

Dietary factors and risk of rheumatoid arthritis using data linkage

Yuanyuan Dong

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
School of Food Science and Nutrition

May 2026

Intellectual property and Publication Statements

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work have been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Chapter 2 incorporates the work of one jointly-authored publication:

Dong Y, Greenwood DC, Webster J, Uzokwe C, Tao J, Hardie LJ, Cade JE. Dose-Response Associations Between Diet and Risk of Rheumatoid Arthritis: A Meta-Analysis of Prospective Cohort Studies. *Nutrients*, 2024;16(23), 4050. <https://doi.org/10.3390/nu16234050>

The candidate was responsible for designing the study protocol and statistical analysis plan; conducting the literature search; screening; data extraction; quality assessment; interpreting the data; writing the initial draft; revising the paper critically for important intellectual content; and managing the publication process as the corresponding author. JC, LJH and DCG conceived and supervised the work, and all authors provided input on the study design, data analysis, and interpretation of results; revised the paper critically for important intellectual content; and approved the final version.

Chapter 3 incorporates the work of one jointly-authored publication:

Dong Y, Greenwood DC, Hardie LJ, Cade JE. Measures of Adiposity and Risk of Rheumatoid Arthritis in Middle-Aged UK Women: A Prospective Cohort Study. *Nutrients*. 2025; 17(9):1557. <https://doi.org/10.3390/nu17091557>

This publication was a secondary analysis of the UK Women's Cohort Study (UKWCS) dataset. The candidate was responsible for designing the study protocol and statistical analysis plan; data cleaning and analysis; and drafting of the manuscript. JC, LJH, and DCG were responsible for the conception of the

UKWCS and supervised the secondary data analysis, and all authors provided input on the study design, data analysis, and interpretation of results; revised the paper critically for important intellectual content; and approved the final manuscript version.

Chapter 4 incorporates the work of one jointly-authored publication:

Dong Y, Greenwood DC, Cade JE. Association of Mediterranean diet with rheumatoid arthritis risk: evidence from the UK Women's Cohort Study. *PLOS One* (Accepted)

This publication was a secondary analysis of the UKWCS dataset. The candidate was responsible for designing the study protocol and statistical analysis plan; data cleaning and analysis; and drafting of the manuscript. JC and DCG were responsible for the conception of the UKWCS and supervised the secondary data analysis, and all authors provided input on the study design, data analysis, and interpretation of results; revised the paper critically for important intellectual content; and approved the final manuscript version.

Chapter 5 incorporates the work of one jointly-authored publication:

Dong Y, Greenwood DC, Cade JE. Consumption of alcoholic beverages and risk of rheumatoid arthritis: findings from the UK Women's Cohort Study.

This publication was a secondary analysis of the UKWCS dataset. The candidate was responsible for designing the study protocol and statistical analysis plan; data cleaning and analysis; and drafting of the manuscript. JC and DCG were responsible for the conception of the UKWCS and supervised the secondary data analysis, and all authors provided input on the study design, data analysis, and interpretation of results; revised the paper critically for important intellectual content; and approved the final manuscript version.

Chapter 6 incorporates the work of one jointly-authored publication:

Dong Y, Greenwood DC, Cade JE. Healthy lifestyle score, C-reactive protein levels, and incident rheumatoid arthritis: a prospective cohort study of 354,512 UK Biobank participants.

This publication was a secondary analysis of the UK Biobank dataset. The candidate was responsible for designing the study protocol and statistical analysis plan; data cleaning and analysis; and drafting of the manuscript. All authors provided input on the study design, data analysis, and interpretation of results; revised the paper critically for important intellectual content; and approved the final manuscript version.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Yuanyuan Dong to be identified as author of this work has been asserted by her in accordance with the Copyright, Designs and Patents Act 1988.

This thesis is submitted in accordance with the requirements of the “alternative route via publication” due to the volume of published material produced from the analyses conducted in this thesis.

© 2026The University of Leeds and Yuanyuan Dong.

Acknowledgments

I am truly grateful to my supervisors, Prof Janet Cade, Prof Laura Hardie, and Dr Darren Greenwood. They gave me the freedom to explore ambitious ideas while gently guiding me back to the core of the research when needed. Janet supported with unwavering guidance and encouragement from my scholarship application. Darren offered invaluable methodological insight and Laura provided careful scientific critique. I deeply appreciate the time and dedication they invested in helping me think clearly, solve problems, and present my research with confidence.

I thank my colleagues in the Nutritional Epidemiology Group for their companionship and support. The lunchtime seminars enriched my early development as a researcher and fostered a stimulating academic environment. I am particularly grateful to Dr James Webster for his input on the systematic review, and to Professor Jinhui Tao for his expert clinical advice on rheumatology. Many thanks also to Professor Michael Zulyniak and Dr Charlotte Evans for their feedback during my transfer.

I thank all those involved in establishing and managing the UK Women's Cohort Study and UK Biobank, as well as the participants who made this work possible. I am grateful to the Leeds Institute for Data Analytics, especially Ifeanyi Chukwu for technical support, and to George Winder, Jenna King, and Gita Winder for their support in managing my PhD activities. I am grateful to the China Scholarship Council and the University of Leeds who funded this work.

Finally, I thank to my family, who have stood by me with understanding and love. Their support made it possible for me to pursue my PhD abroad and strengthened my commitment to research and discovery.

Contraria sunt complementa.

— Niels Bohr

Abstract

Background: Rheumatoid arthritis (RA) is the most common inflammatory arthritis in the UK, leading to irreversible joint damage, disability, and reduced quality of life. Modifiable risk factors in the pre-diagnostic phase, particularly diet, remain poorly understood. This thesis aimed to investigate associations between dietary and lifestyle factors and the risk of incident RA.

Methods: A dose-response meta-analysis of 30 prospective cohort studies was conducted. Data were obtained from the UK Women's Cohort Study (UKWCS) and UK Biobank, with RA cases ascertained through hospital linkage. Analyses examined central adiposity, adherence to the Mediterranean diet, alcohol intake (frequency and amount). In UK Biobank, a composite healthy lifestyle index was further investigated, and mediation analysis was performed to test whether C-reactive protein (CRP) mediates the association.

Results: The meta-analysis showed that higher intakes of alcohol, fruit, and cereals were associated with reduced RA risk, while tea and caffeinated coffee were linked to increased risk. In the UKWCS, central adiposity measures (particularly waist-to-height ratio) were stronger predictors of RA risk than BMI. High adherence to the Mediterranean diet was associated with a 33% lower RA risk, and moderate wine consumption was also protective, with effect modification by smoking. In UK Biobank, a higher healthy lifestyle score was associated with lower RA risk, an effect stronger in men and leaner individuals. CRP partially mediated this association, explaining approximately 30% of the protective effect.

Conclusion: Across multiple analytical scales and two independent British cohorts, this thesis is the first to demonstrate that various aspects of diet, from individual foods and drinks to overall dietary patterns and lifestyle, are associated with RA risk. The protective effects of the Mediterranean diet, wine consumption and healthy living were quantified, with systemic inflammation accounting for around 30% of the association between healthy lifestyle score and RA incidence.

Table of Contents

Intellectual property and Publication Statements	1
Acknowledgments	4
Abstract	5
List of Tables	17
List of Figures	18
List of Abbreviations	19
Thesis outline	22
Chapter 1 Literature review	23
1.1 Rheumatoid arthritis epidemiology	23
1.1.1 Prevalence	23
1.1.2 Clinical and socioeconomic burden in the UK	24
1.2 Risk factors for rheumatoid arthritis	24
1.2.1 Non-modifiable factors	25
1.2.2 Modifiable lifestyle factors	26
1.2.3 Diet as a key modifiable factor	29
1.2.4 Temporal trends in lifestyle behaviours and policy context	30
1.3 Potential mechanisms for diet and rheumatoid arthritis risk factors ..	30
1.3.1 The gut–joint axis and intestinal barrier dysfunction	31
1.3.2 Dietary modulation of immune responses and inflammation	31
1.3.3 Metabolic pathways, adiposity, and autoimmunity	32
1.4 Aims, hypotheses, and objectives	33
1.4.1 Aim 33	
1.4.2 Hypotheses and objectives	33
1.5 References	35
Chapter 2 Dose–response associations between diet and risk of rheumatoid arthritis: A meta-analysis of prospective cohort studies	44
2.1 Abstract	45
2.2 Introduction	46
2.3 Materials and methods	47
2.3.1. Search Strategy	47
2.3.2 Inclusion criteria and study selection	47
2.3.3 Data extraction and quality assessment	48
2.3.4 Statistical methods	49
2.4 Results	50

2.4.1 Literature search	50
2.4.2 Characteristics of included studies	51
2.4.3 Alcohol consumption and RA risk	52
2.4.4 Non-alcoholic beverages and RA risk	57
2.4.5 Fruit, vegetables, and RA risk	57
2.4.6 Meat, fish, and RA risk	58
2.4.7 Other food components and RA risk	58
2.4.8 Nutrients and RA risk	58
2.4.9 Subgroup analysis and bias analysis	59
2.5 Discussion	60
2.5.1 Main findings	60
2.5.2 Comparison with previous studies	60
2.5.3 Strengths and limitations	64
2.6 Conclusions	65
2.7 References	65
Chapter 3 Measures of adiposity and risk of rheumatoid arthritis in middle-aged UK women: A prospective cohort study	75
3.1 Abstract	76
3.2 Introduction	78
3.3 Materials and methods	79
3.3.1 Study design and participants	79
3.3.2 Anthropometric assessments	80
3.3.3 Assessment of lifestyle factors and other covariates	80
3.3.4 Ascertainment of RA cases and mortality from CVD /cancer ...	81
3.3.5 Statistical methods	81
3.4 Results	83
3.4.1 Participants	83
3.4.2 Associations of BMI, WC, WHR, WHtR, blouse size, skirt size with RA risk	87
3.4.3 Subgroup analysis	91
3.4.4 Sensitivity analysis	91
3.5 Discussion	92
3.5.1 Main Findings	92
3.5.2 Comparison with previous studies	92
3.5.3 Interpretation and implications	94
3.5.4 Strengths and limitaions	95

3.6 Conclusions	96
3.7 References	96
Chapter 4 Association of Mediterranean diet with rheumatoid arthritis risk: Evidence from the UK Women's Cohort Study	100
4.1 Abstract	101
4.2 Introduction	102
4.3 Subjects and methods	103
4.3.1 Study population	103
4.3.2 Dietary assessment	104
4.3.3 Case definition	104
4.3.4 Statistical methods	104
4.4 Results	106
4.4.1 Participants	106
4.4.2 Descriptive data	106
4.4.3 Adherence to the Mediterranean diet	110
4.4.4 Subgroup analysis	112
4.4.5 Sensitivity analysis	113
4.5 Discussion	113
4.5.1 Main findings	113
4.5.2 Comparison with previous studies	113
4.5.3 Interpretation and implications	115
4.5.4 Strengths and limitations	117
4.6 Conclusions	118
4.7 References	118
Chapter 5 Consumption of alcoholic beverages and risk of rheumatoid arthritis: findings from the UK Women's Cohort Study	125
5.1 Abstract	126
5.2 Introduction	126
5.3 Methods	128
5.3.1 Study population	128
5.3.2 Alcohol consumption	129
5.3.3 Incident RA and mortality outcomes	129
5.3.4 Additional covariates	130
5.3.5 Statistical analysis	130
5.4 Results	132
5.4.1 Participants	132

5.4.2	Observational associations with RA incidence and lifestyle factors	132
5.4.3	Total alcohol consumption and risk of RA	137
5.4.4	Consumption of specific alcoholic beverages and risk of RA ..	137
5.4.5	Subgroup analyses	140
5.4.6	Sensitivity analyses	140
5.5	Discussion	140
5.5.1	Principal findings	140
5.5.2	Comparison with previous studies	141
5.5.3	Interpretation and implications	142
5.5.4	Strengths and limitations	143
5.6	Conclusions	144
5.7	References	145
Chapter 6 Healthy lifestyle score, C-reactive protein levels, and incident rheumatoid arthritis: a prospective cohort study of 354,512 UK Biobank participants		150
6.1	Abstract	151
6.2	Introduction	152
6.3	Methods	153
6.3.1	Study population	153
6.3.2	Assessment of healthy lifestyle score	153
6.3.3	Outcome ascertainment	154
6.3.4	Statistical analyses	155
6.4	Results	156
6.4.1	Participants	156
6.4.2	Descriptive data	157
6.4.3	Lifestyle score groups	158
6.4.4	Subgroup analyses	163
6.4.5	Mediation analyses	165
6.4.6	Sensitivity analyses	165
6.5	Discussion	166
6.5.1	Main findings	166
6.5.2	Comparison with previous studies	166
6.5.3	Interpretation and implications	167
6.5.4	Strengths and limitations	168
6.6	Conclusions	169

6.7 References	169
Chapter 7 Summary and Conclusions	174
7.1 Summary of principal findings	175
7.2 Highlights of what is novel in this thesis	179
7.3 Comparison with previous literature	179
7.3.1 The meta-analysis	179
7.3.2 Dietary assessment	181
7.3.3 Cohort differences	183
7.3.4 Dietary exposure: from single food component to dietary pattern 185	
7.4 Critique of potential mechanisms	187
7.4.1 Alcohol intake	187
7.4.2 Mediterranean diet	189
7.4.3 Central obesity	190
7.4.4 Lifestyle, inflammation, and RA risk	191
7.5 Critique of potential limitations	192
7.5.1 The meta-analysis design	192
7.5.2 UKWCS and UK Biobank designs	192
7.5.3 Dietary assessment	193
7.5.4 Case ascertainment using hospital records	193
7.5.5 Statistical approaches and residual confounding	194
7.5.6 Generalisability	195
7.6 Critique of study strengths	195
7.6.1 The meta-analysis design	195
7.6.2 UKWCS and UK Biobank designs	196
7.6.3 Dietary assessment	196
7.6.4 Case ascertainment using hospital records	196
7.6.5 Statistical approaches	197
7.6.6 Generalisability	197
7.7 Recommendations for further research	197
7.7.1 Establish causality and underpinning mechanisms	198
7.7.2 Improve dietary measurement accuracy and precision	198
7.7.3 Investigate RA subtypes and other outcomes	198
7.7.4 Expand evidence to diverse populations	199
7.7.5 Address dietary exposures with insufficient evidence	199
7.7.6 Elucidate the role of diet across the preclinical stages of RA ..	200

7.8 Public health messages	200
7.8.1 Fruit, cereals, and tea	201
7.8.2 Maintain a healthy central adiposity	201
7.8.3 Adherence to a Mediterranean-style diet	201
7.8.4 Moderate alcohol consumption	202
7.8.5 Adopt a healthy lifestyle	202
7.8.6 Balancing benefits and risks	203
7.9 Conclusions	203
7.9.1 What was already known	203
7.9.2 What this thesis adds	204
7.10 References	204
Appendix A Chapter 2 Supplementary material	213
Table A1. Search strategy	213
Table A2. PICOS criteria for inclusion and exclusion of studies	219
Table A3. Standard conversions for food groups and nutrients	220
Table A4. Studies reviewed in full text for eligibility (excluded reasons for meta-analysis)	221
Table A5. Characteristics of the included studies on the consumption of food, and nutrient and risk of rheumatoid arthritis	225
Table A6. Characteristics of 14 included cohort studies	232
Table A7 (a). Relationship between food groups and beverages consumption and risk of rheumatoid arthritis: a pooled analysis comparing the highest with the lowest categories and dose-response analyses	235
Table A7 (b). Relationship between nutrient consumption and risk of RA: a pooled analysis comparing the highest with the lowest categories and dose-response analyses	239
Table A8. Scoring for the food groups, beverages and food components of NutriGrade for the risk of rheumatoid arthritis	242
Table A9. Newcastle Ottawa Scale scores of the included prospective cohort studies	245
Table A10 (a). Subgroup analyses for total alcohol consumption and risk of rheumatoid arthritis	249
Table A10 (b). Subgroup analyses for total alcohol consumption and risk of seropositive rheumatoid arthritis	250
Table A10 (c). Subgroup analyses for total coffee consumption and risk of rheumatoid arthritis	250
Table A10 (d). Subgroup analyses for decaffeinated coffee consumption and risk of rheumatoid arthritis	252

Table A10 (e). Subgroup analyses for tea consumption and risk of rheumatoid arthritis.	253
Table A10 (f). Subgroup analyses for fruit consumption and risk of rheumatoid arthritis.	254
Table A10 (g). Subgroup analyses for vegetables consumption and risk of rheumatoid arthritis.	255
Table A11 (a). Sensitivity analyses for food group consumption and risk of rheumatoid arthritis: excluding one study at a time ¹	256
Table A11 (b). Sensitivity analyses for nutrient intake and risk of rheumatoid arthritis: excluding one study at a time ¹	262
Table A12 (a). Trim and fill tests for food group consumption and risk of rheumatoid arthritis.	264
Table A12 (b). Trim and fill tests for nutrient intake and risk of rheumatoid arthritis.	265
Figure A1. Non-linear dose-response meta-analyses for consumption of food group, and beverages and risk of rheumatoid arthritis using restricted cubic splines.	266
Figure A2. Non-linear dose-response meta-analyses for nutrient intake and risk of rheumatoid arthritis using restricted cubic splines.	267
Figure A3. Linear dose-response analyses for alcohol consumption and risk of rheumatoid arthritis.	268
Figure A4. Meta-analysis of total alcohol consumption and risk of seropositive rheumatoid arthritis comparing highest with the lowest categories.	269
Figure A5. Linear dose-response analyses for total alcohol consumption and risk of seropositive rheumatoid arthritis.	270
Figure A6. Meta-analysis of non-alcoholic beverages (tea and coffee, soda) consumption and risk of rheumatoid arthritis comparing highest with the lowest categories.	271
Figure A7. Linear dose-response analyses for non-alcoholic beverages (tea, coffee, and soda) consumption and risk of rheumatoid arthritis.	272
Figure A8. Meta-analysis of decaffeinated coffee, and tea consumption and risk of RF-positive rheumatoid arthritis comparing highest with the lowest categories.	273
Figure A9. Linear dose-response analyses for decaffeinated coffee, and tea consumption and risk of RF-positive rheumatoid arthritis.	273
Figure A10. Meta-analysis of fruit, and vegetable consumption and risk of rheumatoid arthritis comparing highest with the lowest categories.	274
Figure A11. Linear dose-response analyses for fruit, and vegetable consumption and risk of rheumatoid arthritis.	275

Figure A12. Meta-analysis of meat, and fish consumption and risk of rheumatoid arthritis comparing highest with the lowest categories.	276
Figure A13. Linear dose-response analyses for meat, and fish consumption and risk of rheumatoid arthritis.	277
Figure A14. Meta-analysis of other food group consumption and risk of rheumatoid arthritis comparing highest with the lowest categories.	278
Figure A15. Linear dose-response analyses for other food group consumption and risk of rheumatoid arthritis.	279
Figure A16. Meta-analysis of dietary n-3 PUFA and risk of rheumatoid arthritis comparing highest with the lowest categories.	279
Figure A17. Linear dose-response analyses for dietary n-3 PUFA and risk of rheumatoid arthritis.	281
Figure A18. Meta-analysis of vitamin intake and risk of rheumatoid arthritis comparing highest with the lowest categories.	281
Figure A19. Linear dose-response analyses for vitamins intake and risk of rheumatoid arthritis.	282
Figure A20. Meta-analysis of dietary carotenoids intake and the risk of rheumatoid arthritis comparing highest with the lowest categories.	283
Figure A21. Linear dose-response analyses for dietary carotenoids intake and risk of rheumatoid arthritis.	284
Figure A22. Funnel plots for food group, and beverages consumption and risk of rheumatoid arthritis.	285
Figure A23. Funnel plots for nutrients intake and risk of rheumatoid arthritis.	286
References	286
Appendix B Chapter 3 Supplementary material	290
Figure B1. Flow chart of UKWCS participants.	290
Table B1. Strengthening the reporting of observational studies in nutritional epidemiology (STROBE-Nut) checklist.	291
Table B2. Adiposity group categorisation and definitions.	296
Table B3. Covariates at recruitment and their derivation.	297
Table B4. The proportional hazards assumption test for obesity indicators and rheumatoid arthritis risk (in the fully-adjusted model).	299
Table B5. Baseline characteristics of 27,968 UKWCS participants stratified by body mass index, waist circumference, waist-to-hip ratio, and waist-to-height ratio.	300
Table B6. Baseline characteristics of 27,968 UKWCS participants stratified by clothing size.	301

Table B7. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by age.	303
Table B8. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by weight change.	303
Table B9. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by AHEI-2010.	304
Table B10. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by physical activity.	305
Table B11. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by menopausal status.	305
Table B12. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by smoking status.	306
Table B13. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by age.	307
Table B14. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by weight change.	309
Table B15. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by AHEI-2010.	311
Table B16. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by physical activity.	313
Table B17. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by smoking status.	316
Table B18. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by menopausal status.	317
Table B19. Risk of rheumatoid arthritis by adiposity indicators with varying restrictions in the UKWCS.	319
Supplementary methods	329
Appendix C Chapter 4 Supplementary material	330
Figure C1. Flow chart of UKWCS participants.	330
Table C1. Strengthening the reporting of observational studies in nutritional epidemiology (STROBE-Nut) checklist.	331

Table C2. Covariates at recruitment and their derivation.	335
Table C3. Grading of the Mediterranean diet score.	337
Table C4. Dietary characteristics of the UKWCS by adherence to Mediterranean diet at recruitment.	338
Table C5. Associations between individual Mediterranean diet components and risk of rheumatoid arthritis in the UK Women's Cohort Study.	340
Table C6. Association between adherence to the Mediterranean diet and risk of rheumatoid arthritis by potential effect modifiers in the UKWCS.	343
Table C7. Sensitivity analysis for the association between risk of rheumatoid arthritis and the Mediterranean diet adherence with varying restrictions in the UKWCS.	346
Supplementary Methods	347
References	347
Appendix D Chapter 5 Supplementary material	349
Table D1. Strengthening the reporting of observational studies in nutritional epidemiology (STROBE-Nut) checklist.	349
Table D2. Conversion of alcohol nutrient intake ^[1]	353
Table D3. Components of death causes.	353
Table D4. Covariates at recruitment and their derivation.	354
Table D5. Baseline characteristics of 23,699 UKWCS participants stratified by drinking frequency.	355
Table D6. Risk of rheumatoid arthritis according to drinking status stratified by potential effect modifiers	357
in the UKWCS.	357
Table D7. Risk of rheumatoid arthritis according to amount of alcohol drinking stratified by potential effect	358
modifiers in the UKWCS.	358
Table D8. Sensitivity analyses of alcohol consumption and rheumatoid arthritis incidence in the UKWCS.	361
Appendix E Chapter 6 Supplementary material	364
Figure E1. Flow chart of UK Biobank participants.	364
Figure E2. Time-varying coefficient ($\beta(t)$) of log-transformed C-reactive protein (logCRP) in the Cox model.	365
Table E1. Strengthening the reporting of observational studies in nutritional epidemiology (STROBE-Nut) checklist.	366
Table E2. Diet group categorisation and definitions.	370
Table E3. Healthy lifestyle Score definition.	372

Table E4. Summary of mediation analyses using the product-of-coefficients method with bootstrap resampling (UK Biobank).	373
Table E5. Characteristics of UK Biobank participants by lifestyle score group at recruitment, stratified by sex.	374
Table E6. Risk of rheumatoid arthritis by lifestyle score group in the UK Biobank with varying restrictions.	377

List of Tables

Table3.1	Baseline characteristics of 27,968 UKWCS participants stratified by rheumatoid arthritis incidence.	84
Table4.1	Baseline characteristics of 27,766 UKWCS participants.	109
Table4.2	Association between Mediterranean diet adherence and risk of rheumatoid arthritis in the UK Women’s Cohort Study.	111
Table5. 1	Baseline characteristics of 23,699 UKWCS participants by drinking status.	135
Table5.2	Rheumatoid arthritis risk by drinking status and amount of total alcohol consumption.	136
Table5.3	Consumption of specific alcoholic beverages and incident rheumatoid arthritis risk.	138
Table5.4	Risk of rheumatoid arthritis according to alcohol drinking status stratified by smoking status in the UKWCS.	139
Table6.1	Baseline characteristics of participants in the UK Biobank according to C-reactive protein levels.	1599
Table6.2	Risk of rheumatoid arthritis among life score groups in UK Biobank participants, stratified by sex, region, and body mass index.	164
Table7.1	Summary of key findings, potential explanations, and novelty in this thesis.	176

List of Figures

Figure 1.1	A conceptual framework for the accumulation of risk in rheumatoid arthritis development	26
Figure 2.1	PRISMA flow chart summarising article retrieval and screening process. .	51
Figure 2.2	The summary of pooled HRs (95% CIs) for different dietary exposures and incident RA comparing the highest with the lowest categories	54
Figure 2.3	Meta-analysis of alcohol consumption and risk of RA comparing the highest with the lowest categories	55
Figure 2.4	Non-linear dose–response association of food component consumption with risk of RA using restricted cubic splines	56
Figure 3.1	Association between indicators of adiposity and risk of rheumatoid arthritis in the UK Women’s Cohort Study	89
Figure 3.2	Association between indicators of adiposity with rheumatoid arthritis incidence	90
Figure 4.1	Directed Acyclic Graph illustrating the hypothesised relationship between the Mediterranean diet score, rheumatoid arthritis incidence, and related factors	108
Figure 4.2	Restricted cubic spline for the association between rheumatoid arthritis incidence and the Mediterranean diet score.	112
Figure 5.1	Flow chart of UKWCS participants	133
Figure 5.2	Directed Acyclic Graph	134
Figure 6.1	Risk of rheumatoid arthritis by healthy lifestyle score	162
Figure 6.2	Mediation analysis of C-reactive protein	165

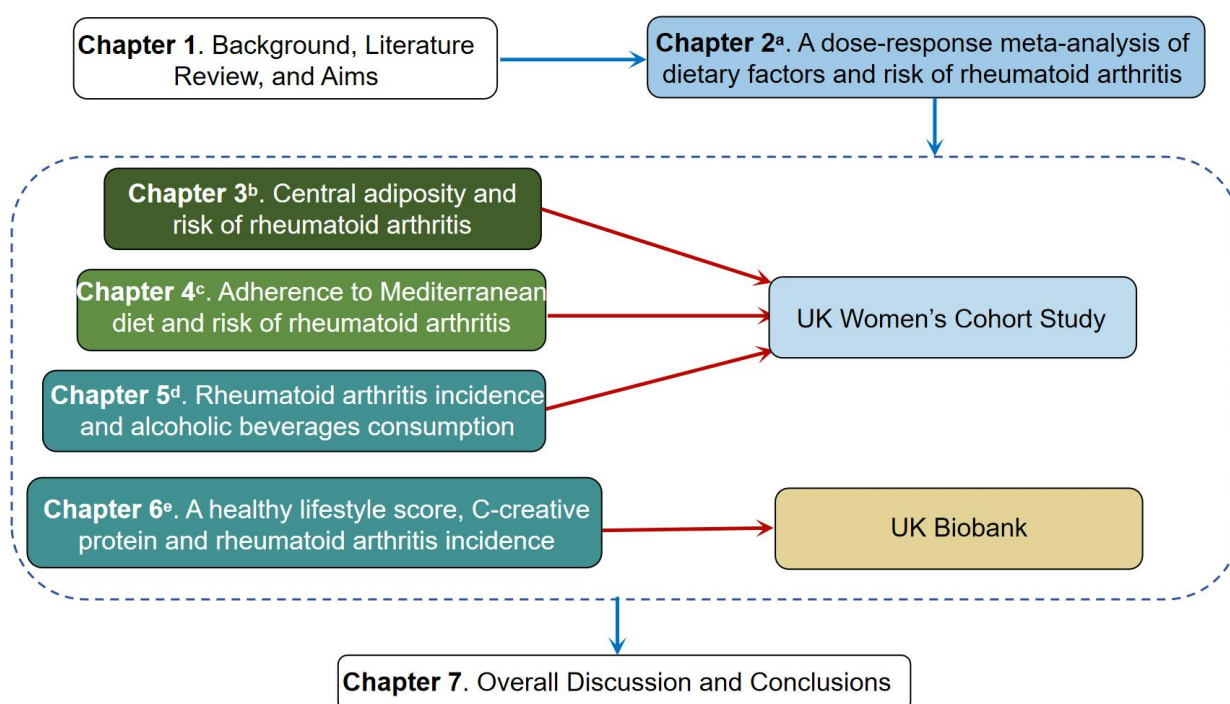
List of Abbreviations

ACPA	Anti-citrullinated protein antibody
ACR	American College of Rheumatology
AGEs	Advanced glycation end products
AHEI	Alternative Healthy Eating Index
BMI	Body mass index
CI	Confidence interval
CRP	C-reactive protein
CVD	Cardiovascular disease
DAG	Directed acyclic graph
EPIC	European Prospective Investigation into Cancer and Nutrition
EULAR	European League against Rheumatism
FDRs	First-degree relatives
FFQ	Food frequency questionnaire
FSR	French Society for Rheumatology
GBD	Global Burden of Disease
GPR43	G protein-coupled receptor 43
HES	Hospital Episode Statistics
HLA	Human leukocyte antigen
HLA-DRB1	Human leukocyte antigen - DR beta 1
HR	Hazard ratio
HRT	Hormone replacement therapy

HSCIC	Health and Social Care Information Centre
ICD	International Classification of Diseases
IL-6	Interleukin-6
IRR	Incidence rate ratio
IU	International units
MEDRA	Mediterranean diet intervention in Rheumatoid Arthritis (trial name)
METs	Metabolic equivalents of tasks
METS-VF	Metabolic Score for Visceral Fat
MUFA	Monounsaturated fatty acids
NHANES	National Health and Nutrition Examination Survey
NHSCR	National Health Service Central Register
NICE	National Institute for Health and Care Excellence
NOS	Newcastle–Ottawa Scale
NRES	National Research Ethics Service
PAF	Population attributable fraction
PH	Proportional hazards
PICO	Population, Intervention, Comparison, Outcome
PM2.5	Particulate matter $\leq 2.5 \mu\text{m}$
PRISMA	Preferred Reporting Items for Systematic review and Meta-Analysis
PUFA	Polyunsaturated fatty acids
PUFAs	Polyunsaturated fatty acids
RA	Rheumatoid arthritis

RAGE	Receptor for advanced glycation end products
RCS	Restricted cubic spline
RCTs	Randomised controlled trials
RF	Rheumatoid factor
RR	Relative risk
SCFA	Short-chain fatty acid
SD	Standard deviation
SE	Shared epitope
SES	Socio-economic status
SFA	Saturated fatty acids
SPMs	Specialized pro-resolving mediators
STROBE-nut	Strengthening the Reporting of Observational Studies in Epidemiology—Nutritional Epidemiology
Tfh	T follicular helper
TNF- α	Tumour necrosis factor-alpha
Tregs	Regulatory T cells
UK	United Kingdom
UKWCS	UK Women's Cohort Study
WC	Waist circumference
WCRF	World Cancer Research Fund
WHI	Women's Health Initiative
WHO	World Health Organisation
WHR	Waist-to-hip ratio
WHtR	Waist-to-height ratio

Thesis outline



^aDong Y, Greenwood DC, Webster J, Uzokwe C, Tao J, Hardie LJ, Cade JE. Dose-Response Associations Between Diet and Risk of Rheumatoid Arthritis: A Meta-Analysis of Prospective Cohort Studies. *Nutrients*, 2024;16(23), 4050. <https://doi.org/10.3390/nu16234050>

^bDong Y, Greenwood DC, Hardie LJ, Cade JE. Measures of Adiposity and Risk of Rheumatoid Arthritis in Middle-Aged UK Women: A Prospective Cohort Study. *Nutrients*. 2025; 17(9):1557. <https://doi.org/10.3390/nu17091557>

^cDong Y, Greenwood DC, Cade JE. Association of Mediterranean diet with rheumatoid arthritis risk: evidence from the UK Women's Cohort Study.

^dDong Y, Greenwood DC, Cade JE. Consumption of alcoholic beverages and risk of rheumatoid arthritis: findings from the UK Women's Cohort Study.

^eHealthy lifestyle score, C-reactive protein levels, and incident rheumatoid arthritis: a prospective cohort study of 354,512 UK Biobank participants.

Chapter 1 Literature review

1.1 Rheumatoid arthritis epidemiology

1.1.1 Prevalence

Rheumatoid arthritis (RA) is the most common chronic inflammatory arthritis, affecting 0.5-1.0% of the global population [1]. Worldwide, an estimated 18 million people live with RA [2], over half of whom are of working age [3], underlining the societal and economic burden of premature work loss. Between 1990 and 2019, the age-standardised incidence rates rose from approximately 12.2 to 13.0 per 100,000, while prevalence rate increased from 183.0 to 208.8 per 100,000 [3]. This upward trend is set to continue, with the total number of global cases nearly doubling from 17.6 million in 2020 to an estimated 31.7 million by 2050 [3]. The burden is disproportionately skewed towards women, who are two to three times more likely to be affected than men across all populations [4]. Data from a 2020 study of English primary care records show that the point-prevalence of diagnosed RA in England was 0.78%, with prevalence in women approximately 2.4 times higher than in men [5]. This consistent sex disparity points to the critical roles of hormonal, genetic, and immunological factors in disease development, and underscores the importance of examining sex-specific risk factors in RA research.

RA is diagnosed according to the 2010 American College of Rheumatology/European League Against Rheumatism (ACR/EULAR) classification criteria, which incorporate joint involvement, serological markers, acute-phase reactants, and symptom duration [6]. Seropositive RA, defined by the presence of rheumatoid factor (RF) and/or anti-cyclic citrullinated peptide (anti-CCP) antibodies, is generally associated with more aggressive disease progression and worse long-term outcomes compared with seronegative RA, which lacks these autoantibodies and is considered to have distinct genetic and prognostic features [7]. Throughout this thesis, a distinction is made between prevalent RA (existing cases at baseline) and incident RA (new cases arising

during follow-up), reflecting the different study designs employed across the empirical chapters.

1.1.2 Clinical and socioeconomic burden in the UK

The clinical course of RA extends beyond joint damage, evolving into a complex systemic syndrome with multi-organ involvement. While characterised by symmetrical synovial inflammation and subsequent bone erosion, the persistent systemic inflammatory state in RA serves as a catalyst for diverse extra-articular comorbidities [8]. Notably, RA is associated with an approximately 50% increase in cardiovascular disease (CVD) risk, which remains the primary driver of inpatient admissions and premature mortality among RA patients [9,10]. It causes significant disability and reduces life expectancy by three to ten years [11]. For affected individuals, routine tasks requiring fine motor dexterity can become painful tests of endurance, due to severely compromised hand and wrist function.

