Clinical Rheumatology (2019) 38:243–250
https://doi.org/10.1007/s10067-018-4261-5

BRIEF REPORT

Inflammatory dietary pattern and risk of developing rheumatoid arthritis in women

Jeffrey A. Sparks[1,2] · Medha Barbhaiya[3] · Sara K. Tedeschi[1,2] · Cianna L. Leatherwood[1,2] · Fred K. Tabung[4,5] · Cameron B. Speyer[1] · Susan Malspeis[1] · Karen H. Costenbader[1,2] · Elizabeth W. Karlson[1,2] · Bing Lu[1,2]

Received: 19 May 2018 / Revised: 1 August 2018 / Accepted: 8 August 2018 / Published online: 14 August 2018

Abstract

Our objective was to investigate whether a dietary pattern derived using inflammatory biomarkers is associated with rheumatoid arthritis (RA) risk. We prospectively followed 79,988 women in the Nurses' Health Study (NHS, 1984–2014) and 93,572 women in the NHSII (1991–2013); incident RA was confirmed by medical records. Food frequency questionnaires (FFQ) were completed at baseline and approximately every 4 years. Inflammatory dietary pattern was assessed from FFQ data using the Empirical Dietary Inflammatory Pattern (EDIP), including 18 anti-/pro-inflammatory food/beverage groups weighted by correlations with plasma inflammatory biomarkers (interleukin-6, C-reactive protein, and tumor necrosis factor-α receptor 2). We investigated associations between EDIP and RA using Cox regression. We identified 1185 incident RA cases over 4,425,434 person-years. EDIP was not associated with overall RA risk (p trend = 0.21 across EDIP quartiles). Among women ≤ 55 years, increasing EDIP was associated with increased overall RA risk; HRs (95% CIs) across EDIP quartiles were 1.00 (reference), 1.14 (0.86–1.51), 1.35 (1.03–1.77), and 1.38 (1.05–1.83; p for trend = 0.01). Adjusting for BMI attenuated this association. Increasing EDIP was associated with increased seropositive RA risk among women ≤ 55 years (p for trend = 0.04). There was no association between EDIP and RA among women > 55 years (EDIP-age interaction, p = 0.03). An inflammatory dietary pattern was associated with increased seropositive RA risk with onset ≤ 55 years old, and this association may be partially mediated through BMI.

Keywords Diet · Epidemiology · Inflammation · Rheumatoid arthritis

Introduction

Systemic inflammation is important in the development of rheumatoid arthritis (RA)-related autoimmunity such as autoantibodies prior to clinical presentation [1]. Identifying RA risk factors may help to elucidate pathogenesis and could identify lifestyle changes to lower RA risk. Individual foods and beverages are associated with RA, perhaps by contributing to systemic inflammation; fish and alcohol intake may lower RA risk, while red meat and sugar-sweetened soda consumption may increase RA risk [2–5].

Dietary pattern analyses combine multiple foods/beverages, a preferable method for investigation since these are not consumed in isolation. We previously reported that a healthier dietary pattern was associated with lower RA risk, particularly seropositive RA risk for women aged ≤ 55 years [6]. However, the dietary quality measure investigated was developed for other chronic diseases [6].

Therefore, we investigated the association between dietary inflammatory potential assessed using the empirical dietary inflammatory pattern (EDIP) score and RA risk. We used two Nurses' Health Study (NHS) cohorts, characterized by rich data on lifestyle factors including repeated dietary assessments. We hypothesized that higher EDIP scores (indicating pro-inflammatory diets) would be

✉ Jeffrey A. Sparks
jsparks@bwh.harvard.edu

1 Department of Medicine, Division of Rheumatology, Immunology, and Allergy, Brigham and Women's Hospital, 60 Fenwood Road, #6016U, Boston, MA 02115, USA

2 Harvard Medical School, Boston, MA, USA

3 Hospital for Special Surgery, New York City, NY, USA

4 Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA

5 Department of Internal Medicine, Division of Medical Oncology, The Ohio State University College of Medicine, Columbus, OH, USA

associated with increased RA risk, particularly seropositive RA diagnosed at a younger age.

Methods

Study population

The NHS and NHSII are prospective cohorts of US registered nurses. The NHS enrolled 121,700 women aged 30–55 years in 1976; the NHSII enrolled 116,670 women aged 25–42 years in 1989. Participants answered questionnaires at each study's inception and every 2 years during follow-up. Each time point where questionnaires are collected is referred to as a cycle of follow-up. The baseline for this analysis is 1984 for NHS and 1991 for NHSII when a comprehensive Food Frequency Questionnaire (FFQ) was introduced into each cohort. Excluding participants with prevalent RA and FFQ non-responders at baseline in each cohort, 79,988 NHS participants, followed from 1984 to 2014 and 93,572 NHSII participants followed from 1991 to 2013, were analyzed. All aspects of the study, including obtaining informed consent from participants, content/mailing/data management of questionnaires, and identification of incident RA cases, were approved by the Partners HealthCare Institutional Review Board.

