*, G *vs.* A, OR=1.15, 95%CI=1.02-1.29) was associated with an increased risk of CRC whereas rs653765 (*ADAM10*, C *vs.* T, OR=0.88; 95%CI=0.80-0.97) was associated with a lower CRC risk, in univariate models. After multivariate adjustments, the association remained statistically significant for rs653765 (*ADAM10*, C *vs.* T, OR=0.90; 95%CI=0.82-0.99), but not for rs1800625 (*AGER*, G *vs.* A, OR=1.11, 95%CI=0.99-1.25).

Sensitivity analysis

Exclusion of the cases that occurred within the first two years of follow-up did not change the associations between sRAGE concentrations and CRC (**Table 1**). The associations between SNPs and CRC in participants with overlapping genetic and sRAGE data showed similar, but no statistically significant associations for rs653765 (*ADAM10*, OR=0.90, 95%CI=0.78-1.05) or rs1800625 (*AGER*, G *vs.* A, OR=1.00, 95%CI=0.83-1.19) (**Supplementary Table 4**).

## Discussion

In this large, case-control study nested within a European prospective cohort, we found that pre-diagnostic circulating sRAGE levels were inversely associated with CRC risk in men but not in women. The associations observed between sRAGE and CRC did not vary by age, or by lifestyle factors including obesity and smoking status, suggesting that sex is the main effect modifier in the association between sRAGE and CRC. With respect to the SNP analyses, we found that the minor allele of rs653765 (*ADAM10*) was inversely associated with risk of CRC, whereas an increased risk was suggested for rs1800625 *(AGER)*. However, we did not observe the association between rs653765 and levels of sRAGE.

RAGE is a pattern recognition receptor that recognizes multiple ligands such as S100, high mobility group box 1 protein (HMGB1), amyloid-β peptide, in addition to the AGEs (44). RAGE is overexpressed in several diseases of the colon, including inflammatory bowel diseases (45). RAGE action in colon tissues may participate in CRC tumour initiation, progression and invasion (46-48). sRAGE by acting as a decoy of RAGE, binds to AGEs in the circulation and clears them by decreasing interaction with full-length cell-surface RAGE. The evidence from mouse studies shows that injection of sRAGE is associated with a reduction in the expression of inflammatory mediators such as TNF-α (49). Evidence from case-control studies also shows that elevated sRAGE levels are associated with a lower risk of several cancers including liver (50) and pancreatic cancer (51). This suggests that higher concentrations of sRAGE are protective against AGEs-induced inflammation which is involved in the aetiology of various chronic diseases such as diabetes and cancers, but the mechanisms need further exploration.

The underlying reasons for the observed difference between men and women in the association between sRAGE and CRC risk are unclear. Several previously published studies that compared sRAGE levels between men and women suggest higher circulating levels in women (14,15,17), which we also observed in our study. One explanation of the sex difference in sRAGE levels may be that oestrogens stimulate sRAGE expression and production (52). Oestrogens have also been reported to reduce AGEs production and AGEs-related inflammation (53). In our study, women with higher sRAGE levels have lower CRP concentrations (Spearman rho=-0.156, p=0.004) and lower CRC risk, suggesting that sRAGE may possibly reduce CRC risk in women, by mitigating overall inflammation. However, analysis by menopausal status showed no differences across strata in our study population. Our findings suggest that additional studies are needed to understand the physiological sex differences in sRAGE levels and how they may translate into the differential CRC risk associations that we have observed in this study.

Interestingly, the two previous publications on sRAGE and CRC in prospective cohorts have been conducted in men (22) and in women (23) only. The Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC) study reported high serum sRAGE to be associated with low CRC risk in Finnish male smokers (22). We expanded this observation by showing that such an inverse association was also

observed in male never smokers. We expected to observe a greater reduction in CRC risk in non-smokers compared to smokers, but our findings did not differ by smoking status. Smoking may be a source of AGEs exposure (2), but the magnitude of the contribution of smoking to overall AGEs exposures remains to be explored. sRAGE levels have been reported to be higher, lower or unchanged in smokers compared to non-smokers (54-56). It is still unknown whether smoking could induce an adaptive mechanism of sRAGE synthesis to cope with sustained formation of AGEs from glycotoxins contained in cigarettes. In a previous nested case-control study on a subsample of 1,249 postmenopausal women in the Women's Health Initiative (WHI) study, higher sRAGE levels were observed to be associated with lower CRC risk in individuals with overweight and obesity, but not among normal weight postmenopausal women (23). Overall, our findings showed that sRAGE levels were associated with an inverse risk of CRC only in men, with no difference in magnitude across smoking status or any other lifestyle factor.

We found that rs653765 located within *ADAM10* (C vs. T) was associated with lower risk for CRC. However, rs653765 (*ADAM10*) was not associated with sRAGE levels in our study, in contrast to previous studies in which the minor allele of rs653765 was associated with lower sRAGE levels (21). Another SNP, rs1800625, located in the promoter region of *AGER* is involved in the initiation of the production of the RAGE or its isomers (39). Xu et al. (57) reported in a meta-analysis of 18 case-control genetic studies that the recessive model of rs1800625 was associated with an increase of overall cancer risk, while analysing case-controls studies of 6246 cases of renal, lung, breast, cervical, liver, oral, breast and CRC cancers. Although our findings with genetic variants are intriguing, they may be attributed to the diversity of functions associated with the *AGER* and *ADAM10* genes. The production of sRAGE through the shedding of RAGE is dependant of ADAM10 levels. Thus, the overexpression of *AGER* coupled with lower ADAM10 activity will result in higher transmembrane RAGE and lower circulating sRAGE levels. This suggests that the interactions between AGER and ADAM10 may provide a better understanding of the genetic implications of RAGE and sRAGE in CRC development. In addition, the associations observed with the genetic data could be explained by other functions of the SNPs examined, particularly in the case of ADAM10 when considering its multiple actions such as the formation of amyloid inclusions and the cleavage of a range of proteins (58). We did not observe a significant association between rs2070600 (*AGER*) and CRC, albeit our study showed that the major allele (C allele) of this SNP associates with higher sRAGE levels. A meta-analysis of 15 case-control studies showed that homozygous minor allele of this SNP was associated with an increased risk of all cancers (59). The absence of association of this SNP with CRC may be due to low statistical power, particularly as carriers of the minor allele are rare. Additional studies, using genetic data from larger research consortia, are needed to explore the link between the expression of *AGER*, *ADAM10,* and *RNF5* genes, and levels of sRAGE and CRC initiation and development.