RA imposes a substantial and unequal socioeconomic toll on the UK population [12,13]. It is fundamentally an unequal condition, disproportionately affecting the most vulnerable. Individuals from lower-income or deprived areas are more likely to develop RA, experience more severe symptoms, and face poorer long-term clinical outcomes [12,14]. For the estimated 700,000 people living with RA in the UK, biological and social vulnerability often leads to early exit from the workforce [13]. Approximately one-third of patients stop working within five years of diagnosis [15]. In 2024, musculoskeletal conditions, predominantly RA, accounted for an estimated 23 million lost working days in the UK [16]. Furthermore, the financial strain on the healthcare system is immense, with the combined cost of RA and osteoarthritis to the National Health Service (NHS) projected to reach £120 billion over the next decade [17]. These trends emphasise an urgent need for affordable, lifestyle-based strategies to reduce disease risk and alleviate its broader societal impact.

1.2 Risk factors for rheumatoid arthritis

Although the aetiology of RA is not fully understood, it is clear that the disease development arises from a complex interplay of genetic, environmental, and

lifestyle determinants that drive immune dysregulation and chronic systemic inflammation [18,19].

1.2.1 Non-modifiable factors

Genetic factors are estimated to account for 60% of RA risk [20]. The strongest association is with immune system, human leukocyte antigen (HLA)-DRB1 alleles that encode the shared epitope (SE), a five amino acid sequence motif in residues 70-74 of the HLA-DR β chain. This SE confers particular risk for anti-citrullinated protein antibody (ACPA)-positive RA [21]. The SE is thought to increase disease susceptibility by enhancing the presentation of citrullinated autoantigens to T cells, thereby lowering the threshold for breaking immune tolerance [22]. This heritable component is reflected in the 2-4 times higher risk of RA in first-degree relatives (FDRs) of patients [23]. Even higher prevalence and greater seropositivity are observed in certain ethnic groups, such as North American native populations, reaching up to three to five times for White populations worldwide [24,25]. Research shows that compared with White RA patients, African American, Hispanic patients perform worse on higher disease activity, more severe symptoms [26]. Female sex, older age and hormonal changes also influence disease susceptibility. However, the concordance rate in monozygotic twins is only 12-15% [27,28], implying that environmental and lifestyle factors play a critical role in triggering disease onset. Taken together, these non-modifiable factors establish a biological baseline of risk, but rarely trigger RA in isolation (as shown in **Figure 1.1**).

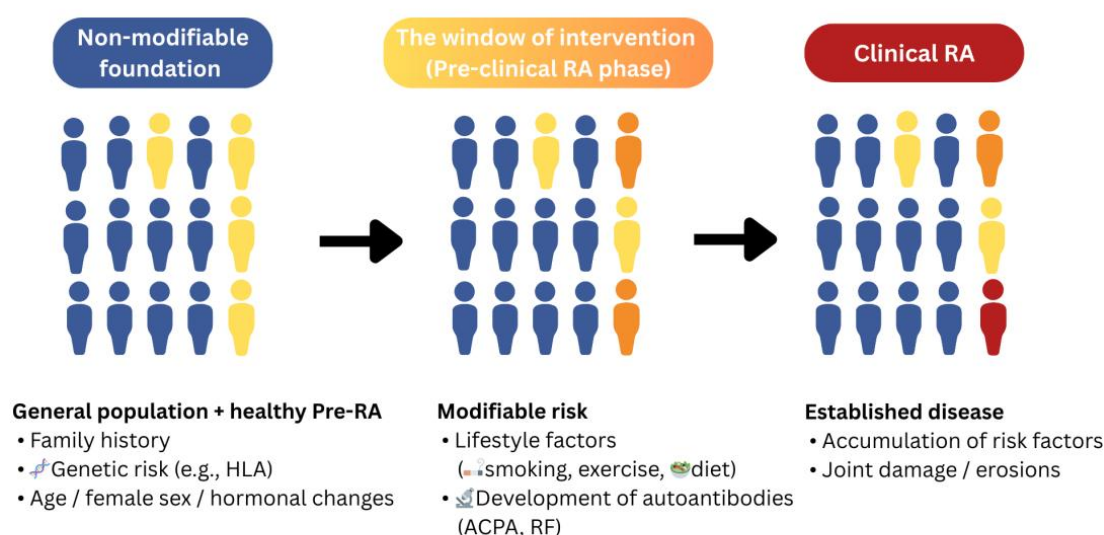


Figure 1.1 A conceptual framework for the accumulation of risk in rheumatoid arthritis development. It shows three sequential phases: genetic predisposition (yellow), pre-clinical RA (orange), and clinical disease onset (red). The framework emphasises the pre-clinical phase as a critical window for lifestyle intervention. RA, rheumatoid arthritis; HLA, human leukocyte antigen; ACPA, anti-citrullinated protein antibody; RF, rheumatoid factor.

1.2.2 Modifiable lifestyle factors

Smoking

The association between cigarette smoking and RA risk is supported by methodologically robust evidence from large prospective cohorts with consistent findings [29,30] and biological plausibility [31]. Consequently, this thesis treats smoking as an established causal factor and adjusts for it as a confounder.

A body of evidence from prospective cohorts, mostly from the US and Europe, showed that smoking has been consistently linked to RA risk [31-33], an association driven primarily by ACPA-positive cases. The biological plausibility is supported by experimental studies showing that smoking induces protein citrullination in the lungs, generating neoantigens that trigger ACPA production in genetically susceptible individuals [31,34,35]. The interaction between smoking and HLA-DRB1 SE alleles therefore is one of the most well-characterised examples of gene-environment interplay in autoimmune disease [36-38]. The dose-response relationship has been reported in a meta-analysis of three cohort studies and seven case-control studies, which found that smoking was

associated with RA incidence even at low levels (<10 pack-years). The risk doubled at >20 pack-years and plateaued thereafter, with a stronger effect observed for rheumatoid factor (RF)-positive than RF-negative RA [39]. However, this meta-analysis was limited by the predominance of case-control designs, which are more susceptible to recall bias than prospective cohorts. Also, the confidence intervals for the highest exposure category were wide, reflecting imprecise risk estimates at very high smoking intensities.

More recent large-scale prospective studies have addressed these limitations. A UK Biobank analysis of over 200,000 participants demonstrated that in utero tobacco exposure was associated with an 18% increased risk of RA, with earlier smoking initiation further elevating risk in a dose-response manner [40]. This study also highlighted the synergistic effect of smoking with systemic inflammation, providing direct evidence from a contemporary British setting [40]. Similar findings have been reported in other large cohorts, including the Nurses' Health Study [30] and the Swedish Epidemiological Investigation of Rheumatoid Arthritis (EIRA) [41], confirming the robustness of the smoking-RA association across different populations.

Obesity

Obesity is a consistent risk factor for RA, though its impact may depend on sex, body composition, and RA serological subtype. A recent systematic review and meta-analysis of 20 observational studies (including 8 cohorts and 12 case-control studies) with 568,889 participants found that overweight (BMI 25–30 kg/m²) was associated with a 13% increased risk of RA, while obesity (BMI ≥30 kg/m²) was associated with a 25% greater risk, with the association being more pronounced in women [42]. These findings align with population attributable fraction (PAF) estimates from US adults, which indicate that excess BMI accounts for approximately 14.7% of RA incidence [43]. Adipose tissue functions as an active endocrine organ, secreting pro-inflammatory adipokines and cytokines, including tumour necrosis factor-alpha (TNF-α) and interleukin-6 (IL-6), which contribute to systemic low-grade inflammation [44] and promote autoimmunity [45]. Mendelian randomisation studies provided a causal evidence, supporting a link between higher BMI and RA risk thorough a shared genetic basis at the HLA region (chromosome 6, positions 28–35M and 50–52M) [46].

However, BMI captures only general obesity and ignores the distribution of body fat. Central adiposity is more metabolically active and pro-inflammatory than subcutaneous fat because secreted factors from visceral fat (such as IL-6 and free fatty acids) directly reach the liver via the portal circulation, promoting systemic inflammation [47,48]. A recent UK Biobank study showed that stable abdominal obesity and transition from non-abdominal to abdominal obesity were both associated with increased RA risk [49]. Additionally, a study using data from National Health and Nutrition Examination Survey (NHANES) found that the Metabolic Score for Visceral Fat (METS-VF, a composite of insulin resistance, waist-to-height ratio, age, and sex) showed superior predictive ability for RA compared with BMI only [50]. The 2025 Lancet Commission on clinical obesity has called for a paradigm shift away from BMI-only diagnostic criteria, explicitly incorporating measures of central adiposity [51]. Given the limitations of single adiposity indices, Chapter 3 provides the first comparison of multiple adiposity measures in relation to incident RA risk in a large British cohort.

Other factors

Other modifiable factors have been implicated in RA risk, such as occupational dust exposure (e.g., silica, asbestos) [52,53], air pollution (particularly particulate matter $\leq 2.5 \mu\text{m}$, PM_{2.5}) [54], periodontal disease [55], and gut microbiome dysbiosis [56,57]. For example, studies on silica exposure have reported elevated RA risk among miners and stone workers, whereas findings on air pollution have been inconsistent across geographic regions [58,59]. Critically, the evidence is generally less robust due to prevalent methodological limitations: most studies are small, cross-sectional, or rely on imprecise exposure measures, making it difficult to exclude reverse causation or residual confounding. Furthermore, for factors such as gut microbiome dysbiosis, the direction of causality remains ambiguous, as the alterations may be a consequence of RA or its treatment rather than a cause [60].

Given these uncertainties and the lack of consistent prospective evidence, these environmental factors were not the primary focus of this thesis.

1.2.3 Diet as a key modifiable factor

Current guidelines from the ACR and the EULAR emphasise the prevention of rheumatic and musculoskeletal diseases through smoking cessation, regular exercise (both aerobic and strengthening), maintenance of a healthy weight, and a balanced diet [61,62]. Strong evidence demonstrates that smoking not only increases RA risk but also worsens symptoms, functional outcomes, disease activity, and comorbidities [63]. Similarly, the benefits of regular exercise are supported by robust evidence for improving muscular endurance and strength, reducing pain, and mitigating systemic manifestations [64].

In contrast, evidence for dietary adjustments is less developed. Among the various dietary approaches investigated, only the Mediterranean diet has accumulated low-to-moderate certainty evidence for improving pain and reducing disease activity [61] and is more likely to reduce the RA risk [65]. However, the existing literature on diet and RA risk has several limitations. First, most studies are observational design, which limits causal inference due to the potential for residual confounding, reverse causation, and recall bias in dietary assessment. Second, study populations vary in demographic composition and geographical location, raising questions about the generalisability of findings to the UK population. Third, many case-control studies have relied on unvalidated or retrospectively collected dietary assessment methods, which are prone to differential recall between cases and controls. Fourth, earlier investigations focused on single nutrients or individual food groups rather than dietary patterns. This approach fails to account for the synergistic and antagonistic interactions among dietary components that usually occur in real-world eating behaviours. Although a paradigm shift from considering single nutrients/foods towards examining whole dietary patterns has occurred, prospective evidence linking dietary patterns to RA risk remains rare and the results have been mixed.

Even among cohort-based studies, heterogeneity persists in the definitions of dietary exposures, the timing of dietary assessment relative to RA diagnosis, and the handling of potential confounders. A recent systematic literature review of diet and RA risk including 19 cohort-based studies and 8 case-control studies, and concluded the lack of evidence on the role of diet in the development of RA, partly

due to differences in study quality and methodology [66]. Consequently, clinical practice guidelines acknowledge the importance of diet, but concede that the evidence base to support specific, individualised dietary recommendations for RA prevention is currently insufficient.

Current NICE guideline NG100 [67] recommend a healthy balanced diet and weight management for RA patients, but do not specify targeted dietary interventions for disease modification.

1.2.4 Temporal trends in lifestyle behaviours and policy context

Over the past three decades, the UK has seen major changes in RA-related lifestyle behaviours, alongside shifts in public health policy and social norms. Smoking prevalence has fallen sharply, from around 45% of adults in the mid-1970s to 10.6% by 2024 [68], driven by smoke-free legislation (implemented in England in 2007), advertising bans, and plain packaging. In contrast, obesity prevalence has more than doubled, rising from about 14–15% in the early 1990s to around 30% by 2024 [69], reflecting an increasingly obesogenic environment with larger portion sizes, greater availability of ultra-processed foods, and more sedentary lifestyles [70]. Alcohol guidance has also shifted considerably: the 1995 UK recommendations allowed up to 2–3 units/day for women and 3–4 units/day for men, whereas the 2016 guidelines reduced the low-risk threshold to 14 units/week for both sexes, with greater emphasis on cancer risk [71]. These contextual differences may partly explain variations in lifestyle–RA associations observed between the UKWCS (recruited 1995–1998, included a validated 217-item FFQ) [72] and UK Biobank (recruited 2006–2010). These factors are revisited in the discussion (Chapter 7).

1.3 Potential mechanisms for diet and rheumatoid arthritis risk factors

Mechanistic studies offer plausible biological pathways that may explain how dietary factors influence disease risk. These mechanisms, centred on the gut microbiome, immune regulation, and metabolic function, are interconnected and provide a biological rationale for the role of diet in the transition from genetic risk to clinical disease.

1.3.1 The gut–joint axis and intestinal barrier dysfunction

Systemic immune responses that promote joint inflammation can be triggered by alterations in the composition and function of the gut microbiota [73]. Long-term dietary intake is a key factor in human microbial ecology [74], and alterations in gut microbiota have been documented in pre-clinical RA [73,75]. Western diets, high in red meat, processed foods and refined sugars, and low in fibre, can promote a pro-inflammatory microbial profile [76,77]. This includes a reduction in short-chain fatty acid (SCFA)-producing bacteria (e.g., *Faecalibacterium prausnitzii*) and an overgrowth of potentially pathogenic species [78]. In contrast, dietary patterns rich in fibre, such as the Mediterranean diet, improve the growth of SCFA-producing bacteria. SCFAs (e.g., butyrate, propionate, acetate) are key metabolites that exert pleiotropic anti-inflammatory effects. They strengthen the intestinal epithelial barrier, reducing gut permeability (“leaky gut”) [79]. Furthermore, SCFAs promote the differentiation of regulatory T cells (Tregs) while inhibiting pro-inflammatory Th17 cells, thereby helping to maintain systemic immune homeostasis [80]. The breakdown of this homeostatic balance, characterised by increased gut permeability and the translocation of bacterial products into the body’s systemic circulation, is an initiating event in the breach of immune tolerance as clinical RA onset [81].

However, most existing cross-sectional studies cannot establish temporality [82,83]. Also, observed microbial alterations may reflect disease activity, medication use, or post-diagnosis dietary changes rather than causal precursors [60].

1.3.2 Dietary modulation of immune responses and inflammation

Certain dietary components can modulate immune cell function and the inflammatory milieu. High intake of n-6 polyunsaturated fatty acids (PUFAs), common in vegetable oils and processed foods, can be metabolised into pro-inflammatory eicosanoids [84]. In contrast, n-3 PUFAs, found in oily fish, are precursors to specialised pro-resolving mediators (SPMs) such as resolvins and protectins, which actively resolve inflammation rather than merely suppressing it [85,86]. This pro-resolving capacity is relevant to RA, where the failure to resolve chronic inflammation is a central question. Observational studies have linked

higher n-3 PUFA intake with a lower risk of ACPA-positive RA, suggesting a role in the early, pre-clinical stages of autoimmunity [87]. Vitamin D can modulate both innate and adaptive immunity. It suppresses the differentiation of Th17 cells and promotes Treg differentiation [88]. Low vitamin D status has been consistently associated with an increased risk of RA development and higher disease activity [89,90]. Similarly, dietary antioxidants, including carotenoids and vitamins C and E found in fruits and vegetables, can mitigate oxidative stress, a state of cellular damage implicated in the pathogenesis of RA [91,92].

For n-3 PUFAs, several large prospective cohorts support an inverse association with RA risk [93], but randomised controlled trials have only tested supplementation in established RA [94], not prevention. For vitamin D and antioxidants, most human data come from cross-sectional or case-control studies [95], making the evidence less conclusive.

1.3.3 Metabolic pathways, adiposity, and autoimmunity

Having established obesity as a risk factor in Section 1.2.2, this section examines the mechanistic pathways through which adiposity may promote autoimmunity. Adipose tissue is not merely an inert energy store but an active endocrine organ that secretes pro-inflammatory cytokines, including tumour necrosis factor- α (TNF- α), interleukin-6 (IL-6), and leptin [96]. This chronic, low-grade inflammatory state driven by adiposity may contribute to the systemic inflammation that predisposes to autoimmunity. The mechanisms linking obesity to RA risk may also involve enhanced peripheral aromatisation of androgens to oestrogens in adipose tissue, contributing to the sex disparity observed in RA [97]. Diet can also influence post-translational protein modifications that generate autoantigens. Driven by high-glycaemic diets and obesity, chronic hyperglycaemia and insulin resistance can promote non-enzymatic glycation, leading to the formation of advanced glycation end products (AGEs) [98]. AGEs can induce protein modifications and activate inflammatory pathways via their receptor (RAGE), contributing to RA development and related autoimmune condition [99].

1.4 Aims, hypotheses, and objectives

The previous sections of this chapter have outlined the epidemiology, the clinical burden of RA, described unmodifiable and modifiable risk factors, and considered plausible biological mechanisms linking diet to immune regulation and inflammation. Despite a substantial body of epidemiological research investigating dietary factors and RA risk, which I review in detail in Chapter 2, several important gaps remain unaddressed.

1.4.1 Aim

The overarching aim of this thesis was:

To investigate associations between dietary factors and rheumatoid arthritis risk in British adults using data linkage.

The rationale for this work was to support the formulation of dietary recommendations for reducing RA risk, providing updated information on the adequacy of specific dietary patterns (including Mediterranean diet and alcohol consumption), as well as the role of diet within a broader healthy lifestyle context. This could result in fewer cases and delayed RA onset.

1.4.2 Hypotheses and objectives

This thesis tested the following objectives and associated hypotheses, organised by chapter:

Chapter 2 Dose-Response Associations Between Diet and Risk of Rheumatoid Arthritis: A Meta-Analysis of Prospective Cohort Studies.

Objective: To synthesise prospective evidence on associations between dietary factors and RA risk, quantify dose-response relationships for key exposures, and evaluate the quality of existing evidence.

Hypothesis: Prospective cohort studies show consistent associations between specific dietary factors (e.g., fish, fibre, red meat) and RA risk, with dose-response relationships detectable for key exposures.

Chapter 3 Measures of Adiposity and Risk of Rheumatoid Arthritis in Middle-Aged UK Women: A Prospective Cohort Study.

Objective: To investigate associations between multiple adiposity measures and incident RA risk, and to compare their predictive performance.

Hypothesis: Central adiposity measures (waist circumference, waist-to-hip ratio, waist-to-height ratio) are more strongly associated with RA risk than BMI alone.

Chapter 4 Association of Mediterranean diet with rheumatoid arthritis risk: evidence from the UK Women's Cohort Study.

Objective: To examine whether adherence to the Mediterranean diet is associated with incident RA risk, independent of other lifestyle factors.

Hypothesis: Higher adherence to the Mediterranean diet is associated with a lower risk of RA in British middle-aged women.

Chapter 5 Consumption of alcoholic beverages and risk of rheumatoid arthritis: findings from the UK Women's Cohort Study.

Objective: To investigate whether the association between alcohol consumption (frequency and amount) and RA risk differs by type of alcoholic beverage.

Hypothesis: The association between alcohol frequency, and consumption and RA risk differs by beverage type, even after accounting for competing risks.

Chapter 6 Healthy lifestyle score, C-reactive protein levels, and incident rheumatoid arthritis: a prospective cohort study of 354,512 UK Biobank participants.

Objective: To construct a composite healthy lifestyle index and examine its association with incident RA risk, including mediation by C-reactive protein.

Hypothesis: A composite healthy lifestyle score is inversely associated with RA risk, and C-reactive protein level account for part of this association.

In the next chapter (Chapter 2), I review the prospective evidence on dietary factors and RA risk through a systematic review and meta-analysis.

1.5 References

1. Chen L, Wu B, Mo L, et al: Associations between biological ageing and the risk of, genetic susceptibility to, and life expectancy associated with rheumatoid arthritis: a secondary analysis of two observational studies. *Lancet Healthy Longev.* 2024;5(1):e45–e55.
2. WHO: Rheumatoid arthritis. 2023;[Available from: <https://www.who.int/news-room/fact-sheets/detail/rheumatoid-arthritis>]
3. Global, regional, and national burden of rheumatoid arthritis, 1990-2020, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. *Lancet Rheumatol.* 2023;5(10):e594–e610.
4. Carmona L, Aurrecoechea E, García de Yébenes MJ: Tailoring Rheumatoid Arthritis Treatment through a Sex and Gender Lens. *J Clin Med.* 2023;13(1)
5. Scott IC, Whittle R, Bailey J, et al: Rheumatoid arthritis, psoriatic arthritis, and axial spondyloarthritis epidemiology in England from 2004 to 2020: An observational study using primary care electronic health record data. *Lancet Reg Health Eur.* 2022;23:100519.
6. Aletaha D, Neogi T, Silman AJ, et al: 2010 Rheumatoid arthritis classification criteria: An American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis & Rheumatism.* 2010;62(9):2569–2581.
7. Matthijssen XME, Huizinga TWJ, Niemantsverdriet E, van der Helm-van Mil AHM: Early intensive treatment normalises excess mortality in ACPA-negative RA but not in ACPA-positive RA. *Annals of the Rheumatic Diseases.* 2020;79(10):e124.
8. Figus FA, Piga M, Azzolin I, McConnell R, Iagnocco A: Rheumatoid arthritis: Extra-articular manifestations and comorbidities. *Autoimmun Rev.* 2021;20(4):102776.

9. Avina-Zubieta JA, Thomas J, Sadatsafavi M, Lehman AJ, Lacaille DJAotrd: Risk of incident cardiovascular events in patients with rheumatoid arthritis: a meta-analysis of observational studies. 2012;71(9):1524–1529.
10. Iyer P, Gao Y, Field EH, et al: Trends in Hospitalization Rates, Major Causes of Hospitalization, and In-Hospital Mortality in Rheumatoid Arthritis in the United States From 2000 to 2014. *ACR Open Rheumatol*. 2020;2(12):715–724.
11. Alamanos Y, Drosos AA: Epidemiology of adult rheumatoid arthritis. *Autoimmun Rev*. 2005;4(3):130–136.
12. Izadi Z, Li J, Evans M, et al: Socioeconomic Disparities in Functional Status in a National Sample of Patients With Rheumatoid Arthritis. *JAMA Network Open*. 2021;4(8):e2119400–e2119400.
13. Whittaker W, Higgerson J, Eden M, Payne K, Wilkie R, Verstappen SM: Effects of employees living with an 'arthritis' on sickness absence and transitions out of employment: a comparative observational study in the UK. *RMD Open*. 2024;10(4)
14. Venetsanopoulou AI, Alamanos Y, Voulgari PV, Drosos AA: Epidemiology and Risk Factors for Rheumatoid Arthritis Development. *Mediterr J Rheumatol*. 2023;34(4):404–413.
15. Connolly D, Fitzpatrick C, O'Toole L, Doran M, O'Shea F: Impact of Fatigue in Rheumatic Diseases in the Work Environment: A Qualitative Study. *Int J Environ Res Public Health*. 2015;12(11):13807–13822.
16. ONS.: Sickness absence in the UK labour market: 2023 and 2024. 2025 [Available from: <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/labourproductivity/articles/sicknessabsenceinthelabourmarket/2023and2024>]
17. Economics. YH: The Cost of Arthritis: Calculation conducted on behalf of Arthritis Research UK. 2017;
18. Romão VC, Fonseca JE: Etiology and Risk Factors for Rheumatoid Arthritis: A State-of-the-Art Review. *Front Med (Lausanne)*. 2021;8:689698.
19. Firestein GS, McInnes IB: Immunopathogenesis of Rheumatoid Arthritis. *Immunity*. 2017;46(2):183–196.
20. Dedmon LE: The genetics of rheumatoid arthritis. *Rheumatology (Oxford)*. 2020;59(10):2661–2670.

21. Shakeel L, Shaukat A, Khaliq N, et al: Rheumatoid arthritis: a comprehensive overview of genetic markers, emerging therapies, and personalized medicine. *Ann Med Surg (Lond)*. 2025;87(2):696–710.
22. Song Y, Li J, Wu Y: Evolving understanding of autoimmune mechanisms and new therapeutic strategies of autoimmune disorders. *Signal Transduct Target Ther*. 2024;9(1):263.
23. Bunnewell S, Wells I, Zemedikun D, et al: Predictors of perceived risk in first-degree relatives of patients with rheumatoid arthritis. *RMD Open*. 2022;8(2)
24. Hitchon CA, L ON, Peschken CA, Robinson DB, Fowler-Woods A, El-Gabalawy HS: Disparities in rheumatoid arthritis outcomes for North American Indigenous populations. *Int J Circumpolar Health*. 2023;82(1):2166447.
25. El-Gabalawy H: The Impact of Rheumatoid Arthritis on First Nations and How We Can Work With Communities to Prevent It. *J Rheumatol*. 2024;
26. Greenberg JD, Spruill TM, Shan Y, et al: Racial and ethnic disparities in disease activity in patients with rheumatoid arthritis. 2013;126(12):1089–1098.
27. Silman AJ, MacGregor AJ, Thomson W, et al: Twin concordance rates for rheumatoid arthritis: results from a nationwide study. *Br J Rheumatol*. 1993;32(10):903–907.
28. Aho K, Koskenvuo M, Tuominen J, Kaprio J: Occurrence of rheumatoid arthritis in a nationwide series of twins. *J Rheumatol*. 1986;13(5):899–902.
29. Hedström AK, Klareskog L, Alfredsson L: Exposure to passive smoking and rheumatoid arthritis risk: results from the Swedish EIRA study. *Ann Rheum Dis*. 2018;77(7):970–972.
30. Liu X, Tedeschi SK, Barbhuiya M, et al: Impact and Timing of Smoking Cessation on Reducing Risk of Rheumatoid Arthritis Among Women in the Nurses' Health Studies. *Arthritis Care Res (Hoboken)*. 2019;71(7):914–924.
31. Klareskog L, Malmström V, Lundberg K, Padyukov L, Alfredsson L: Smoking, citrullination and genetic variability in the immunopathogenesis of rheumatoid arthritis. *Semin Immunol*. 2011;23(2):92–98.
32. De Rooy D, van Nies J, Kapetanovic MC, et al: Smoking as a risk factor for the radiological severity of rheumatoid arthritis: a study on six cohorts. 2014;73(7):1384–1387.

33. Di Giuseppe D, Orsini N, Alfredsson L, Askling J, Wolk AJAr, therapy: Cigarette smoking and smoking cessation in relation to risk of rheumatoid arthritis in women. 2013;15(2):R56.
34. Lugli EB, Correia RE, Fischer R, et al: Expression of citrulline and homocitrulline residues in the lungs of non-smokers and smokers: implications for autoimmunity in rheumatoid arthritis. 2015;17(1):9.
35. Catrina A, Krishnamurthy A, Rethi BJRo: Current view on the pathogenic role of anti-citrullinated protein antibodies in rheumatoid arthritis. 2021;7(1)
36. Lee YH, Bae SC, Song GG: Gene-environmental interaction between smoking and shared epitope on the development of anti-cyclic citrullinated peptide antibodies in rheumatoid arthritis: a meta-analysis. *Int J Rheum Dis*. 2014;17(5):528–535.
37. Ishikawa Y, Terao C: The Impact of Cigarette Smoking on Risk of Rheumatoid Arthritis: A Narrative Review. *Cells*. 2020;9(2)
38. Baka Z, Buzás E, Nagy G: Rheumatoid arthritis and smoking: putting the pieces together. *Arthritis Res Ther*. 2009;11(4):238.
39. Di Giuseppe D, Discacciati A, Orsini N, Wolk A: Cigarette smoking and risk of rheumatoid arthritis: a dose-response meta-analysis. *Arthritis Res Ther*. 2014;16(2):R61.
40. Dou X, He M, Wang Y, et al: Early-life exposure to tobacco, inflammation, and risk of rheumatic diseases: a prospective cohort study. *J Public Health (Oxf)*. 2025;47(4):710–720.
41. Hedström AK, Stawiarz L, Klareskog L, Alfredsson L: Smoking and susceptibility to rheumatoid arthritis in a Swedish population-based case-control study. *Eur J Epidemiol*. 2018;33(4):415–423.
42. Cao F, Kang X, Wang W, Huang S: Association of body mass index with the risk of rheumatoid arthritis: a systematic review and meta-analysis. *Front Med (Lausanne)*. 2025;12:1750640.
43. Ye D, Mao Y, Xu Y, Xu X, Xie Z, Wen C: Lifestyle factors associated with incidence of rheumatoid arthritis in US adults: analysis of National Health and Nutrition Examination Survey database and meta-analysis. *BMJ Open*. 2021;11(1):e038137.
44. Uti DE, Atangwho IJ, Omang WA, et al: Cytokines as key players in obesity low grade inflammation and related complications. 2025;54:100585.

45. Procaccini C, De Rosa V, Galgani M, et al: Role of adipokines signaling in the modulation of T cells function. 2013;4:332.
46. Tang B, Shi H, Alfredsson L, Klareskog L, Padyukov L, Jiang X: Obesity-Related Traits and the Development of Rheumatoid Arthritis: Evidence From Genetic Data. *Arthritis Rheumatol*. 2021;73(2):203–211.
47. Maikaew P, Phattanakiatsakul T, Sitticharoon C, Keadkraichaiwat I: Distinct metabolic associations of subcutaneous and visceral adipocyte morphology in women with or without obesity. *Sci Rep*. 2025;15(1):41287.
48. Item F, Konrad D: Visceral fat and metabolic inflammation: the portal theory revisited. *Obes Rev*. 2012;13 Suppl 2:30–39.
49. Zhao X, Sun X, Wu F, et al: Longitudinal patterns of obesity index changes and risk of incident rheumatoid arthritis: evidence from a population-based cohort. *Arthritis Res Ther*. 2025;27(1):186.
50. Li Y, Zhu Y, Tang X, et al: Association of visceral fat metabolism score with risk of rheumatoid arthritis in US adults. *Front Nutr*. 2025;12:1544624.
51. Rubino F, Cummings DE, Eckel RH, et al: Definition and diagnostic criteria of clinical obesity. *Lancet Diabetes Endocrinol*. 2025;13(3):221–262.
52. Stolt P, Källberg H, Lundberg I, Sjögren B, Klareskog L, Alfredsson L: Silica exposure is associated with increased risk of developing rheumatoid arthritis: results from the Swedish EIRA study. *Ann Rheum Dis*. 2005;64(4):582–586.
53. Ilar A, Klareskog L, Saevarsdottir S, et al: Occupational exposure to asbestos and silica and risk of developing rheumatoid arthritis: findings from a Swedish population-based case-control study. *RMD Open*. 2019;5(2):e000978.
54. Cheng B, Pan C, Cai Q, et al: Long-term ambient air pollution and the risk of musculoskeletal diseases: A prospective cohort study. *J Hazard Mater*. 2024;466:133658.
55. Kaur S, White S, Bartold MJ: Periodontal disease as a risk factor for rheumatoid arthritis: a systematic review. 2012;10(42):1–12.
56. Horta-Baas G, Romero-Figueroa MdS, Montiel-Jarquín AJ, Pizano-Zárate ML, García-Mena J, Ramírez-Durán NJ: Intestinal dysbiosis and rheumatoid arthritis: a link between gut microbiota and the pathogenesis of rheumatoid arthritis. 2017;2017(1):4835189.

57. Wells PM, Adebayo AS, Bowyer RC, et al: Associations between gut microbiota and genetic risk for rheumatoid arthritis in the absence of disease: a cross-sectional study. 2020;2(7):e418–e427.
58. Wrangel O, Graff P, Bryngelsson IL, Fornander L, Wiebert P, Vihlborg P: Silica Dust Exposure Increases Risk for Rheumatoid Arthritis: A Swedish National Registry Case-Control Study. *J Occup Environ Med*. 2021;63(11):951–955.
59. Essouma M, Noubiap JJ: Is air pollution a risk factor for rheumatoid arthritis? *J Inflamm (Lond)*. 2015;12:48.
60. Wang Y, Wei J, Zhang W, et al: Gut dysbiosis in rheumatic diseases: A systematic review and meta-analysis of 92 observational studies. *EBioMedicine*. 2022;80:104055.
61. Gwinnutt JM, Wieczorek M, Balanescu A, et al: 2021 EULAR recommendations regarding lifestyle behaviours and work participation to prevent progression of rheumatic and musculoskeletal diseases. *Ann Rheum Dis*. 2023;82(1):48–56.
62. England BR, Smith BJ, Baker NA, et al: 2022 American College of Rheumatology Guideline for Exercise, Rehabilitation, Diet, and Additional Integrative Interventions for Rheumatoid Arthritis. *Arthritis Care Res (Hoboken)*. 2023;75(8):1603–1615.
63. Chang K, Yang SM, Kim SH, Han KH, Park SJ, Shin JI: Smoking and rheumatoid arthritis. *Int J Mol Sci*. 2014;15(12):22279–22295.
64. Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH: Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews. *Cochrane Database Syst Rev*. 2017;4(4):Cd011279.
65. Mosalmanzadeh N, Jandari S, Soleimani D, et al: Major dietary patterns and food groups in relation to rheumatoid arthritis in newly diagnosed patients. *Food Sci Nutr*. 2020;8(12):6477–6486.
66. Bäcklund R, Drake I, Bergström U, Compagno M, Sonestedt E, Turesson C: Diet and the risk of rheumatoid arthritis - A systematic literature review. *Semin Arthritis Rheum*. 2023;58:152118.
67. NICE. Rheumatoid arthritis in adults: management. 2020.[Available from: <https://www.nice.org.uk/guidance/ng100>]

68. ONS. Adult smoking habits in the UK: 2024. 2025.[Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandlifeexpectancies/bulletins/adultsmokinghabitsingreatbritain/2024>]
69. NHS Digital. Health Survey for England, 2024: adults' overweight and obesity. 2026.[Available from: <https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/2024/adults-overweight-and-obesity>]
70. Martínez Steele E, Baraldi LG, Louzada ML, Moubarac JC, Mozaffarian D, Monteiro CA: Ultra-processed foods and added sugars in the US diet: evidence from a nationally representative cross-sectional study. *BMJ Open*. 2016;6(3):e009892.
71. Care DoHaS. UK Chief Medical Officers' alcohol guidelines review: report from the guidelines development group to the UK Chief Medical Officers. 2016.[Available from: https://assets.publishing.service.gov.uk/media/5a7507d8ed915d3c7d529d6f/GD_G_report-Jan2016.pdf]
72. Cade JE, Burley VJ, Alwan NA, et al: Cohort Profile: The UK Women's Cohort Study (UKWCS). *International Journal of Epidemiology*. 2017;46(2):e11–e11.
73. Zaiss MM, Joyce Wu HJ, Mauro D, Schett G, Ciccia F: The gut-joint axis in rheumatoid arthritis. *Nat Rev Rheumatol*. 2021;17(4):224–237.
74. David LA, Maurice CF, Carmody RN, et al: Diet rapidly and reproducibly alters the human gut microbiome. *Nature*. 2014;505(7484):559–563.
75. Tajik N, Frech M, Schulz O, et al: Targeting zonulin and intestinal epithelial barrier function to prevent onset of arthritis. *Nat Commun*. 2020;11(1):1995.
76. Thorburn AN, Macia L, Mackay CR: Diet, metabolites, and "western-lifestyle" inflammatory diseases. *Immunity*. 2014;40(6):833–842.
77. Tilg H, Moschen AR: Food, immunity, and the microbiome. *Gastroenterology*. 2015;148(6):1107–1119.
78. Smith PM, Howitt MR, Panikov N, et al: The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science*. 2013;341(6145):569–573.
79. Matei DE, Menon M, Alber DG, et al: Intestinal barrier dysfunction plays an integral role in arthritis pathology and can be targeted to ameliorate disease. *Med*. 2021;2(7):864–883.e869.