Dietary assessments

FFQs assessed dietary intake over the previous year, ranking frequency of each food/beverage on a scale from never or < 1/month to ≥ 6 servings/day [6, 7]. FFQs were administered in 1984, 1986, and every 4 years until 2010 in the NHS and 1991 and every 4 years until 2011 in the NHSII.

Empirical dietary inflammatory pattern score

The empirical dietary inflammatory pattern (EDIP) was developed as an agnostic method to classify food/beverage groups as anti- or pro-inflammatory, as described in more detail elsewhere [8]. In the previous validation study, the NHS was used as the discovery dataset, while the NHSII and the Health Professional Follow-up Study (a similar large prospective cohort study, but performed among only men) were used as the validation datasets [8]. Briefly, intake of each food/beverage group was tested for correlations with plasma inflammatory biomarker levels (interleukin-6 [IL-6], C-reactive protein [CRP], and tumor necrosis factor-α receptor 2 [TNFαR2]) using reduced rank regression [8]. Reduced rank regression is a statistical method used to derive a set of dietary patterns associated with biomarkers on the causal pathway to an outcome. The advantages of this method are that the predictive potential for the outcome is maximized while simultaneously accounting for the collinearity of predictors. Each food/beverage group related to these biomarkers was then weighted by the beta coefficient in the multivariable linear regression model. The summary EDIP score was the weighted sum for all intake of food groups. The raw EDIP scores were rescaled by dividing by 1000 to reduce the magnitude of scores.

The 9 pro-inflammatory groups (positively associated with IL-6/CRP/TNFαR2) previously identified were processed meat, red meat, organ meat, non-dark meat fish, other vegetables (other than green leafy and dark-yellow vegetables), refined grains, high-energy beverages, low-energy beverages, and tomatoes [8]. The 9 anti-inflammatory groups (inversely associated with IL-6/CRP/TNFαR2) previously identified were beer, wine, tea, coffee, dark-yellow vegetables, leafy green vegetables, snacks, fruit juice, and pizza [8].

We pooled both cohorts for statistical efficiency due to planned subgroup analyses and in anticipation of a modest effect size for a dietary exposure. The EDIP for each cycle, as a time-varying exposure, was calculated using cumulative average dietary intake to reflect long-term diet, by averaging the dietary intake from baseline until that follow-up cycle. The EDIP values in our analysis ranged between − 3.6 (most anti-inflammatory) and + 2.5 (most pro-inflammatory). Since there are no known clinically meaningful EDIP cutpoints, individuals were ranked from lowest to highest EDIP score and then placed in quartiles at the baseline of each cohort. The lowest (first) quartile represents the group with the least inflammatory dietary intake (the reference group) and the highest (fourth) quartile represents the group with the most inflammatory dietary intake. This type of analysis accounts for all previous and current dietary measures for each cycle predicting RA risk in the subsequent period. For example, we used the averaged dietary intake at 1984, 1986, and 1990 to calculate EDIP, and predict the subsequent RA incidence from 1990 to 1994, similarly, using the averaged diet up to 1994 to predict RA incidence from 1994 to 1998, and so on.

Identification of incident RA

Women who self-reported incident RA were mailed a questionnaire [9]. Medical records were obtained to verify RA diagnosis and determine diagnosis date/serologic status. All cases were reviewed by two rheumatologists and confirmed RA according to either 1987 American College of Rheumatology or 2010 ACR/European League Against Rheumatism criteria [10, 11]. Seropositive RA was defined as positive rheumatoid factor (RF) or anti-cyclic citrullinated peptide (anti-CCP).

Covariates

Time-varying covariate data were obtained through biennial questionnaires. Age and household income, derived from US Census-tract median income by zip code, were continuous

variables. Smoking was categorized by status (never/past/current) and pack-years. Body mass index (BMI) was categorized according to the World Health Organization recommendations into underweight/normal, overweight, and obese or considered as a continuous variable. Self-reported physical activity was a continuous variable measured as weekly metabolic equivalents. We included reproductive/hormonal factors such as oral contraceptive use, parity, breastfeeding duration, and menopausal status/postmenopausal hormone (PMH) use.

Statistical analysis

We reported baseline characteristics in each cohort according to EDIP quartile. We investigated EDIP and RA risk, overall and stratified by serologic status and by pre-specified age groups. We previously found differences in metabolic/dietary RA risk factors related to age of onset ≤ 55 or > 55 years old [5, 6, 12]. We used that age cutpoint as an a priori hypothesis due to different RA clinical phenotypes based on age of onset and as an approximation of completion of menopausal transition.