The strengths of our study include the large number of cases and controls, the prospective design and the availability of dietary and lifestyle factors and genetic variants. Our study was, however,

limited by the fact that we did not differentiate between endogenous secretory RAGE (esRAGE), and proteolytically cleaved RAGE (cRAGE), the two components of sRAGE. esRAGE is formed by alternative splicing of RAGE mRNA, and cRAGE is produced by the shedding of the ectodomain of RAGE par metalloproteinases located at the surface of the cells. esRAGE is stable throughout the life course whereas cRAGE levels vary with age and with environmental factors (60). Because we have measured the total pool of plasma sRAGE we therefore cannot discern whether the different variants of sRAGE have specific and potentially opposite associations with study outcomes. Although the variability of cRAGE makes it a poor biomarker for a prospective study, cRAGE levels data would have permitted us to explore the association between SNPs from the *ADAM10* gene, levels of cRAGE and CRC risk. Our study was also limited by the fact that lifestyle factors and blood samples were collected at the recruitment, and may not necessarily reflect changes over years. Moreover, we cannot rule out residual confounding or unmeasured confounders such as lifetime history of anti-inflammatory medication use.

In conclusion, we observed that pre-diagnostic circulating sRAGE levels were inversely associated with CRC risk in men, but not among women. We also found that the minor allele of rs653765 (*ADAM10*) was inversely associated with CRC risk. Additional studies are, however, required to further investigate how genetic variation and sex may affect sRAGE levels or modify its association with CRC risk.

**Acknowledgements:** The authors would like to thank the EPIC study participants and staff for their valuable contribution to this research. The authors would also like to especially thank Mr. Bertrand Hemon and Dr. Aurelie Moskal for their support in preparing the databases and providing technical support pertaining to the data analysis; along with Ms. Audrey Brunat-Manquat for her assistance with the laboratory analyses for sRAGE.

The coordination of EPIC is financially supported by the European Commission (DG-SANCO); and the International Agency for Research on Cancer. The national cohorts are supported by Danish Cancer Society (Denmark); Ligue Contre le Cancer; Institut Gustave Roussy; Mutuelle Générale de l'Education Nationale; and Institut National de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Aid, German Cancer Research Center (DKFZ), and Federal Ministry of Education and Research (BMBF) (Germany); Italian Association for Research on Cancer (AIRC); National Research Council; and Associazione Iblea per la Ricerca Epidemiologica (AIRE-ONLUS) Ragusa, Associazione Volontari Italiani Sangu (AVIS) Ragusa, Sicilian Government (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS); Netherlands Cancer Registry (NKR); LK Research Funds; Dutch Prevention Funds; Dutch ZON (Zorg Onderzoek Nederland); World Cancer Research Fund (WCRF); and Statistics Netherlands (the Netherlands); Health Research Fund (FIS); Regional Governments of Andalucía, Asturias, Basque Country, Murcia (No. 6236) and Navarra; and the Centro de Investigación Biomédica en Red en Epidemiología y Salud Pública and Instituto de Salud Carlos II (ISCIII RETIC) (RD06/0020) (Spain); Health Research Fund (FIS) - Instituto de Salud Carlos III (ISCIII), Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, and the Catalan Institute of Oncology - ICO (Spain); Swedish Cancer Society; Swedish Scientific Council; and Regional Government of Skåne and Västerbotten (Sweden); Cancer Research UK; Medical Research Council; Stroke Association; British Heart Foundation; Department of Health; Food Standards Agency; and the Wellcome Trust (UK). Cancer Research UK (14136 to EPIC-Norfolk; C570/A16491 and C8221/A19170 and C8221/A29017 to EPIC-Oxford), Medical Research Council (1000143 to EPIC-Norfolk, MR/M012190/1 to EPIC-Oxford) (United Kingdom). The EPIC-Norfolk study (DOI 10.22025/2019.10.105.00004) has received funding from the Medical Research Council (MR/N003284/1 and MC-UU_12015/1) and Cancer Research UK (C864/A14136). We are grateful to all the participants who have been part of the project and to the many members of the study teams at the University of Cambridge who have enabled this research. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

This work was partially financially supported by the Fondation de France (FDF, grant, no 00081166; to HF and RC, and FDF grant no. 00089811, ALM).