80. Kim JY, Lim K, Kim KH, Kim JH, Choi JS, Shim SC: N-3 polyunsaturated fatty acids restore Th17 and Treg balance in collagen antibody-induced arthritis. *PLoS One*. 2018;13(3):e0194331.
81. Holers VM, Demoruelle MK, Kuhn KA, et al: Rheumatoid arthritis and the mucosal origins hypothesis: protection turns to destruction. *Nat Rev Rheumatol*. 2018;14(9):542–557.
82. Wells PM, Adebayo AS, Bowyer RCE, et al: Associations between gut microbiota and genetic risk for rheumatoid arthritis in the absence of disease: a cross-sectional study. *Lancet Rheumatol*. 2020;2(7):e418–e427.
83. Alpizar-Rodriguez D, Lesker TR, Gronow A, et al: *Prevotella copri* in individuals at risk for rheumatoid arthritis. *Ann Rheum Dis*. 2019;78(5):590–593.
84. Patterson E, Wall R, Fitzgerald GF, Ross RP, Stanton C: Health implications of high dietary omega-6 polyunsaturated Fatty acids. *J Nutr Metab*. 2012;2012:539426.
85. Oppedisano F, Macrì R, Gliozzi M, et al: The Anti-Inflammatory and Antioxidant Properties of n-3 PUFAs: Their Role in Cardiovascular Protection. *Biomedicines*. 2020;8(9)
86. Norris PC, Skulas-Ray AC, Riley I, et al: Identification of specialized pro-resolving mediator clusters from healthy adults after intravenous low-dose endotoxin and omega-3 supplementation: a methodological validation. *Sci Rep*. 2018;8(1):18050.
87. Gan RW, Demoruelle MK, Deane KD, et al: 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*. 2017;76(1):147–152.
88. E LB, Ismailova A, Dimeloe S, Hewison M, White JH: Vitamin D and Immune Regulation: Antibacterial, Antiviral, Anti-Inflammatory. *JBMR Plus*. 2021;5(1):e10405.
89. Merlino LA, Curtis J, Mikuls TR, Cerhan JR, Criswell LA, Saag KG: Vitamin D intake is inversely associated with rheumatoid arthritis: results from the Iowa Women's Health Study. *Arthritis Rheum*. 2004;50(1):72–77.
90. Lin J, Liu J, Davies ML, Chen W: Serum Vitamin D Level and Rheumatoid Arthritis Disease Activity: Review and Meta-Analysis. *PLoS One*. 2016;11(1):e0146351.

91. Low CE, Loke S, Chew NSM, Lee A, Tay SH: Vitamin, antioxidant and micronutrient supplementation and the risk of developing incident autoimmune diseases: a systematic review and meta-analysis. *Front Immunol.* 2024;15:1453703.
92. Mateen S, Moin S, Khan AQ, Zafar A, Fatima N: Increased Reactive Oxygen Species Formation and Oxidative Stress in Rheumatoid Arthritis. *PLoS One.* 2016;11(4):e0152925.
93. Di Giuseppe D, Wallin A, Bottai M, Askling J, Wolk A: Long-term intake of dietary long-chain n-3 polyunsaturated fatty acids and risk of rheumatoid arthritis: a prospective cohort study of women. *Ann Rheum Dis.* 2014;73(11):1949–1953.
94. Wang W, Xu Y, Zhou J, Zang Y: Effects of omega-3 supplementation on lipid metabolism, inflammation, and disease activity in rheumatoid arthritis: a meta-analysis of randomized controlled trials. *Clin Rheumatol.* 2024;43(8):2479–2488.
95. Kriegel MA, Manson JE, Costenbader KH: Does vitamin D affect risk of developing autoimmune disease?: a systematic review. *Semin Arthritis Rheum.* 2011;40(6):512–531.e518.
96. Feijóo-Bandín S, Aragón-Herrera A, Moraña-Fernández S, et al: Adipokines and Inflammation: Focus on Cardiovascular Diseases. *Int J Mol Sci.* 2020;21(20)
97. Balistreri CR, Caruso C, Candore G: The role of adipose tissue and adipokines in obesity-related inflammatory diseases. *Mediators Inflamm.* 2010;2010:802078.
98. Gu MJ, Lee YR, Kim D, Kim Y, Ha SK: Comprehensive research on the properties of advanced glycation end products in food and biological samples and their harmful role in inducing metabolic diseases. *Compr Rev Food Sci Food Saf.* 2024;23(5):e13412.
99. Zhou M, Zhang Y, Shi L, et al: Activation and modulation of the AGEs-RAGE axis: Implications for inflammatory pathologies and therapeutic interventions - A review. *Pharmacol Res.* 2024;206:107282.

Chapter 2 Dose–response associations between diet and risk of rheumatoid arthritis: A meta-analysis of prospective cohort studies

Yuanyuan Dong ¹, Darren C. Greenwood ², James Webster ³, Chinwe Uzokwe ¹, Jinhui Tao⁴, Laura J. Hardie ⁵ and Janet E. Cade ¹

1 Nutritional Epidemiology Group, School of Food Science and Nutrition, University of Leeds, Leeds LS2 9JT, UK;

2 Leeds Institute for Data Analytics, Faculty of Medicine and Health, University of Leeds, Leeds LS2 9JT, UK;

3 Applied Health Research Unit, Nuffield Department of Population Health, University of Oxford, Oxford OX3 7LF, UK;

4 Department of Rheumatology and Immunology, The First Affiliated Hospital of University of Science and Technology of China, Hefei 230001, China;

5 Division of Clinical and Population Sciences, Leeds Institute of Cardiovascular and Metabolic Medicine, School of Medicine, University of Leeds, Leeds LS2 9JT, UK.

Published: Nutrients. 2024. 16(23):4050. <https://doi.org/10.3390/nu16234050>

This Chapter is an exact copy of the journal paper referred to above.

2.1 Abstract

Background: To provide a systematic and quantitative summary of dietary factors and rheumatoid arthritis (RA) risk.

Methods: A systematic review and meta-analysis included prospective cohort studies from 2000 to 2024 reporting relative risks (RRs) with 95% confidence intervals (CIs) for RA incidence relating to 32 different dietary exposures. Linear and non-linear dose–response analyses were conducted.

Results: Thirty studies were included, involving 2,986,747 participants with 9,677 RA cases. Linear dose–response analysis suggested that each 2-unit per week increase in total alcohol intake was linked to 4% risk reduction (RR (95%-CI), heterogeneity (I^2), NutriGrade score: 0.96 (0.94, 0.98), 58%, moderate certainty), and beer consumption was associated with a 10% reduction per 2 units/week increase (0.90 (0.84, 0.97), 0%, very low certainty). Each 2-unit/week increase in total alcohol intake was associated with a 3% decrease in seropositive RA risk (0.97 (0.96, 0.99), 28%, moderate certainty). Increased intakes of fruit (per 80 g/day) and cereals (per 30 g/day) were associated with 5% (0.95 (0.92, 0.99), 57%, moderate certainty) and 3% (0.97 (0.96, 0.99), 20%, moderate certainty) reduced risk, respectively. Conversely, tea consumption showed a 4% increased risk per additional cup/day (1.04 (1.02, 1.05), 0%, moderate certainty). Non-linear associations were observed for total coffee, vegetables, oily fish, and vitamin D supplementation. Data on dietary patterns and specific micronutrients were limited.

Conclusions: The findings suggest that moderate alcohol consumption and a higher intake of fruits, oily fish, and cereals are associated with a reduced risk of RA, while tea and coffee may be linked to an increased risk. Optimising dietary intake of certain food components may reduce RA risk, despite moderate-quality evidence.

Keywords: Rheumatoid arthritis; diet; meta-analysis; dose–response; nutrient intake

2.2 Introduction

Rheumatoid arthritis (RA) is the most common systemic autoimmune disease, characterised by the presence of autoantibodies, like rheumatoid factor (RF) and anti-citrullinated protein antibody (ACPA) [1]. While it can lead to long-term disability and increased morbidity [2], early diagnosis and guideline-based treatment significantly mitigate these effects. The Global Burden of Disease (GBD) 2021 study stated that the incidence of RA was set to continue increasing, with an estimated global prevalence of 17.6 million cases in 2020 and notable regional variations [3]. Only a few established risk elements for RA are recognised, including older age, female sex, infections, air pollution, and cigarette smoking [4].

Diet has been associated with other autoimmune diseases [5], and both genetic and environmental factors can influence the development of RA [6]. A better understanding of how diet affects RA could help inform more effective prevention strategies. Several dietary elements, particularly plant-based foods rich in fibre, vitamins, and antioxidants, can reduce systemic inflammation by lowering levels of C-reactive protein (CRP), interleukin-6 (IL-6), and tumour necrosis factor- α (TNF- α) [7]. Other components, such as short-chain fatty acids and immune-enhancing nutrients like vitamin D [8] and folate [9], may contribute to inflammation regulation.

Nevertheless, the roles of dietary patterns and individual components remain unclear. For instance, the Mediterranean diet (MD) has been associated with lower inflammatory markers and reduced risks of chronic diseases like type 2 diabetes and cardiovascular disease [10,11], but direct evidence for its effect on RA risk is still limited. The benefit of alcohol consumption for RA risk has been the subject of previous studies [12]. A 2014 meta-analysis demonstrated a non-linear trend indicating that low-to-moderate alcohol intake reduced the risk of RA [13], but it included non-peer-reviewed literature. In addition, data on specific alcoholic beverages are scarce. Although several case–control studies have suggested that tea consumption may reduce RA disease activity [14,15], findings from the UK Biobank indicated that it could increase RA risk [16]. The French Society for Rheumatology (FSR) recommends omega-3 supplementation for inflammatory

rheumatic diseases [17]; however, a recent meta-analysis of randomised controlled trials (RCTs) found no reduction in RA symptoms or inflammation [18]. There is also evidence regarding the impact of anti-rheumatic therapies on periodontal health, highlighting the complex relationship between lifestyle factors and RA progression [19]. The inconsistency in these findings points to the need for further research to clarify the effects of food groups and nutrients on RA risk and assess the credibility of the evidence.

Two recent systematic reviews examined the role of diet in preventing RA, but did not include individual dietary components in their search strategies [20,21]. To address these gaps, this study aims to synthesise the existing data and provide high-quality evidence on the associations of dietary components with the risk of RA in a dose–response analysis.

2.3 Materials and methods

The findings from this systematic review and dose–response meta-analysis are reported based on the Preferred Reporting Items for Systematic review and Meta-Analysis (PRISMA) guidelines [22]. The protocol for this review was registered on PROSPERO as CRD42022320959.

2.3.1. Search Strategy

We conducted a systematic search of all articles published from 1 January 2000 to 30 April 2024 using databases, including Medline, Embase, Web of Science, and the Cochrane Library. In brief, we searched for prospective cohort studies that reported associations of dietary factors (food groups, beverages, food components, nutrients, and dietary patterns) with RA incidence. Full details of the search strategy are provided in **Appendix A: Table A1**. The search followed the PICO framework [23], as shown in **Appendix A: Table A2**. To avoid missing any publications, the references of relevant papers were checked.

2.3.2 Inclusion criteria and study selection

The eligibility criteria for inclusion in the meta-analysis were as follows: (1) original research studies; (2) human studies conducted on adults; (3) prospective studies that provided data on dietary consumption factors; (4) studies reporting the effect

sizes, including hazard ratio (HR) and relative ratio (RR) with 95% confidence intervals (CIs), where the outcome was RA incidence; (5) English-language studies; and (6) studies for which full texts were available. The exclusion criteria were: (1) reviews and book chapters, or secondary research evidence such as meta-analyses; (2) animal studies, letters, comments, cross-sectional studies, and ecological studies; (3) overlapping studies or data; (4) studies without dietary exposure or measurement (e.g., studies that only report smoking and RA risk); (5) studies focusing on RA-related symptoms, life quality, functional status, or disease activity; and (6) Mendelian randomisation studies.

The publications were independently assessed by YD, JC, DG, LH, and CU, and any discrepancies were resolved through discussion with the senior authors (YD and JC). After deduplication, articles were screened for eligibility based on their titles and abstracts, followed by a detailed assessment using the *Systematic Review Accelerator* [24]. If there were duplicate publications from the same cohort, we included the study with the longer follow-up or the larger number of participants or incident cases, provided the data were sufficient to generate effect estimates for meta-analysis.

Our research was based on the 1987 American Rheumatism Association (ACR) classification criteria for RA [25], expecting that prospective cohort studies would be incorporated into the research.

2.3.3 Data extraction and quality assessment

JW and YD extracted data from each eligible article using a predesigned form, including the study details (first author, cohort name, publication year, study location, follow-up duration, dietary exposure, cohort size, and the number of RA cases), participant characteristics (age range, health status at entry, and sex), assessment methods (mean/median or range of food/nutrient intake, and RA case ascertainment), statistical analysis (models and confounders), and effect sizes (incidence rate ratio [IRR], HR, and RR) from the most fully adjusted models. A meta-analysis was conducted only for exposure–outcome associations with data from at least two studies. For the dose–response meta-analysis, studies were excluded if they lacked data on non-cases or person-years.

The study quality was assessed using the Newcastle–Ottawa Scale (NOS), a tool commonly employed for evaluating observational studies based on their selection, comparability, and outcome assessment [26]. Studies scoring above six were considered high-quality, while those scoring six or below were considered low-quality. We used the NutriGrade tool (developed in Nuthetal, Brandenburg, Germany) to evaluate the certainty of the meta-evidence for each dietary exposure’s association with RA [27]. NutriGrade assigns a score (out of 10) by considering factors including study design, risk of bias, precision, heterogeneity, directness, publication bias, funding bias, effect size, and dose–response. The meta-evidence was classified as high (8~10), moderate (6~7.99), low (4~5.99), or very low (0~3.99). Moderate-quality evidence indicates robust findings with some limitations, while low-quality evidence reflects greater uncertainty.

2.3.4 Statistical methods

Associations for both highest versus lowest category comparisons and dose–response analyses were investigated. RR or HR estimates from individual studies were combined using a random effects model. The methods described by Greenland and Longnecker [28] and Orsini et al. [29] were used for the dose–response meta-analysis. To estimate dose amounts, the mean, median, or midpoint values of the upper and lower limits were used. When the ranges were open-ended, we estimated the limits based on adjacent intervals. Restricted cubic splines with three knots (at 10%, 50%, and 90% of exposure distribution) to handle potential non-linear relationships were used. For the non-linear dose–response analysis, we included studies that reported the RR with 95% CIs for at least three categories of exposure to estimate a dose–response trend.

To standardise the dietary exposures across studies, we converted the consumption data into a common scale, such as grams per week for alcoholic beverages. The standard conversions are shown in **Appendix A: Table A3**. To ensure consistency in the analysis, we used the lowest exposure as the reference category. If the reference group was different from the lowest category, we recalculated the effect size using the method of Orsini et al. [30]. Meta-analyses were performed for food groups, beverages (e.g., total alcohol, wine, beer, coffee types, tea, fruits, vegetables, meats, fish, dairy products, cereals), nutrients, and

phytochemicals (e.g., vitamin C, D, A, E, *n*-3 polyunsaturated fatty acids, PUFAs, carotenoids).

Heterogeneity across studies was assessed by I^2 statistic, which quantifies the percentage of total variation due to heterogeneity rather than chance, with $I^2 > 50\%$ indicating substantial heterogeneity. When substantial heterogeneity was detected, we conducted subgroup analyses for dietary factors with data from at least 5 studies, categorised by study characteristics such as age (≥ 50 vs. < 50), gender (both sexes vs. female only), location (Europe, America, Asia), follow-up duration (< 10 years vs. ≥ 10 years), dietary assessment method (food frequency questionnaires, FFQs vs. food diaries), case ascertainment method (self-reported vs. registry-based data), and number of RA cases (< 500 vs. ≥ 500).

The potential effects of small-study biases, including publication bias, were evaluated using funnel plots, Egger's test [31], and Begg's test [32], provided there were enough studies to perform the test. Sensitivity analyses were conducted to evaluate the robustness of the results by excluding 1 study at a time when more than 3 studies were available. All the statistical analyses were performed using *R* version 4.2.0 (R Foundation for Statistical Computing, Vienna, Austria). A two-tailed *P* value of < 0.05 was regarded as statistically significant.

2.4 Results

2.4.1 Literature search

The search strategy identified 3525 unique citations, of which 55 full-text articles were assessed in detail (**Figure 2.1**). The reasons for exclusion during full-text screening are provided in **Appendix A: Table A4**.

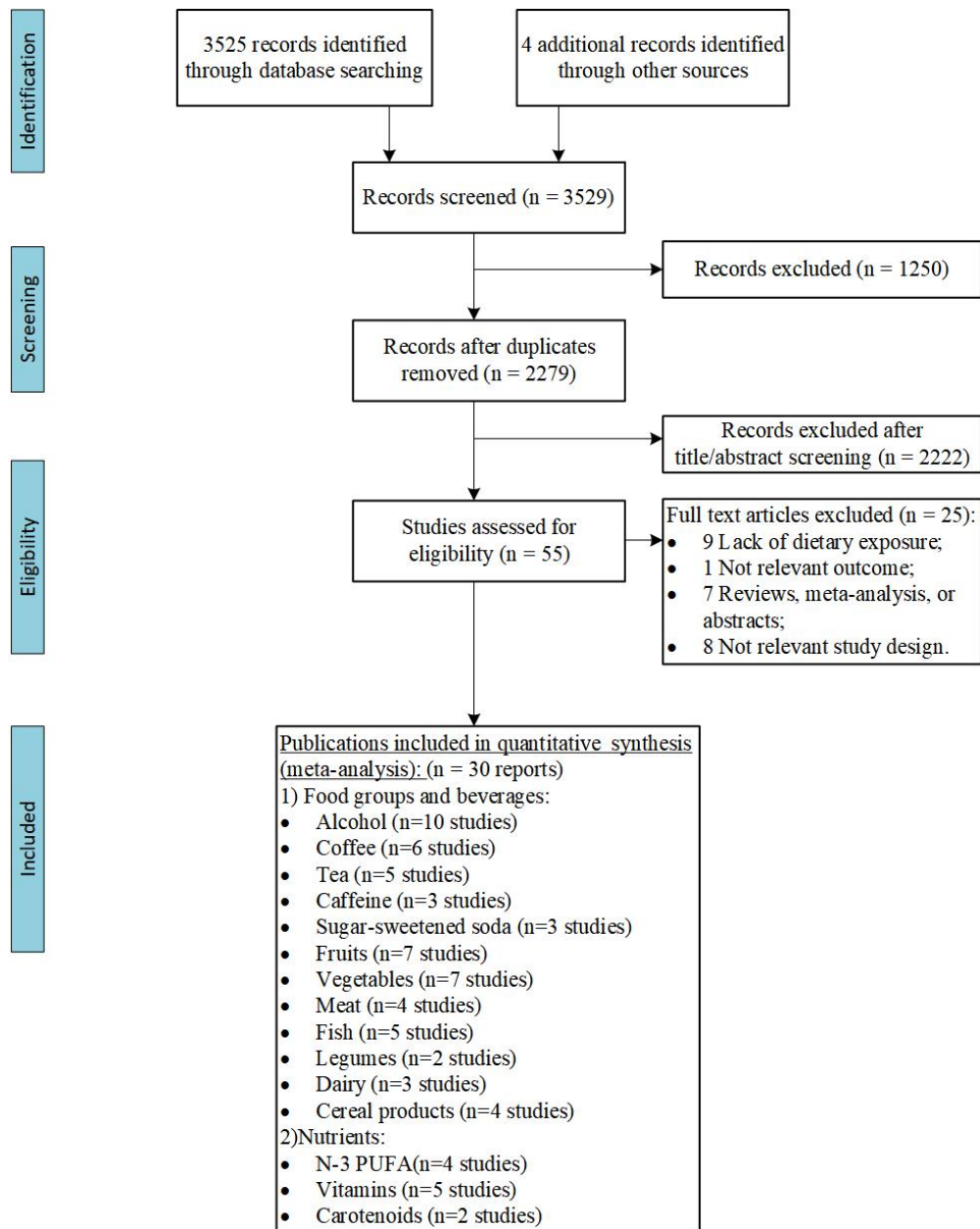


Figure 2.1 PRISMA flow chart summarising article retrieval and screening process.

2.4.2 Characteristics of included studies

The characteristics of the 30 studies included in the analysis, which investigated the relationship between dietary factors and the risk of developing RA [12,16,33-58], are summarised in **Figure 2.2** and **Appendix A: Table A5**. Of these, 22 articles focused exclusively on women [81, 85, 92-103], and 8 included both sexes [104-111]. The studies were primarily nationally representative: 12 were conducted in Europe [85, 97, 98, 104-108, 110, 112-114], 15 in America [81, 92-96, 99-103, 115-118], and 3 in Asia [50-53,58,60].

Overall, during the follow-up periods ranging from 4 to 30 years, 14 prospective studies included 9677 RA cases and 8772 seropositive RA cases among 2,986,747 participants with an age range spanning from 20 to 98 years (**Appendix A: Table A6**). Key covariates included smoking (n = 29), age (n = 26), body mass index (BMI) (n = 22), total energy intake (n = 15), physical activity (n = 13), and alcohol consumption (n = 12).

Ten prospective studies were included in the meta-analysis for the consumption of alcohol (nine reports) [16,33,34,37,41,48,51,54,57], five studies for wine (four reports) [12,41,54,57], five for beer (four reports) [12,41,54,57], five for liquor (four reports) [12,41,54,57], six for total coffee [16,46,47,52,54,55], four for caffeinated coffee [46,47,54,55], five for decaffeinated coffee [100, 101, 104, 113, 118], five for tea [100, 101, 104, 113, 118], three for caffeine [100, 101, 118], three for sugar-sweetened soda (two reports) [54,59], seven for fruits (six reports) [16,33-35,38,49], seven for vegetables (six reports) [16,33-35,38,49], four for total meat [36,42,50,58], three for red meat [33,42,58], two for processed meat [16,42], three for poultry [16,42,58], five for total fish (four reports) [36,43,49,56], three for oily fish [16,36,49], two for legumes [35,36], three for total dairy [36,42,45], two for milk [42,45], two for cheese [16,42], four for cereal products (three reports) [16,34,36], four for *n*-3 PUFAs (three reports) [33,34,43], four for vitamin C (three reports) [35,40,49], four for vitamin D (three reports) [39,45,49], three for vitamin D supplements (two reports) [39,45], three for vitamin A (two reports) [40,49], four for vitamin E (three reports) [35,40,49], and three for carotenoids (two reports) [35,40] (**Appendix A: Table A7(a,b)**).

2.4.3 Alcohol consumption and RA risk

An inverse association was observed between total alcohol consumption and RA risk across eight studies ((RR (95%-CI), heterogeneity (I^2): 0.76 (0.68, 0.85), 56%, P-heterogeneity = 0.03; **Figure 2.3**) when comparing the highest and lowest categories. This inverse relationship was seen with a total alcohol intake of up to 60 g/week (P-non-linearity = 0.004, **Figure 2.4**). Similarly, a non-linear dose-response association for beer was detected (P-non-linearity = 0.013, **Appendix A: Figure A1**). Linear analysis showed a 4% risk reduction for each additional 2 units of total alcohol intake (0.96 (0.94,0.98), 58%, P-heterogeneity = 0.01), and

a 10% risk reduction for each additional 2 units of beer consumption (0.90 (0.84,0.97), 0%, P-heterogeneity = 0.98; **Appendix A: Figure A3**). While there was some evidence of an association for each additional 2 units/week of wine consumption (0.98 (0.97,1.00); **Appendix A: Figure A3**), this was borderline non-significant. No evidence of association was observed for liquor intake. The NutriGrade scores showed moderate-quality evidence for total alcohol, low-quality evidence for wine and liquor, and very-low-quality evidence for beer (**Appendix A: Table A8**).

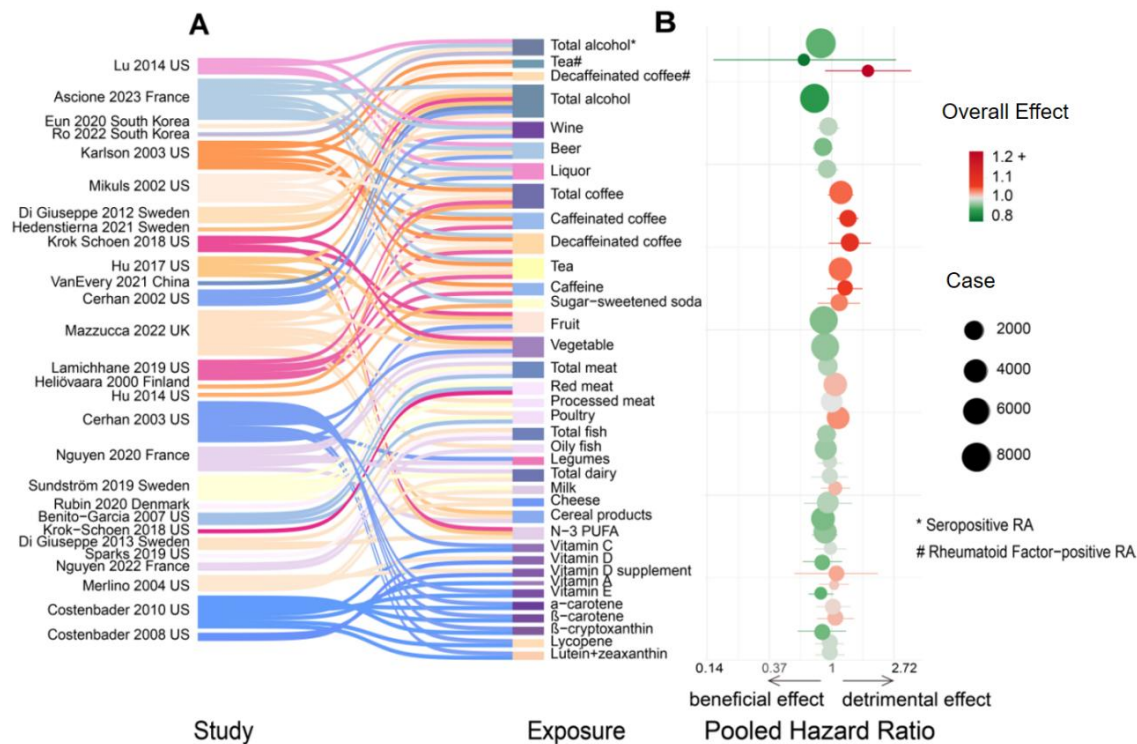


Figure 2.2 The summary of pooled HRs (95% CIs) for different dietary exposures and incident RA comparing the highest with the lowest categories. (A) Sankey diagram illustrating the 21 food groups, beverages, and 11 nutrients across 28 included cohort studies. (B) OncoPrint chart displaying the pooled hazard ratios (HRs) for incident RA, incident seropositive RA, and incident RF-positive RA, according to 37 food components and nutrients. Green colour represents a beneficial effect on RA incidence; red colour represents a detrimental effect on RA incidence.

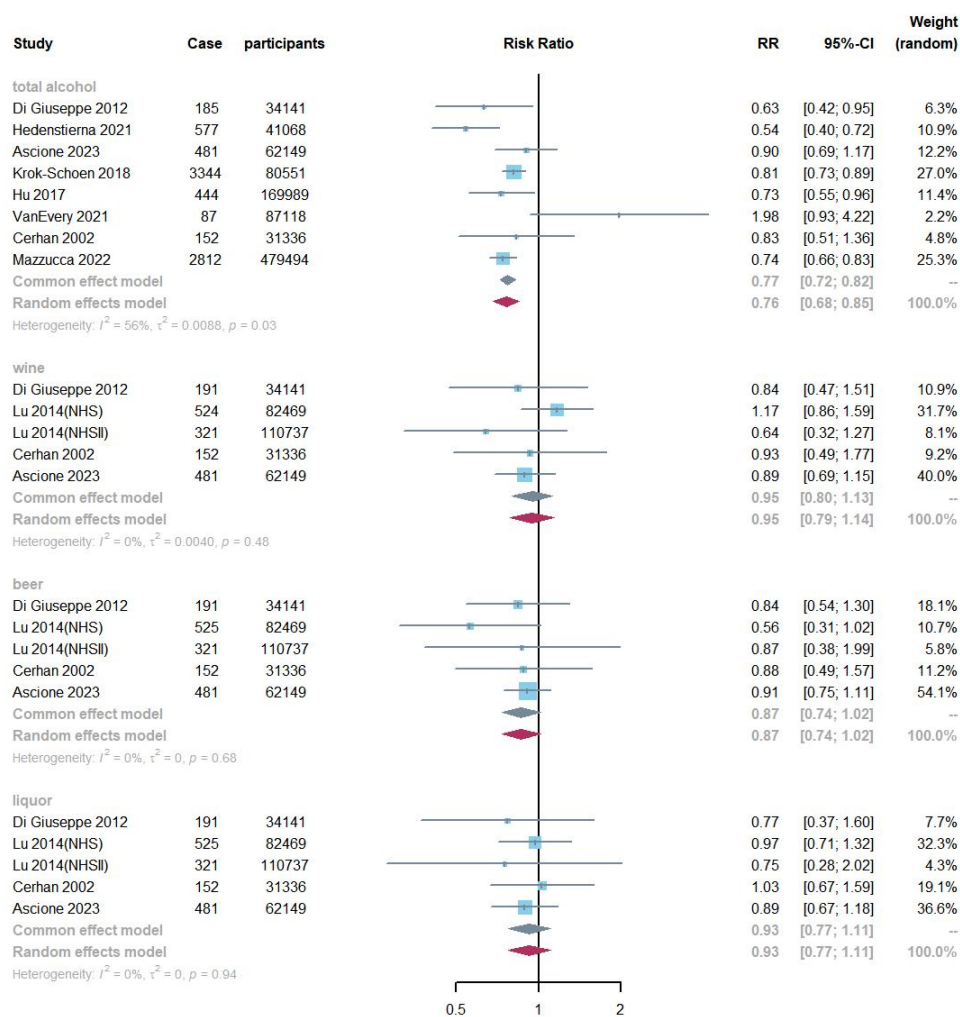


Figure 2.3 Meta-analysis of alcohol consumption and risk of RA comparing the highest with the lowest categories. Diamonds represent pooled estimates from random effects analysis.

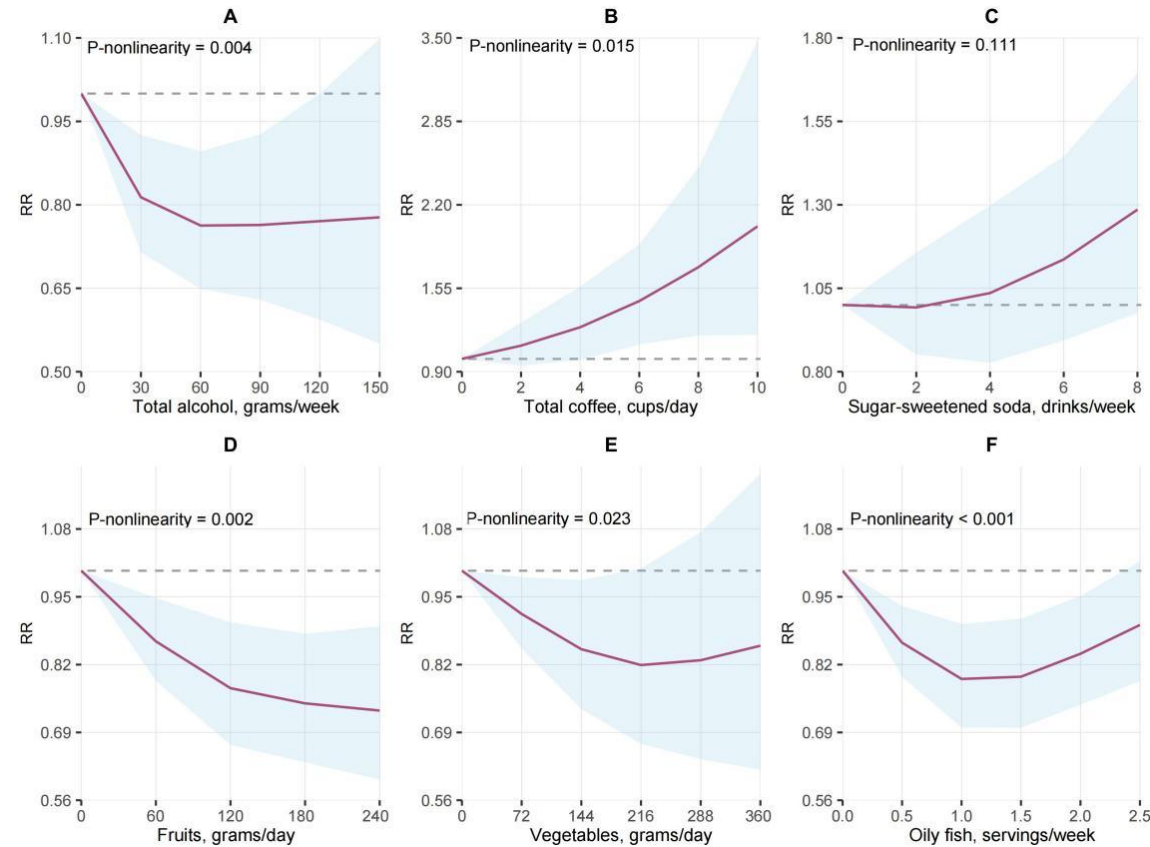


Figure 2.4 Non-linear dose–response association of food component consumption with risk of RA using restricted cubic splines. Relations for food components of (A) total alcohol, (B) total coffee, (C) sugar-sweetened soda, (D) fruits, (E) vegetables, and (F) oily fish. Solid lines represent the fitted non-linear trend and blue shade represents the pointwise 95% CIs.

A pooled analysis of four studies on seropositive RA found an RR of 0.84 ((0.73, 0.96); **Appendix A: Figure A4**) [12,53,54,60] for total alcohol when comparing extreme categories. Non-linear dose–response analysis supported this inverse association (P-non-linearity < 0.001; **Appendix A: Figure A1**). Each 2 unit/week increase in total alcohol intake was associated with a 3% reduction in the risk of seropositive RA (0.97(0.96, 0.99); **Appendix A: Figure A5**).