We used a prospective cohort design wherein dietary intake at each cycle of follow-up was used to predict incident RA occurring in the subsequent period prior to the next cycle of dietary data. In all analyses, the sample at each cycle that we analyzed was free of RA or other connective tissue diseases.

We hypothesized that a long-term measure of dietary intake was most important for risk of a chronic disease such as RA. Therefore, we used cumulative average updated EDIP as a measure of long-term dietary intake. This type of analysis accounts for all previous and current dietary measures for each cycle predicting RA risk in the subsequent 2-year cycle.

Time-varying Cox regression models estimated hazard ratios (HR) and 95% confidence intervals (CI) for the association of EDIP quartiles with RA (reference, first quartile). We calculated *p* values for trend using the median EDIP score for each quartile as a continuous variable in the model. Person-years commenced from the return date of the baseline questionnaire to date of last follow-up/death/censor, whichever came first.

In base models, we adjusted for age, questionnaire period, cohort, and total energy intake. We built multivariable models based on the primary analysis and included covariates that were associated with EDIP and RA. The final multivariable model included the base factors and census-tract household income, smoking pack-years, and menopausal status/hormone use (premenopausal, postmenopausal/never PMH use, postmenopausal/current PMH use, and postmenopausal/past PMH use). We further adjusted for BMI, and performed formal mediation analysis to examine and quantify the mediating effect of BMI on the association of EDIP with RA [13]. Since the imaging and laboratory tests used to diagnose RA, as well as treatment strategies for those newly diagnosed, have changed over time, we performed a sensitivity analysis examining whether EDIP effects on RA varied by calendar time. We tested whether the association of EDIP with RA risk differed before or after the year 2000 by including an interaction between EDIP and an indicator variable for calendar time (before 2000 vs. 2000 or later) in a separate model and reporting the *p* value.

Results

Baseline characteristics of both cohorts (total $n = 173{,}560$) according to EDIP quartiles are shown in Table 1. Those with lower inflammatory dietary pattern were older, smoked more, had higher income, had higher physical activity, and had lower BMI than women with higher inflammatory dietary pattern. We identified 1185 incident RA cases (62.6% seropositive) over 4,425,434 person-years of follow-up (mean follow-up per participant of 25.5 years).

Table 2 shows the association of EDIP with RA risk. There was no association of EDIP with all RA (*p* for trend = 0.21), seropositive RA (*p* for trend = 0.38), or seronegative RA (*p* for trend = 0.37).

Table 3 shows the association of EDIP with RA by age (≤ 55 years and > 55 years). Among women ≤ 55 years, increasing EDIP was associated with increased overall RA risk. HRs (95% CIs) across EDIP quartiles were 1.00 (reference), 1.14 (0.86–1.51), 1.35 (1.03–1.77), and 1.38 (1.05–1.83) with a significant linear trend (*p* for trend = 0.01). These results were attenuated when additionally adjusting for BMI ($HR_{Q4vs.Q1}$ 1.25, 95% CI 0.94–1.65; *p* for trend = 0.09). In mediation analyses, the estimated proportion of EDIP effect mediated by BMI was 41.8% (95% CI 10.3–81.8%). Among women ≤ 55 years, intake of pro-inflammatory diets was associated with increased seropositive RA risk (*p* for trend = 0.04), but EDIP was not associated with seronegative RA (*p* for trend = 0.13). Additionally, we did not observe the association of EDIP with RA risk differs between before or after the year 2000 (EDIP-calendar time interaction $p = 0.27$). Among women > 55 years old, there was no association of EDIP with all RA (*p* for trend = 0.55; EDIP-age interaction $p = 0.03$), seropositive RA (*p* for trend = 0.47), or seronegative RA (*p* for trend = 0.96).

Discussion

In this study of 173,560 women with up to 30 years of follow-up, an inflammatory dietary pattern was associated with increased seropositive RA risk among women aged ≤ 55 years. However, we found no association of inflammatory dietary pattern for overall RA risk or among women age > 55 years. In addition, the association between inflammatory diet and RA may be partially mediated through BMI.