## References


1. Gugliucci A. Formation of Fructose-Mediated Advanced Glycation End Products and Their Roles in Metabolic and Inflammatory Diseases. Advances in nutrition (Bethesda, Md) **2017**;8(1):54-62 doi 10.3945/an.116.013912.
2. Cerami C, Founds H, Nicholl I, Mitsuhashi T, Giordano D, Vanpatten S*, et al.* Tobacco smoke is a source of toxic reactive glycation products. Proc Natl Acad Sci U S A **1997**;94(25):13915-20.
3. Scheijen J, Clevers E, Engelen L, Dagnelie PC, Brouns F, Stehouwer CDA*, et al.* Analysis of advanced glycation endproducts in selected food items by ultra-performance liquid chromatography tandem mass spectrometry: Presentation of a dietary AGE database. Food Chem **2016**;190:1145-50 doi 10.1016/j.foodchem.2015.06.049.
4. Takeuchi M, Takino J-I, Furuno S, Shirai H, Kawakami M, Muramatsu M*, et al.* Assessment of the Concentrations of Various Advanced Glycation End-Products in Beverages and Foods That Are Commonly Consumed in Japan. PloS one **2015**;10:e0118652 doi 10.1371/journal.pone.0118652.
5. Rabbani N, Thornalley PJ. Advanced glycation end products in the pathogenesis of chronic kidney disease. Kidney international **2018**;93(4):803-13 doi 10.1016/j.kint.2017.11.034.
6. Semba RD, Nicklett EJ, Ferrucci L. Does accumulation of advanced glycation end products contribute to the aging phenotype? J Gerontol A Biol Sci Med Sci **2010**;65(9):963-75 doi 10.1093/gerona/glq074.
7. Ott C, Jacobs K, Haucke E, Navarrete Santos A, Grune T, Simm A. Role of advanced glycation end products in cellular signaling. Redox biology **2014**;2:411-29 doi 10.1016/j.redox.2013.12.016.
8. Riehl A, Nemeth J, Angel P, Hess J. The receptor RAGE: Bridging inflammation and cancer. Cell communication and signaling : CCS **2009**;7:12 doi 10.1186/1478-811x-7-12.
9. Hsieh HL, Schafer BW, Sasaki N, Heizmann CW. Expression analysis of S100 proteins and RAGE in human tumors using tissue microarrays. Biochem Biophys Res Commun **2003**;307(2):375-81 doi 10.1016/s0006-291x(03)01190-2.
10. Turner DP. Advanced glycation end-products: a biological consequence of lifestyle contributing to cancer disparity. Cancer Res **2015**;75(10):1925-9 doi 10.1158/0008-5472.can-15-0169.
11. Yamagishi S, Matsui T. Soluble form of a receptor for advanced glycation end products (sRAGE) as a biomarker. Frontiers in bioscience (Elite edition) **2010**;2:1184-95 doi 10.2741/e178.
12. Ramasamy R, Yan SF, Schmidt AM. RAGE: therapeutic target and biomarker of the inflammatory response--the evidence mounts. Journal of leukocyte biology **2009**;86(3):505-12 doi 10.1189/jlb.0409230.
13. Duan Z, Chen G, Chen L, Stolzenberg-Solomon R, Weinstein SJ, Mannisto S*, et al.* Determinants of concentrations of N(epsilon)-carboxymethyl-lysine and soluble receptor for advanced glycation end products and their associations with risk of pancreatic cancer. International journal of molecular epidemiology and genetics **2014**;5(3):152-63.
14. Prakash J, Pichchadze G, Trofimov S, Livshits G. Age and genetic determinants of variation of circulating levels of the receptor for advanced glycation end products (RAGE) in the general human population. Mechanisms of ageing and development **2015**;145:18-25 doi 10.1016/j.mad.2015.01.001.
15. Norata GD, Garlaschelli K, Grigore L, Tibolla G, Raselli S, Redaelli L*, et al.* Circulating soluble receptor for advanced glycation end products is inversely associated with body mass index and waist/hip ratio in the general population. Nutrition, metabolism, and cardiovascular diseases : NMCD **2009**;19(2):129-34 doi 10.1016/j.numecd.2008.03.004.
16. Moriya S, Yamazaki M, Murakami H, Maruyama K, Uchiyama S. Two soluble isoforms of receptors for advanced glycation end products (RAGE) in carotid atherosclerosis: the