2.4.4 Non-alcoholic beverages and RA risk

Tea and caffeinated coffee consumption were associated with increased RA risk (tea: 1.15 (1.00, 1.33), 53%, P-heterogeneity = 0.07; caffeinated coffee: 1.30 (1.09,1.54), 0%, P-heterogeneity = 0.67) when comparing extreme categories, while total coffee, decaffeinated coffee, caffeine, and soda showed no evidence of association with RA risk (**Appendix A: Figure A6**). We found a positive trend for tea with a pooled RR of 1.04 per cup/d (95% CI: 1.02, 1.05, I^2 = 0%, P-heterogeneity = 0.53) in the linear dose–response analysis, but no conclusive dose–response association was observed for other beverages (**Appendix A: Figure A7**). A positive association between total coffee consumption and RA risk was noted in the non-linear trend analysis (P-non-linearity = 0.015, **Appendix A: Figure A1**). Additionally, there was no evidence of associations for decaffeinated coffee, tea, and RF-positive RA (**Appendix A: Figures A8 and A9**). The NutriGrade evaluations showed moderate-quality evidence for tea, lower-quality evidence for caffeinated coffee and sugar-sweetened soda, and very-low-quality evidence for total coffee, decaffeinated coffee, and caffeine (**Appendix A: Table A8**).

2.4.5 Fruit, vegetables, and RA risk

Higher fruit intake was associated with reduced RA risk (0.88 (0.79, 0.97), 50%, P-heterogeneity = 0.09; **Appendix A: Figure A10**), with evidence of a non-linear dose–response trend (P-non-linearity = 0.002, **Figure 2.4**). A linear analysis of six studies also indicated that an increase in fruit intake of 80 g/d increments was inversely associated with a reduced RA risk of 5% (95% CI: 0.92, 0.99, I^2 = 57%, P-heterogeneity = 0.04; **Appendix A: Figure A11**). No evidence of association was found for vegetables (0.90 (0.76,1.06), 81%, P-heterogeneity < 0.01; **Appendix A: Figure A10**), though there was evidence of a non-linear dose–

response association (P-non-linearity = 0.023, **Figure 2.4**). The NutriGrade evaluations indicated moderate-quality evidence for fruits, and lower-quality evidence for vegetables (**Appendix A: Table A8**).

2.4.6 Meat, fish, and RA risk

No evidence of RA risk associations emerged for total meat, red meat, processed meat, poultry, or total fish consumption (**Appendix A: Figures A12 and A13**). A non-linear inverse association between RA risk and oily fish intake was observed (P-non-linearity < 0.001; **Figure 2.4**). The NutriGrade assessments rated the evidence for total meat, processed meat, poultry, total fish, and oily fish as low-quality, and the evidence for red meat as very-low-quality (**Appendix A: Table A8**).

2.4.7 Other food components and RA risk

Higher cereal intake was associated with a reduced risk of RA (0.87 (0.80, 0.96), 0%, P-heterogeneity = 0.86; **Appendix A: Figure A14**) when comparing extreme categories. The dose–response analysis indicated a 3% risk reduction for each additional 30 g/d increase (0.97(0.96, 0.99), 20%, P-heterogeneity = 0.29; **Appendix A: Figure A15**). However, there was no evidence of associations between legumes, dairy, or cheese and RA risk (**Appendix A: Figures A14 and A15**). The NutriGrade scores revealed moderate-quality evidence for cereal products, low-quality evidence for legumes, dairy, and milk, and very-low-quality evidence for cheese (**Appendix A: Table A8**).

2.4.8 Nutrients and RA risk

Dietary Macronutrients

No evidence of an association between *n*-3 PUFA intake and RA risk was observed when comparing the highest with the lowest categories (**Appendix A: Figure A16**), although a borderline inverse non-linear association was noted (P-non-linearity = 0.063; **Appendix A: Figure A2**). A slight reduction in RA risk was seen for each additional 100 mg/d of *n*-3 PUFAs, but the evidence remains insufficient (0.96 (0.90,1.03), 74%, P-heterogeneity = 0.13; **Appendix A: Figure A17**). A meta-analysis for other macronutrients (e.g., energy, protein,

carbohydrates, total fat) was not possible due to limited data, despite alcohol having previously been analysed. The NutriGrade scores revealed very-low-quality evidence for n-3 PUFAs (**Appendix A: Table A8**).

Dietary Micronutrients

Vitamin D supplement intake showed an inverse non-linear association with RA risk ($P_{\text{non-linearity}} = 0.003$, **Appendix A: Figure A2**). The risk of RA decreased by 20% with an increase in vitamin D supplement intake of ≤ 320 IU/d. No other evidence of association was found between dietary vitamins or carotenoids and RA risk (**Appendix A: Figures A2 and A18–A21**).

The NutriGrade evaluations rated the evidence quality as low for vitamins, β -cryptoxanthin, and lycopene, and very low for α -carotenoids, β -carotenoids, and lutein/zeaxanthin (**Appendix A: Table A8**).

2.4.9 Subgroup analysis and bias analysis

The stratified analyses showed evidence of heterogeneity between subgroups. An inverse association between total alcohol consumption and risk of RA was confirmed in most of the subgroup analyses, except for studies conducted in Asia (**Appendix A: Table A10(a)**). Significant associations with RA risk were observed for the consumption of decaffeinated coffee, fruit, and vegetables in studies involving women, and in those conducted in America, but no association was found in studies including both sexes or those conducted in other regions (**Appendix A: Table A10 (b–g)**). The quality assessment, detailed in **Appendix A: Table A9**, showed that 80% (24 out of 30) of the included studies were of good quality. Sensitivity analyses confirmed that excluding any single study did not change the pooled effect sizes (**Appendix A: Table A11 (a,b)**). Although the funnel plots and Egger's test indicated a publication bias for coffee intake ($P < 0.01$), no bias was detected for other foods (**Appendix A: Figures A22 and A23**). Trim-and-fill adjustments showed consistent results, suggesting a minimal impact from publication bias (**Appendix A: Table A12 (a,b)**).

2.5 Discussion

2.5.1 Main findings

This is the first comprehensive meta-analysis of the associations between RA risk in adults and the consumption of a range of foods, beverages, and nutrients, conducted to concurrently examine the potential protective and harmful effects of diet on RA risk in a dose-specific association.

2.5.2 Comparison with previous studies

Our results identified associations between a higher intake of alcohol, fruits, and cereals and a lower risk of RA, while higher consumption of tea and caffeinated coffee was linked to an increased risk. Dose–response analyses further clarified these associations, revealing J-shaped trends for total alcohol and beer, protective linear relationships for fruit and cereals, and non-linear patterns for vegetables, oily fish, and vitamin D supplements. Notably, tea intake showed a positive linear association with RA risk, while total coffee consumption followed a non-linear trend. Evidence regarding dietary patterns and micronutrients was limited. Although based on low-to-moderate-quality meta-evidence, the findings provide new insights into the potential role of diet in RA risk, supporting their integration into prevention strategies.

To our knowledge, this is the first meta-analysis to focus on the associations of fruit, vegetables, and cereal intake with RA risk. Although the World Health Organisation recommends consuming over 400 g (about five servings) of fruits and vegetables daily to prevent non-communicable diseases [61], evidence regarding the direct effect of fruit and vegetable intake on RA is rare. Our analysis, which included 7294 RA cases from Europe and America, found moderate-certainty evidence of an association between fruit intake and RA risk. This aligns with previous meta-analyses that linked fruit intake to reduced risks of inflammatory bowel disease, multiple sclerosis, and type 2 diabetes [62-64]. In addition, two RCTs have demonstrated the anti-inflammatory effects of fruit- and vegetable-rich diets compared to placebos [65,66]. A diet rich in fruits and vegetables is biologically plausible for enhancing immune function [67], reducing inflammation [68], and preventing autoimmune diseases [69]. An alternative

explanation could be that individuals consuming more fruit and vegetables are more likely to adhere to a healthy diet and lifestyle, like engaging in regular physical activity, which has been shown to have a protective effect on RA risk [70].

The inverse non-linear association between vegetable intake and RA risk may be influenced by preparation methods and nutrient composition. Despite vegetables generally appearing to be beneficial, some studies indicate that cooked vegetables may have negative health effects due to harmful components such as acrylamide, advanced glycation end-products, and a loss of water-soluble vitamins. A recent study involving 400,000 adults found that raw vegetable intake was inversely associated with both cardiovascular disease incidence and mortality, unlike cooked vegetables [71]. Moreover, vegetable oils high in omega-6 PUFAs may promote low-grade inflammation and oxidative stress, unlike fresh vegetables containing omega-3 PUFAs with anti-inflammatory properties [72]. Another explanation could be that BMI mediates the relationship between vegetable intake and RA risk. Since the inverse association between vegetable intake and RA risk was over-adjusted for BMI as a confounder, the potential effect of vegetables on RA risk was diminished. This suggests that obesity [73] may play a role in RA development.

Regarding cereals, our study showed a 3% reduction in RA risk with each additional 30 g/d of intake. This inverse association is consistent with recent large-scale cohort studies, including one from the UK Biobank showing a protective effect of breakfast cereals [16], as well as studies from America and France which linked whole grains and MD components with a reduced risk [34,55]. Despite differences in cereal types between the included studies, the low heterogeneity and robust results were confirmed by sensitivity analyses. The protective mechanism by which cereals may reduce RA risk could be similar to their effects on cardiovascular disease and type 2 diabetes [74,75], possibly attributable to the anti-inflammatory properties of cereal phytochemicals, fibre, vitamins, and minerals [76].

Associations between alcohol consumption and a lower risk of RA have been reported in previous meta-analyses [13,77]. In our study, the risk estimate for total alcohol was 0.76, slightly lower than earlier estimates of 0.86. This difference

could be attributed to our stricter inclusion criteria, as we only considered prospective cohort studies, which are less prone to bias compared to prior meta-analyses that included both retrospective and prospective data [13,78]. This study includes more recent data, featuring over four times the number of RA cases and an average follow-up duration of 22.5 years, providing more robust evidence. Furthermore, this inverse association was consistently seen in the analysis of seropositive RA. Alcohol may exert a protective effect, particularly in individuals with auto-antibodies. A Swedish population-based study has shown that low-to-moderate alcohol consumption may reduce the risk of developing ACPA-positive RA by mitigating the combined effects of smoking and the human leucocyte antigen (HLA)-DRB1 shared epitope [78]. Among alcoholic beverages, beer intake was negatively associated with RA risk, potentially contributing to the overall protective effect of total alcohol. A possible mechanism involves polyphenols like resveratrol, which may modulate the immune response by inhibiting pro-inflammatory cytokines [79], altering immune cell activity [80], and reducing oxidative stress through free-radical neutralisation [81]. Further intervention studies are warranted to clarify this beverage-specific relationship.

The dose–response analysis revealed a J-shaped relationship, suggesting that low-to-moderate alcohol consumption is relatively safe, with benefits diminishing beyond 60 g/week. These findings could be used to refine the alcohol consumption guidelines, especially for individuals with rheumatic and musculoskeletal diseases, by identifying the optimal level of consumption for maximum benefit. We observed this inverse association between alcohol consumption and RA risk in studies from both America and Europe, but not in the limited data from Asia. This discrepancy may result from insufficient statistical power in the Asian studies or differences in the types of alcohol commonly consumed in Asia, which may influence the association with RA risk. The subgroup analysis indicated that the protective effect was consistent across both female and mixed-sex groups, suggesting that the effect may be applicable broadly. Future studies with sex-specific analyses and more extensive geographic coverage could strengthen the evidence.

The relationship between tea consumption and RA risk appears to be complex and controversial. Two meta-analyses found no association [82,83], while some

studies suggested an inverse relationship [15,54]. Our analysis identified, for the first time, a positive linear association between higher tea consumption and increased RA risk, particularly in individuals with over 10 years of follow-up. Unlike previous meta-analyses [78,79], we found no significant association of total coffee with RA risk in extreme intake comparisons. While green tea's antioxidant content may suggest differential effects, our study did not distinguish tea types; this merits further investigation. However, we did find a non-linear positive relationship for total coffee and an apparent significant association for caffeinated coffee, suggesting that higher intake may elevate RA risk. This may be due to the pro-inflammatory effect of caffeine when consumed in excess, despite its known anti-inflammatory properties in small amounts.

Caffeine, a common component of both tea and coffee, could play a role in influencing RA risk. Supporting our findings, evidence from animal studies has supported the possibility that caffeine combined with sucrose can exacerbate inflammation [84]. Moreover, tea and coffee consumption is often linked to Western diets high in refined sugars and unhealthy fats, which are known to promote inflammation [81]. The inconsistencies across studies may be due to variations in tea and coffee types, intake levels, and study duration. More well-controlled trials are needed to clarify these associations.

In contrast, no evidence was found of an association for decaffeinated coffee or sugar-sweetened soda. The lack of findings for decaffeinated coffee may be attributed to its lower caffeine content, which likely contributes to the association observed with caffeinated coffee. Interestingly, a positive association between decaffeinated coffee and increased risk was found in both linear dose-response and extreme intake comparison meta-analyses under the common-effects model but not the random-effects model. This discrepancy may be due to differences in study demographics or methodologies, which are more apparent in random-effects models. For soda, despite the well-documented link between excessive dietary sugar intake and inflammation [85], the absence of significant findings may reflect heterogeneity across studies or differences in soda formulations. Further research is needed to investigate the specific components of these beverages and their interactions with other dietary factors.

Our study found that oily fish and vitamin D supplements were associated with a reduced RA risk in the non-linear dose–response analysis, but no evidence of association was observed in either the comparison of extreme categories or the linear dose–response meta-analysis. Because of this inconsistency, there is a need for further research to provide more conclusive evidence.

2.5.3 Strengths and limitations

This study has several limitations. First, although we focused on prospective cohort studies, observational studies are still prone to selection bias, reverse causality, publication bias, and small-study effects. Second, dietary assessment errors are unavoidable. Most studies relied on a single dietary measurement, which may not accurately reflect long-term intake. The use of different dietary assessment methods (e.g., FFQs, dietary records) and inconsistent intake units could also introduce heterogeneity. Under-reporting of alcohol consumption and changes in dietary habits over time may further weaken the observed associations with RA risk.

Third, unmeasured or residual confounding cannot be avoided. Low-to-moderate alcohol drinkers may have other beneficial lifestyle factors or a higher socioeconomic status compared to non-drinkers or heavy drinkers. For example, a cross-sectional study of the American population indicated that alcohol consumers tend to exercise more vigorously [86]. Additionally, dietary nutrients are consumed as part of an overall diet, making the total effect potentially greater than the sum of the individual components, and interactions between nutrients should be considered. Finally, our findings may not be fully generalisable to low- or middle-income countries due to differences in diets, food availability, or to males, as most participants were female and from America and Europe.

Our results indicate significant relative risk reductions for RA associated with higher consumption of total alcohol and certain food groups. However, a limitation is the absence of data on adjusted absolute risk differences, which restricts our ability to estimate the actual number of RA cases that could be prevented by dietary changes. Most of the included studies reported relative risks without providing baseline incidence rates or adjusted absolute risks. Future research should prioritise the reporting of both relative and absolute measures of risk,

enabling a more comprehensive assessment of the public health impact of dietary factors on RA risk. Furthermore, while dietary factors may influence the severity of RA, this study did not address that aspect, leaving the role of diet in disease progression as an important area for future investigation.

This meta-analysis has some strengths. It included a large number of participants and RA cases, allowing for well-powered analyses of 32 dietary factors in relation to RA risk. This meta-analysis included only prospective cohort studies, which are less prone to recall bias compared to case-control studies. Dose-response analyses enabled exploration of the direction of the relationships and possible threshold effects, as well as an assessment of the credibility and reliability of the evidence.

2.6 Conclusions

This meta-analysis identified that total alcohol, fruit, cereals, oily fish, and vitamin D supplements were associated with a reduced risk of RA, supported by evidence of low-to-moderate credibility. Tea and coffee were associated with an increased risk. There is no evidence of an association between sugar-sweetened soda and RA risk. Further research is needed to confirm these associations and explore whether specific dietary patterns or nutrients could emerge as a viable strategy for RA prevention.

2.7 References

1. van Delft, M.A.M.; Huizinga, T.W.J. An overview of autoantibodies in rheumatoid arthritis. *J. Autoimmun.* 2020, *110*, 102392.
2. Finckh, A.; Gilbert, B.; Hodgkinson, B.; Bae, S.C.; Thomas, R.; Deane, K.D.; Alpizar-Rodriguez, D.; Lauper, K. Global epidemiology of rheumatoid arthritis. *Nat. Rev. Rheumatol.* 2022, *18*, 591–602.
3. Ren, J.; Ren, Y.; Mu, Y.; Zhang, L.; Chen, B.; Li, S.; Fang, Q.; Zhang, Z.; Zhang, K.; Li, S.; et al. Microbial imbalance in Chinese children with diarrhea or constipation. *Sci. Rep.* 2024, *14*, 13516.
4. Romão, V.C.; Fonseca, J.E. Etiology and Risk Factors for Rheumatoid Arthritis: A State-of-the-Art Review. *Front. Med.* 2021, *8*, 689698.

5. Manzel, A.; Muller, D.N.; Hafler, D.A.; Erdman, S.E.; Linker, R.A.; Kleinewietfeld, M. Role of "Western diet" in inflammatory autoimmune diseases. *Curr. Allergy Asthma Rep.* 2014, *14*, 404.
6. Pedersen, M.; Jacobsen, S.; Klarlund, M.; Pedersen, B.V.; Wiik, A.; Wohlfahrt, J.; Frisch, M. Environmental risk factors differ between rheumatoid arthritis with and without auto-antibodies against cyclic citrullinated peptides. *Arthritis Res. Ther.* 2006, *8*, R133.
7. Menzel, J.; Biemann, R.; Longree, A.; Isermann, B.; Mai, K.; Schulze, M.B.; Abraham, K.; Weikert, C. Associations of a vegan diet with inflammatory biomarkers. *Sci. Rep.* 2020, *10*, 1933.
8. Al-Saoodi, H.; Kolahdooz, F.; Andersen, J.R.; Jalili, M. Effect of vitamin D on inflammatory and clinical outcomes in patients with rheumatoid arthritis: A systematic review and dose-response meta-analysis of randomized controlled trials. *Nutr. Rev.* 2024, *82*, 600–611.
9. Mölzer, C.; Wilson, H.M.; Kuffova, L.; Forrester, J.V. A Role for Folate in Microbiome - Linked Control of Autoimmunity. *J. Immunol. Res.* 2021, *2021*, 9998200.
10. Schwingshackl, L.; Missbach, B.; König, J.; Hoffmann, G. Adherence to a Mediterranean diet and risk of diabetes: A systematic review and meta-analysis. *Public Health Nutr.* 2015, *18*, 1292–1299.
11. Shen, J.; Wilmot, K.A.; Ghasemzadeh, N.; Molloy, D.L.; Burkman, G.; Mekonnen, G.; Gongora, M.C.; Quyyumi, A.A.; Sperling, L.S. Mediterranean dietary patterns and cardiovascular health. *Annu. Rev. Nutr.* 2015, *35*, 425–449.
12. Lu, B.; Solomon, D.H.; Costenbader, K.H.; Karlson, E.W. Alcohol consumption and risk of incident rheumatoid arthritis in women: A prospective study. *Arthritis Rheumatol.* 2014, *66*, 1998–2005.
13. Jin, Z.; Xiang, C.; Cai, Q.; Wei, X.; He, J. Alcohol consumption as a preventive factor for developing rheumatoid arthritis: A dose-response meta-analysis of prospective studies. *Ann. Rheum. Dis.* 2014, *73*, 1962–1967.
14. Jin, J.; Li, J.; Gan, Y.; Liu, J.; Zhao, X.; Chen, J.; Zhang, R.; Zhong, Y.; Chen, X.; Wu, L. et al. Tea consumption is associated with decreased disease activity of rheumatoid arthritis in a real-world, large-scale study. *Ann. Nutr. Metab.* 2020, *76*, 54–61.

15. Westerlind, H.; Palmqvist, I.; Saevarsdottir, S.; Alfredsson, L.; Klareskog, L.; Di Giuseppe, D. Is tea consumption associated with reduction of risk of rheumatoid arthritis? A Swedish case-control study. *Arthritis Res. Ther.* 2021, *23*, 209.
16. Mazzucca, C.B.; Scotti, L.; Cappellano, G.; Barone-Adesi, F.; Chiocchetti, A. Nutrition and Rheumatoid Arthritis Onset: A Prospective Analysis Using the UK Biobank. *Nutrients* 2022, *14*, 1554.
17. Daien, C.; Czernichow, S.; Letarouilly, J.G.; Nguyen, Y.; Sanchez, P.; Sigaux, J.; Beauvais, C.; Desouches, S.; Le Puillandre, R.; Rigalleau, V.; et al. Dietary recommendations of the French Society for Rheumatology for patients with chronic inflammatory rheumatic diseases. *Jt. Bone Spine* 2022, *89*, 105319.
18. Gkiouras, K.; Grammatikopoulou, M.G.; Myrogiannis, I.; Papamitsou, T.; Rigopoulou, E.I.; Sakkas, L.I.; Bogdanos, D.P. Efficacy of n-3 fatty acid supplementation on rheumatoid arthritis' disease activity indicators: A systematic review and meta-analysis of randomized placebo-controlled trials. *Crit. Rev. Food Sci. Nutr.* 2024, *64*, 16–30.
19. Martu, M.A.; Maftai, G.A.; Luchian, I.; Stefanescu, O.M.; Scutariu, M.M.; Solomon, S.M. The Effect of Acknowledged and Novel Anti-Rheumatic Therapies on Periodontal Tissues-A Narrative Review. *Pharmaceuticals* 2021, *14*, 1209.
20. Bäcklund, R.; Drake, I.; Bergström, U.; Compagno, M.; Sonestedt, E.; Turesson, C. Diet and the risk of rheumatoid arthritis—A systematic literature review. *Semin. Arthritis Rheum.* 2023, *58*, 152118.
21. Guan, C.M.; Beg, S. Diet as a risk factor for rheumatoid arthritis. *Cureus* 2023, *15*, e39273.
22. Stroup, D.F.; Berlin, J.A.; Morton, S.C.; Olkin, I.; Williamson, G.D.; Rennie, D.; Moher, D.; Becker, B.J.; Sipe, T.A.; Thacker, S.B. Meta-analysis of observational studies in epidemiology: A proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000, *283*, 2008–2012.
23. Schardt, C.; Adams, M.B.; Owens, T.; Keitz, S.; Fontelo, P. Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Med. Inform. Decis. Mak.* 2007, *7*, 16.

24. Clark, J.; Glasziou, P.; Del Mar, C.; Bannach-Brown, A.; Stehlik, P.; Scott, A.M. A full systematic review was completed in 2 weeks using automation tools: A case study. *J. Clin. Epidemiol.* 2020, *121*, 81–90.
25. Arnett, F.C.; Edworthy, S.M.; Bloch, D.A.; McShane, D.J.; Fries, J.F.; Cooper, N.S.; Healey, L.A.; Kaplan, S.R.; Liang, M.H.; Luthra, H.S.; et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum.* 1988, *31*, 315–324.
26. Stang, A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur. J. Epidemiol.* 2010, *25*, 603–605.
27. Schwingshackl, L.; Knüppel, S.; Schwedhelm, C.; Hoffmann, G.; Missbach, B.; Stelmach-Mardas, M.; Dietrich, S.; Eichelmann, F.; Kontopantelis, E.; Iqbal, K.; et al. Perspective: NutriGrade: A Scoring System to Assess and Judge the Meta-Evidence of Randomized Controlled Trials and Cohort Studies in Nutrition Research. *Adv. Nutr.* 2016, *7*, 994–1004.
28. Greenland, S.; Longnecker, M.P. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *Am. J. Epidemiol.* 1992, *135*, 1301–1309.
29. Orsini, N.; Li, R.; Wolk, A.; Khudyakov, P.; Spiegelman, D. Meta-analysis for linear and nonlinear dose-response relations: Examples, an evaluation of approximations, and software. *Am. J. Epidemiol.* 2012, *175*, 66–73.
30. Orsini, N. From floated to conventional confidence intervals for the relative risks based on published dose-response data. *Comput. Methods Programs Biomed.* 2010, *98*, 90–93.
31. Egger, M.; Davey Smith, G.; Schneider, M.; Minder, C. Bias in meta-analysis detected by a simple, graphical test. *BMJ (Clin. Res. Ed.)* 1997, *315*, 629–634.
32. Begg, C.B.; Mazumdar, M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994, *50*, 1088–1101.
33. Krok-Schoen, J.L.; Brasky, T.M.; Hunt, R.P.; Rohan, T.E.; Baker, T.A.; Li, W.; Carbone, L.; Mackey, R.H.; Snetselaar, L.; Lustberg, M.B.; et al. Dietary Long-Chain n-3 Fatty Acid Intake and Arthritis Risk in the Women's Health Initiative. *J. Acad. Nutr. Diet.* 2018, *118*, 2057–2069.

34. Hu, Y.; Sparks, J.A.; Malspeis, S.; Costenbader, K.H.; Hu, F.B.; Karlson, E.W.; Lu, B. Long-term dietary quality and risk of developing rheumatoid arthritis in women. *Ann. Rheum. Dis.* 2017, *76*, 1357–1364.
35. Cerhan, J.R.; Saag, K.G.; Merlino, L.A.; Mikuls, T.R.; Criswell, L.A. Antioxidant micronutrients and risk of rheumatoid arthritis in a cohort of older women. *Am. J. Epidemiol.* 2003, *157*, 345–354.
36. Nguyen, Y.; Salliot, C.; Mariette, X.; Boutron-Ruault, M.-C.; Seror, R. Fish consumption and risk of rheumatoid arthritis: Findings from the E3N cohort study. *Nutrients* 2022, *14*, 861.
37. Hedenstierna, L.; Bellocco, R.; Ye, W.; Adami, H.O.; Åkerstedt, T.; Trolle Lagerros, Y.; Hedström, A.K. Effects of alcohol consumption and smoking on risk for RA: Results from a Swedish prospective cohort study. *RMD Open* 2021, *7*, e001379.
38. Nguyen, Y.; Salliot, C.; Gelot, A.; Gambaretti, J.; Mariette, X.; Boutron - Ruault, M.C.; Seror, R. Mediterranean diet and risk of rheumatoid arthritis: Findings from the French E3N - EPIC cohort study. *Arthritis Rheumatol.* 2021, *73*, 69–77.
39. Costenbader, K.H.; Feskanich, D.; Holmes, M.; Karlson, E.W.; Benito-Garcia, E. Vitamin D intake and risks of systemic lupus erythematosus and rheumatoid arthritis in women. *Ann. Rheum. Dis.* 2008, *67*, 530–535.
40. Costenbader, K.H.; Kang, J.H.; Karlson, E.W. Antioxidant intake and risks of rheumatoid arthritis and systemic lupus erythematosus in women. *Am. J. Epidemiol.* 2010, *172*, 205–216.
41. Di Giuseppe, D.; Alfredsson, L.; Bottai, M.; Askling, J.; Wolk, A. Long term alcohol intake and risk of rheumatoid arthritis in women: A population based cohort study. *BMJ (Clin. Res. Ed.)* 2012, *345*, e4230.
42. Sundström, B.; Ljung, L.; Di Giuseppe, D. Consumption of Meat and Dairy Products Is Not Associated with the Risk for Rheumatoid Arthritis among Women: A Population-Based Cohort Study. *Nutrients* 2019, *11*, 2825.
43. Di Giuseppe, D.; Wallin, A.; Bottai, M.; Askling, J.; Wolk, A. Long-term intake of dietary long-chain n-3 polyunsaturated fatty acids and risk of rheumatoid arthritis: A prospective cohort study of women. *Ann. Rheum. Dis.* 2014, *73*, 1949–1953.

44. Hiraki, L.T.; Munger, K.L.; Costenbader, K.H.; Karlson, E.W. Dietary intake of vitamin D during adolescence and risk of adult-onset systemic lupus erythematosus and rheumatoid arthritis. *Arthritis Care Res.* 2012, *64*, 1829–1836.
45. Merlino, L.A.; Curtis, J.; Mikuls, T.R.; Cerhan, J.R.; Criswell, L.A.; Saag, K.G. Vitamin D intake is inversely associated with rheumatoid arthritis: Results from the Iowa Women's Health Study. *Arthritis Rheum.* 2004, *50*, 72–77.
46. Karlson, E.W.; Mandl, L.A.; Aweh, G.N.; Grodstein, F. Coffee consumption and risk of rheumatoid arthritis. *Arthritis Rheum.* 2003, *48*, 3055–3060.
47. Lamichhane, D.; Collins, C.; Constantinescu, F.; Walitt, B.; Pettinger, M.; Parks, C.; Howard, B.V. Coffee and Tea Consumption in Relation to Risk of Rheumatoid Arthritis in the Women's Health Initiative Observational Cohort. *J. Clin. Rheumatol. Pract. Rep. Rheum. Musculoskelet. Dis.* 2019, *25*, 127–132.
48. Lahiri, M.; Luben, R.N.; Morgan, C.; Bunn, D.K.; Marshall, T.; Lunt, M.; Verstappen, S.M.; Symmons, D.P.; Khaw, K.T.; Wareham, N.; et al. Using lifestyle factors to identify individuals at higher risk of inflammatory polyarthritis (results from the European Prospective Investigation of Cancer-Norfolk and the Norfolk Arthritis Register--the EPIC-2-NOAR Study). *Ann. Rheum. Dis.* 2014, *73*, 219–226.
49. Pedersen, M.; Stripp, C.; Klarlund, M.; Olsen, S.F.; Tjønneland, A.M.; Frisch, M. Diet and risk of rheumatoid arthritis in a prospective cohort. *J. Rheumatol.* 2005, *32*, 1249–1252.
50. Rubin, K.H.; Rasmussen, N.F.; Petersen, I.; Kopp, T.I.; Stenager, E.; Magyari, M.; Hetland, M.L.; Bygum, A.; Glinborg, B.; Andersen, V. Intake of dietary fibre, red and processed meat and risk of late-onset Chronic Inflammatory Diseases: A prospective Danish study on the diet, cancer and health cohort. *Int. J. Med. Sci.* 2020, *17*, 2487–2495.
51. VanEvery, H.; Yang, W.; Olsen, N.; Bao, L.; Lu, B.; Wu, S.; Cui, L.; Gao, X. Alcohol Consumption and Risk of Rheumatoid Arthritis among Chinese Adults: A Prospective Study. *Nutrients* 2021, *13*, 2231.
52. Heliövaara, M.; Aho, K.; Knekt, P.; Impivaara, O.; Reunanen, A.; Aromaa, A. Coffee consumption, rheumatoid factor, and the risk of rheumatoid arthritis. *Ann. Rheum. Dis.* 2000, *59*, 631–635.

53. Ro, J.; Kim, S.H.; Kim, H.R.; Lee, S.H.; Min, H.K. Impact of lifestyle and comorbidities on seropositive rheumatoid arthritis risk from Korean health insurance data. *Sci. Rep.* 2022, *12*, 2201.
54. Ascione, S.; Barde, F.; Artaud, F.; Nguyen, Y.; Macdonald, C.; Mariette, X.; Boutron-Ruault, M.C.; Salliot, C.; Seror, R. Association between beverage consumption and risk of rheumatoid arthritis: A prospective study from the French E3N Cohort. *Rheumatology* 2023, *62*, 1814–1823.
55. Mikuls, T.R.; Cerhan, J.R.; Criswell, L.A.; Merlino, L.; Mudano, A.S.; Burma, M.; Folsom, A.R.; Saag, K.G. Coffee, tea, and caffeine consumption and risk of rheumatoid arthritis: Results from the Iowa Women's Health Study. *Arthritis Rheum.* 2002, *46*, 83–91.
56. Sparks JA, O'Reilly É J, Barbhaiya M, et al. Association of fish intake and smoking with risk of rheumatoid arthritis and age of onset: a prospective cohort study. *BMC musculoskeletal disorders* 2019;20(1):2.
57. Cerhan, J.R.; Saag, K.G.; Criswell, L.A.; Merlino, L.A.; Mikuls, T.R. Blood transfusion, alcohol use, and anthropometric risk factors for rheumatoid arthritis in older women. *J. Rheumatol.* 2002, *29*, 246–254.
58. Benito-Garcia, E.; Feskanich, D.; Hu, F.B.; Mandl, L.A.; Karlson, E.W. Protein, iron, and meat consumption and risk for rheumatoid arthritis: A prospective cohort study. *Arthritis Res. Ther.* 2007, *9*, R16.
59. Hu, Y.; Costenbader, K.H.; Gao, X.; Al-Daabil, M.; Sparks, J.A.; Solomon, D.H.; Hu, F.B.; Karlson, E.W.; Lu, B. Sugar-sweetened soda consumption and risk of developing rheumatoid arthritis in women. *Am. J. Clin. Nutr.* 2014, *100*, 959–967.
60. Eun, Y.; Jeon, K.H.; Han, K.; Kim, D.; Kim, H.; Lee, J.; Lee, D.Y.; Yoo, J.E.; Shin, D.W. Menopausal factors and risk of seropositive rheumatoid arthritis in postmenopausal women: A nationwide cohort study of 1.36 million women. *Sci. Rep.* 2020, *10*, 20793.
61. WHO. Increasing Fruit and Vegetable Consumption to Reduce the Risk of Noncommunicable Diseases. 2023. [Available online: <https://www.who.int/tools/elena/interventions/fruit-vegetables-ncds>](accessed on 1 May 2024).

62. Li, F.; Liu, X.; Wang, W.; Zhang, D. Consumption of vegetables and fruit and the risk of inflammatory bowel disease: A meta-analysis. *Eur. J. Gastroenterol. Hepatol.* 2015, 27, 623–630.
63. Fotros, D.; Noormohammadi, M.; Razeghi Jahromi, S.; Abdolkarimi, M. Fruits and vegetables intake may be associated with a reduced odds of multiple sclerosis: A systematic review and dose–response meta-analysis of observational studies. *Nutr. Neurosci.* 2024, 27, 887–898.
64. Carter, P.; Gray, L.J.; Troughton, J.; Khunti, K.; Davies, M. Fruit and vegetable intake and incidence of type 2 diabetes mellitus: Systematic review and meta-analysis. *BMJ* 2010, 341, c4229.
65. Gualtieri, P.; Marchetti, M.; Frank, G.; Smeriglio, A.; Trombetta, D.; Colica, C.; Cianci, R.; De Lorenzo, A.; Di Renzo, L. Antioxidant-Enriched Diet on Oxidative Stress and Inflammation Gene Expression: A Randomized Controlled Trial. *Genes* 2023, 14, 206.
66. van der Merwe, M.; Moore, D.; Hill, J.L.; Keating, F.H.; Buddington, R.K.; Bloomer, R.J.; Wang, A.; Bowman, D.D. The Impact of a Dried Fruit and Vegetable Supplement and Fiber Rich Shake on Gut and Health Parameters in Female Healthcare Workers: A Placebo-Controlled, Double-Blind, Randomized Clinical Trial. *Microorganisms* 2021, 9, 843.
67. Gibson, A.; Edgar, J.D.; Neville, C.E.; Gilchrist, S.E.; McKinley, M.C.; Patterson, C.C.; Young, I.S.; Woodside, J.V. Effect of fruit and vegetable consumption on immune function in older people: A randomized controlled trial. *Am. J. Clin. Nutr.* 2012, 96, 1429–1436.
68. Hosseini, B.; Berthon, B.S.; Saedisomeolia, A.; Starkey, M.R.; Collison, A.; Wark, P.A.B.; Wood, L.G. Effects of fruit and vegetable consumption on inflammatory biomarkers and immune cell populations: A systematic literature review and meta-analysis. *Am. J. Clin. Nutr.* 2018, 108, 136–155.
69. Ecarnot, F.; Maggi, S. The impact of the Mediterranean diet on immune function in older adults. *Aging Clin. Exp. Res.* 2024, 36, 117.
70. Liu, X.; Tedeschi, S.K.; Lu, B.; Zaccardelli, A.; Speyer, C.B.; Costenbader, K.H.; Karlson, E.W.; Sparks, J.A. Long - term physical activity and subsequent risk for rheumatoid arthritis among women: a prospective cohort study. *Arthritis Rheumatol.* 2019, 71, 1460–1471.