Table 1 Age-standardized baseline characteristics of Nurses' Health Study (1984) and Nurses' Health Study II (1991) participants by Empirical Dietary Inflammatory Pattern (EDIP) quartiles ($n = 173,560$)

	Nurses' Health Study ($n = 79,988$)				Nurses' Health Study II ($n = 93,572$)			
EDIP quartiles*	Q1 (least inflammatory) ($n = 23,174$)	Q2 ($n = 20,884$)	Q3 ($n = 19,234$)	Q4 (most inflammatory) ($n = 16,696$)	Q1 (least inflammatory) ($n = 20,217$)	Q2 ($n = 22,508$)	Q3 ($n = 24,153$)	Q4 (most inflammatory) ($n = 26,694$)
Mean age, years (SD)**	51.2 (6.9)	51.3 (7.2)	51.0 (7.2)	50.1 (7.3)	37.6 (4.4)	36.9 (4.6)	36.3 (4.7)	35.8 (4.7)
Mean US Census-tract household income ($USD × 1000)	68.3 (27.6)	65.4 (26.0)	63.0 (24.5)	60.6 (23.6)	64.6 (24.4)	63.1 (23.4)	60.9 (21.9)	57.8 (20.3)
Mean BMI, kg/m^2 (SD)	24.0 (3.8)	24.6 (4.3)	25.4 (4.9)	26.6 (5.8)	23.7 (4.4)	24.0 (4.7)	24.6 (5.2)	25.9 (6.3)
Mean total METs/ week (SD)	15.9 (23.9)	14.4 (20.5)	13.3 (18.7)	12.1 (17.8)	24.6 (31.4)	21.6 (27.2)	19.9 (25.6)	18.3 (25.1)
Current smoking, %	29.1	23.7	20.8	21.7	15.7	11.7	10.0	12.3
Menarche at < 12 years of age, %	23.1	22.8	23.0	24.3	24.9	23.9	23.8	25.2
Oral contraceptive use, %	49.6	48.5	47.5	47.5	84.9	84.6	84.6	84.0
Parous, %	92.4	92.7	92.5	92.3	69.9	71.8	73.4	72.8
Breastfed ≥ 12 months, %	17.7	17.3	16.8	15.7	30.6	30.5	29.3	24.5
Postmenopausal, %	59.1	60.2	60.7	62.0	3.1	3.0	3.3	3.8
Current postmenopausal hormone use, %	22.7	22.6	22.9	21.0	2.5	2.4	2.7	3.0
Mean energy intake, kcal (SD)	1705 (509)	1662 (501)	1705 (514)	1938 (564)	1787 (541)	1689 (508)	1715 (518)	1944 (577)

*The cutpoints for EDIP quartiles were based on baseline data pooled from both cohorts

**Not age-adjusted

Missing values are not shown

These results add to the literature suggesting that diet plays an important role in RA pathogenesis, perhaps by altering systemic inflammation and autoimmunity [2–6]. Fish intake may exert a protective effect on RA by the anti-inflammatory effects of omega-3 polyunsaturated fatty acids [2]. Among at-risk individuals, erythrocyte membrane-bound omega-3 PUFA levels were inversely associated with anti-CCP/RF positivity [14]. Previous studies investigating EDIP found robust associations between EDIP scores and colorectal cancer risk but no association with ovarian cancer risk [15, 16]. Similar to previous studies investigating dietary/metabolic factors in RA, the inflammatory dietary pattern was important specifically in younger-onset seropositive RA, defined as ≤ 55 years [5, 6, 12]. Thus, metabolic/dietary lifestyle factors may affect RA risk differently based on age of onset as well as serologic status.

We found different results based on an age cutpoint of ≤ 55 or > 55 years. Menopause has known effects on both metabolism and adipose tissue distribution such that women may both gain adiposity while also losing bone mass. Therefore, dietary intake behaviors and metabolism change after menopause as well as the reliability of anthropometric measures such as body mass index. The median age for menopause completion is about 55 years of age. While the NHS and NHSII collected data on date of menopause, there is typically a period of perimenopause lasting for months or even years, such that a clear transition point is typically not obvious. We previously showed that menopausal status has a strong impact on RA risk [17]. Therefore, we relied on the age of 55 years to clearly classify women as younger or older in this analysis.

Younger- and older-onset RA also have clinical phenotypic differences related to presentation and response to treatment. Patients with older-onset RA may be more likely to be seronegative and have a less severe onset compared to younger-onset RA. Older-onset RA may also be more likely to be responsive to glucocorticoids and may not be as likely to require combinations of drugs compared to younger-onset RA. While there is no standard age cutpoint to classify younger- vs. older-onset RA, we chose 55 years given the biologic plausibility related to menopause and clinical RA differences based on age at onset. It is possible other age cutpoints might have yielded different results.