difference of soluble and endogenous secretory RAGE. Journal of stroke and cerebrovascular diseases : the official journal of National Stroke Association **2014**;23(10):2540-6 doi 10.1016/j.jstrokecerebrovasdis.2014.05.037.
17. Maruthur NM, Li M, Halushka MK, Astor BC, Pankow JS, Boerwinkle E*, et al.* Genetics of Plasma Soluble Receptor for Advanced Glycation End-Products and Cardiovascular Outcomes in a Community-based Population: Results from the Atherosclerosis Risk in Communities Study. PloS one **2015**;10(6):e0128452 doi 10.1371/journal.pone.0128452.
18. Salonen KM, Ryhanen SJ, Forbes JM, Harkonen T, Ilonen J, Laine AP*, et al.* Circulating concentrations of soluble receptor for AGE are associated with age and AGER gene polymorphisms in children with newly diagnosed type 1 diabetes. Diabetes care **2014**;37(7):1975-81 doi 10.2337/dc13-3049.
19. Gaens KH, Ferreira I, van der Kallen CJ, van Greevenbroek MM, Blaak EE, Feskens EJ*, et al.* Association of polymorphism in the receptor for advanced glycation end products (RAGE) gene with circulating RAGE levels. J Clin Endocrinol Metab **2009**;94(12):5174-80 doi 10.1210/jc.2009-1067.
20. Lim SC, Dorajoo R, Zhang X, Wang L, Ang SF, Tan CSH*, et al.* Genetic variants in the receptor for advanced glycation end products (RAGE) gene were associated with circulating soluble RAGE level but not with renal function among Asians with type 2 diabetes: a genome-wide association study. Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association **2017**;32(10):1697-704 doi 10.1093/ndt/gfw263.
21. Huang WH, Chen W, Jiang LY, Yang YX, Yao LF, Li KS. Influence of ADAM10 Polymorphisms on Plasma Level of Soluble Receptor for Advanced Glycation End Products and The Association With Alzheimer's Disease Risk. Front Genet **2018**;9:540 doi 10.3389/fgene.2018.00540.
22. Jiao L, Taylor PR, Weinstein SJ, Graubard BI, Virtamo J, Albanes D*, et al.* Advanced glycation end products, soluble receptor for advanced glycation end products, and risk of colorectal cancer. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology **2011**;20(7):1430-8 doi 10.1158/1055-9965.epi-11-0066.
23. Chen L, Duan Z, Tinker L, Sangi-Haghpeykar H, Strickler H, Ho GY*, et al.* A prospective study of soluble receptor for advanced glycation end-products and colorectal cancer risk in postmenopausal women. Cancer epidemiology **2016**;42:115-23 doi 10.1016/j.canep.2016.04.004.
24. Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M*, et al.* European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. Public Health Nutr **2002**;5(6b):1113-24 doi 10.1079/phn2002394.
25. Wareham NJ, Jakes RW, Rennie KL, Mitchell J, Hennings S, Day NE. Validity and repeatability of the EPIC-Norfolk Physical Activity Questionnaire. International journal of epidemiology **2002**;31(1):168-74 doi 10.1093/ije/31.1.168.
26. Wu F, Afanasyeva Y, Zeleniuch-Jacquotte A, Zhang J, Schmidt AM, Chen Y. Temporal reliability of serum soluble and endogenous secretory receptors for advanced glycation end-products (sRAGE and esRAGE) in healthy women. Cancer causes & control : CCC **2018**;29(10):901-5 doi 10.1007/s10552-018-1066-4.
27. Huyghe JR, Bien SA, Harrison TA, Kang HM, Chen S, Schmit SL*, et al.* Discovery of common and rare genetic risk variants for colorectal cancer. Nature genetics **2019**;51(1):76-87 doi 10.1038/s41588-018-0286-6.
28. Sessa L, Gatti E, Zeni F, Antonelli A, Catucci A, Koch M*, et al.* The receptor for advanced glycation end-products (RAGE) is only present in mammals, and belongs to a family of cell adhesion molecules (CAMs). PloS one **2014**;9(1):e86903 doi 10.1371/journal.pone.0086903.
29. Raucci A, Cugusi S, Antonelli A, Barabino SM, Monti L, Bierhaus A*, et al.* A soluble form of the receptor for advanced glycation endproducts (RAGE) is produced by proteolytic cleavage of

the membrane-bound form by the sheddase a disintegrin and metalloprotease 10 (ADAM10). FASEB journal : official publication of the Federation of American Societies for Experimental Biology **2008**;22(10):3716-27 doi 10.1096/fj.08-109033.
30. Schalkwijk CG, Stehouwer CDA. Methylglyoxal, a Highly Reactive Dicarbonyl Compound, in Diabetes, Its Vascular Complications, and Other Age-Related Diseases. Physiological reviews **2020**;100(1):407-61 doi 10.1152/physrev.00001.2019.
31. Chocholaty M, Jachymova M, Schmidt M, Havlova K, Krepelova A, Zima T*, et al.* Polymorphisms of the receptor for advanced glycation end-products and glyoxalase I in patients with renal cancer. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine **2015**;36(3):2121-6 doi 10.1007/s13277-014-2821-0.
32. Krechler T, Jachymova M, Mestek O, Zak A, Zima T, Kalousova M. Soluble receptor for advanced glycation end-products (sRAGE) and polymorphisms of RAGE and glyoxalase I genes in patients with pancreas cancer. Clinical biochemistry **2010**;43(10-11):882-6 doi 10.1016/j.clinbiochem.2010.04.004.
33. Li T, Qin W, Liu Y, Li S, Qin X, Liu Z. Effect of RAGE gene polymorphisms and circulating sRAGE levels on susceptibility to gastric cancer: a case-control study. Cancer cell international **2017**;17:19 doi 10.1186/s12935-017-0391-0.
34. Adams JN, Raffield LM, Martelle SE, Freedman BI, Langefeld CD, Carr JJ*, et al.* Genetic analysis of advanced glycation end products in the DHS MIND study. Gene **2016**;584(2):173-9 doi 10.1016/j.gene.2016.02.029.
35. Daborg J, von Otter M, Sjolander A, Nilsson S, Minthon L, Gustafson DR*, et al.* Association of the RAGE G82S polymorphism with Alzheimer's disease. Journal of neural transmission (Vienna, Austria : 1996) **2010**;117(7):861-7 doi 10.1007/s00702-010-0437-0.
36. Yamaguchi K, Iwamoto H, Horimasu Y, Ohshimo S, Fujitaka K, Hamada H*, et al.* AGER gene polymorphisms and soluble receptor for advanced glycation end product in patients with idiopathic pulmonary fibrosis. Respirology (Carlton, Vic) **2017**;22(5):965-71 doi 10.1111/resp.12995.
37. Miyashita M, Watanabe T, Ichikawa T, Toriumi K, Horiuchi Y, Kobori A*, et al.* The regulation of soluble receptor for AGEs contributes to carbonyl stress in schizophrenia. Biochem Biophys Res Commun **2016**;479(3):447-52 doi 10.1016/j.bbrc.2016.09.074.
38. Peculis R, Konrade I, Skapare E, Fridmanis D, Nikitina-Zake L, Lejnieks A*, et al.* Identification of glyoxalase 1 polymorphisms associated with enzyme activity. Gene **2013**;515(1):140-3 doi 10.1016/j.gene.2012.11.009.
39. Serveaux-Dancer M, Jabaudon M, Creveaux I, Belville C, Blondonnet R, Gross C*, et al.* Pathological Implications of Receptor for Advanced Glycation End-Product (AGER) Gene Polymorphism. Disease markers **2019**;2019:2067353 doi 10.1155/2019/2067353.
40. Jabaudon M, Berthelin P, Pranal T, Roszyk L, Godet T, Faure JS*, et al.* Receptor for advanced glycation end-products and ARDS prediction: a multicentre observational study. Sci Rep **2018**;8(1):2603 doi 10.1038/s41598-018-20994-x.
41. Kalousova M, Jachymova M, Germanova A, Kubena AA, Tesar V, Zima T. Genetic predisposition to advanced glycation end products toxicity is related to prognosis of chronic hemodialysis patients. Kidney & blood pressure research **2010**;33(1):30-6 doi 10.1159/000285845.
42. Ivancovsky-Wajcman D, Zelber-Sagi S, Fliss Isakov N, Webb M, Zemel M, Shibolet O*, et al.* Serum Soluble Receptor for AGE (sRAGE) Levels Are Associated With Unhealthy Lifestyle and Nonalcoholic Fatty Liver Disease. Clinical and translational gastroenterology **2019**;10(5):1-10 doi 10.14309/ctg.0000000000000040.
43. WHO. Waist circumference and waist–hip ratio: report of a WHO expert consultation. Geneva: World Health Organization; 2008. 39 p.