71. Feng, Q.; Kim, J.H.; Omiyale, W.; Bešević, J.; Conroy, M.; May, M.; Yang, Z.; Wong, S.Y.; Tsoi, K.K.; Allen, N.; et al. Raw and Cooked Vegetable Consumption and Risk of Cardiovascular Disease: A Study of 400,000 Adults in UK Biobank. *Front. Nutr.* 2022, 9, 831470.
72. DiNicolantonio, J.J.; O'Keefe, J.H. Importance of maintaining a low omega-6/omega-3 ratio for reducing inflammation. *Open Heart* 2018, 5, e000946.
73. Arnotti, K.; Bamber, M. Fruit and vegetable consumption in overweight or obese individuals: A meta-analysis. *West. J. Nurs. Res.* 2020, 42, 306–314.
74. Aune, D.; Keum, N.; Giovannucci, E.; Fadnes, L.T.; Boffetta, P.; Greenwood, D.C.; Tonstad, S.; Vatten, L.J.; Riboli, E.; Norat, T. Whole grain consumption and risk of cardiovascular disease, cancer, and all cause and cause specific mortality: Systematic review and dose-response meta-analysis of prospective studies. *BMJ* 2016, 353, i2716.
75. Aune, D.; Norat, T.; Romundstad, P.; Vatten, L.J. Whole grain and refined grain consumption and the risk of type 2 diabetes: A systematic review and dose-response meta-analysis of cohort studies. *Eur. J. Epidemiol.* 2013, 28, 845–858.
76. Garutti, M.; Nevola, G.; Mazzeo, R.; Cucciniello, L.; Totaro, F.; Bertuzzi, C.A.; Caccialanza, R.; Pedrazzoli, P.; Puglisi, F. The impact of cereal grain composition on the health and disease outcomes. *Front. Nutr.* 2022, 9, 888974.
77. Scott, I.C.; Tan, R.; Stahl, D.; Steer, S.; Lewis, C.M.; Cope, A.P. The protective effect of alcohol on developing rheumatoid arthritis: A systematic review and meta-analysis. *Rheumatology* 2013, 52, 856–867.
78. Hedström, A.K.; Hössjer, O.; Klareskog, L.; Alfredsson, L. Interplay between alcohol, smoking and HLA genes in RA aetiology. *RMD Open* 2019, 5, e000893.
79. Boatca, R.M.; Scutariu, M.M.; Rudnic, I.; Stefanache, M.A.M.; Hurjui, L.; Rezus, E.; Martu, S. Evolution of inflammatory biochemical markers within periodontal therapy to patients with rheumatoid arthritis. *Rev. De Chim.* 2016, 67, 741–744.
80. Carvalho, J.F.; Lerner, A. Resveratrol in Rheumatological Diseases: A Systematic Review. *Eur. J. Rheumatol.* 2023, 10, 163–168.

81. Dey, M.; Cutolo, M.; Nikiphorou, E. Beverages in Rheumatoid Arthritis: What to Prefer or to Avoid. *Nutrients* 2020, *12*, 3155.
82. Lee, Y.H.; Bae, S.C.; Song, G.G. Coffee or tea consumption and the risk of rheumatoid arthritis: A meta-analysis. *Clin. Rheumatol.* 2014, *33*, 1575–1583.
83. Asoudeh, F.; Dashti, F.; Jayedi, A.; Hemmati, A.; Fadel, A.; Mohammadi, H. Caffeine, Coffee, Tea and Risk of Rheumatoid Arthritis: Systematic Review and Dose-Response Meta-analysis of Prospective Cohort Studies. *Front. Nutr.* 2022, *9*, 822557.
84. Mizoguchi, E.; Sadanaga, T.; Okada, T.; Minagawa, T.; Akiba, J. Does caffeine have a double-edged sword role in inflammation and carcinogenesis in the colon? *Intest. Res.* 2023, *21*, 306–317.
85. Ma, X.; Nan, F.; Liang, H.; Shu, P.; Fan, X.; Song, X.; Hou, Y.; Zhang, D. Excessive intake of sugar: An accomplice of inflammation. *Front. Immunol.* 2022, *13*, 988481.
86. French, M.T.; Popovici, I.; Maclean, J.C. Do alcohol consumers exercise more? Findings from a national survey. *Am. J. Health Promot. AJHP* 2009, *24*, 2–10.

Chapter 3 Measures of adiposity and risk of rheumatoid arthritis in middle-aged UK women: A prospective cohort study

Yuanyuan Dong ¹, Darren C. Greenwood ², Laura J. Hardie ³ and Janet E. Cade ¹

1 Nutritional Epidemiology Group, School of Food Science and Nutrition, University of Leeds,

Leeds LS2 9JT, UK;

2 Leeds Institute for Data Analytics, Faculty of Medicine and Health, University of Leeds, Leeds LS2 9JT, UK;

3 Division of Clinical and Population Sciences, Leeds Institute of Cardiovascular and Metabolic Medicine, School of Medicine, University of Leeds, Leeds LS2 9JT, UK.

Published: Nutrients. 2025. 17(9): 1557. <https://doi.org/10.3390/nu17091557>

This Chapter is an exact copy of the journal paper referred to above.

3.1 Abstract

Background: To estimate the association between various indicators of obesity-related health risk and the incidence of rheumatoid arthritis (RA) in a large cohort of women.

Methods: The UK Women's Cohort Study is a prospective cohort of 35,372 middle-aged women (aged 35–69 at recruitment) initiated in 1995–1998. Obesity was assessed using body mass index (BMI), waist circumference (WC), waist-to-hip ratio (WHR), waist-to-height ratio (WHtR), categorised according to WHO and NICE guidelines, as well as clothing size. Incident RA cases were identified via Hospital Episode Statistics (HES) linkage up to March 2019. Cox regression models were used to estimate RA risk, adjusting for demographics, reproductive factors, and lifestyle factors. Non-linear associations were examined using restricted cubic splines.

Results: Among 27,968 eligible subjects with complete data linkage (625,269 person-years of follow-up), there were 255 incident RA cases. Obesity (≥ 30.0 kg/m²) was associated with increased RA risk (HR (95% CI) 1.48 (1.02, 2.17)), as were abdominal obesity (WC > 88 cm: 1.58 (1.10, 2.27)), WHR ≥ 0.85 (1.56 (1.03, 2.36)), and WHtR ≥ 0.6 (2.25 (1.34, 3.80)). Each 2.5 kg/m² increase in BMI was associated with a 9% higher risk of RA; each 5 cm increase in WC with 6%; each 0.1 increase in WHR with 20%, and each 0.1 increase in WHtR with 27%. Larger clothing sizes were associated with a greater RA risk: for each onesize increment in blouse size and skirt size, the HRs were 1.13 (95% CI: 1.04, 1.22) and 1.13 (95% CI: 1.05, 1.22), respectively. Notably, skirt size ≥ 20 was associated with a 2.36-fold increased risk of RA. There was evidence of effect modification by weight change and menopausal status in obesity-related RA risk.

Conclusions: Our findings suggest that managing obesity and central adiposity in middle-aged women may be associated with the risk of developing RA. WHtR may serve as a practical alternative to BMI in assessing RA risk. Clothing size, particularly skirt size, could provide a simple, cost-effective proxy for identifying at high risk of RA.

Keywords: Rheumatoid arthritis; obesity; body mass index; waist circumference; diet quality; waist-to-hip ratio; waist-to-height ratio

3.2 Introduction

Rheumatoid arthritis (RA) is the most common inflammatory arthritis, affecting approximately 17.6 million people worldwide [1]. Middle-aged women are disproportionately affected and are more likely to suffer from mechanical pain and fibromyalgia [2]. Obesity, a major modifiable risk factor for various chronic diseases, is highly prevalent in this demographic, with 69% of women aged 55 to 74 in the UK classified as overweight or obese [3]. Given the increasing burden of obesity, understanding its role in RA development is crucial for targeted prevention strategies.

Recent scientific consensus has highlighted the limitations of body mass index (BMI) as a measure of obesity, particularly its inability to account for the distribution of body fat and total fat mass, which is strongly associated with metabolic and inflammatory pathways [4]. Alternative anthropometric indices, such as waist circumference (WC), waist-to-hip ratio (WHR), and waist-to-height ratio (WHtR), provide a more precise assessment of obesity-related health risks, as they better capture visceral fat distribution and body shape [4]. These indices have been shown to improve early identification of individuals at metabolic risk, account for population diversity and sex, and enhance targeted interventions for conditions such as cardiovascular disease (CVD) and type 2 diabetes [5]. Similarly, the National Institute for Health and Care Excellence in the UK (NICE) guidelines have suggested that WHtR could be a replacement for obesity measurement [6].

Although public health and clinical advantages of alternatives to BMI are well known, few studies have investigated their role in RA development. Previous epidemiological studies examining the relationship between obesity and RA risk have relied on BMI-based classification, with inconsistent findings. Two meta-analyses, primarily based on case-control studies, showed that being overweight or obese is associated with a higher risk of RA [7,8]; other studies reported no evidence of any association [9,10]. One possible reason could be that these studies overlooked the role of body fat distribution, which could be distinct across regions and sexes. Limited evidence suggests that WC could be more strongly associated with the risk of RA [11], particularly seropositive RA [12].

However, existing studies have not yet examined whether individuals with obesity or abdominal obesity can lower their RA risk by adopting a healthier lifestyle. To our knowledge, it remains unclear whether significant weight changes modify the association between obesity and RA risk, particularly in individuals with high visceral adiposity. Addressing this gap is critical for informing evidence-based public health strategies. The United Kingdom Women's Cohort Study (UKWCS) provides a unique opportunity to investigate this question in a large, prospective sample of middle-aged UK women. Our study aims to assess the association between multiple anthropometric indices (e.g., BMI, WC, WHR, WHtR, clothing size) and RA risk, while investigating the potential modifying effects of weight change and lifestyle factors.

3.3 Materials and methods

We followed the Strengthening the Reporting of Observational Studies in Epidemiology—Nutritional Epidemiology (STROBE-nut) guidelines for the reporting of cohort studies (**Appendix B: Table B1**) [13].

3.3.1 Study design and participants

We used data from the UKWCS, a large-scale prospective cohort with 35,372 women aged 35–69 years enrolled at baseline [14]. In brief, 500,000 women from England, Scotland, and Wales, who responded to a direct mail questionnaire from the World Cancer Research Fund (WCRF), were eligible for inclusion in the UKWCS. All participants completed a postal questionnaire that collected dietary, lifestyle, demographic, and anthropometric data at recruitment (1995–1998). Dietary intake in the UKWCS was assessed using a validated 217-item FFQ. The FFQ was validated against 4-day weighed food diaries and a repeat FFQ administration among a subsample of 283 women in 2001 [14]. Individuals living outside of England ($n = 2314$) and those without a National Health Service number to facilitate data linkage ($n = 695$) were excluded. We also excluded participants who had a diagnosis of RA before or at recruitment ($n = 38$), had outlier anthropometric index ($\text{BMI} < 10$ or $> 60 \text{ kg/m}^2$, or waist < 50 or $> 140 \text{ cm}$, or hip < 60 or $> 160 \text{ cm}$, or weight < 30 or $> 200 \text{ kg}$, $n = 160$), and those with missing data for anthropometric index ($n = 1302$). After exclusion, a sample size of 30,863 participants remained eligible for this analysis.

The study protocol was approved by the National Research Ethics Service Committee for Yorkshire & the Humber—Leeds East (reference 15/YH/0027) in 1993 and was updated in 2017 to include linkage to Hospital Episode Statistics (HES) data (reference 17/YH/0144).

3.3.2 Anthropometric assessments

Anthropometric indices were derived from self-reported data collected during the baseline survey, following detailed guidance and standardised procedures provided to all participants. Measurements included body weight (kg), height (m), WC (cm), hip circumference (cm), and clothing size (blouse and skirt). These data were used to calculate key adiposity indicators: BMI as body weight (kg) divided by height squared (m^2); WHR as WC (cm) divided by hip circumference (cm); and WHtR as WC (cm) divided by height (cm). Further details are provided in **Appendix B: Table B2**.

Clothing size has previously been associated with central adiposity and used to predict chronic disease risk [15]. It is considered a practical alternative when direct anthropometric measurements (e.g., WC or hip circumference) are missing or incomplete, and it enhances the assessment of body shape in self-reported surveys.

3.3.3 Assessment of lifestyle factors and other covariates

Lifestyle factors, including smoking, drinking, physical activity, dietary quality, and energy intake, were obtained through the baseline questionnaire. Smoking status and alcohol consumption were evaluated through frequency, type, and daily quantity. Physical activity was provided using metabolic equivalent tasks (METs), calculated as MET-hours per week by multiplying the duration of activities by their respective MET values [16]. Dietary quality was measured using a validated 217-item FFQ, which evaluated the intakes of fruits, vegetables, whole grains, nuts and legumes, fish and long-chain omega-3 fatty acids, red and processed meats, sugar-sweetened beverages and fruit juices, trans fat, alcohol, and sodium. Each component was scored from 0 to 10, and the Alternative Healthy Eating Index (AHEI)-2010 score (ranging from 0 to 110) was derived as the sum of these scores, with higher values indicating a healthier diet [17]. Total energy intake was

estimated from portion sizes reported in the FFQ and linked to nutrient composition tables (McCance and Widdowson's Composition of Foods database), reflecting the caloric content of consumed foods rather than their nutritional quality.

Weight change was computed as the difference between weight at enrolment and weight at early adulthood (age 20), and categorised into weight loss (< -2.5 kg), stable weight (-2.5 to 2.5 kg), and weight gain (moderate: 2.5 to 10 kg; large: > 10 kg). Additional covariates included age (continuous, years); socio-economic status (SES; categorical: routine/manual, intermediate, professional/managerial); marital status (categorical: married/living as married, separated/divorced, single/widowed); menopausal status (binary: premenopausal, postmenopausal), hormone replacement therapy (HRT; binary: yes, no); and the prevalence of CVD, cancer, or diabetes at recruitment (binary: yes, no). Details on covariate definitions and assessments are provided in **Appendix B: Table B3**.

3.3.4 Ascertainment of RA cases and mortality from CVD and cancer

The primary outcome was RA incidence (International Classification of Diseases, ICD-10 codes M05–M06), ascertained through linkage to HES data up to 31 March 2019. The secondary outcome was mortality from CVD or cancer, obtained through linkage to the National Health Service Central Register (NHSCR) and the Health and Social Care Information Centre (HSCIC), now NHS Digital. Death data were classified according to ICD-9 and ICD-10 codes. The first recorded event (either RA diagnosis or mortality from CVD/cancer) from the combined data sources was used as the endpoint in the time-to-event analysis.

3.3.5 Statistical methods

All statistical methods were registered in advance on [ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT06670144) (NCT06670144).

Participant characteristics were summarised using descriptive statistics, with means or percentages calculated by obesity indicator group. Cox proportional hazards (PH) regression models, stratified by age group, were used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) for BMI, WC, WHR, WHtR, blouse and skirt size in relation to RA incidence. All exposures were

modelled as continuous variables to assess the linear relationships with RA risk. We also used WHO and NICE guidelines as the basis for classification: BMI ($18.5 < 25.0$, $25.0 < 30.0$, ≥ 30.0 kg/m²), WC (≤ 88 , > 88 cm), WHR (< 0.85 , ≥ 0.85), WHtR (< 0.5 , 0.5 to 0.6 , and ≥ 0.6). Since the sample size of blouse and skirt sizes allowed for a more detailed investigation, the participants were divided into six garment size groups (≤ 10 , 12 , 14 , 16 , 18 , ≥ 20). Reference groups included normal weight (BMI: $18.5 < 25.0$ kg/m²), abdominal leanness (WC ≤ 88 cm), WHR < 0.85 , WHtR < 0.5 , and blouse/skirt size $\leq$ UK 10, which were assumed to have the lowest risk.

The person-years were calculated from enrolment to the earliest of the following events: RA diagnosis, censorship due to CVD- or cancer-related death, or the end of follow-up, using age as the timescale. Adjustment for potential confounders was applied progressively in three steps: Model 1 was unadjusted; Model 2 adjusted for age, SES, marital status, menopausal status, HRT, and prevalence of chronic diseases; Model 3 further adjusted for lifestyle factors, including smoking status, alcohol consumption, dietary quality (AHEI-2010 score) and energy intake, physical activity (METs). The PH assumptions were checked using Schoenfeld residuals (**Appendix B: Table B4**).

To estimate the effect of weight change and age on these associations, we added the weight change group (< 2.5 , -2.5 to 2.5 , 2.5 to 10 , > 10 kg) and age group (≤ 55 , > 55 years) to the fully-adjusted model as an interaction term. Further exploratory analyses included testing for interaction effects of lifestyle factors: dietary quality (AHEI-2010 tertiles: low, moderate, high), physical activity (METs: < 600 , 600 – 1500 , > 1500 min/week), and current smoker (yes, no). In addition, we conducted subgroup analysis by menopausal status (premenopausal, postmenopausal) to investigate its modifying effect among middle-aged women. The adjustment set in each subgroup analysis omitted the potential effect modifier.

We conducted a sensitivity analysis to test the robustness of the main results: (1) excluding participants with a survival time < 5 years to check for reverse causation; (2) excluding current smokers, given its known effect on RA risk; (3) excluding participants who developed chronic diseases (e.g., CVD, cancer, or diabetes); (4) excluding participants with weight change greater than 10 kg to

minimise the influence of subclinical conditions; (5) additionally adjusting the model for vitamin D intake; (6) applying a competing risk regression model (Fine-Gray method), considering death as a competing event. All statistical analyses were performed using Stata (version 18).

3.4 Results

Of the 30,863 participants potentially eligible in the UKWCS, we excluded those with missing data for marital status ($n = 280$), smoking status ($n = 382$), menopausal status ($n = 438$), physical activity ($n = 1389$) or SES ($n = 406$), leaving 27,968 middle-aged women for analyses. The study flowchart is provided in **Appendix B: Figure B1**.

3.4.1 Participants

Over a median follow-up of 22.3 years (625,269 person-years), 255 incident RA cases (1% of this cohort) were identified. At baseline, participants had a mean age of 51.9 years (standard deviation [SD], 9.2). Mean BMI increased from 21.4 kg/m² (SD 2.9) at age 20 to 24.4 kg/m² (SD 4.2) at recruitment. Compared to non-RA participants, those diagnosed with RA were more likely to be older, have lower educational attainment; hold non-professional occupations; be unmarried; use HRT; and have a higher prevalence of chronic disease. RA cases also exhibited higher baseline BMI and waist circumference, lower physical activity levels, and were more likely to be current or former smokers, with slightly lower alcohol consumption compared to non-cases, though total energy intake and AHEI-2010 scores were comparable between the two groups. **Table 3.1** summarises these characteristics by RA incidence among the 27,968 participants.

Table 3.1 Baseline characteristics of 27,968 UKWCS participants stratified by rheumatoid arthritis incidence.

Characteristics	Total	Cases	Non-cases
Participants (%)	27,968	255 (1.0)	27,713 (99.0)
Age, years			
Mean (SD)	51.9 (9.2)	53.9 (9.0)	51.9 (9.2)
Range (Min, Max)	(35, 69)	(37, 62)	(35, 69)
Age of diagnosis, years			
Mean (SD)	-	65 (10)	-
Range (Min, Max)	-	(38, 86)	-
Degree-level education (%)	7057 (25.2)	48 (18.8)	7009 (25.3)
Socio-economic status (%)			
Professional or managerial	17,753 (63.5)	146 (57.3)	17,607 (63.5)
Intermediate	7735 (27.7)	80 (31.4)	7655 (27.6)
Routine or manual	2480 (8.9)	29 (11.4)	2451 (8.8)
Married (%)	21,207 (75.8)	189 (74.1)	21,018 (75.8)
Anthropometric indices			
Height, cm			
Mean (SD)	164 (6.6)	163 (6.8)	164 (6.6)

Range (Min, Max)	(145, 182)	(143, 180)	(145, 182)
Weight, kg			
Mean (SD)	65 (11.6)	67 (14.3)	65 (11.5)
Range (Min, Max)	(42, 118)	(45, 118)	(42, 115)
Waist circumference, cm (SD)	74.2 (9.6)	76.2 (11.4)	74.2 (9.6)
BMI at year 20, kg/m ²			
Mean (SD)	21.4(2.9)	21.7(3.0)	21.4(2.9)
Range (Min, Max)	(12.0, 32.0)	(16.0, 33.0)	(12.0, 32.0)
BMI at recruitment, kg/m ²			
Mean (SD)	24.4 (4.2)	25.4 (5.3)	24.4 (4.2)
Range (Min. Max)	(15.5, 41.1)	(17.2, 41.1)	(15.5, 38.1)
Waist-to-hip ratio (SD)	0.8 (0.1)	0.8 (0.1)	0.8 (0.1)
Waist-to-height ratio (SD)	0.5 (0.1)	0.5 (0.1)	0.5 (0.1)
Blouse size (SD)	13.9 (2.9)	14.7 (3.6)	13.9 (2.9)
Skirt size (SD)	14.5 (3.2)	15.4 (3.7)	14.5 (3.2)
Physical activity (METs), hours/week (SD)	16.6 (11.5)	15.4 (9.8)	16.6 (11.5)
Smoking status (%)			

Current	3081 (11.0)	41 (16.1)	3040 (11.0)
Former	8722 (31.2)	89 (34.9)	8633 (31.2)
Never	16,165 (57.8)	125 (49.0)	16,040 (57.9)
Alcohol consumption, grams/week			
Mean (SD)	9.0 (10.5)	7.1 (8.4)	9.0 (10.5)
Range (Min, Max)	(0, 224)	(0, 183)	(0, 224)
Total energy intake, kcal/day (SD)	2334.2 (701.5)	2352.8 (847.3)	2334.1 (700.0)
AHEI-2010 score (SD)	65.6 (11.0)	65.8 (10.1)	65.6 (11.0)
Menopausal status (%)			
Premenopausal	13,571 (48.5)	93 (36.5)	13,571 (49.0)
Postmenopausal	14,397 (51.5)	162 (63.5)	14,397 (52.0)
Hormone replacement therapy (%)	7698 (27.5)	82 (32.2)	7616 (27.5)
Prevalence of CVD, cancer, or diabetes (%)	1908 (6.8)	30 (11.8)	1878 (6.8)

SD, standard deviation; BMI, body mass index; WC, waist circumference; WHR, waist-to-hip ratio; WHtR, waist-to-height ratio; MET, Metabolic Equivalent of Task; AHEI-2010, Alternative Healthy Eating Index-2010; CVD, cardiovascular disease. BMI, WC, and WHR were categorised using WHO guidelines, and WHtR was categorised using NICE guidelines.

Baseline characteristics of the 27,968 participants by indicators of obesity-related health are presented in **Appendix B: Tables B5 and B6**. Different adiposity indicators showed consistent associations with demographic and lifestyle factors. Specifically, participants with higher BMI, WC, WHR, WHtR, and larger blouse or skirt size were more likely to be older, postmenopausal, have lower education levels, be unmarried, and hold non-professional occupations. Also, chronic diseases and HRT use were more prevalent among participants with higher adiposity, with the highest occurrence observed in those with a WHtR ≥ 0.6 ($n = 115$ [15.6%]). Physical activity levels were highest in participants with a skirt size ≤ 10 (17.7 h/week, SD 12.0), a blouse size ≤ 10 (17.1 h/week, SD 11.5), and those in the normal weight group (16.8 h/week, SD 11.4), and lowest among individuals living with obesity (BMI ≥ 30 kg/m², 15.8 h/week, SD 12.9), a WHtR ≥ 0.6 (15.8 h/week, SD 12.6), and a skirt size ≥ 20 (15.4 h/week, SD 12.0). Participants with higher BMI, WC, WHR, WHtR, blouse size, and skirt size were generally less likely to smoke, consume alcohol, or have a high diet quality. Energy intake was highest in the group with a WHR of 0.80–0.85 (2355 kcal/day, SD 697.8) and lowest among those with a skirt size ≤ 10 (2287 kcal/day, SD 714.8).

3.4.2 Associations of BMI, WC, WHR, WHtR, blouse size, skirt size with RA risk

Figure 3.1 demonstrates the associations between BMI, WC, WHR, WHtR, blouse size, and skirt size and the incidence of RA. After adjustment for lifestyle factors (model 3), a linear association showed an increased risk of incident RA with BMI (per 2.5 kg/m² increase: HR 1.09 (95% CI: 1.02, 1.16)), WC (per 5 cm increase): 1.06 (1.00, 1.13), WHR (per 0.1 increase: 1.20 (1.01, 1.43)), WHtR (per 0.1 increase: 1.27 (1.05, 1.53)), blouse size (per one size increase: 1.13 (1.04, 1.22)), and skirt size (per one size increase: 1.13 (1.05, 1.22)). Obesity (BMI ≥ 30.0 kg/m²) was associated with a 48% higher risk of RA (1.48 (1.02, 2.17)), while abdominal obesity (WC: per 5 cm increase) conferred a 58% higher risk (1.58 (1.10, 2.17)). WHR ≥ 0.85 was associated with a 56% greater risk of RA (1.56 (1.03, 2.36)) and WHtR ≥ 0.6 with a 2.25-fold higher risk (2.25 (1.34, 3.80)). Similarly, skirt size ≥ 20 was linked to a 2.36-fold increased risk of RA (2.36 (1.25, 4.43)). **Figure 3.2** presents non-linear associations, revealing a U-shaped

relationship between BMI and RA risk, with the lowest risk around 22 kg/m², though the evidence of association was lacking (P-non-linearity = 0.5). The risk of RA increased monotonically with WC, WHtR, blouse size and skirt size, but there was no evidence of non-linearity in these associations (P-non-linearity > 0.05).

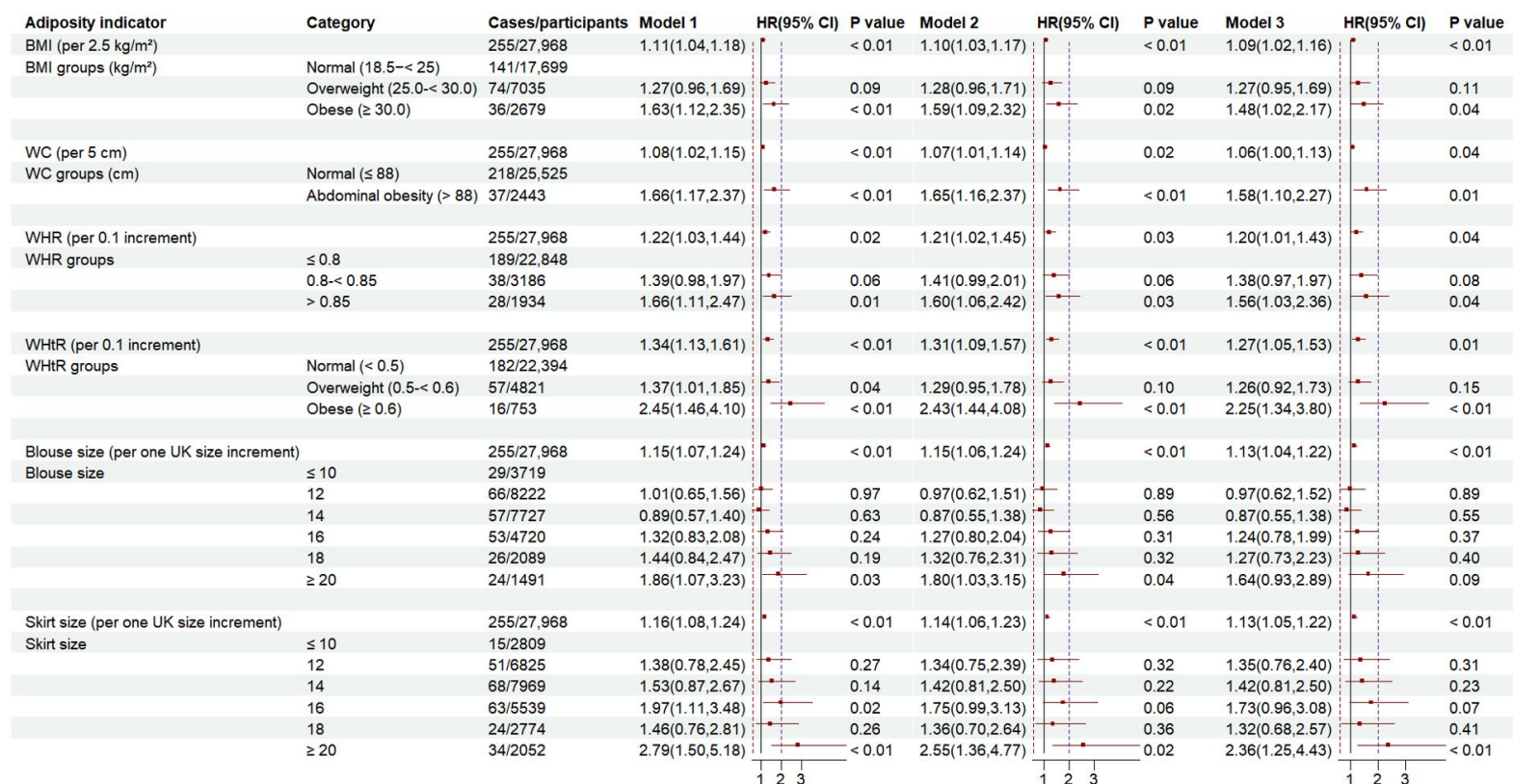


Figure 3.1 Association between indicators of adiposity and risk of rheumatoid arthritis in the UK Women's Cohort Study. BMI, body mass index; WC, waist circumference; WHR, waist-to-hip ratio; WHtR, waist-to-height ratio; Models are defined as follows: model 1: unadjusted model; model 2: adjusted for age, socio-economic status, marital status, menopausal status, hormone replacement therapy, and prevalence of cardiovascular diseases, cancer, or diabetes at recruitment; model 3: model 2 + physical activity, smoking status, alcohol consumption, total energy intake, and Alternate Healthy Eating Index-2010; HR, hazard ratio; CI, confidence interval.

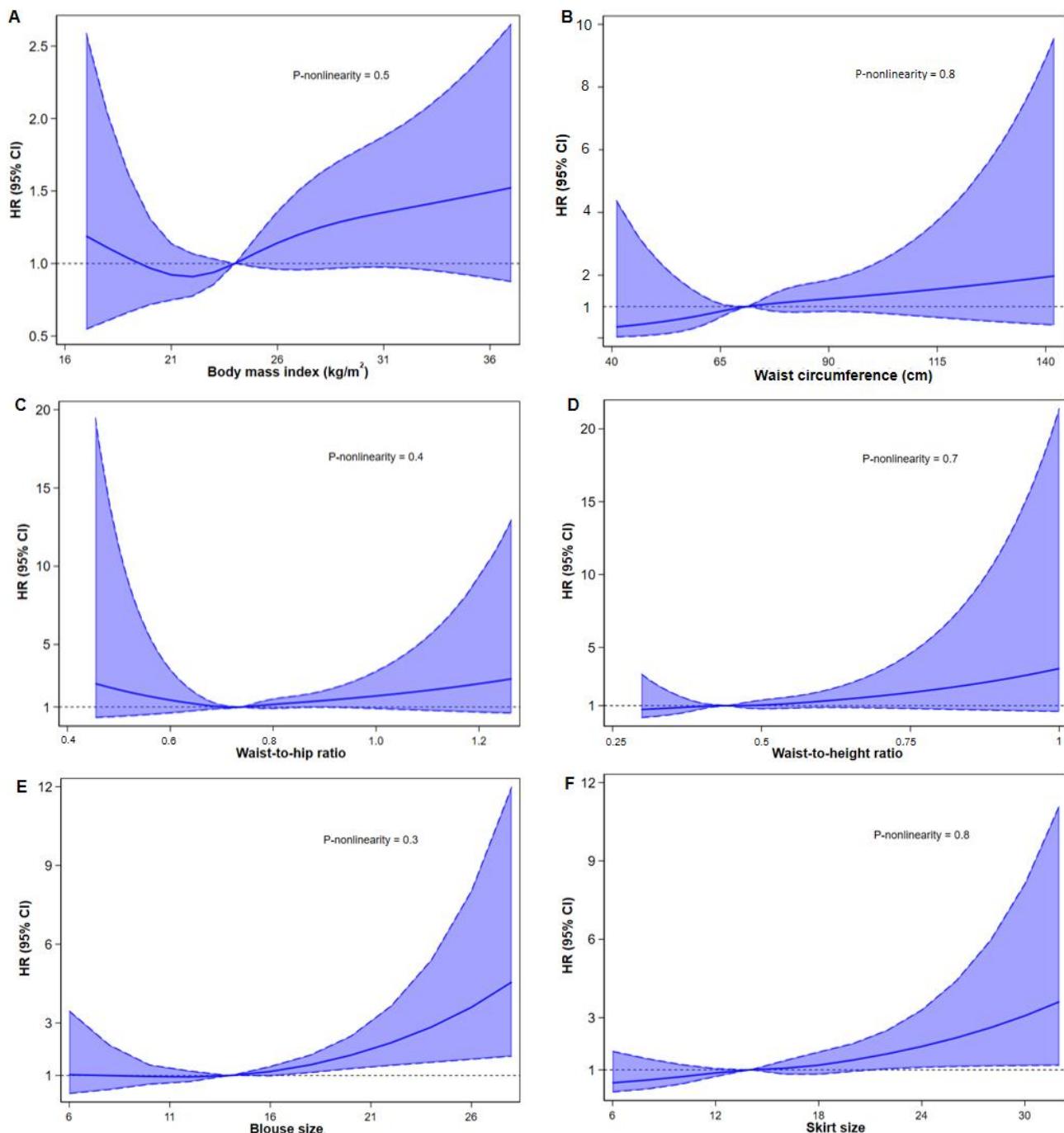


Figure 3.2 Association between indicators of adiposity with rheumatoid arthritis incidence. Graphs show the dose-response associations between the risk of rheumatoid arthritis and body mass index (A), waist circumference (B), waist-to-hip ratio (C), waist-to-height ratio (D), blouse size (E), and skirt size (F). Estimates were adjusted for age, socio-economic status, marital status, physical activity, smoking status, alcohol consumption, total energy intake, the Alternate Healthy Eating Index-2010, menopausal status, hormone replacement therapy, and the prevalence of CVD, cancer, or diabetes at recruitment. Solid lines are multivariable-adjusted HRs, with shaded areas showing 95% confidence intervals derived from restricted cubic spline regressions with four knots. The dashed horizontal line indicates a reference for no association at an HR 1.0. HR, hazard ratio; CI, confidence interval.