We previously reported a protective effect on RA for a healthier dietary quality based on the 2010 Alternative Healthy Eating Index (2010-AHEI). The 2010-AHEI was derived based on consensus from experts related to risk of developing diabetes, cancer, and cardiovascular disease [18]. Therefore, the 2010-

Table 2 Hazard ratios (95% CI) for rheumatoid arthritis according to EDIP quartiles in the Nurses' Health Study and Nurses' Health Study II

	Empirical Dietary Inflammatory Pattern (EDIP) quartiles				
	Q1 (least inflammatory)	Q2	Q3	Q4 (most inflammatory)	*p* trend‡
All RA (*n* = 1185 cases)					
Cases/person-years	252/960,107	367/1,316,534	327/1,218,122	239/930,671	
Age-adjusted*	1.00 (Ref)	1.08 (0.92 to 1.27)	1.07 (0.91 to 1.26)	1.10 (0.92 to 1.32)	0.31
Multivariable model 1 (main)**	1.00 (Ref)	1.11 (0.95 to 1.31)	1.12 (0.95 to 1.32)	1.12 (0.94 to 1.35)	0.21
Multivariable model 2†	1.00 (Ref)	1.09 (0.93 to 1.28)	1.08 (0.91 to 1.28)	1.07 (0.89 to 1.28)	0.54
Seropositive RA (*n* = 741 cases)					
Cases/person-years	151/958,050	231/1,314,010	222/1,215,646	137/928,344	
Age-adjusted*	1.00 (Ref)	1.13 (0.92 to 1.39)	1.20 (0.98 to 1.48)	1.03 (0.82 to 1.31)	0.61
Multivariable model 1 (main)**	1.00 (Ref)	1.18 (0.96 to 1.45)	1.28 (1.04 to 1.58)0	1.07 (0.85 to 1.36)	0.38
Multivariable model 2†	1.00 (Ref)	1.16 (0.94 to 1.42)	1.24 (1.00 to 1.52)	1.01 (0.79 to 1.29)	0.72
Seronegative RA (*n* = 444 cases)					
Cases/person-years	101/957,570	136/1,313,598	105/1,215,052	102/928,321	
Age-adjusted*	1.00 (Ref)	1.00 (0.78 to 1.30)	0.87 (0.66 to 1.15)	1.22 (0.92 to 1.61)	0.32
Multivariable model 1 (main)**	1.00 (Ref)	1.01 (0.78 to 1.31)	0.88 (0.67 to 1.16)	1.20 (0.91 to 1.60)	0.37
Multivariable model 2†	1.00 (Ref)	0.99 (0.77 to 1.29)	0.85 (0.64 to 1.12)	1.15 (0.86 to 1.53)	0.59

The cutpoints for EDIP quartiles were based on baseline data pooled from both cohorts

*Adjusted for age, questionnaire cycle, and cohort

**Adjusted for age, questionnaire cycle, cohort, total energy intake, census-tract household income, smoking pack-years, hormone use (premenopausal, postmenopausal/never postmenopausal hormone use, postmenopausal/current postmenopausal hormone use, and postmenopausal/past postmenopausal hormone use)

† Additionally adjusted for BMI categories (< 25, 25–29.9, ≥ 30 kg/m^2)

‡ *p* trend was derived from tests of linear trend across categories of EDIP using the median value of each category as a continuous variable

AHEI was not derived related to RA or other autoimmune diseases which may have different pathogenesis. Further, the methods for selecting components of the 2010-AHEI do not consider the correlation structure of food and nutrient intake. The EDIP was derived using reduced rank regression, which offers some advantages over the 2010-AHEI [6]. First, this method uses objective inflammatory plasma biomarkers that are known to be important in the causal pathway for development of RA [1]. Therefore, the associations are more likely be directly applicable to RA development compared to the 2010-AHEI. Second, the statistical methods for deriving the EDIP take into account the correlation structure of food/nutrient groups lowering the possibility that the association is confounded by intake from other measured or unmeasured components of the diet [6]. However, it is possible that both of these methods classifying patterns of dietary intake are measuring a similarly healthy or unhealthy diet for RA risk. Future comparative studies may be needed to directly compare the predictive ability for RA for these and other dietary measures in order to inform public health recommendations.

Our study was limited by including only women who were well-educated and working at cohort entry so may not be generalizable. While detailed time-varying data on covariates such as smoking and BMI were available, other unmeasured confounders could have affected our results. While the relationship between diet and BMI is complex, we considered BMI as a mediator since diet affects BMI but not vice versa. In our study, over 40% of the EDIP association with RA may be mediated through BMI. However, another interpretation of these results would be that BMI confounds the association between the EDIP and RA risk such that BMI explains most of the observed excess risk. Overall, our results provide further support for dietary modifications to maintain healthy weight to possibly decrease RA risk [12]. EDIP was developed using the population studied by correlating food/beverage groups with levels of plasma inflammatory markers [8]. While many of the food/beverage groups were classified as expected (e.g., beer/wine as anti-inflammatory), there were some unexpected classifications (e.g., pizza as anti-inflammatory, which may be due to the higher bioavailable lycopene in pizza than in fresh tomatoes), perhaps related to the dietary habits of a population of older health professionals [8]. EDIP scores may be interpreted as reflecting the potential of diet to contribute to systemic inflammation. EDIP can be considered as a surrogate long-term measure contributing to IL-6/CRP/TNFαR2 levels, but the external generalizability to non-health professionals of EDIP is currently unclear. The biomarkers used to develop EDIP are important in RA pathogenesis, so this dietary pattern may be more specific for RA than others [6].