44. Leclerc E, Fritz G, Vetter SW, Heizmann CW. Binding of S100 proteins to RAGE: an update. Biochimica et biophysica acta **2009**;1793(6):993-1007 doi 10.1016/j.bbamcr.2008.11.016.
45. Ciccocioppo R, Vanoli A, Klersy C, Imbesi V, Boccaccio V, Manca R*, et al.* Role of the advanced glycation end products receptor in Crohn's disease inflammation. World journal of gastroenterology **2013**;19(45):8269-81 doi 10.3748/wjg.v19.i45.8269.
46. Kuniyasu H, Chihara Y, Kondo H. Differential effects between amphoterin and advanced glycation end products on colon cancer cells. International journal of cancer **2003**;104(6):722-7 doi 10.1002/ijc.11016.
47. Sakellariou S, Fragkou P, Levidou G, Gargalionis AN, Piperi C, Dalagiorgou G*, et al.* Clinical significance of AGE-RAGE axis in colorectal cancer: associations with glyoxalase-I, adiponectin receptor expression and prognosis. BMC cancer **2016**;16:174 doi 10.1186/s12885-016-2213-5.
48. Yilmaz Y, Yonal O, Eren F, Atug O, Hamzaoglu HO. Serum levels of soluble receptor for advanced glycation endproducts (sRAGE) are higher in ulcerative colitis and correlate with disease activity. Journal of Crohn's & colitis **2011**;5(5):402-6 doi 10.1016/j.crohns.2011.03.011.
49. Chen Y, Yan SS, Colgan J, Zhang HP, Luban J, Schmidt AM*, et al.* Blockade of late stages of autoimmune diabetes by inhibition of the receptor for advanced glycation end products. J Immunol **2004**;173(2):1399-405 doi 10.4049/jimmunol.173.2.1399.
50. Moy KA, Jiao L, Freedman ND, Weinstein SJ, Sinha R, Virtamo J*, et al.* Soluble receptor for advanced glycation end products and risk of liver cancer. Hepatology (Baltimore, Md) **2013**;57(6):2338-45 doi 10.1002/hep.26264.
51. White DL, Hoogeveen RC, Chen L, Richardson P, Ravishankar M, Shah P*, et al.* A prospective study of soluble receptor for advanced glycation end products and adipokines in association with pancreatic cancer in postmenopausal women. Cancer medicine **2018**;7(5):2180-91 doi 10.1002/cam4.1426.
52. Mukherjee TK, Reynolds PR, Hoidal JR. Differential effect of estrogen receptor alpha and beta agonists on the receptor for advanced glycation end product expression in human microvascular endothelial cells. Biochimica et biophysica acta **2005**;1745(3):300-9 doi 10.1016/j.bbamcr.2005.03.012.
53. Lin J, Zuo G-Y, Xu Y, Xiong J, Zheng Z, Wang S*, et al.* Estrogen Reduces Advanced Glycation End Products Induced HUVEC Inflammation Via NF-kappa B Pathway. Latin American Journal of Pharmacy **2014**;33:93-100.
54. Pouwels SD, Klont F, Kwiatkowski M, Wiersma VR, Faiz A, van den Berge M*, et al.* Reply to Biswas: Acute and Chronic Effects of Cigarette Smoking on sRAGE. Am J Respir Crit Care Med **2019**;199(6):806-7 doi 10.1164/rccm.201812-2257LE.
55. Pouwels SD, Klont F, Kwiatkowski M, Wiersma VR, Faiz A, van den Berge M*, et al.* Cigarette Smoking Acutely Decreases Serum Levels of the Chronic Obstructive Pulmonary Disease Biomarker sRAGE. Am J Respir Crit Care Med **2018**;198(11):1456-8 doi 10.1164/rccm.201807-1249LE.
56. Biswas SK. Acute and Chronic Effects of Cigarette Smoking on sRAGE. Am J Respir Crit Care Med **2019**;199(6):805 doi 10.1164/rccm.201811-2169LE.
57. Xu Y, Lu Z, Shen N, Wang X. Association of RAGE rs1800625 Polymorphism and Cancer Risk: A Meta-Analysis of 18 Case-Control Studies. Medical science monitor : international medical journal of experimental and clinical research **2019**;25:7026-34 doi 10.12659/msm.916260.
58. Przemyslaw L, Boguslaw HA, Elzbieta S, Malgorzata SM. ADAM and ADAMTS family proteins and their role in the colorectal cancer etiopathogenesis. BMB reports **2013**;46(3):139-50 doi 10.5483/bmbrep.2013.46.3.176.
59. Huang Q, Mi J, Wang X, Liu F, Wang D, Yan D*, et al.* Genetically lowered concentrations of circulating sRAGE might cause an increased risk of cancer: Meta-analysis using Mendelian

randomization. The Journal of international medical research **2016**;44(2):179-91 doi 10.1177/0300060515617869.