3.4.3 Subgroup analysis

In subgroup analysis by weight change, we observed effect modification, with WC, WHR, and WHtR showing a linear increase in RA risk only among participants who gained more than 10 kg (P-interaction < 0.05; **Appendix B: Table B8**). Also, menopausal status appeared to influence the associations, with BMI, WC, and WHtR (both continuous and categorical measures) showing an increased risk only in premenopausal women (P-interaction < 0.05; **Appendix B: Tables B11 and B18**). While the risk of RA was significantly higher in individuals younger than 55 years compared to those aged 55 or older, there was no evidence that age modified the associations between obesity indicator and RA risk (**Appendix B: Table B7**). No evidence of effect modification was found according to AHEI-2010 tertiles, physical activity, SES, or smoking status (**Appendix B: Tables B7–B18**).

3.4.4 Sensitivity analysis

Most of the sensitivity analyses supported the robustness of our main findings. However, when participants with substantial weight gain (> 10 kg) were excluded, traditional obesity indicators (BMI, WC, WHR, WHtR) lost their clear associations with the risk of RA, whereas blouse size and skirt size maintained significant associations (blouse size: 1.23 (1.03, 1.46); skirt size: 1.19 (1.04, 1.36); **Appendix B: Table B19**). Additionally, the exclusion of current smokers strengthened the obesity-RA associations: the association of BMI (per 2.5 kg/m² increase) with RA risk increased from 1.09 (1.02, 1.16) to 1.13 (1.06, 1.21); WC (per 5 cm increment) from 1.06 (1.00, 1.13) to 1.11 (1.04, 1.19); WHR (per 0.1 increment) from 1.20 (1.01, 1.43) to 1.30 (1.11, 1.53); WHtR (per 0.1 increment) from 1.27 (1.05, 1.53) to 1.42 (1.17, 1.74); blouse size (per one size increment) from 1.13 (1.04, 1.22) to 1.17 (1.07, 1.29); and skirt size (per one size increment) from 1.13 (1.05, 1.22) to 1.17 (1.08, 1.28), indicating that the dominant effect of smoking had previously obscured obesity-related risk (**Appendix B: Table B19**).

3.5 Discussion

3.5.1 Main Findings

This study is the first to simultaneously assess BMI, WC, WHR, WHtR, and clothing size to RA risk in a population of UK women. Our findings support a linear and consistently risk-increasing association between all indicators of obesity-related health and the risk of RA, independent of lifestyle factors, with even stronger positive associations observed in younger women (≤ 55 years). Participants who did not meet the WHO or NICE guidelines, with a BMI ≥ 30 kg/m², WC > 88 cm, WHR ≥ 0.85 , or WHtR ≥ 0.6 had a greater incidence of RA, with WHtR showing the strongest association (a 2.25-fold increased risk). Similarly, a skirt size ≥ 20 was associated with more than a doubling in RA risk. These findings highlight the impact of central obesity on RA risk and emphasise the need for lifestyle interventions to mitigate obesity-related RA risk.

3.5.2 Comparison with previous studies

Our findings align with prior prospective cohort studies in US and European populations [18–20]. The Danish cohort reported a 46% higher RA risk in individuals with obesity (BMI ≥ 30 kg/m²) [20], and our study found a similar 48% increase in middle-aged women, in contrast to the mixed-sex sample in the Danish study. The Nurses' Health Studies I and II also reported a 37% higher RA risk in overweight US women (BMI: 25.0 –< 30.0 kg/m²), though their confidence intervals were wide [18]. The French E3N cohort reported a 26% higher risk in individuals with obesity (BMI ≥ 30 kg/m²), but this was marginally significant [19]. Divergence in risk estimates across studies may be due to differences in case identification: the E3N cohort used self-reported data, whereas our study leveraged the UKWCS-HES linkage for clinically verified diagnoses. A study using the UK General Practice Research Database found no association between obesity and RA risk in over 1 million men and women, potentially due to reliance on physician-recorded BMI data [21]. Compared to our results, this discrepancy may also arise from differences in study design, including the absence of smoking-stratified analyses and the exclusive use of BMI rather than WTHR, which may better capture central obesity. Consistent with the results from a meta-

analysis showing a 13% higher RA risk per 10 cm increase in WC [11], we found a 6% higher risk for every 5 cm increase in WC. This meta-analysis included the Danish cohort [20], but also combined data with a case-control study [22], which may affect the generalisability of the findings.

In this study, our findings permit several important inferences. First, the association of central obesity with RA incidence may be stronger than that of overall obesity as measured by BMI. Indicators such as WC, WHR, WHtR demonstrated stronger associations with RA incidence than BMI. Notably, exceeding the NICE guideline threshold for WHtR (≥ 0.6) was linked to a substantially greater effect than BMI-defined obesity ($\geq 30 \text{ kg/m}^2$). Skirt size, which reflects central fat distribution (waist and hips), showed a nearly twofold increase in RA risk when ≥ 20 , whereas blouse size (more reflective of upper body size) showed a weaker association. When excluding participants with substantial weight gain, only blouse size and skirt size retained significant associations with RA risk. This suggests that clothing size may reflect long-term, stable patterns of obesity that are less affected by short-term weight changes. However, clothing size may also be influenced by factors such as fashion trends, socioeconomic status, and cultural norms, which could affect its reliability as a consistent adiposity proxy across populations. These findings suggest that WHtR ≥ 0.6 and skirt size ≥ 20 could serve as practical indicators for identifying individuals at higher risk of RA in primary care settings. Risk assessments and public health guidelines should, therefore, expand beyond BMI to include central obesity informed by clothing size, which better captures adiposity-related disease risk, even in individuals with “normal” BMI. For example, in our cohort, women with WHtR ≥ 0.6 had 2.25-fold higher RA risk, comparable to that associated with smoking. This includes women with normal BMI but elevated WHtR, a high-risk subgroup often missed by traditional screening.

While the HR for BMI of 1.09 corresponds to a relatively small increment (per 2.5 kg/m^2), the cumulative risk for RA associated with a 10 kg/m^2 increase in BMI is substantial, with a 41% increased risk. The clinical significance of these small changes becomes more apparent when BMI is categorised.

Second, our sensitivity analyses provide insights into the underlying mechanisms. Excluding current smokers consistently strengthened the associations between all obesity indicators and RA risk, highlighting the possibility that smoking may obscure the role of obesity. This is consistent with findings from the Danish cohort, where central adiposity (assessed by WC or body shape trajectories) predicted RA risk more robustly than BMI among never-smokers [20]. Given the well-established and dominant role of smoking in RA pathogenesis, its exclusion from analyses helps uncover the independent contribution of obesity, likely mediated through distinct metabolic and inflammatory pathways. This finding needs further mechanistic investigation into how adiposity-related inflammation may play a role in RA development independently of smoking-related autoimmune pathways.

Third, after adjusting for lifestyle factors such as smoking, alcohol intake, diet quality, energy intake, and physical activity, the associations between adiposity indicators and RA were attenuated. This suggests that a healthy lifestyle may be associated with a weaker association between obesity and RA. These findings support the rationale for developing preventive strategies targeting modifiable behaviours in individuals with elevated central obesity that could lower their risk of RA.

3.5.3 Interpretation and implications

Although excess weight may contribute to joint stress, this mechanism is more pertinent to osteoarthritis [23]. In this study, the stronger associations with central adiposity indicators rather than BMI suggest that systemic inflammation driven by visceral fat is a more plausible pathway linking obesity to RA risk. Excess body fat has been linked to RA risk through chronic low-grade systemic inflammation, insulin resistance, metabolic hormones, gut microbiota alterations [24], and oxidative stress [25]. Enlarged adipocytes secrete pro-inflammatory cytokines such as TNF- α , IL-6, and CRP, while adipose tissue recruits macrophages, further amplifying inflammation. Adipokines like leptin and adiponectin also regulate immune responses; leptin promotes pro-inflammatory T-cell activity [26], while elevated adiponectin exacerbates joint inflammation by activating synovial fibroblasts and promoting cartilage degradation [6]. Additionally, insulin resistance, a well-established consequence of excess abdominal fat, may disrupt

immune function and promote systemic inflammation [27]. Other potential mechanisms, such as DNA methylation [28] and telomere shortening [29], could further mediate the link between obesity and RA, independent of direct metabolic changes.

Beyond their biological effects, obesity-related indices also have psychological implications, especially waist-related indicators, influencing individual's awareness of their body shape and weight status. This "self-perception of obesity"[30] may affect health behaviours, such as dietary choices and physical activity, which in turn impact metabolic health and inflammation.

3.5.4 Strengths and limitations

This study benefits from a large sample size, a prospective design, and a long follow-up period, strengthening the reliability of these findings. RA cases were identified using ICD-10 codes from hospital records, minimising misclassification and loss of follow-up. Also, we adjusted for a comprehensive set of confounders, including sociodemographic factors (e.g., age, socioeconomic status), lifestyle factors (physical activity, smoking, alcohol intake, diet quality), reproductive history (menopausal status, hormone replacement therapy), and total energy intake, reducing the risk of residual confounding.

However, residual confounding remains a possibility, particularly from unmeasured factors such as family history and genetic predisposition if they increase both adiposity and RA risk other than through increased adiposity, medication use that increases both adiposity and RA risk other than through adiposity, and specific dietary patterns that independently influence both. Consequently, as in all observational studies, causality cannot be inferred. The absence of inflammatory biomarkers (e.g., CRP, IL-6) limits our ability to investigate the underlying mechanisms linking obesity to RA. Future studies integrating genetic and clinical data may further clarify these associations. Secondly, the generalisability of our findings may be constrained since our cohort comprised predominantly White, middle-aged women in England. This limits the extrapolation of results to men or non-European populations. Third, anthropometric data were based on self-reported measures, which may be subject to reporting bias. This limitation may be particularly relevant for obese

participants, who tend to under-report weight and over-report height. However, participants were given standardised instructions, and self-reported anthropometry has been shown to be acceptably accurate in epidemiological research [31]. Any misclassification is likely non-differential, which could attenuate the observed associations rather than exaggerate them. Finally, BMI, WC, WHR, and WHtR are likely to be correlated. While we assessed the association of each measure independently, these indices may jointly influence RA risk. Although we did not formally test interaction effects between obesity indicators, future analyses using combined or interaction models could offer insights into whether combinations of anthropometric traits elevate RA risk.

3.6 Conclusions

The findings of this study suggest the importance of recognising central obesity as a distinct and potentially stronger predictor of RA than BMI. It also highlights the value of alternative, practical indicators such as clothing size in long-term risk assessment. Future research should explore the biological pathways linking adiposity and RA, particularly in non-smoking populations, and evaluate whether lifestyle interventions can modify this relationship.

3.7 References

1. Global, regional, and national burden of rheumatoid arthritis, 1990–2020, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. *Lancet Rheumatol.* 2023, 5, e594–e610.
2. Casale, R.; Atzeni, F.; Bazzichi, L.; Beretta, G.; Costantini, E.; Sacerdote, P.; Tassorelli, C. Pain in Women: A Perspective Review on a Relevant Clinical Issue that Deserves Prioritization. *Pain Ther.* 2021, 10, 287–314.
3. NHS Digital: Health Survey for England 2021: Overweight and Obesity (Part 2). 2022. [Available online: <https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/2021/part-2-overweight-and-obesity>].
4. Rubino, F.; Cummings, D.E.; Eckel, R.H.; Cohen, R.V.; Wilding, J.P.H.; Brown, W.A.; Stanford, F.C.; Batterham, R.L.; Farooqi, I.S.; Farpour-Lambert, N.J.; et al. Definition and diagnostic criteria of clinical obesity. *Lancet. Diabetes Endocrinol.* 2025, 13, 221–262.

5. Powell-Wiley, T.M.; Poirier, P.; Burke, L.E.; Després, J.P.; Gordon-Larsen, P.; Lavie, C.J.; Lear, S.A.; Ndumele, C.E.; Neeland, I.J.; Sanders, P.; et al. Obesity and Cardiovascular Disease: A Scientific Statement From the American Heart Association. *Circulation* 2021, *143*, e984–e1010.
6. NICE. Keep the Size of Your Waist to Less Than Half of Your Height. 2022. [Available from: <https://www.nice.org.uk/news/articles/keep-the-size-of-your-waist-to-less-than-half-of-your-height-nice--recommends>].
7. Zhou, Y.; Sun, M. A meta-analysis of the relationship between body mass index and risk of rheumatoid arthritis. *EXCLI J.* 2018, *17*, 1079–1089.
8. Qin, B.; Yang, M.; Fu, H.; Ma, N.; Wei, T.; Tang, Q.; Hu, Z.; Liang, Y.; Yang, Z.; Zhong, R. Body mass index and the risk of rheumatoid arthritis: a systematic review and dose-response meta-analysis. *Arthritis Res Ther.* 2015, *17*, 86.
9. Rodríguez, L.A.; Tolosa, L.B.; Ruigómez, A.; Johansson, S.; Wallander, M.A. Rheumatoid arthritis in UK primary care: incidence and prior morbidity. *Scand. J. Rheumatol.* 2009, *38*, 173–177.
10. Lee, J.S.; Oh, J.S.; Kim, S.; Kim, Y.J.; Hong, S.; Kim, Y.G.; Lee, C.K.; Yoo, B. The association of obesity and the risk of rheumatoid arthritis according to abdominal obesity status: a nationwide population-based study in Korea. *Rheumatol. Int.* 2024, *44*, 2863–2871.
11. Ohno, T.; Aune, D.; Heath, A.K. Adiposity and the risk of rheumatoid arthritis: a systematic review and meta-analysis of cohort studies. *Sci. Rep.* 2020, *10*, 16006.
12. Marchand, N.E.; Sparks, J.A.; Tedeschi, S.K.; Malspeis, S.; Costenbader, K.H.; Karlson, E.W.; Lu, B. Abdominal Obesity in Comparison with General Obesity and Risk of Developing Rheumatoid Arthritis in Women. *J. Rheumatol.* 2021, *48*, 165–173.
13. Lachat, C.; Hawwash, D.; Ocké, M.C.; Berg, C.; Forsum, E.; Hörnell, A.; Larsson, C.; Sonestedt, E.; Wirfält, E.; Åkesson, A.; et al. Strengthening the Reporting of Observational Studies in Epidemiology-Nutritional Epidemiology (STROBE-nut): An Extension of the STROBE Statement. *PLoS Med.* 2016, *13*, e1002036.
14. Cade, J.E.; Burley, V.J.; Alwan, N.; Hutchinson, J. Cohort profile: the UK Women's cohort study (UKWCS). *Int. J. Epidemiol.* 2017, *46*, e11.

15. Moy, F.M.; Greenwood, D.C.; Cade, J.E. Associations of clothing size, adiposity and weight change with risk of postmenopausal breast cancer in the UK Women's Cohort Study (UKWCS). *BMJ Open* 2018, *8*, e022599.
16. Threapleton, D.E.; Greenwood, D.C.; Burley, V.J.; Aldwairji, M.; Cade, J.E. Dietary fibre and cardiovascular disease mortality in the UK Women's Cohort Study. *Eur. J. Epidemiol.* 2013, *28*, 335–346.
17. Wang, T.; Heianza, Y.; Sun, D.; Huang, T.; Ma, W.; Rimm, E.B.; Manson, J.E.; Hu, F.B.; Willett, W.C.; Qi, L. Improving adherence to healthy dietary patterns, genetic risk, and long term weight gain: gene-diet interaction analysis in two prospective cohort studies. *BMJ* 2018, *360*, j5644.
18. Lu, B.; Hiraki, L.T.; Sparks, J.A.; Malspeis, S.; Chen, C.Y.; Awosogba, J.A.; Arkema, E.V.; Costenbader, K.H.; Karlson, E.W. Being overweight or obese and risk of developing rheumatoid arthritis among women: a prospective cohort study. *Ann. Rheum. Dis.* 2014, *73*, 1914–1922.
19. Salliot, C.; Nguyen, Y.; Mariette, X.; Boutron-Ruault, M.C.; Seror, R. Anthropometric Measures and Risk of Rheumatoid Arthritis in the French E3N Cohort Study. *Nutrients* 2022, *14*, 934.
20. Linauskas, A.; Overvad, K.; Symmons, D.; Johansen, M.B.; Stengaard-Pedersen, K.; de Thurah, A. Body Fat Percentage, Waist Circumference, and Obesity As Risk Factors for Rheumatoid Arthritis: A Danish Cohort Study. *Arthritis Care Res.* 2019, *71*, 777–786.
21. Garcia Rodriguez, L.A.; Tolosa, L.B.; Ruigomez, S.; Johansson, S.; Wallander, M.-A.J.S. Rheumatoid arthritis in UK primary care: incidence and prior morbidity. *Scand. J. Rheumatol.* 2009, *38*, 173–177.
22. Ljung, L.; Rantapää-Dahlqvist, S. Abdominal obesity, gender and the risk of rheumatoid arthritis - a nested case-control study. *Arthritis Res. Ther.* 2016, *18*, 277.
23. Gao, T.; Chen, Z.Y.; Li, T.; Lin, X.; Hu, H.G.; Tang, J.D.; Wu, C. Association between body roundness index and osteoarthritis/rheumatoid arthritis: a cross-sectional study. *Sci. Rep.* 2025, *15*, 9682.
24. Zaiss, M.M.; Joyce Wu, H.-J.; Mauro, D.; Schett, G.; Ciccia, F.J.N.R.R. The gut–joint axis in rheumatoid arthritis. *Nat. Rev. Rheumatol.* 2021, *17*, 224–237.

25. Rohm, T.V.; Meier, D.T.; Olefsky, J.M.; Donath, M.Y. Inflammation in obesity, diabetes, and related disorders. *Immunity* 2022, *55*, 31–55.
26. Kiernan, K.; MacIver, N.J. The Role of the Adipokine Leptin in Immune Cell Function in Health and Disease. *Front. Immunol.* 2020, *11*, 622468.
27. Nicolau, J.; Lequerré, T.; Bacquet, H.; Vittecoq, O. Rheumatoid arthritis, insulin resistance, and diabetes. *Jt. Bone Spine* 2017, *84*, 411–416.
28. Payet, M.; Dargai, F.; Gasque, P.; Guillot, X. Epigenetic regulation (including micro-RNAs, DNA methylation and histone modifications) of rheumatoid arthritis: a systematic review. *Int. J. Mol. Sci.* 2021, *22*, 12170.
29. Costenbader, K.H.; Prescott, J.; Zee, R.Y.; De Vivo, I. Immunosenescence and rheumatoid arthritis: does telomere shortening predict impending disease? *Autoimmun. Rev.* 2011, *10*, 569–573.
30. Robinson, E.; Haynes, A.; Sutin, A.; Daly, M. Practice: Self - perception of overweight and obesity: A review of mental and physical health outcomes. *Obes. Sci. Pract.* 2020, *6*, 552–561.
31. Wright, F.L.; Green, J.; Reeves, G.; Beral, V.; Cairns, B.J. Validity over time of self-reported anthropometric variables during follow-up of a large cohort of UK women. *BMC Med. Res. Methodol.* 2015, *15*, 81.

Chapter 4 Association of Mediterranean diet with rheumatoid arthritis risk: Evidence from the UK Women's Cohort Study

Yuanyuan Dong¹, Darren C. Greenwood², Janet E. Cade¹

1 Nutritional Epidemiology Group, School of Food Science and Nutrition, University of Leeds, Leeds LS2 9JT, UK;

2 Leeds Institute for Data Analytics, Faculty of Medicine and Health, University of Leeds, Leeds, LS2 9JT, UK.

4.1 Abstract

Background: Uncertainty remains regarding the role of overall dietary patterns, particularly the Mediterranean diet, in the development of rheumatoid arthritis (RA). We investigated the association between Mediterranean diet adherence and RA risk in middle-aged UK women.

Methods: Dietary intake at baseline (1995–1998) was assessed via a validated 217-item food frequency questionnaire, and Mediterranean diet adherence was scored using a ten-point index. Incident RA cases were identified through linkage to NHS Hospital Episode Statistics up to March 2019. Multivariable-adjusted Cox regression models were used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) for RA incidence. Directed acyclic graphs informed covariate selection and helped visualise the structure of the composite dietary exposure.

Results: Over a median follow-up of 22.4 years among 27,766 women, 255 incident RA cases were identified. High adherence to the Mediterranean diet (score > 6) was associated with a lower risk of RA compared to low adherence (score < 4) (HR (95% CI) 0.67(0.45, 1.00)). No evidence of associations was observed between individual food components and RA risk.

Conclusions: Higher adherence to the Mediterranean diet was associated with reduced RA risk in this cohort of UK women. These findings support the potential role of overall dietary patterns in RA prevention. Further research to assess generalisability to men and non-European populations is recommended.

Trial registration: ClinicalTrials.gov, NCT06670144

Keywords: Mediterranean diet; nutrition; body mass index; rheumatoid arthritis; cohort

4.2 Introduction

Rheumatoid arthritis (RA) is one of the most common immune mediated inflammatory diseases characterised by persistent joint inflammation, pain, stiffness, and progressive functional decline [1]. It affects approximately 1% of the UK population [2], with women disproportionately affected (3:1 female-to-male ratio) [3]. Despite therapeutic advances [4], RA remains a lifelong condition with a substantial impact on quality of life and healthcare systems[5].

While genetic susceptibility is a major determinant and cigarette smoking is a well-established modifiable risk factor, few other lifestyle factors have been conclusively confirmed[6]. Diet is increasingly recognised as relevant to autoimmune disease risk, given its influence on immune regulation, metabolic pathways, and the gut microbiome [7-9]. The Mediterranean diet, a dietary pattern characterised by high consumption of vegetables, legumes, whole grains, fruits, fish, unsaturated fats (primarily from olive oil), and low intake of red and processed meat, is widely perceived as beneficial to health [10] and is associated with a reduced risk of cardiovascular disease[11], diabetes [12], and certain types of cancer [13, 14]. Its adherence is typically assessed using a composite score based on population-specific medians of ten dietary components [15], with health benefits thought to derive from closer alignment with this pattern. Notably, the Mediterranean diet is recommended for RA prevention by the American College of Rheumatology[16].

Despite biological plausibility, evidence linking the Mediterranean diet to RA risk is limited and inconclusive, due to the small number of studies and conflicting findings. Two Swedish case-control studies reported no evidence of association in women [17, 18], while a recent UK Biobank analysis did find a beneficial association with high adherence in a relatively short follow-up [19]. The French E3N cohort observed a modest risk reduction only among ever-smokers [20]. In contrast, the US-based Nurses' Health Studies (NHS I and II) reported no association [21]. To our knowledge, no large-scale UK prospective study has examined this relationship.

We drew on studies identified through our previous systematic review and dose-response meta-analysis of prospective studies of diet and RA [22], and updated to December 2025. We aimed to address this evidence gap by investigating the association between Mediterranean diet adherence and incident RA in the UK Women's Cohort Study (UKWCS), which is well-suited for this investigation given the higher incidence of RA in women, with over two decades of follow-up.

4.3 Subjects and methods

4.3.1 Study population

The UKWCS is a population-based prospective study of 35,325 women aged 35–69 years. It was designed to investigate the link between diet and cancer, and the women were recruited via postal questionnaires distributed by the World Cancer Research Fund (WCRF) across England, Scotland and Wales between 1995 and 1998. Baseline data included demographics, lifestyle, health-related information, anthropometrics, and dietary intake. Further details on the cohort design and recruitment procedures have been previously published [23].

We excluded participants who could not be linked to English Hospital Episode Statistics (HES) data ($n = 2314$ living outside England, $n = 695$ without NHS number), had prevalent RA at baseline ($n = 38$), had outlier body mass index values ($\text{BMI} < 10$ or $> 60 \text{ kg/m}^2$; $n = 1186$), or missing data on age ($n = 38$). The final analytical sample included 30,887 women (**Appendix C: Figure C1**).

Ethical approval was initially granted at inception in 1993 by the National Research Ethics Service (NRES) Committee Yorkshire & the Humber–Leeds East (Ref: 15/YH/0027), with participants providing implied consent by returning a baseline questionnaire (a form of written implied consent) between 01/01/1995 and 31/12/1998. Further approvals under Section 251 were obtained from the National Information Governance Board and the Health Research Authority Confidentiality Advisory Group (17/CAG/0081), permitting ongoing data retention and linkage. The latest amendment (REC Ref: 17/YH/0144; IRAS project ID: 213210) was approved on 28 March 2018.

4.3.2 Dietary assessment

Dietary intake at baseline were assessed using a validated 217-item food frequency questionnaire (FFQ), adapted from the European Prospective Investigation into Cancer and Nutrition (EPIC) study [24] and validated through comparison with 4-day weighed food diaries and a repeated FFQ administration among a subsample of 283 women in 2001 [25]. Missing responses were interpreted as non-consumers.

Mediterranean diet adherence was assessed using a modified 10-point score [15, 26], assigning one point for intake above the cohort-specific median for beneficial components (vegetables [excluding potatoes], legumes, fruit and nuts [excluding fruit juice], cereals, fish, and the ratio of monounsaturated and polyunsaturated to saturated fats ($MUFA + PUFA: SFA$)), and one point for intake below the median for less healthy components (meat, poultry, dairy). Moderate alcohol intake (5–25 g/day) scored one point. Total scores ranged from 0 to 10 and were categorised as low (≤ 3), moderate (4–6), and high (≥ 7) adherence. Full scoring details are provided in **Appendix C: Table C3 and Supplementary Methods**.

4.3.3 Case definition

The primary outcome was incident RA, identified using International Classification of Diseases (ICD-10) codes M05–M06 through the linkage to HES data. Follow-up time was calculated in person-years from baseline to the first event: RA diagnosis, death, or the study end (31 March 2019), whichever occurred first. Calendar date was used as the timescale.

4.3.4 Statistical methods

All statistical methods were pre-registered on [ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT06670144) (NCT06670144). Participant characteristics were summarised across Mediterranean diet adherence categories. Cox proportional hazards regression models were used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) for the association between Mediterranean diet adherence and RA incidence.

In analyses categorising Mediterranean diet adherence into low, moderate, and high levels, the low-adherence group served as the reference. We also modelled

adherence as a continuous variable, estimating the HR per 2-point increase in the score to assess linear associations. Potential non-linear associations were explored using multivariable restricted cubic spline (RCS) models with four knots placed at the 5th, 35th, 65th, and 95th percentiles of the Mediterranean diet score[27], following Harrell's recommendations for distributions with low skewness [28]. In order to account for the stratified sampling at recruitment, which over-sampled vegetarians and fish-eaters, Cox models incorporated inverse probability weights to generate estimates representative of the UK population [29].

We fitted two models: an age-adjusted model and a fully-adjusted model. Both controlled for age at baseline (continuous). Covariates in the multivariable model were selected based on prior knowledge and a directed acyclic graph (DAG), incorporating recent guidance on deterministic and derived variables [30] shown in **Figure 4.1**. The Mediterranean diet score was treated as a deterministic composite of dietary intake variables. The fully adjusted model included the following covariates: education level ('O level', 'A level', 'degree'); BMI (continuous); socio-economic status (SES, professional/managerial, intermediate, routine/manual); marital status (married/living as married, separated/divorced, single/widowed); smoking status (current, former, never); energy intake (continuous); physical activity (continuous), which was derived from metabolic equivalents (METs) [31], were calculated from baseline questionnaire data; alcohol consumption (continuous); menopausal status (premenopausal, postmenopausal); Hormone replacement therapy (HRT, yes or no); prevalence of cardiovascular disease, cancer, or diabetes at baseline (yes or no); nutritional supplement use (yes or no). Covariates at recruitment and their derivation were described in **Appendix C: Table C2**. The proportional hazards assumptions were checked using Schoenfeld residuals.

To assess potential effect modification, we conducted exploratory subgroup analyses by including interaction terms between Mediterranean diet adherence (modelled both categorically and continuously) and the following five key variables in the fully adjusted model: BMI (<25.0 vs. ≥25.0 kg/m²), smoking status (non-smoker, ex-smoker, or current smoker), age group (<55 vs. ≥55 years), physical activity (≤10 vs. >10 MET-hours/week), and menopausal status

(premenopausal vs. postmenopausal). In each subgroup analysis, the corresponding effect modifier was omitted from the adjustment set.

We conducted sensitivity analysis to assess the robustness of findings: (1) excluding participants with a survival time < 5 years to check for reverse causation; (2) excluding participants who developed other chronic diseases (e.g. CVD, cancer, or diabetes) to reduce the influence of preclinical disease on dietary habits; (3) excluding current smokers, given its known effect on RA risk[32]; (4) additionally adjusting the model for vitamin D intake, due to its potential immunomodulatory role and possible confounding effect on the relationship between diet and RA; (5) excluding participants who reported no use of nutrient supplements to test whether supplement use may modify dietary reporting or nutrient exposure; (6) applying a competing risk regression model using the Fine-Gray method[33], considering death as a competing event. All analyses were conducted by Stata version 18, and R version 4.3.1 (R Foundation). A two-sided P-value < 0.05 indicated statistical significance.

4.4 Results

4.4.1 Participants

Of the 30,877 women eligible for inclusion, those lacking data on the following essential covariates were excluded from multi-variable models requiring complete covariate data: ethnicity (n = 786), physical activity (n = 561), marital status (n = 237), socio-economic status (n = 221), menopausal status (n = 546), and smoking status (n = 770). A total of 26,766 cohort participants were included in the final analysis, among whom 255 developed incident RA. The study flow chart is presented in **Appendix C: Figure C1**.

4.4.2 Descriptive data

Among 26,766 cohort participants with complete data, the mean age at baseline was 51.9 years (SD 9.2). **Table 4.1** presents participant characteristics at baseline across the Mediterranean diet adherence categories. During a median follow-up of 22.4 years (620,634 person-years), 255 participants (1.0%) developed RA. Women with higher adherence category were generally younger,

had lower BMI, higher physical activity and energy intake, were more likely to hold a degree education and professional/managerial jobs, and were less likely to smoke, use HRT, or have chronic disease. They were also more likely to be premenopausal. In contrast, those with the lowest Mediterranean diet adherence tended to smoke, have higher BMI, hold routine/manual jobs, be postmenopausal, use HRT, and have a higher prevalence of chronic conditions, while being less likely to consume alcohol, exercise, have higher energy intake, or hold a university degree. Food and nutrient intakes by adherence category are detailed in **Appendix C: Table C4**.

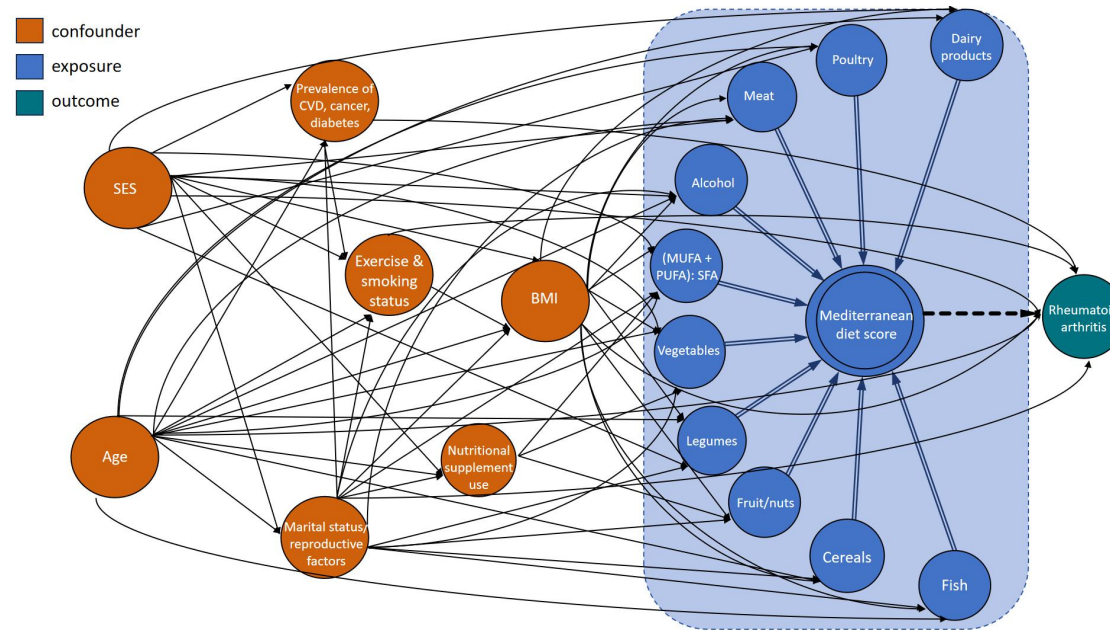


Figure 4.1 Directed Acyclic Graph illustrating the hypothesised relationship between the Mediterranean diet score, rheumatoid arthritis incidence, and related factors. The exposure (Mediterranean diet score, a composite derived variable based on 10 food components) is depicted by the blue circle and the outcome (rheumatoid arthritis incidence) by the green circle. Variables represented as orange circles are considered common causes (confounders) of the exposure and outcome. These include socio-demographic factors (age, socioeconomic status, educational level, marital status), reproductive factors (menopausal status, hormone replacement therapy), and established or potential risk factors (smoking status, physical activity, nutritional supplements use, and body mass index). SES: socioeconomic status, BMI: body mass index, CVD: cardiovascular diseases., MUFA: monounsaturated fatty acids, PUFA: polyunsaturated fatty acids, SFA: saturated fatty acids.

Table 4.1 Baseline characteristics of 27,766 UKWCS participants.