Table 3 Hazard ratios (95% CI) for rheumatoid arthritis according to EDIP quartiles in the Nurses' Health Study and Nurses' Health Study II stratified by age ≤ 55 or > 55 years

	Empirical Dietary Inflammatory Pattern (EDIP) quartiles				
	Q1 (least inflammatory)	Q2	Q3	Q4 (most inflammatory)	*p* trend‡
Age ≤ 55 years***	All RA (*n* = 490 cases)				
Cases/person-years	85/455,964	123/610,841	144/630,515	138/598,330	
Age-adjusted*	1.00 (Ref)	1.11 (0.84 to 1.46)	1.31 (1.00 to 1.71)	1.40 (1.07 to 1.84)	0.01
Multivariable model 1 (main)**	1.00 (Ref)	1.14 (0.86 to 1.51)	1.35 (1.03 to 1.77)	1.38 (1.05 to 1.83)	0.01
Multivariable model 2†	1.00 (Ref)	1.11 (0.84 to 1.46)	1.27 (0.97 to 1.67)	1.25 (0.94 to 1.65)	0.09
Seropositive RA (*n* = 319 cases)					
Cases/person-years	53/454,800	78/609,357	104/628,972	84/596,561	
Age-adjusted*	1.00 (Ref)	1.12 (0.79 to 1.59)	1.51 (1.08 to 2.10)	1.36 (0.96 to 1.92)	0.03
Multivariable model 1 (main)**	1.00 (Ref)	1.16 (0.82 to 1.65)	1.58 (1.13 to 2.21)	1.37 (0.96 to 1.94)	0.04
Multivariable model 2†	1.00 (Ref)	1.12 (0.79 to 1.59)	1.46 (1.04 to 2.05)	1.19 (0.83 to 1.71)	0.21
Seronegative RA (*n* = 171 cases)					
Cases/person-years	32/454,512	45/609,227	40/628,581	54/596,511	
Age-adjusted*	1.00 (Ref)	1.13 (0.72 to 1.78)	1.01 (0.63 to 1.61)	1.53 (0.98 to 2.38)	0.07
Multivariable model 1 (main)**	1.00 (Ref)	1.14 (0.72 to 1.80)	1.01 (0.63 to 1.62)	1.45 (0.92 to 2.28)	0.13
Multivariable model 2†	1.00 (Ref)	1.12 (0.71 to 1.77)	0.98 (0.61 to 1.57)	1.38 (0.87 to 2.18)	0.20
Age ≥> 55 years***	All RA (*n* = 695 cases)				
Cases/person-years	167/508,826	244/713,996	183/595,759	101/338,264	
Age-adjusted*	1.00 (Ref)	1.06 (0.87 to 1.29)	0.95 (0.77 to 1.17)	0.90 (0.70 to 1.15)	0.29
Multivariable model 1 (main)**	1.00 (Ref)	1.10 (0.90 to 1.34)	1.00 (0.81 to 1.23)	0.94 (0.73 to 1.21)	0.55
Multivariable model 2†	1.00 (Ref)	1.09 (0.89 to 1.33)	0.99 (0.80 to 1.22)	0.93 (0.72 to 1.20)	0.49
Seropositive RA (*n* = 422 cases)					
Cases/person-years	98/507,922	153/712,941	118/594,817	53/337,689	
Age-adjusted*	1.00 (Ref)	1.13 (0.88 to 1.46)	1.03 (0.79 to 1.35)	0.79 (0.57 to 1.11)	0.20
Multivariable model 1 (main)**	1.00 (Ref)	1.19 (0.92 to 1.54)	1.12 (0.85 to 1.47)	0.85 (0.60 to 1.19)	0.47
Multivariable model 2†	1.00 (Ref)	1.19 (0.92 to 1.54)	1.12 (0.85 to 1.47)	0.86 (0.61 to 1.21)	0.52
Seronegative RA (*n* = 273 cases)					
Cases/person-years	69/507,726	91/712,650	65/594,594	48/337,715	
Age-adjusted*	1.00 (Ref)	0.96 (0.71 to 1.32)	0.83 (0.59 to 1.16)	1.06 (0.73 to 1.53)	0.90
Multivariable model 1 (main)**	1.00 (Ref)	0.97 (0.71 to 1.33)	0.83 (0.59 to 1.17)	1.08 (0.74 to 1.56)	0.96
Multivariable model 2†	1.00 (Ref)	0.95 (0.69 to 1.30)	0.80 (0.57 to 1.14)	1.03 (0.70 to 1.50)	0.76