60. Rebholz CM, Astor BC, Grams ME, Halushka MK, Lazo M, Hoogeveen RC*, et al.* Association of plasma levels of soluble receptor for advanced glycation end products and risk of kidney disease: the Atherosclerosis Risk in Communities study. Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association **2015**;30(1):77-83 doi 10.1093/ndt/gfu282.

694 **Table 1**: Selected baseline demographic and lifestyle characteristics of study participants by
695 colorectal cancer status, EPIC study 1992-2012

| | Cases (n=1,361) | Controls (n=1,361) | P-value* |
|---|---|---|---|
| Women, % | 51.5 | 51.7 | |
| Age, years, mean±SD | 58.4±7.35 | 58.3±7.38 | *0.877* |
| Anthropometry, mean±SD | | | |
| BMI, kg/m² | 26.7±4.25 | 26.2±3.74 | *0.004* |
| Waist circumference, cm | 90.4±13.0 | 88.3±12.1 | *<0.001* |
| Waist-to-hip ratio | 0.88±0.10 | 0.87±0.10 | *0.001* |
| Lifestyle variables, n (%) | | | |
| Smoking status and intensity | | | |
| Never | 514 (37.9) | 542 (39.8) | *0.703* |
| Current, 1-<=15 cig/day | 129 (9.51) | 139 (10.2) | |
| Current, 16-<=25 cig/day | 87 (6.40) | 94 (6.91) | |
| Current, >26 cig/day | 20 (1.47) | 23 (1.69) | |
| Former, quit <= 10 years | 139 (10.3) | 129 (9.48) | |
| Former, quit 11-<=20 years | 144 (10.6) | 123 (9.04) | |
| Former, quit >20 years | 166 (12.2) | 177 (13.0) | |
| Current, pipe/cigar/occasional | 125 (9.22) | 102 (7.49) | |
| Physical activity | | | |
| Inactive | 343 (25.4) | 307 (22.6) | *0.057* |
| Moderately inactive | 439 (32.4) | 446 (32.3) | |
| Moderately active | 307 (22.7) | 282 (20.8) | |
| Active | 264 (19.5) | 321 (23.7) | |
| Highest education level attained | | | |
| None | 68 (5.01) | 66 (4.85) | *0.275* |
| Primary school completed | 453 (33.4) | 490 (36.0) | |
| Technical/professional school | 324 (23.9) | 343 (25.2) | |
| Secondary school | 217 (16.0) | 184 (13.5) | |
| Higher education | 247 (18.2) | 244 (17.9) | |
| Dietary intake, mean (SD) | | | |
| Energy, Kcal/day | 2124±620 | 2127±609 | *0.764* |
| Alcohol, g/day | 17.0±22.1 | 15.4±19.7 | *0.040* |
| Red and processed meats, g/day | 87.6±53.1 | 85.1±52.0 | *0.215* |
| Fruits and vegetables, g/day | 396±233 | 421±248 | *0.007* |
| Cereals, g/day | 216±121 | 216±119 | *0.941* |
| Dairy products, g/day | 331±251 | 351±244 | *0.042* |
| Fish, g/day | 28.2±28.8 | 29.6±30.6 | *0.226* |
| Sugar and confectionaries, g/day | 48.7±66.6 | 48.7±68.9 | *0.995* |
| Fat, g/day | 28.3±15.6 | 27.9±16.0 | *0.536* |
| Protein, g/day | 89.3±27.9 | 90.3±27.5 | *0.337* |
| Biomarkers | | | |
| CRP, ng/mL† | 4013±6011 | 3433±5607 | *0.026* |
| sRAGE levels, mean±SD, pg/mL | | | |
| All participants | 1086±469 | 1130±470 | *0.015* |
| Men | 982±431 | 1066±438 | *<0.001* |

| | | | | |
|---|---|---|---|---|
| | Women | 1185±483 | 1191±490 | *0.831* |

Frequencies may not add up to 100% due to missing data

Abbreviations: AGE, Advanced glycation end products; BMI, body mass index; sRAGE, soluble receptor for advanced glycation end-products

*Student's paired t-test and Wilcoxon's signed-rank test for continuous variables and Kruskal–Wallis test for categorical variables

[†]CRP was available for 1103 cases and 925 controls

**Table 2:** Odds ratios (OR) and 95% confidence intervals for colorectal cancer risk associated with circulating sRAGE (Quintiles and continuous), EPIC study 1992-2012