Characteristics, mean (SD) or n (%)	Diet group			Total
	Low adherence (≤3 points)	Moderate adherence (4-6 points)	High adherence (≥7 points)	
Participants (%)	7940 (28.6)	14,069 (50.7)	5757 (20.7)	27,766
Age, years				
Mean (SD)	53.2 (9.3)	52.0 (9.2)	50.0 (8.7)	51.9 (9.2)
Range (Min, Max)	(41, 67)	(35, 65)	(45, 63)	(35, 69)
Age at diagnosis, years				
Mean (SD)	66.4 (9.5)	65.0 (9.0)	62.9 (8.5)	65.0 (10.0)
Range (Min, Max)	(40, 86)	(38, 85)	(38, 84)	(38, 86)
Socio-demographics				
Degree-level education (%)	1513 (21.1)	3482 (26.8)	1953 (35.9)	6948 (27.1)
Socio-economic status (%)				
Professional or managerial	4539 (57.2)	8943 (63.6)	4081 (70.9)	17,563 (63.3)
Intermediate	2538 (32.0)	3886 (27.6)	1293 (22.5)	7717 (27.8)
Routine or manual	863 (10.9)	1240 (8.8)	383 (6.7)	2486 (9.0)
Married (%)	6152 (77.5)	10,737 (76.3)	4402 (76.5)	21,291 (76.7)
Anthropometric indices				
Height, cm				
Mean (SD)	164 (6.9)	164 (6.7)	164 (6.6)	164(6.6)
Range (Min, Max)	(144, 183)	(145, 182)	(146, 182)	(145, 182)
Weight, kg				
Mean (SD)	67 (12.4)	66 (11.8)	63 (10.5)	65 (10.5)
Range (Min, Max)	(44, 120)	(42, 115)	(40, 105)	(42, 120)
BMI at recruitment, kg/m ²				
Mean (SD)	25.1 (4.5)	24.5 (4.2)	23.5 (3.7)	24.4 (4.2)
Range (Min, Max)	(16.4, 41.1)	(14.5, 40.0)	(16.0, 40.2)	(14.5, 41.1)
Lifestyle				
Physical activity (METs), hours/week				
Mean (SD)	16.7 (11.6)	15.7 (11.3)	16.8 (11.3)	18.1 (12.3)
Range (Min, Max)	(0, 72.5)	(0, 70.0)	(0, 65.5)	(0, 72.5)

Smoking status (%)				
Current	1103 (13.9)	1409 (10.0)	533 (9.3)	3045 (11.0)
Former	2144 (27.0)	4480 (31.8)	2049 (35.6)	8673 (31.2)
Never	4693 (59.1)	8180 (58.1)	3175 (55.2)	3175 (55.2)
Alcohol consumption, grams/week				
Mean (SD)	9.0 (12.9)	10.2 (11.8)	10.5 (10.0)	9.9 (11.8)
Range (Min, Max)	(0, 170)	(0, 210)	(0, 224)	(0, 224)
Total energy intake, kcal/day				
Mean (SD)	2121.1 (599)	2336.8 (684)	2569.1 (670)	2323.3 (676)
Range (Min, Max)	(652, 4378)	(732, 4945)	(815, 4901)	(652, 4945)
Other				
Menopausal status (%)				
Premenopausal	3394 (42.8)	6783 (48.2)	3292 (57.2)	13,469 (48.5)
Postmenopausal	4546 (57.3)	7286 (51.8)	2465 (42.8)	14,297 (51.5)
Hormone replacement therapy (%)	2435 (30.7)	4029 (28.6)	1442 (25.1)	7906 (28.5)
Nutrient supplements use (%)	3683 (50.5)	7587 (59.0)	3447 (66.2)	14717 (58.1)
Prevalence of CVD, cancer, or diabetes at baseline (%)	598 (7.5)	980 (7.0)	329 (5.7)	1907 (6.9)

SD standard deviation, BMI body mass index, METs Metabolic Equivalent of Tasks, CVD cardiovascular disease. Diet adherence groups were defined based on Mediterranean diet score: Low (≤ 3 points), Moderate (4–6 points), and High (≥ 7 points).

4.4.3 Adherence to the Mediterranean diet

We assessed RA risk according to levels of adherence to the Mediterranean diet. After adjusting for age, participants with high adherence had a reduced risk of RA (HR=0.61 (0.42, 0.90)). This association remained after multivariate adjustment (HR=0.67 (0.45, 1.00)). However, in the continuous Mediterranean diet score, a 2-point increase was not statistically associated with RA risk after multivariable adjustment (HR=0.90 (0.79, 1.03); **Table 4.2**). No evidence of associations was observed between individual dietary components and RA incidence (**Appendix C: Table C5**).

The restricted cubic spline analysis was used to visually explore the dose-response relationship. While the figure suggested a downward slope with increasing adherence, there is no evidence for a linear trend ($P = 0.4$) or non-

linearity ($P = 0.7$; **Figure 4.2**). The primary evidence for a protective association is derived from the categorical comparison of high versus low adherence (**Table 4.2**).

Table 4.2 Association between Mediterranean diet adherence and risk of rheumatoid arthritis in the UK Women's Cohort Study.

Mediterranean diet	Cases/participants	Person-years	Age-adjusted model HR (95%CI)	Multivariable-adjusted model HR (95%CI) *
Low adherence	83/7940	175,537	1 (reference)	1 (reference)
Moderate adherence	134/14,069	314,489	0.90(0.68, 1.18)	0.94(0.71, 1.25)
High adherence	38/5757	130,707	0.61(0.42, 0.90)	0.67(0.45, 1.00)
Per 2-point increment	255/27,766	620,634	0.88(0.78, 0.99)	0.90(0.79, 1.03)

* Adjusted for age, education, socio-economic status, marital status, BMI, smoking status, energy intake, prevalence of cardiovascular disease, cancer, or diabetes, menopausal status, and hormone replacement therapy use, nutritional supplement use. HR, hazard ratio; CI, confidence interval.

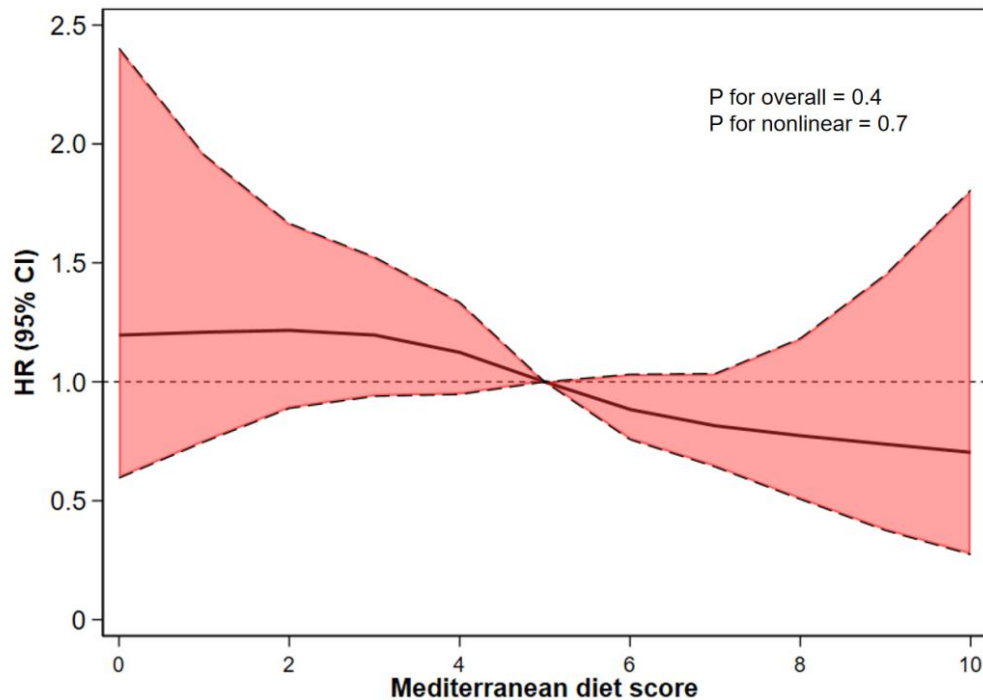


Figure 4.2 Restricted cubic spline for the association between rheumatoid arthritis incidence and the Mediterranean diet score. HR was estimated using a Cox proportional hazards model, adjusted for age, education, socio-economic status, marital status, BMI, smoking status, energy intake, prevalence of cardiovascular disease, cancer, diabetes, menopausal status, hormone replacement therapy use, and nutritional supplements use. HR, hazard ratio; CI, confidence interval.

4.4.4 Subgroup analysis

Although the inverse association between high adherence to the Mediterranean diet and RA risk appeared stronger in certain subgroups, no clear evidence of effect modification for any of the stratifying variables was seen (all P for interaction > 0.4 ; **Appendix C: Table C6**). For example, the protective effect of high adherence to the Mediterranean diet on outcome was stronger among women aged ≤ 55 years ($HR=0.57$ (0.34, 0.95)) than among those aged >55 years ($HR=0.78$ (0.43, 1.42); P -interaction = 0.7). Similarly, high adherence to the Mediterranean diet was associated with a reduced RA risk among women who exercised adequately (>10 MET-hrs/week) ($HR=0.58$ (0.37, 0.91)) compared to those who were less active ($HR=0.67$ (0.31, 1.46); P -interaction = 0.8). By BMI, high adherence was only associated with a lower RA risk among normal weight

women (BMI ≤ 25.0 kg/m²) (HR=0.62 (0.38, 1.00)), but not among women who were overweight or obese (HR=0.67 (0.35, 1.28); P-interaction = 0.9). No variation in the associations was observed by menopausal or smoking status.

4.4.5 Sensitivity analysis

Appendix C: Table C7 showed that the relationship between high adherence to the Mediterranean diet and lower RA risk remained stable across most sensitivity analyses, including the exclusion of participants with less than 5 years of follow-up or chronic diseases. Competing risk analysis using the modified Fine and Gray method yielded consistent results (HR=0.67(0.44,0.99)), reinforcing the primary findings. Additional adjustments for overall dietary supplement use slightly strengthened the association (HR=0.61 (0.40, 0.95)), while adjustment for dietary vitamin D intake had no appreciable impact (HR = 0.67 (0.45, 1.00)). However, the association was not robust to the exclusion of current smokers (HR=0.61(0.34,1.09)).

4.5 Discussion

4.5.1 Main findings

In this large prospective cohort of UK women, greater adherence to the Mediterranean diet was associated with a lower risk of developing RA, independent of established confounders. This inverse association persisted after accounting for baseline health status, supplement use, and follow-up duration. The association weakened after excluding current smokers, suggesting a possible interaction with smoking. No effect modification was observed by BMI. The absence of strong associations with individual food components underscores the importance of overall dietary patterns in RA prevention.

4.5.2 Comparison with previous studies

This study adds to the evidence linking greater adherence to the Mediterranean diet with a reduced RA risk. Our findings align with a UK Biobank study, which reported a 29% lower RA risk among women with high Mediterranean diet adherence [19], comparable to the 33% risk reduction observed in our cohort. However, differences in scoring systems may partly explain variations in results.

The Biobank study used the MEDI-LITE score (range 0–18), which allocates points based on absolute intake with finer gradation [19], while our study used the alternative Trichopoulou method (range 0–10) [15], based on population-specific medians. The binary nature of the latter may reduce sensitivity to detect weaker associations. Neither study observed benefit at moderate adherence levels. This indicates that the stronger association may be limited to higher adherence.

Although our continuous model showed a 10% risk reduction per 2-point increment in diet score, but no strong evidence for this association. Several factors might explain this. Continuous scores may be more vulnerable to random measurement error in self-reported dietary data, whereas categorical comparisons (e.g. high vs. low adherence) offer greater contrast and are more robust to noise data. Furthermore, high adherence likely reflects a broader set of healthy behaviours, such as nonsmoking, regular exercise, and lower adiposity, not fully captured by the score itself, contributing to the stronger associations in categorical analyses.

A growing body of evidence supports the role of individual dietary components in RA development, with red and processed meat identified as a potential risk factors and plant-based foods as protective[22, 34, 35]. A recent systematic review, synthesising data from prospective studies and case-control studies, concluded that high consumption of red and processed meat may increase RA risk, while higher intakes of vegetables and fruits appear beneficial [36]. For example, a large cohort study from the Women's Health Initiative (8 years of follow-up) [35] reported a protective effect of fruit and vegetables and increased risk of red meat intake. This cohort had shorter follow-up periods than the UKWCS, indicating that the follow-up was shorter makes UKWCS stand out as better in that regard. A Swedish nested case-control study further demonstrated that adherence to overall dietary guidelines, which inherently encourage higher plant food consumption and limit meat intake, was associated with reduced RA risk [37], aligning with our findings on the Mediterranean diet pattern. The biological plausibility of these associations is supported by proposed mechanisms, including the pro-inflammatory effects of advanced glycation end-products and haem iron in meat [38], and the anti-inflammatory effects of dietary

fibre, antioxidants, and bioactive components from plants [39]. Taken together, these studies support the hypothesis that a dietary pattern characterised by a high intake of plant foods and a low intake of meat may protect against RA, reinforcing the observed inverse association between Mediterranean diet adherence and RA risk.

Our results also align with findings from the EPIC-E3N cohort, where a reduced RA risk was observed among ever-smokers. In contrast, no evidence of an association was found in the NHS I and II [21] studies, potentially due to methodological differences (e.g., exclusion of dairy in the score) and limited behavioural variation within a homogenous population of US nurses. Compared to these studies, our cohort was younger, and we observed stronger associations among younger women, possibly reflecting the importance of early-life exposures in autoimmune disease development.

As a contemporary mainstream dietary pattern, small-scale clinical trials have explored the Mediterranean diet's impact on RA disease activity and quality of life, reporting improvements in disease activity score and physical function [40, 41]. However, these studies often suffer from limitations, such as short follow-up periods, inevitable lack of rigorous blinding in the design, and variability in dietary compliance [42]. Well-designed, adequately powered trials involving standardised dietary interventions and long-term follow-up are required to establish causality and guide evidence-based dietary recommendations for prevention and management.

4.5.3 Interpretation and implications

Women with high Mediterranean diet adherence tended to lead healthier lifestyles [43], which could partly explain the observed association with reduced RA risk. Whilst we found no clear effect modification by BMI, the association was more apparent in women with lower BMI. The attenuated association among overweight/obese women may reflect greater measurement error in this subgroup, as well as obesity-related inflammation that could dilute the protective effect of Mediterranean diet adherence. This supports previous evidence linking adiposity to systemic inflammation and immune dysregulation through cytokine secretion such as TNF- α , IL-6, and CRP [44]. Recent findings from randomised

controlled trials have emphasised the effectiveness of a personalised Mediterranean diet combined with physical activity as a practical strategy for improving disease management in patients with RA[45]. Given the interactive nature of diet, body weight, and health behaviours, it is likely that the Mediterranean diet exerts a direct protective effect or operates through broader lifestyle patterns.

Smoking is a major established risk factor for RA and may confound or interact with dietary patterns in RA risk. The attenuation of associations after excluding current smokers suggests that smoking could influence the anti-inflammatory effects of diet. Some experimental studies have proposed mechanisms by which smoking alters gut microbiota [46] or modulates immune responses [47]. Notably, evidence suggests that nicotine may have anti-inflammatory effects in arthritis [48], though this line of evidence is still limited and controversial.

Our analysis is the first to demonstrate restricted cubic splines to explore potential nonlinear dose–response associations in the UK adult female population. Restricted cubic splines is a flexible and widely used method in medical research that not only accommodates nonlinearity but also visualises performance to capture the variation of the effect of a change. There was no strong evidence of nonlinearity. Nonetheless, the greatest risk reductions were observed among women with the highest adherence, supporting the benefit of comprehensive, sustained dietary patterns for RA prevention, through synergistic biological mechanisms such as modulation of gut microbiome, suppression of inflammatory cytokines, and improved metabolic regulation [49-51].

Although the Mediterranean diet score reflects overall dietary habits, similar scores can result from different combinations of foods, making interpretation complicated [52]. In our analysis, no individual food components were independently associated with RA risk, likely due to confounding structures that introduce tautological associations when 10 food components are analysed in isolation. As illustrated by the DAG, the deterministic nature of the diet score constrains variation across components, potentially leading to the fixed-effect estimates that obscure the total dietary pattern effect. Nonetheless, participants with higher adherence consumed more protein, fibre, folate, calcium, and vitamin

C, nutrients implicated in immune regulation [53]. Vitamin D intake did not vary across adherence levels, and additional adjustment did not affect the observed association. While a meta-analysis of US cohorts reported a modest reduction in RA risk with vitamin D supplementation[22], causality remains unclear, and our study could not evaluate supplement intake due to non-specific data. This supports the view that dietary patterns, rather than single foods, better capture the synergistic effects of diet on outcomes. For example, the Empirical Dietary Inflammatory Pattern, characterised by anti- and pro-inflammatory food components, has been associated with an increased risk of seropositive RA [54].

4.5.4 Strengths and limitations

This study has several limitations that should be considered. First, diet and lifestyle were collected only once at baseline. The long follow-up period might lead to “regression dilution bias” [55], as participants’ eating habits could change over time and the protective effect of the Mediterranean diet is underestimated. However, the sensitivity analysis excluding the first 5 years of follow-up showed stable results, and there was no evidence of non-proportional hazards, demonstrating that the association was not diluted over time. Second, while the Mediterranean diet score offers a useful summary of dietary patterns, it assumes equal and interchangeable effects of all components, which may oversimplify complex biological pathways. As mentioned, the use of population-specific medians to generate a relative adherence score may reduce its applicability to public health interventions or dietary guidelines based on absolute intake. Furthermore, interpreting a 2-point increase in the score assumes equal effects across components, which may not reflect biological reality and limits interpretability, although no individual components were independently linked with RA risk in this analysis.

The relatively small number of incident RA cases constrained the power to perform detailed subgroup analyses, such as stratification by seropositive and seronegative RA, which may have distinct etiologies and risk profiles. Moreover, the demographic homogeneity (predominantly white women) may constrain generalisability to other populations. Furthermore, we lacked data on specific occupational exposures, which may confound the relationship. We also did not

have genetic data to account for underlying genetic susceptibility via polygenic risk scores. While the extent to which these factors are strong confounders in the relationship between Mediterranean diet and RA risk remains uncertain. Future studies should explore dietary effects in men and ethnically diverse groups and consider alternative scoring approaches that incorporate absolute intake and biomarker-based validation.

Main strengths of this study include its prospective design, long follow-up, and robust RA case ascertainment via linked hospital records, reducing recall bias and loss to follow-up. The use of a validated FFQ allowed for reasonably accurate dietary classification, though measurement error is an inherent limitation. Exclusion of participants with pre-existing RA supports a temporal relationship between diet and disease onset. Furthermore, the use of a competing risk model accounting for mortality in sensitivity analyses strengthens the reliability of this observed association.

4.6 Conclusions

This study supports the potential benefits of close adherence to the Mediterranean diet, a composite score reflecting the synergistic effects of an overall dietary pattern, in this cohort of UK women. Further research is needed to confirm these findings in broader populations and to elucidate the mechanism responsible.

4.7 References

1. England BR, Thiele GM, Anderson DR, Mikuls TR: Increased cardiovascular risk in rheumatoid arthritis: mechanisms and implications. *Bmj* 2018, 361:k1036.
2. Ashai S, Harvey NC: Rheumatoid arthritis and bone health. *Clin Med (Lond)* 2020, 20(6):565-567.
3. van Vollenhoven RF: Sex differences in rheumatoid arthritis: more than meets the eye. *BMC Med* 2009, 7:12.
4. Guo Q, Wang Y, Xu D, Nossent J, Pavlos NJ, Xu J: Rheumatoid arthritis: pathological mechanisms and modern pharmacologic therapies. *Bone Res* 2018, 6:15.

5. Global, regional, and national burden of gout, 1990-2020, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. *Lancet Rheumatol* 2024, 6(8):e507-e517.
6. Maisha JA, El-Gabalawy HS, O'Neil LJ: Modifiable risk factors linked to the development of rheumatoid arthritis: evidence, immunological mechanisms and prevention. *Front Immunol* 2023, 14:1221125.
7. Bonaccio M, Costanzo S, Di Castelnuovo A, Gialluisi A, Ruggiero E, De Curtis A, Persichillo M, Cerletti C, Donati MB, de Gaetano G, Iacoviello L: Increased Adherence to a Mediterranean Diet Is Associated With Reduced Low-Grade Inflammation after a 12.7-Year Period: Results From the Moli-sani Study. *J Acad Nutr Diet* 2023, 123(5):783-795.e787.
8. Merra G, Noce A, Marrone G, Cintoni M, Tarsitano MG, Capacci A, De Lorenzo A: Influence of Mediterranean Diet on Human Gut Microbiota. *Nutrients* 2020, 13(1).
9. Li J, Guasch-Ferré M, Chung W, Ruiz-Canela M, Toledo E, Corella D, Bhupathiraju SN, Tobias DK, Tabung FK, Hu J, Zhao T, Turman C, Feng YA, Clish CB, Mucci L, Eliassen AH, Costenbader KH, Karlson EW, Wolpin BM, Ascherio A, Rimm EB, Manson JE, Qi L, Martínez-González M, Salas-Salvadó J, Hu FB, Liang L: The Mediterranean diet, plasma metabolome, and cardiovascular disease risk. *Eur Heart J* 2020, 41(28):2645-2656.
10. Dinu M, Pagliai G, Casini A, Sofi F: Mediterranean diet and multiple health outcomes: an umbrella review of meta-analyses of observational studies and randomised trials. *Eur J Clin Nutr* 2018, 72(1):30-43.
11. Martínez-González MA, Gea A, Ruiz-Canela M: The Mediterranean Diet and Cardiovascular Health. *Circ Res* 2019, 124(5):779-798.
12. Schwingshackl L, Missbach B, König J, Hoffmann G: Adherence to a Mediterranean diet and risk of diabetes: a systematic review and meta-analysis. *Public Health Nutr* 2015, 18(7):1292-1299.
13. Bai X, Li X, Ding S, Dai D: Adherence to the Mediterranean Diet and Risk of Gastric Cancer: A Systematic Review and Meta-Analysis. *Nutrients* 2023, 15(17).
14. González-Palacios Torres C, Barrios-Rodríguez R, Muñoz-Bravo C, Toledo E, Dierssen T, Jiménez-Moleón JJ: Mediterranean diet and risk of breast cancer: An umbrella review. *Clin Nutr* 2023, 42(4):600-608.

15. Trichopoulou A, Bamia C, Norat T, Overvad K, Schmidt EB, Tjønneland A, Halkjaer J, Clavel-Chapelon F, Vercambre MN, Boutron-Ruault MC, Linseisen J, Rohrmann S, Boeing H, Weikert C, Benetou V, Psaltopoulou T, Orfanos P, Boffetta P, Masala G, Pala V, Panico S, Tumino R, Sacerdote C, Bueno-de-Mesquita HB, Ocke MC, Peeters PH, Van der Schouw YT, González C, Sanchez MJ, Chirlaque MD, Moreno C, Larrañaga N, Van Guelpen B, Jansson JH, Bingham S, Khaw KT, Spencer EA, Key T, Riboli E, Trichopoulos D: Modified Mediterranean diet and survival after myocardial infarction: the EPIC-Elderly study. *Eur J Epidemiol* 2007, 22(12):871-881.
16. England BR, Smith BJ, Baker NA, Barton JL, Oatis CA, Guyatt G, Anandarajah A, Carandang K, Chan KK, Constien D, Davidson E, Dodge CV, Bemis-Dougherty A, Everett S, Fisher N, Fraenkel L, Goodman SM, Lewis J, Menzies V, Moreland LW, Navarro-Millan I, Patterson S, Phillips LR, Shah N, Singh N, White D, AlHeresh R, Barbour KE, Bye T, Guglielmo D, Haberman R, Johnson T, Kleiner A, Lane CY, Li LC, Master H, Pinto D, Poole JL, Steinbarger K, Sztubinski D, Thoma L, Tsaltzkan V, Turgunbaev M, Wells C, Turner AS, Treadwell JR: 2022 American College of Rheumatology Guideline for Exercise, Rehabilitation, Diet, and Additional Integrative Interventions for Rheumatoid Arthritis. *Arthritis Rheumatol* 2023, 75(8):1299-1311.
17. Johansson K, Askling J, Alfredsson L, Di Giuseppe D: Mediterranean diet and risk of rheumatoid arthritis: a population-based case-control study. *Arthritis Res Ther* 2018, 20(1):175.
18. Sundström B, Johansson I, Rantapää-Dahlqvist S: Diet and alcohol as risk factors for rheumatoid arthritis: a nested case-control study. *Rheumatol Int* 2015, 35(3):533-539.
19. Hu P, Lee EK, Li Q, Tam LS, Wong SY, Poon PK, Yip BH: Mediterranean diet and rheumatoid arthritis: A nine-year cohort study and systematic review with meta-analysis. *Eur J Clin Nutr* 2025.
20. Nguyen Y, Salliot C, Gelot A, Gambaretti J, Mariette X, Boutron-Ruault MC, Seror R: Mediterranean Diet and Risk of Rheumatoid Arthritis: Findings From the French E3N-EPIC Cohort Study. *Arthritis Rheumatol* 2021, 73(1):69-77.

21. Hu Y, Costenbader KH, Gao X, Hu FB, Karlson EW, Lu B: Mediterranean diet and incidence of rheumatoid arthritis in women. *Arthritis Care Res (Hoboken)* 2015, 67(5):597-606.
22. Dong Y, Greenwood DC, Webster J, Uzokwe C, Tao J, Hardie LJ, Cade JE: Dose-Response Associations Between Diet and Risk of Rheumatoid Arthritis: A Meta-Analysis of Prospective Cohort Studies. *Nutrients* 2024, 16(23).
23. Cade JE, Burley VJ, Alwan NA, Hutchinson J, Hancock N, Morris MA, Threapleton DE, Greenwood DC: Cohort Profile: The UK Women's Cohort Study (UKWCS). *Int J Epidemiol* 2017, 46(2):e11.
24. Riboli E, Kaaks R: The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* 1997, 26 Suppl 1:S6-14.
25. Cade JE, Burley VJ, Greenwood DC: The UK Women's Cohort Study: comparison of vegetarians, fish-eaters and meat-eaters. *Public Health Nutr* 2004, 7(7):871-878.
26. Trichopoulou A, Kouris-Blazos A, Wahlqvist ML, Gnardellis C, Lagiou P, Polychronopoulos E, Vassilakou T, Lipworth L, Trichopoulos D: Diet and overall survival in elderly people. *Bmj* 1995, 311(7018):1457-1460.
27. Spencer SM, Trower AJ, Jia X, Scott DJA, Greenwood DC: Meta-analysis of the association between alcohol consumption and abdominal aortic aneurysm. *Br J Surg* 2017, 104(13):1756-1764.
28. Schuster NA, Rijnhart JJM, Twisk JWR, Heymans MW: Modeling non-linear relationships in epidemiological data: The application and interpretation of spline models. *Front Epidemiol* 2022, 2:975380.
29. Rada-Fernandez de Jauregui D, Evans CEL, Jones P, Greenwood DC, Hancock N, Cade JE: Common dietary patterns and risk of cancers of the colon and rectum: Analysis from the United Kingdom Women's Cohort Study (UKWCS). *Int J Cancer* 2018, 143(4):773-781.
30. Arnold KF, Berrie L, Tennant PWG, Gilthorpe MS: A causal inference perspective on the analysis of compositional data. *Int J Epidemiol* 2020, 49(4):1307-1313.

31. Threapleton DE, Greenwood DC, Burley VJ, Aldwairji M, Cade JE: Dietary fibre and cardiovascular disease mortality in the UK Women's Cohort Study. *Eur J Epidemiol* 2013, 28(4):335-346.
32. Källberg H, Ding B, Padyukov L, Bengtsson C, Rönnelid J, Klareskog L, Alfredsson L: Smoking is a major preventable risk factor for rheumatoid arthritis: estimations of risks after various exposures to cigarette smoke. *Ann Rheum Dis* 2011, 70(3):508-511.
33. Austin PC, Fine JP: Practical recommendations for reporting Fine-Gray model analyses for competing risk data. *Stat Med* 2017, 36(27):4391-4400.
34. Jin J, Li J, Gan Y, Liu J, Zhao X, Chen J, Zhang R, Zhong Y, Chen X, Wu L, Xiang X, Zhou Y, He J, Guo J, Liu X, Li Z: Red meat intake is associated with early onset of rheumatoid arthritis: a cross-sectional study. *Sci Rep* 2021, 11(1):5681.
35. Krok-Schoen JL, Brasky TM, Hunt RP, Rohan TE, Baker TA, Li W, Carbone L, Mackey RH, Snetselaar L, Lustberg MB, Neuhauser ML: Dietary Long-Chain n-3 Fatty Acid Intake and Arthritis Risk in the Women's Health Initiative. *J Acad Nutr Diet* 2018, 118(11):2057-2069.
36. Bäcklund R, Drake I, Bergström U, Compagno M, Sonestedt E, Turesson C: Diet and the risk of rheumatoid arthritis - A systematic literature review. *Semin Arthritis Rheum* 2023, 58:152118.
37. Bäcklund RT, Drake I, Bergström U, Compagno M, Sonestedt E, Turesson C: Adherence to dietary guidelines, and the risk of developing rheumatoid arthritis: results from a nested case-control study. *Rheumatology (Oxford)* 2024, 63(2):407-413.
38. Demirer B, Fisunoğlu M: Evaluation of the effects of dietary advanced glycation end products on inflammation. *Nutr Bull* 2024, 49(1):6-18.
39. Sutcliffe JT, Wilson LD, de Heer HD, Foster RL, Carnot MJ: C-reactive protein response to a vegan lifestyle intervention. *Complement Ther Med* 2015, 23(1):32-37.
40. Forsyth C, Kouvari M, D'Cunha NM, Georgousopoulou EN, Panagiotakos DB, Mellor DD, Kellett J, Naumovski N: The effects of the Mediterranean diet on rheumatoid arthritis prevention and treatment: a systematic review of human prospective studies. *Rheumatol Int* 2018, 38(5):737-747.

41. Raad T, George E, Griffin A, Larkin L, Fraser A, Kennedy N, Tierney A: Effects of a telehealth-delivered Mediterranean diet intervention in adults with Rheumatoid Arthritis (MEDRA): a randomised controlled trial. *BMC Musculoskelet Disord* 2024, 25(1):631.
42. Kjeldsen-Kragh J: Mediterranean diet intervention in rheumatoid arthritis. *Ann Rheum Dis* 2003, 62(3):193-195.
43. Van den Brandt PAJTajocn: The impact of a Mediterranean diet and healthy lifestyle on premature mortality in men and women—. 2011, 94(3):913-920.
44. Ouchi N, Parker JL, Lugus JJ, Walsh K: Adipokines in inflammation and metabolic disease. *Nat Rev Immunol* 2011, 11(2):85-97.
45. Papandreou P, Gioxari A, Daskalou E, Grammatikopoulou MG, Skouroliahou M, Bogdanos DP: Mediterranean Diet and Physical Activity Nudges versus Usual Care in Women with Rheumatoid Arthritis: Results from the MADEIRA Randomized Controlled Trial. *Nutrients* 2023, 15(3).
46. Gui X, Yang Z, Li MD: Effect of Cigarette Smoke on Gut Microbiota: State of Knowledge. *Front Physiol* 2021, 12:673341.
47. Gui X, Yang Z, Li MDJFiP: Effect of cigarette smoke on gut microbiota: state of knowledge. 2021, 12:673341.
48. Zhang W, Lin H, Zou M, Yuan Q, Huang Z, Pan X, Zhang W: Nicotine in Inflammatory Diseases: Anti-Inflammatory and Pro-Inflammatory Effects. *Front Immunol* 2022, 13:826889.
49. Meslier V, Laiola M, Roager HM, De Filippis F, Roume H, Quinquis B, Giacco R, Mennella I, Ferracane R, Pons N, Pasolli E, Rivellese A, Dragsted LO, Vitaglione P, Ehrlich SD, Ercolini D: Mediterranean diet intervention in overweight and obese subjects lowers plasma cholesterol and causes changes in the gut microbiome and metabolome independently of energy intake. *Gut* 2020, 69(7):1258-1268.
50. Lozada-Mellado M, Llorente L, Hinojosa-Azaola A, García-Morales JM, Ogata-Medel M, Alcocer-Varela J, Pineda-Juárez JA, Castillo-Martínez L: Comparison of the Impacts of a Dynamic Exercise Program vs. a Mediterranean Diet on Serum Cytokine Concentrations in Women With Rheumatoid Arthritis. A Secondary Analysis of a Randomized Clinical Trial. *Front Nutr* 2022, 9:834824.

51. Hoffmann A, Meir AY, Hagemann T, Czechowski P, Müller L, Engelmann B, Haange SB, Rolle-Kampczyk U, Tsaban G, Zelicha H, Rinott E, Kaplan A, Shelef I, Stumvoll M, Blüher M, Liang L, Ceglarek U, Isermann B, von Bergen M, Kovacs P, Keller M, Shai I: A polyphenol-rich green Mediterranean diet enhances epigenetic regulatory potential: the DIRECT PLUS randomized controlled trial. *Metabolism* 2023, 145:155594.
52. Berrie L, Arnold KF, Tomova GD, Gilthorpe MS, Tennant PWG: Depicting deterministic variables within directed acyclic graphs: an aid for identifying and interpreting causal effects involving derived variables and compositional data. *Am J Epidemiol* 2025, 194(2):469-479.
53. Stone J, Doube A, Dudson D, Wallace J: Inadequate calcium, folic acid, vitamin E, zinc, and selenium intake in rheumatoid arthritis patients: results of a dietary survey. *Semin Arthritis Rheum* 1997, 27(3):180-185.
54. Sparks JA, Barbhaiya M, Tedeschi SK, Leatherwood CL, Tabung FK, Speyer CB, Malspeis S, Costenbader KH, Karlson EW, Lu B: Inflammatory dietary pattern and risk of developing rheumatoid arthritis in women. *Clin Rheumatol* 2019, 38(1):243-250.
55. Clarke R, Shipley M, Lewington S, Youngman L, Collins R, Marmot M, Peto R: Underestimation of risk associations due to regression dilution in long-term follow-up of prospective studies. *Am J Epidemiol* 1999, 150(4):341-353.

Chapter 5 Consumption of alcoholic beverages and risk of rheumatoid arthritis: findings from the UK Women's Cohort Study

Yuanyuan Dong¹, Darren C. Greenwood², Janet E. Cade¹

1 Nutritional Epidemiology Group, School of Food Science and Nutrition, University of Leeds, Leeds LS2 9JT, UK;

2 Leeds Institute for Data Analytics, Faculty of Medicine and Health, University of Leeds, Leeds, LS2 9JT, UK.

5.1 Abstract

Background: Whether total alcohol intake and specific alcoholic beverages influence the risk of rheumatoid arthritis (RA) remains unclear, particularly when competing risks are considered.

Methods: The UK Women's Cohort Study recruited women aged 35–69 years at baseline (1995–1998), collecting data on diet, lifestyle, anthropometry, and socioeconomic status. Alcohol consumption was self-reported as drinking status and typical weekly intake of specific alcoholic beverages. Incident cases of RA were identified through linkage to HES up to March 2019. Cox proportional hazards regression was used to estimate hazard ratios (HRs) and 95% confidence intervals (CI) for the association between alcohol consumption (type and amount) and RA incidence, adjusted for potential confounders, both with and without allowance for competing risks.

Results: Among 23,699 women, 255 incident RA cases were observed (529,759 person-years). Compared with non-drinkers, women consuming >14 units/week had a lower risk of RA in both the multivariable-adjusted model (HR (95% CI): 0.57(0.36, 0.89)) and the competing risk model (0.59(0.35, 0.98)). Each 1-unit/week increment in total alcohol intake was associated with a 2% risk reduction (0.98(0.96, 0.99)). Wine showed the most consistent inverse associations after accounting for competing mortality risks (0.99(0.97,1.00)). When compared with regular drinkers (≥ 1 serving/week), occasional drinkers (< 1 serving/week) had a higher risk of RA (1.43 (1.07,1.90)). Non-drinkers showed a similar but non-significant elevated risk (1.45(0.99,2.12)). There was a evidence of interaction between smoking and alcohol consumption ($P_{\text{interaction}} = 0.03$); among non-drinkers, current smokers had a 3.72-fold higher risk of RA (3.72(1.76,7.87)) compared with regular drinkers.

Conclusions: Alcohol intake is inversely associated with RA risk, and this association was modified by smoking.