The cutpoints for EDIP quartiles were based on baseline data pooled from both cohorts

*Adjusted for age, questionnaire cycle, and cohort

**Adjusted for age, questionnaire cycle, cohort, total energy intake, census-tract household income, smoking pack-years, hormone use (premenopausal, postmenopausal/never postmenopausal hormone use, postmenopausal/current postmenopausal hormone use, and postmenopausal/past postmenopausal hormone use)

† Additionally adjusted for BMI categories (< 25, 25–29.9, ≥ 30 kg/m^2)

‡*p* trend was derived from tests of linear trend across categories of EDIP using the median value of each category as a continuous variable

***EDIP-age interaction for all RA: $p = 0.03$

We found statistically significant findings only among the subset with RA diagnosed ≤ 55 years of age. We pre-specified this hypothesis based on previous papers showing that BMI and other dietary factors were important for earlier onset, but not later onset, RA [6]. There were more cases diagnosed at > 55 years of age, so the lack of association in older women is less likely related to reduced power. It is possible that changes in metabolism after menopause may affect the relationship between dietary and metabolic factors and RA risk. The pathogenesis of older-onset RA may be different than younger-onset RA since these subsets have differences in clinical presentation and response to treatment. Our study analyzed two closed cohorts, so we are limited in investigating secular trends based on calendar time since women in both studies were aging in

tandem so there were few younger women in the later years of the analysis. It is possible that the availability of novel diagnostic biomarkers, advanced imaging, treatment options, and the paradigm to diagnose RA and treat early to prevent disease-related damage may have affected the RA phenotype. Therefore, it is possible that the results may have been explained related to secular changes of RA diagnosis in the "biologic era" starting around 2000. Similarly, there have been secular trends over calendar time related to lifestyle factors, such as diet and smoking. However, we did not observe the association of EDIP with RA risk differ before or after the year 2000. Further research is needed to understand possible differences both in the biologic response to dietary exposures and differences in disease phenotypes based on age and calendar time.

Strengths of our study include the prospective cohort study design with large sample size and lengthy follow-up. We had up to 8 repeated measures of dietary intake collected prior to RA onset available and investigated EDIP, a validated dietary measure correlating food/beverage groups with systemic inflammatory biomarkers derived using agnostic methods. We analyzed cumulative average dietary intake to best reflect long-term dietary patterns. We identified incident RA occurring after baseline and all cases met accepted research criteria for RA. We had a relatively large number of incident RA cases so were able to detect modest effect sizes, as expected from dietary factors. We had detailed covariates available to adjust for time-varying confounders such as smoking and quantify the mediating of BMI. We were able to determine serologic status during medical record review to classify cases as seropositive or seronegative RA. However, many of the cases were diagnosed with RA prior to the clinical use of anti-CCP, so there may have been some misclassification of RA serologic status. Since dietary exposures typically have modest effects on chronic disease risk, we pooled the two cohorts to increase the power to be able to detect an association between EDIP score and RA risk, particularly among RA subgroups by serologic status and age at onset. Future studies are needed to replicate our findings in independent datasets.

In conclusion, an inflammatory dietary pattern was associated with increased RA risk, particularly seropositive RA for younger and middle-aged women. These results suggest that a long-term dietary pattern that decreases systemic inflammation may lead to lower RA risk. These findings add to the evidence of metabolic and inflammatory markers contributing to autoimmunity. Further research is needed to understand the timing, impact, and biologic mechanisms of dietary behaviors and RA risk.

Acknowledgements We thank the participants of the NHS and NHSII for their dedicated participation in this longitudinal study as well as the staff members at the Channing Division of Network Medicine (Department of Medicine, Brigham and Women's Hospital and Harvard Medical School).

Author contributions All authors were involved in drafting the article or revising it critically for important intellectual contact, and all authors approved the final version to be published. Dr. Sparks had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design: Sparks, Karlson, Lu

Acquisition of data: Sparks, Barbhaiya, Tedeschi, Leatherwood, Tabung, Costenbader, Karlson, Lu

Analysis and interpretation of data: Sparks, Barbhaiya, Tedeschi, Leatherwood, Tabung, Speyer, Malspeis, Costenbader, Karlson, Lu

Funding This work was supported by the National Institutes of Health (grant numbers R01 AR061362, K24 AR052403, L30 AR066953, R01 AR049880, UM1 CA186107, UM1 CA176726, K99 CA207736, K23 AR069688, T32 AR007530, P30 AR070253, and P30 AR072577).