| | Quintiles of sRAGE (cutpoints, in pg/mL) * | | | | | $P_{trend}$ | Continuous, per SD | Continuous, per SD[†] |
|---|---|---|---|---|---|---|---|---|
| | Quintile 1 (<754) | Quintile 2 (754-<941) | Quintile 3 (941-<1157) | Quintile 4 (1157- <1440) | Quintile 5 (≥1440) | | | |
| All participants | | | | | | | | |
| Cases/controls | 344/273 | 258/272 | 272/271 | 239/272 | 248/273 | | 1361/1361 | 1101/1101 |
| Model 1[‡] | 1.00 (Ref.) | 0.74 (0.58-0.94) | 0.77 (0.61-0.98) | 0.64 (0.50-0.83) | 0.69 (0.54-0.89) | *0.002* | 0.90 (0.83-0.97) | 0.91 (0.82-1.00) |
| Model 2[§] | 1.00 (Ref.) | 0.75 (0.60-0.96) | 0.83 (0.65-1.07) | 0.69 (0.53-0.90) | 0.75 (0.58-0.98) | *0.035* | 0.93 (0.85-1.01) | 0.92 (0.83-1.02) |
| Men | | | | | | | | |
| Cases/controls | 222/156 | 146/138 | 121/140 | 85/124 | 83/99 | | 657/657 | 521/521 |
| Model 1[‡\|] | 1.00 (Ref.) | 0.77 (0.56-1.05) | 0.62 (0.46-0.87) | 0.46 (0.32-0.65) | 0.57 (0.39-0.82) | *<0.001* | 0.81 (0.72-091) | 0.77 (0.65-0.91) |
| Model 2[§\|] | 1.00 (Ref.) | 0.79 (0.57-1.09) | 0.62 (0.44-0.87) | 0.49 (0.33-0.72) | 0.63 (0.42-0.94) | *0.001* | 0.84 (0.74-0.96) | 0.75 (0.63-0.90) |
| Women | | | | | | | | |
| Cases/controls | 122/117 | 115/134 | 151/131 | 152/148 | 164/174 | | 704/704 | 580/580 |
| Model 1[‡\|] | 1.00 (Ref.) | 0.77 (0.53-1.12) | 1.04 (0.73-1.50) | 0.93 (0.65-1.35) | 0.90 (0.63-1.35) | *0.967* | 0.99 (0.88-1.10) | 1.00 (0.88-1.13) |
| Model 2[§\|] | 1.00 (Ref.) | 0.77 (0.52-1.15) | 1.16 (0.79-1.70) | 1.03 (0.70-1.53) | 0.94 (0.63-1.38) | *0.754* | 1.00 (0.89-1.13) | 1.02 (0.89-1.16) |

Abbreviations: BMI, body mass index; sRAGE, soluble receptor for advanced glycation end-products

*Quintiles (in pg/mL) were created based on the distribution of sRAGE in the control group. All the models were run using conditional logistic regression

[†]Analysis excluding cases that occurred within two years of follow-up

[‡]Model 1 was conditioned on the matching factors

§Model 2 is Model 1 further adjusted for body mass index (BMI, continuous), height (continuous), education (none, primary, technical and professional, secondary, higher education), physical activity (inactive, moderately inactive, moderately active, active), smoking status, duration, and intensity (never, 1-<=15 cigarettes/day, 16-<=25 cigarettes/day, >26 cigarettes/day, former smokers who quit <=10 years, former smokers who quit 11-<=20 years, former smokers who quit>20 years, current pipe-cigar and occasional smokers), dietary energy (continuous) and intakes of alcohol, red and processed meat, dietary fibre, and dairy products (all as continuous variables)

|Heterogeneity by sex for sRAGE and colorectal cancer risk association was statistically significant for the two models (*P* for heterogeneity=*0.005*, and *0.006* for the models 1 and 2, respectively)

**Table 3:** Odds ratios (OR) and 95% confidence intervals (CI) for risk of colorectal cancer anatomical subsites associated with circulating sRAGE (Continuous, per SD), EPIC study 1992-2012

| | Colon cancer | | | Rectal cancer |
|---|---|---|---|---|
| | All colon | Proximal colon | Distal colon | |
| All participants | | | | |
| Cases/Controls* | 854/854 | 372/372 | 414/414 | 502/502 |
| OR (95% CI) [†] | 0.94 (0.84 - 1.04) | 0.92 (0.77 - 1.10) | 0.88 (0.75 - 1.03) | 0.90 (0.78 - 1.05) |
| Men | | | | |
| Cases/Controls* | 388/388 | 160/160 | 191/191 | 270/270 |
| OR (95% CI) [†‡] | 0.84 (0.70 – 0.99) | 0.94 (0.69 - 1.29) | 0.61 (0.44 - 0.84) | 0.80 (0.64 - 0.99) |
| Women | | | | |
| Cases/Controls* | 466/466 | 212/212 | 223/223 | 232/232 |
| OR (95% CI) [†‡] | 0.99 (0.85-1.15) | 0.85 (0.64 - 1.13) | 1.05 (0.83 - 1.31) | 1.06 (0.86 - 1.32) |

*Some colorectal cancers cases were not included in the analysis as they were overlapping (5 were neither colon nor rectal tumours, 68 were neither proximal nor distal colon tumours)

[†]Conditional logistic regression models conditioned on matching factors and adjusted for body mass index (BMI, continuous), height (continuous), education (none, primary, technical and professional, secondary, higher education), physical activity (inactive, moderately inactive, moderately active, active), smoking status, duration, and intensity (never, 1-<=15 cigarettes/day, 16-<=25 cigarettes/day, >26 cigarettes/day, former smokers who quit <=10 years, former smokers who quit 11-<=20 years, former smokers who quit>20 years, current pipe-cigar and occasional smokers), dietary energy (continuous) and intakes of alcohol, red and processed meat, dietary fibre, and dairy products (all as continuous variables)

[‡]P for heterogeneity colon cancer vs. rectal cancer were 0.607, 0.091, and 0.291 for all the participants, men and women, respectively

P for heterogeneity proximal colon cancer vs. distal colon cancer were 0.307, 0.671, and 0.870 for all the participants, men and women, respectively

P for heterogeneity by sex were 0.042, 0.832, 0.004, 0.063 for all colon cancer, proximal colon cancer, distal colon cancer, and rectal cancer, respectively