Compliance with ethical standards

Disclosures None.

References

1. Karlson EW, Chibnik LB, Tworoger SS, Lee IM, Buring JE, Shadick NA, Manson JAE, Costenbader KH (2009) Biomarkers of inflammation and development of rheumatoid arthritis in women from two prospective cohort studies. Arthritis Rheum 60(3):641–652
2. Di Giuseppe D, Crippa A, Orsini N, Wolk A (2014) Fish consumption and risk of rheumatoid arthritis: a dose-response meta-analysis. Arthritis Res Ther 16(5):446
3. Lu B, Solomon DH, Costenbader KH, Karlson EW (2014) Alcohol consumption and risk of incident rheumatoid arthritis in women: a prospective study. Arthritis Rheumatol 66(8):1998–2005
4. Benito-Garcia E, Feskanich D, Hu FB, Mandl LA, Karlson EW (2007) Protein, iron, and meat consumption and risk for rheumatoid arthritis: a prospective cohort study. Arthritis Res Ther. 9(1):R16
5. Hu Y, Costenbader KH, Gao X, Al-Daabil M, Sparks JA, Solomon DH et al (2014) Sugar-sweetened soda consumption and risk of developing rheumatoid arthritis in women. Am J Clin Nutr 100(3):959–967
6. Hu Y, Sparks JA, Malspeis S, Costenbader KH, Hu FB, Karlson EW, Lu B (2017) Long-term dietary quality and risk of developing rheumatoid arthritis in women. Ann Rheum Dis 76(8):1357–1364
7. Yuan C, Spiegelman D, Rimm EB, Rosner BA, Stampfer MJ, Barnett JB, Chavarro JE, Subar AF, Sampson LK, Willett WC (2017) Validity of a dietary questionnaire assessed by comparison with multiple weighed dietary records or 24-hour recalls. Am J Epidemiol 185(7):570–584
8. Tabung FK, Smith-Warner SA, Chavarro JE, Wu K, Fuchs CS, Hu FB, Chan AT, Willett WC, Giovannucci EL (2016) Development and validation of an empirical dietary inflammatory index. J Nutr 146(8):1560–1570
9. Karlson EW, Sanchez-Guerrero J, Wright EA, Lew RA, Daltroy LH, Katz JN, Liang MH (1995) A connective tissue disease screening questionnaire for population studies. Ann Epidemiol 5(4):297–302
10. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS et al (1988) The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 31(3):315–324
11. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO 3rd et al (2010) 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 62(9):2569–2581
12. Lu B, Hiraki LT, Sparks JA, Malspeis S, Chen CY, Awosogba JA, Arkema EV, Costenbader KH, Karlson EW (2014) Being

overweight or obese and risk of developing rheumatoid arthritis among women: a prospective cohort study. Ann Rheum Dis 73(11):1914–1922
13. Lin DY, Fleming TR, De Gruttola V (1997) Estimating the proportion of treatment effect explained by a surrogate marker. Stat Med 16(13):1515–1527
14. Gan RW, Demoruelle MK, Deane KD, Weisman MH, Buckner JH, Gregersen PK, Mikuls TR, O'Dell JR, Keating RM, Fingerlin TE, Zerbe GO, Clare-Salzler MJ, Holers VM, Norris JM (2017) Omega-3 fatty acids are associated with a lower prevalence of autoantibodies in shared epitope-positive subjects at risk for rheumatoid arthritis. Ann Rheum Dis 76(1):147–152
15. Tabung FK, Liu L, Wang W, Fung TT, Wu K, Smith-Warner SA, Cao Y, Hu FB, Ogino S, Fuchs CS, Giovannucci EL (2018) Association of dietary inflammatory potential with colorectal cancer risk in men and women. JAMA Oncol 4(3):366–373
16. Tabung FK, Huang T, Giovannucci EL, Smith-Warner SA, Tworoger SS, Poole EM (2017) The inflammatory potential of diet and ovarian cancer risk: results from two prospective cohort studies. Br J Cancer 117(6):907–911
17. Bengtsson C, Malspeis S, Orellana C, Sparks JA, Costenbader KH, Karlson EW (2017) Association between menopausal factors and the risk of seronegative and seropositive rheumatoid arthritis: results from the Nurses' Health Studies. Arthritis Care Res (Hoboken) 69(11):1676–1684
18. Chiuve SE, Fung TT, Rimm EB, Hu FB, McCullough ML, Wang M et al (2012) Alternative dietary indices both strongly predict risk of chronic disease. J Nutr 142(6):1009–1018