**Table 4:** Odds ratios (OR) and 95% confidence intervals (CI) for colorectal cancer risk associated with SNPs associated with sRAGE levels, EPIC study 1992-2012

| SNP | Cases | Controls | OR (95% CI) * | *P*-value‡ | OR (95% CI) † | *P*-value‡ |
|---|---|---|---|---|---|---|
| rs2070600 (*AGER*) | | | | | | |
| CC | 1836 | 2048 | 1.00 (ref.) | | 1.00 (ref.) | |
| CT | 148 | 164 | 1.01 (0.80-1.27) | *0.955* | 1.06 (0.84-1.35) | *0.608* |
| TT | 1 | 8 | 0.14 (0.02-1.12) | *0.063* | 0.17 (0.02-1.36) | *0.095* |
| T *vs.* C | 1985 | 2220 | 0.93 (0.75-1.16) | *0.519* | 0.99 (0.79-1.24) | *0.906* |
| CT+TT *vs.* CC | 1985 | 2220 | 0.97 (0.77-1.21) | *0.768* | 1.03 (0.81-1.30) | *0.835* |
| TT vs. CT+CC | 1985 | 2220 | 0.14 (0.02-1.12) | *0.063* | 0.17 (0.02-1.35) | *0.094* |
| rs1800625 (*AGER*) | | | | | | |
| AA | 1350 | 1584 | 1.00 (ref.) | | 1.00 (ref.) | |
| AG | 574 | 578 | 1.17 (1.02-1.34) | *0.028* | 1.13 (0.98-1.3) | *0.084* |
| GG | 61 | 58 | 1.23 (0.86-1.78) | *0.261* | 1.17 (0.81-1.7) | *0.397* |
| G vs. A | 2135 | 2331 | 1.15 (1.02-1.29) | *0.020* | 1.11 (0.99-1.25) | *0.071* |
| AG+GG *vs.* AA | 2135 | 2331 | 1.17 (1.03-1.34) | *0.019* | 1.13 (0.99-1.3) | *0.067* |
| GG *vs.* AG+AA | 2135 | 2331 | 1.18 (0.82-1.7) | *0.369* | 1.13 (0.78-1.64) | *0.513* |
| rs9469089 (*RNF5)* | | | | | | |
| GG | 1408 | 1619 | 1.00 (ref.) | | 1.00 (ref.) | |
| GC | 532 | 548 | 1.12 (0.97-1.28) | *0.121* | 1.14 (0.99-1.31) | *0.070* |
| CC | 45 | 53 | 0.98 (0.65-1.46) | *0.907* | 0.99 (0.65-1.49) | *0.948* |
| C *vs.* G | 1985 | 2220 | 1.08 (0.95-1.21) | *0.231* | 1.09 (0.97-1.23) | *0.152* |
| GC+CC *vs.* GG | 1985 | 2220 | 1.10 (0.96-1.26) | *0.150* | 1.13 (0.98-1.29) | *0.089* |
| CC *vs.* GC+GG | 1985 | 2220 | 0.95 (0.63-1.42) | *0.796* | 0.95 (0.63-1.43) | *0.813* |
| rs4746 (*GLO1*) | | | | | | |
| TT | 651 | 724 | 1.00 (ref.) | | 1.00 (ref.) | |
| TG | 965 | 1034 | 1.04 (0.90-1.19) | *0.596* | 1.03 (0.9-1.19) | *0.645* |
| GG | 369 | 462 | 0.89 (0.75-1.06) | *0.179* | 0.89 (0.75-1.06) | *0.192* |
| G *vs.* T | 1985 | 2220 | 0.95 (0.88-1.04) | *0.275* | 0.95 (0.88-1.04) | *0.282* |
| TG+GG *vs.* TT | 1985 | 2220 | 0.99 (0.87-1.13) | *0.899* | 0.99 (0.87-1.13) | *0.870* |

| | | | | | | |
|---|---|---|---|---|---|---|
| GG *vs.* TG+ TT | 1985 | 2220 | 0.87 (0.75-1.01) | *0.071* | 0.87 (0.75-1.02) | *0.084* |
| rs653765 (*ADAM10*) | | | | | | |
| TT | 1076 | 1125 | 1.00 (ref.) | | 1.00 (ref.) | |
| TC | 757 | 887 | 0.89 (0.79-1.01) | *0.081* | 0.90 (0.79-1.02) | *0.098* |
| CC | 152 | 208 | 0.76 (0.61-0.96) | *0.019* | 0.83 (0.66-1.04) | *0.109* |
| C *vs.* T | 1985 | 2220 | 0.88 (0.80-0.97) | *0.008* | 0.90 (0.82-0.99) | *0.038* |
| TC+CC *vs.* TT | 1985 | 2220 | 0.87 (0.77-0.98) | *0.022* | 0.88 (0.78-1.00) | *0.051* |
| CC vs. TC+TT | 1985 | 2220 | 0.80 (0.64-1.00) | *0.048* | 0.87 (0.70-1.09) | *0.219* |

*Crude model (unadjusted)

†Adjusted for sex, country, age (1-year categories), BMI (continuous), smoking status (never, former, current) and alcohol intake (continuous)

‡P-values were calculated by considering genetic variant as continuous

**Figure legends:**

**Figure 1:** sRAGE and genetic data available within EPIC

Two endpoints were used for our data; the first ended in June 2003 and included 1361 colorectal cancer cases and 1361 matched controls for the analysis of sRAGE concentrations. December 2012 was considered for the second endpoint, with 1985 samples of colorectal cancer cases, and 2220 controls analysed for genetic data. The overlapping between the two samples was used for sensitivity analysis.

**Figure 2:** Multivariable-adjusted odds ratio and 95%CI of the associations between RAGE and colorectal cancer, stratified by lifestyle, obesity, CRP and menopause status

Multivariable-adjusted OR and 95% CI were computed for the stratified analysis. All the analyses were conditional logistic regression models conditioned on matching factors and adjusted for BMI, education, physical activity, smoking status, dietary energy and intakes of alcohol, red and processed meat, dietary fibre, and dairy products. The analyses stratified by BMI, physical activity, smoking, and alcohol were not adjusted for their respective variables.