Keywords: Alcohol; rheumatoid arthritis; cardiovascular disease; CVD; smoking

5.2 Introduction

Rheumatoid arthritis (RA) is the most prevalent form of chronic inflammatory arthritis and a well-established contributor to cardiovascular disease (CVD) morbidity [1]. According to the Global Burden of Disease Study, 17.6 million people worldwide were living with RA in 2020 [2]. The condition leads to substantial social and economic costs, due to its high disability rates and healthcare expenditure [3]. Beyond genetic susceptibility, growing attention has turned to modifiable lifestyle factors in its development.

Among these lifestyle factors, alcohol consumption could be important[4]. Recent data indicate that about 62% of US adults [5] and 77% of UK adults [6] reported past-year alcohol use in 2024, and prevalence rates are steadily rising across many Asian countries [7]. Even a marginal effect of alcohol on joint health and chronic disease risk could translate into a substantial population-level health effect, reinforcing the need for evidence-based guidelines on safe drinking limits.

To date, the association between alcohol consumption and autoimmune diseases remains complex [8]. Alcohol's effect on systemic inflammation appears to depend on drinking patterns [9]. On the one hand, moderate intake can produce metabolites such as acetate [10], which exhibit anti-inflammatory properties and have been linked to reduced RA risk in experimental studies [11, 12]. On the other hand, alcohol is often associated with excess caloric intake and weight gain [13], and frequently co-occurs with other risk factors such as smoking. However, the interactive effects of these lifestyle clusters on RA pathogenesis remain poorly understood.

Epidemiological evidence remains inconclusive. While several case-control studies suggest a protective effect [14-16], prospective study from Nurses' Health Study (NHS) and NHSII reported a lower risk of RA among only moderate drinkers [17] but failed to establish a clear dose-response relationship. Conversely, the French E3N Cohort Study found no clear association between alcohol intake and RA risk in women [18]. Evidence does not clearly indicate a relationship between specific types of alcoholic beverages and RA risk [19]. A methodological oversight in existing research is not accounting for alcohol-related premature mortality from cancer or CVD [20, 21], which may act as a competing risk and obscure the true association between alcohol and RA incidence.

To address these gaps, we examined the associations of individual alcoholic beverage consumption in relation to RA risk among a large prospective cohort of middle-aged UK women. This study extends previous research in two key ways: first, by evaluating the influence of co-occurring lifestyle and drinking exposures on the observed associations; and second, by using competing risk models to account for alcohol-related premature death.

5.3 Methods

5.3.1 Study population

UK Women's Cohort Study (UKWCS) is a prospective cohort study of 35,372 women aged 35-69 years in the United Kingdom recruited from 1995 to 1998 [22]. At recruitment, information about demographic, anthropometric, dietary, lifestyle, and health-related characteristics were collected by a postal questionnaire. The original study was funded by the World Cancer Research Fund (WCRF). The UKWCS was designed to explore potential associations between diet and chronic diseases and has been described in detail elsewhere [23].

Participants were excluded from this analysis if they: lived outside England (n = 3009) and therefore had no linked hospital episode statistics (HES) data; had a diagnosis of RA recorded in HES on or before the recruitment date (n = 38), had missing age (n = 45) or alcohol data (n = 655); had a daily energy intake < 500 kcal or > 5000 kcal (n = 64); or had a BMI < 10 or > 60 kg/m² (n = 1188). After these exclusions, 30,373 participants remained eligible for the study. A flowchart is shown in **Figure 5.1**.

At its inception in 1993, ethical approval for the UKWCS was obtained from a total of 174 local research ethics committees. The National Research Ethics Service (NRES) Committee for Yorkshire & the Humber – Leeds East took on responsibility for the ongoing cohort (reference 15/YH/0027 in 2015 and was updated to include linkage outcomes, such as RA incidence, in 2017 (reference 17/YH/0144). The UKWCS-HES database subsequently received renewed ethical approval from the Yorkshire & the Humber – Sheffield Research Ethics Committee (reference 25/YH/0144).

We followed the Strengthening the Reporting of Observational Studies in Epidemiology - Nutritional Epidemiology (STROBE-nut) guidelines for the reporting of cohort studies (**Appendix D: Table D1**)[24].

5.3.2 Alcohol consumption

Dietary information over 12 months of cohort participants was collected via a self-administered 217-item food frequency questionnaire (FFQ) at recruitment, which was based on the Oxford branch of the European Prospective Investigation into Cancer and Nutrition (EPIC) study [25]. To check reproducibility of this data the FFQ was validated against four-day weighed food diaries and a repeat FFQ on 283 women, both administered three years after recruitment [26].

Alcohol consumption was assessed via self-reported intake during the preceding year and habitual consumption five years prior, with intake quantified in units per week (1 unit = 8 g of ethanol) [27]. Participants were categorised by drinking status as “regular drinkers” (at least once per week), “occasional drinkers” (less than once per week), or “non-drinkers”. For detailed quantification, participants reported their weekly consumption, recorded in pints, glasses, or measures, across specific beverage types, including beer/cider, wine, spirits, and sherry/fortified wine.

Total weekly alcohol intake was categorised into none, <7, 7-14, and ≥ 14 units. Intake was calculated by summing the products of beverage-specific consumption frequencies, standard portion sizes, and alcohol-by-weight values (shown in **Appendix D: Table D2**). To address partial non-response in drinking frequency (ranging from 18.2% for wine to 52.3% for beer/cider), missing values were conservatively imputed as “none” or “seldom”.

5.3.3 Incident RA and mortality outcomes

Diet and lifestyle data were linked with their HES up to 31 March 2019. Incident RA was defined by the occurrence of the International Classification of Diseases, 10th Revision (ICD-10) M05-M06. Follow-up time was calculated from enrolment to RA diagnosis, death, or the end of study period, whichever occurred first, using calendar dates as the timeline.

Mortality data were ascertained through the national death registry, with causes of death classified by ICD-9 or ICD-10 codes. Mortality data (including deaths from cardiovascular disease and cancer) were obtained via linkage to the Health and Social Care Information Centre (HSCIC). We specifically identified primary causes of death from CVD or cancer to account for competing risks. In this cohort, CVD and cancer mortality rates were higher among non-drinkers (12%) compared to regular drinkers (7%) (**Table 5.1; Appendix D: Table D3**). Participants without a death record by the study's end were censored as alive.

5.3.4 Additional covariates

We identified potential confounders and effect modifiers through a literature review and mapped them using a Directed Acyclic Graph (DAG) to guide our modeling strategy (**Figure 5.2**). Following established guidelines for DAG creation and reporting [28], we adjusted our models for the following baseline variables (all measured at baseline): age (continuous), socio-economic status (SES, routine/manual, intermediate, professional/managerial), marital status (married/living as married, separated/divorced, single/widows), smoking status (current, former, never), total metabolic equivalent task (MET) - physical activity (hours/week, continuous), body mass index (BMI, kg/m²; continuous), menopausal status (premenopausal, postmenopausal), hormone replacement therapy (HRT, current, former, or never), diet quality informed by the Alternate Healthy Eating Index (AHEI-2010, continuous) [29], total energy intake excluding alcohol (kcal/day, continuous). We employed the all-components model, adjusting for all non-alcohol caloric sources [30]. Detailed variable definitions and adjustment sets are provided in **Appendix D: Table D4**. Finally, we confirmed the proportional hazards assumption using Schoenfeld residuals, which showed no violations for any model terms.

5.3.5 Statistical analysis

All statistical methods were pre-registered ([ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT06670144): NCT06670144). Baseline dietary, lifestyle, and demographic characteristics were described by mean (SD) or percentages stratified by drinking status (never, occasional, regular), weekly frequency (none, < 7, 7-14, ≥14 units), and beverage type. Missing values were addressed using median imputation for continuous variables

and a missing indicator category for categorical data. Spearman partial correlation coefficients were used to assess relationships between specific alcoholic beverages.

Cox proportional hazards and Fine-Gray sub-distribution models [31] were used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) for RA incidence. The Fine-Gray model accounts for competing risks. Specifically, deaths from cardiovascular disease or cancer that may occur before RA diagnosis and thereby alter the observed hazard. By modelling the sub-distribution hazard for RA, this approach provides an estimate of the cumulative incidence of RA in the presence of competing events, yielding more unbiased effect estimates than conventional Cox models that censor competing events. Three sequential adjustment sets were used: Model 1 (age-adjusted); Model 2 (Model 1 plus BMI, SES, MET, marital status, smoking, AHEI-2020, menopausal status, and HRT); and Model 3, used exclusively for beverage-specific analyses, further adjusted for the consumption of all other types of alcoholic beverages (mutually adjusted model). Additionally, total and beverage-specific intakes were analysed as continuous variables per 1 unit/week increment.

To examine the role of potential effect modifiers, further exploratory analyses included testing for: age (≤ 55 , > 55), BMI (< 25 , ≥ 25 kg/m²), menopausal status (premenopausal, postmenopausal), smoking status (current, former, never), physical activities (low, moderate, high), SES (routine/manual, intermediate, professional/managerial). In each case, the potential effect modifier was omitted from the relevant adjustment set.

To address potential reverse causation bias, sensitivity analyses were performed excluding RA events diagnosed within the first two years of follow-up, participants reporting long-term treatment for illness at baseline, and those who had already reduced alcohol intake prior to recruitment. All P values were from 2-sided tests and results were deemed statistically significant at $P < 0.05$. Analysis was performed using Stata, version 17.0 (StataCorp).

5.4 Results

5.4.1 Participants

There were 23,699 women from the UKWCS included for the analysis. The mean (SD) age was 51.5 (9.1) years. the mean (SD) alcohol consumption was 9.9 (11.8) units per week (median 9.0); during a median follow-up of 22.4 years, 255 incident RA cases were identified (529,759 person-years), and 1891 events were death from CVD or cancer (**Appendix D: Table D1**). More participants consumed wine (83%; 19,674), than beer (70%; 16,694), sherry (70%; 16,692), or spirits (73%; 17,188). Of the 23,699 participants, 18,955 (80%) consumed a combination of at least two types of alcoholic beverage, and 2130 (10%), 92 (0.4%), 155(0.7%), and 196 (0.9%) participants consumed only wine, beer, sherry, or spirits, respectively. For alcohol status, “non-drinkers”, “occasional drinkers” and “regular drinkers” were observed in 10%, 22%, and 67%, respectively (**Table 5.1**).

5.4.2 Observational associations with RA incidence and lifestyle factors

Compared to non-drinkers, regular drinkers were generally younger, more likely to be married, and attained a degree. This group also exhibited some healthier baseline characteristics, self-reporting better overall health, characterised by lower BMI, superior diet quality, and lower total energy intake. However, smoking prevalence was higher and physical activity levels were lower among regular drinkers (**Table 5.1**). Baseline data from **Appendix D: Table D5** indicates that individuals with high alcohol consumption (>14 units/week) showed a similar picture to regular drinkers compared to non-drinkers. Despite having a lower mean BMI and a decreased prevalence of chronic conditions, this group had significantly higher smoking rates and total energy intake. Also, this cohort showed a distinct trend towards being younger, with a lower proportion of postmenopausal women.

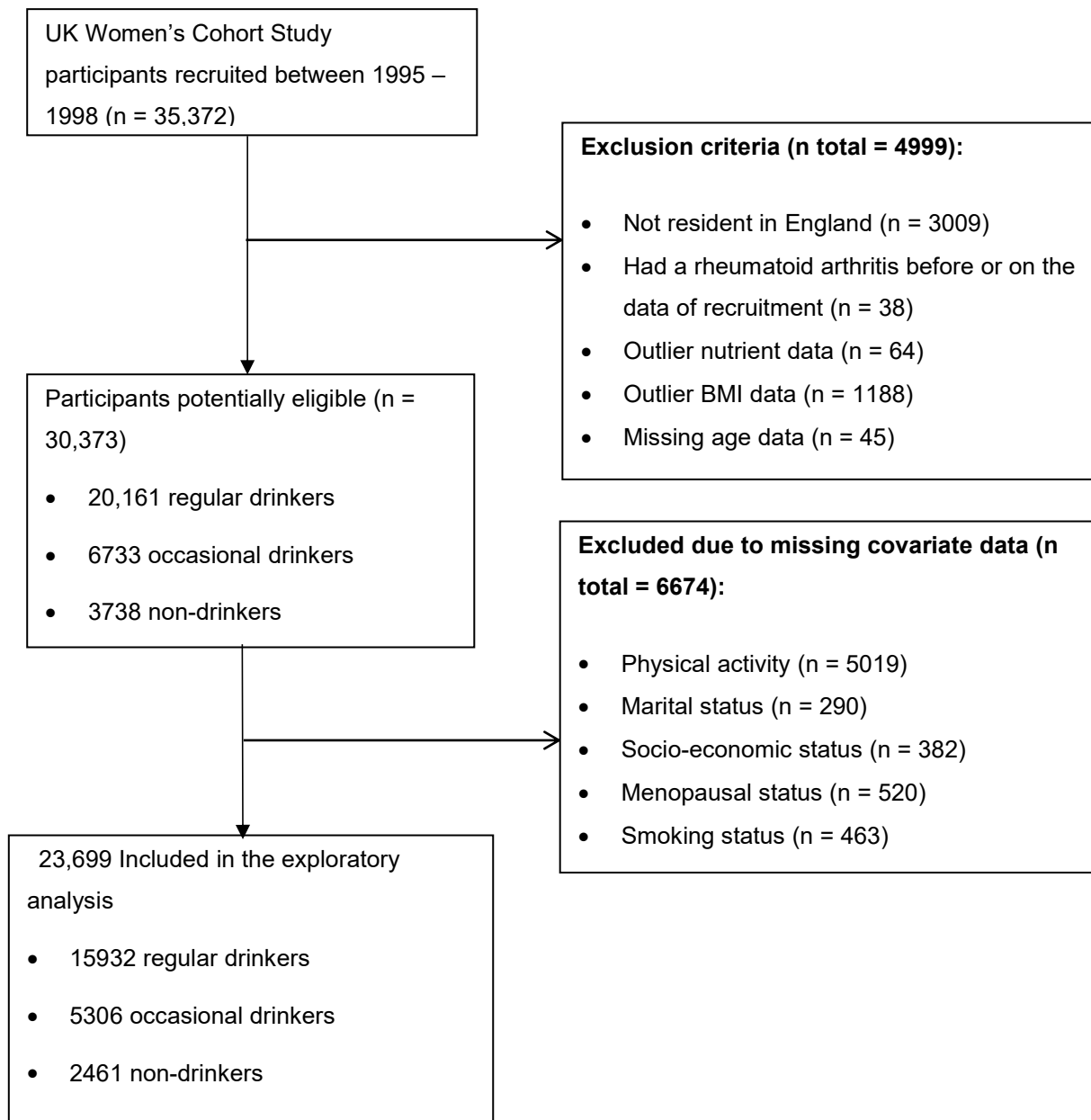


Figure 5.1 Flow chart of UKWCS participants.

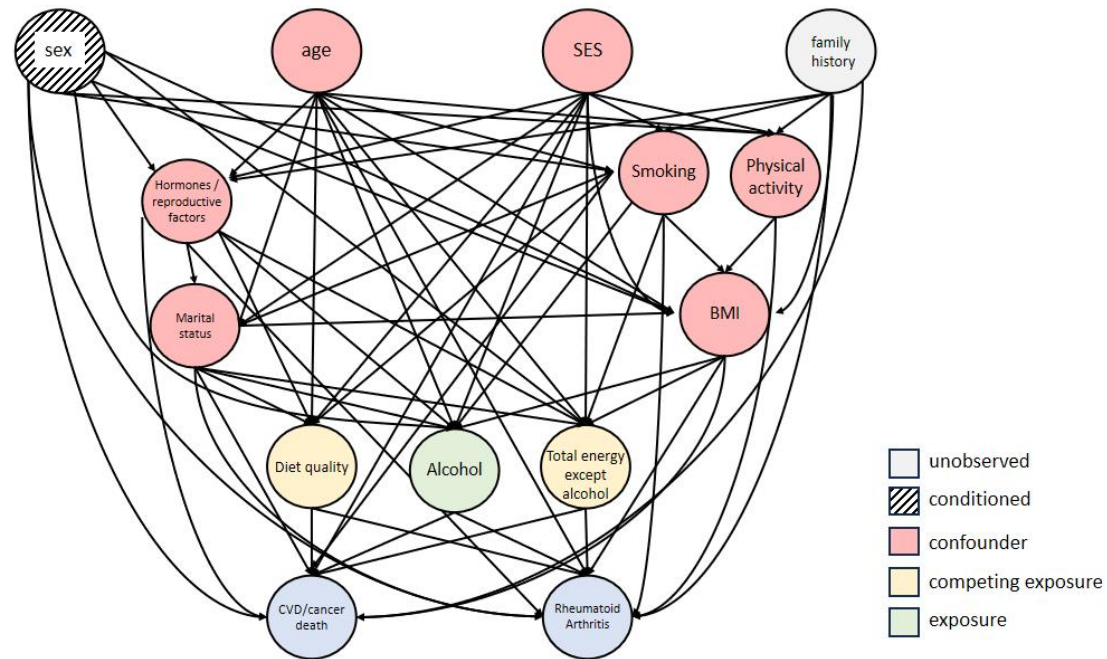


Figure 5.2 Directed Acyclic Graph showing the relationship between alcohol intake, rheumatoid arthritis incidence, and related factors. The exposure is depicted by the green circle and the outcome (rheumatoid arthritis incidence) and competing events are depicted by the blue circle. Variables represented as pink circles are ancestors of the exposure and outcome; variables represented as yellow circles are ancestors of the outcome only (competing exposures); and variables represented as grey circle is unobserved, and hashed circle is conditioned. Known determinants of risk factors include age, socioeconomic status, marital status, hormones/reproductive factors, smoking status, physical activity, diet quality, total energy(except alcohol), and BMI. Hormones/reproductive factors include hormone replacement therapy, breastfeeding, oral contraceptive use. SES: socioeconomic status, BMI: body mass index, CVD: cardiovascular diseases.

Table 5. 1 Baseline characteristics of 23,699 UKWCS participants by drinking status.

Characteristics	Total (n = 23699)	Non-drinkers (n = 2461)	Occasional drinkers (n = 5306)	Regular drinkers (n = 15932)
Participants (%)	23,699	2461 (10)	5306 (22)	15,932 (67)
Age, years				
Mean(SD)	51.5 (9.2)	54.2 (10.0)	52.1 (9.0)	50.9 (9.2)
Range (Min, Max)	(35, 69)	(40, 65)	(39, 66)	(35, 61)
Age at diagnosis, years				
Mean (SD)	65.0 (10.0)	67.8 (10.0)	66.0 (9.0)	64.5 (9.1)
Range (Min, Max)	(38, 86)	(40, 86)	(39, 85)	(38, 85)
Degree-level education (%)	6162 (28)	440 (20)	1086 (23)	4636 (3)
Socio-economic status (%)				
Professional or managerial	15,168 (64)	1438 (58)	3216 (61)	10,514 (66)
Intermediate	6433 (27)	707 (29)	1521 (28)	4205 (26)
Routine or manual	2098 (9)	316 (13)	569 (11)	1213 (8)
Married (%)	17,671 (75)	1679 (68)	3721 (70)	12,271 (77)
White ethnicity (%)	22,874 (99)	2303 (96)	5102 (99)	15,469 (99)
Anthropometric indices				
Height, cm				
Mean (SD)	164 (7.0)	163 (7.0)	163 (7.1)	164 (6.9)
Range (Min, Max)	(145, 182)	(144, 181)	(145, 181)	(146, 183)
Weight, kg				
Mean (SD)	65 (12.4)	66 (14.0)	67 (13.9)	65 (11.9)
Range (Min, Max)	(42, 115)	(43, 118)	(44, 116)	(40, 112)
BMI, kg/m ²				
Mean (SD)	24.4 (4)	24.9 (5)	25.1 (5)	24.1 (4)
Range (Min, Max)	(16.5, 41.0)	(17.0, 40.4)	(17.0, 41.0)	(16.5, 39.3)
Physical activity (METs), hours/week (SD)	14.7 (10)	15.3 (11)	15.3 (10)	14.5 (10)
Smoking status (%)				
Current	2688 (11.3)	236 (9.6)	550 (10.4)	1902 (12.0)
Former	7323 (30.9)	507 (20.6)	1437 (27)	5379 (33.8)
Never	13,688 (58)	1718 (70)	3319 (63)	8651 (54)
Alcohol consumption, units/week				
Mean(SD)	9.0 (11)	0	1.3 (2)	13.0 (11)
Range (Min, Max)	(0, 254)	(0, 0)	(0, 8)	(8, 254)

Characteristics	Total (n = 23699)	Non-drinkers (n = 2461)	Occasional drinkers (n = 5306)	Regular drinkers (n = 15932)
Total energy intake except alcohol, kcal/day				
Mean (SD)	2279 (700)	2291 (750)	2293 (712)	2272 (688)
Range (Min, Max)	(620, 4890)	(620, 4789)	(860, 4690)	(871, 4890)
AHEI-2010 score				
Mean (SD)	67.0 (11)	62.0 (12)	69.9 (11)	66.8 (11)
Range (Min, Max)	(24, 95)	(30, 90)	(38, 95)	(35, 95)
Menopausal status (%)				
Premenopausal	11,994 (51)	931 (38)	2519 (47)	8544(54)
Postmenopausal	11,705 (49)	1530 (62)	2787(53)	7388(46)
HRT use (%)	6230 (27)	642 (26)	1376 (26)	4302 (27)
Prevalence of CVD, cancer, or diabetes (%)	2054 (9)	281 (11)	506 (10)	1267 (8)
Deaths from CVD and cancer (%)	1891 (8)	283 (12)	442 (8)	1166 (7)

SD standard deviation, BMI body mass index, MET Metabolic Equivalent of Task, AHEI-2010 Alternative Healthy Eating Index-2010, HRT hormone replacement therapy, CVD cardiovascular disease. Occasional drinkers: less than once per week (<1 serving/week); Regular drinkers: at least once per week (≥1 serving/week).

Table 5.2 Rheumatoid arthritis risk by drinking status and amount of total alcohol consumption.

Alcohol intake	Participants	n cases	Hazard ratio (95% CI)		
			Adjusted model ^a	Adjusted model ^b	Competing risk model ^c
Alcohol status:					
Regular drinkers	15932	148	1.00	1.00	1.00
Occasional drinkers	5306	72	1.41(1.06,1.87)	1.43(1.07,1.90)	1.33(0.97,1.82)
Non-drinkers	2461	35	1.39(0.96,2.02)	1.45(1.00,2.12)	1.28(0.84,1.98)
Total alcohol (continuous, per 1 unit/week)					
	23699	255	0.98(0.97,1.00)	0.98(0.96,0.99)	0.98(0.97,0.99)
Total alcohol intake (units/week):					
None	2118	33	1.00	1.00	1.00
< 7	10763	121	0.76(0.52,1.12)	0.76(0.50,1.14)	0.74(0.48,1.15)
7-14	5191	55	0.74(0.48,1.15)	0.73(0.46,1.14)	0.80(0.49,1.28)
> 14	5372	46	0.62(0.39,0.96)	0.57(0.36,0.89)	0.59(0.35,0.98)

^a Alcohol intake adjusted for age;

^b Alcohol intake further adjusted for socio-economic status, physical activity, total energy intake, smoking status, diet quality (AHEI-2010 score), hormone replacement therapy, menopausal status, BMI;

^c Competing risk model: Sub-distribution hazard ratios were estimated using the Fine-Gray model, accounting for CVD/cancer mortality as a competing risk for rheumatoid arthritis incidence.

5.4.3 Total alcohol consumption and risk of RA

In the multivariable-adjusted model (Model b), both occasional drinkers and non-drinkers exhibited a significantly higher RA risk compared to regular drinkers (HR (95% CI): 1.43 (1.07, 1.90); 1.45 (1.00, 2.12)). However, in competing risk models accounting for 1891 deaths, these associations were slightly attenuated, where the increased risk for non-drinkers was no longer significant (SHR 1.33(0.97, 1.82)). Notably, women consuming >14 units/week had the lowest risk relative to non-drinkers in the multivariable model (0.57 (0.36, 0.89)), an effect that remained robust in the sub-distribution hazard model (0.59 (0.39, 0.98)). When treated as a continuous variable, alcohol intake showed a consistent inverse association with RA risk (HR 0.98 per 1 unit/week; 95% CI 0.97, 0.99).

5.4.4 Consumption of specific alcoholic beverages and risk of RA

In the categorical analysis, no evidence of associations was found between individual beverage types in relation to risk of RA (**Table 5.3**). In the continuous analysis, wine consumption was associated with a slightly decreased risk of RA per 1 unit/week increment (0.99 (0.97, 1.00)), which remained consistent after accounting for competing risks of mortality. Furthermore, although there was no evidence of association with spirits per unit per week and risk of RA (0.95 (0.90, 1.01)) in the fully-adjusted model (further adjusted for other beverages), it reached borderline significance in the competing risk analysis (0.95 (0.97, 1.00)), suggesting a potential 5% reduction in risk per unit increment after accounting for the competing event of death.

Table 5.3 Consumption of specific alcoholic beverages and incident rheumatoid arthritis risk.

Alcoholic beverages	n cases/Participants	Hazard ratio (95% CI)				<i>P value</i>
		Adjusted model ^a	Adjusted model ^b	Adjusted model ^c	Competing risk model ^d	
Wine only (servings/week):						0.8
0 (reference)	42/3983	1.00	1.00	1.00	1.00	
1-2	61/6112	1.05(0.68,1.61)	1.00(0.64,1.55)	1.05(0.67,1.63)	1.05(0.67,1.63)	
3-4	41/3262	1.32(0.82,2.13)	1.23(0.76,1.99)	1.34(0.81,2.20)	1.35(0.83,2.20)	
> 4	55/6437	0.98(0.64,1.52)	0.94(0.61,1.47)	1.03(0.65,1.64)	1.04(0.66,1.65)	
Wine (continuous, per 1 unit/week):	255/23669	0.98(0.97,1.00)	0.98(0.96,1.00)	0.99(0.97,1.00)	0.99(0.97,1.00)	0.01
Beer only (servings/week):						0.7
0 (reference)	69/6936	1.00	1.00	1.00	1.00	
< 1	10/1151	0.94(0.47,1.90)	0.91(0.45,1.84)	0.96(0.47,1.96)	0.96(0.48,1.91)	
1-2	20/2116	0.98(0.57,1.69)	0.90(0.52,1.55)	1.01(0.58,1.77)	1.01(0.57,1.80)	
≥2	18/2001	1.06(0.61,1.84)	0.91(0.51,1.59)	1.05(0.59,1.86)	1.06(0.58,1.90)	
Beer (continuous, per 1 unit/week):	255/23669	0.99(0.94,1.04)	0.98(0.93,1.03)	0.98(0.93,1.04)	0.98(0.95,1.03)	0.7
Sherry only(servings/week):						0.7
0 (reference)	65/6942	1.00	1.00	1.00	1.00	
1	21/2577	0.81(0.47,1.40)	0.80(0.46,1.38)	0.89(0.51,1.55)	0.91(0.52,1.57)	
2-3	10/1504	0.57(0.26,1.25)	0.55(0.25,1.22)	0.68(0.30,1.53)	0.68(0.31,1.49)	
>3	13/1016	1.27(0.66,2.46)	1.20(0.62,2.33)	1.45(0.74,2.86)	1.46(0.76,2.83)	
Sherry (continuous, per 1 unit/week):	255/23699	0.97(0.89,1.07)	0.98(0.89,1.07)	0.98(0.90,1.07)	0.99(0.91,1.07)	0.5
Spirit only (servings/week):						
0 (reference)	64/6447	1.00	1.00	1.00	1.00	0.8
1	30/2573	1.31(0.83,2.07)	1.20(0.76,1.90)	1.32(0.83,2.10)	1.32(0.83,2.11)	

2	22/1726	1.34(0.79,2.26)	1.19(0.70,2.01)	1.35(0.78,2.32)	1.36(0.80,2.32)	
≥3	30/3098	1.00(0.63,1.60)	0.88(0.55,1.43)	1.02(0.62,1.69)	1.04(0.63,1.73)	
Spirit (continuous, per 1 unit/week):	255/23699	0.96(0.91,1.02)	0.95(0.89,1.01)	0.95(0.90,1.01)	0.95(0.91,1.00)	0.1

^a Alcohol intake adjusted for age;

^b Alcohol intake further adjusted for socio-economic status, physical activity, total energy intake, smoking status, diet quality (AHEI-2010 score), hormone replacement therapy, menopausal status, BMI;

^c further adjusted for other types of alcoholic beverages;

^d Competing risk model: Sub-distribution hazard ratios were estimated using the Fine-Gray model, accounting for all-cause mortality as a competing risk for rheumatoid arthritis incidence.

P value for trend (over categorical intake) and for continuous intake were both derived from **Adjusted Model 3** using the Wald test.

For categorical variables, trend was assessed by assigning the median intake value to each category and treating this as a continuous term in the model.

For continuous variables, the *P* value corresponds to the original continuous intake (per 1 serving/week). HR, hazard ratio; CI, confidence interval.

Table 5.4 Risk of rheumatoid arthritis according to alcohol drinking status stratified by smoking status in the UKWCS.

Drinking status	Never		Former		Current		P-interaction
	cases/participants	HR (95% CI)	cases/participants	HR (95% CI)	cases/participants	HR (95% CI)	
Non-drinkers	23/8651	1.60 (1.05, 2.47)	9/507	1.14 (0.49, 2.63)	16/236	3.72 (1.76, 7.87)	0.03
Occasional drinkers	11/3319	0.91 (0.46, 1.79)	21/1437	1.34 (0.80, 2.23)	40/550	0.95 (0.39, 2.36)	
Regular drinkers (reference)	10/1718	1	58/5379	1	67/1902	1	

Models were adjusted for the following: age, socio-economic status, physical activity, total energy intake, marriage status, body mass index, diet quality (AHEI-2010 score), hormone replacement therapy, and menopausal status. HR, hazard ratio; CI, confidence interval.

5.4.5 Subgroup analyses

Compared to regular drinkers, current smokers who were non-drinkers were found to be at a higher risk of RA (3.72 (1.76, 7.87), with significant effect modification by smoking status (P -interaction = 0.03) (**Table 5.4**). Furthermore, among individuals with low physical activity, occasional drinkers had a higher RA risk than regular drinkers (1.88 (1.06,3.38). Similarly, post-menopausal occasional drinkers showed an elevated risk compared to regular drinkers (1.49 (1.01,2.20)). Correspondingly, effect modification by physical activity was found between categories of alcohol consumption and RA risk (P -interaction = 0.02) (**Appendix D: Table D7**). There was no evidence of an interaction between drinking status or amount by age, menopausal status, SES, or BMI (**Appendix D: Table D6**).

5.4.6 Sensitivity analyses

The association between occasional drinking and RA risk was no longer apparent after excluding individuals who had reduced their alcohol intake for health reasons at baseline ($n = 1,729$) or those who had undergone long-term treatment for illness ($n = 6,396$). The inverse association that we previously observed for higher alcohol intake (≥ 14 units/week) was completely eliminated. While the inverse correlation for total alcohol intake (per 1 unit/week) remained stable in the two-year lag model (0.97 (0.96,0.99)), it was no longer significant after excluding those receiving long-term treatment for illness (0.98 (0.96, 1.01)). In addition, the findings related to total alcohol intake categories (≥ 14 units/week) and wine consumption exhibited no consistent associations across the exclusion of alcohol intake reduced at baseline and those who had long-term treatment.

5.5 Discussion

5.5.1 Principal findings

In this prospective study, weekly alcohol intake (≥ 14 units per week) was associated with a 43% reduction in RA risk compared with non-drinking. This inverse relationship remained consistent when analysing drinking status and continuous alcohol intake. Regarding beverage type, wine showed a marginally stronger inverse association with RA risk. Notably, after accounting for the

competing risk, the elevated risk among non-drinkers was attenuated, suggesting that the association may be partially influenced by the poorer baseline health status of non-drinkers in the UKWCS. The synergistic effect of smoking on these preventive effects was also observed.

5.5.2 Comparison with previous studies

Our findings align with several prospective studies reporting an inverse association between total alcohol consumption and RA risk [17, 19, 32], while highlighting how effect estimates are highly sensitive to reference group selection and unit standardisation. In the UK Biobank, daily intake was associated with a 26% risk reduction [19]; our greater estimate (43%) likely arises from using a strict “none” reference group, whereas the UK Biobank utilised “less than three times per month” [19]. Similarly, the NHS and NHSII cohorts reported stable inverse associations using a “none” reference [17]. In contrast, the French E3N cohort found no evidence of association for total alcohol [33], possibly because the reference group (≤ 1 glass/week) diluted the contrast against higher levels. Some studies relied on self-reported RA (Iowa Women’s Health Study [34] and Kailuan study [35]), whereas our use of hospital episode statistics minimises reporting error and selective loss to follow-up. Furthermore, our data extends evidence by suggesting that the immunomodulatory effects of alcohol may persist beyond the 14 units/week threshold.

Cross-study comparisons are further complicated by unit definitions. US studies often define one unit as 15g of ethanol, nearly double the UK 8g standard. However, this distinction does not fully explain the low exposure levels in the Women’s Health Initiative (WHI) [36], and their highest exposure category was capped at only 3.74 drinks per week, likely because few participants reported higher intakes. So, the WHI may not have reached the biological threshold required to observe immunomodulatory benefits. Our findings also suggest a possible hormonal interaction: while the WHI reported no association in post-menopausal women, we observed a significant risk increase in occasional drinkers within this subgroup. This implies that alcohol’s influence on immune regulation may be modulated by oestrogen levels, with more stable protective associations emerging post-menopause when the baseline immune environment shifts [37].

In line with established evidence [32], we identified a significant interaction between smoking and alcohol intake. Non-drinkers who smoked exhibited a nearly 4-fold increased risk compared to regular drinkers, consistent with the hypothesis that alcohol may mitigate the pro-inflammatory stress of smoking, potentially by inhibiting cytokine production [38]. Specifically, the pathogenic effect of smoking was halved among alcohol consumers compared to non-drinkers. However, a “healthy survivor” bias must be considered within a competing risk framework; it is possible that regular drinkers who survive long enough to be diagnosed with RA possess inherently more favourable genetic or metabolic profiles, reducing their susceptibility to premature mortality.

Beyond ethanol, the heterogeneous composition of alcoholic beverages appears to play a critical role. Our finding that wine is inversely associated with RA risk aligns with its rich profile of polyphenols, flavonoids, and antioxidants. A previous meta-analysis reported a borderline association for wine consumption [39]. In four prospective studies, the association between individual alcoholic beverages and RA risk was examined [17, 18, 33, 34]. Only the French E3N cohort reported wine consumption (4-10 servings per week) was associated with a decreased risk of seropositive RA, but not normal RA [33]. In our cohort, wine was the predominant beverage (reported by >80% of participants), which may explain why our findings differ from cohorts with more diverse preference profiles. This suggests that non-ethanol constituents likely contribute to the observed health benefits. The sample size for non-wine beverages (beer, sherry, spirits) were relatively small, and particularly limited at higher intake levels, which may have prevented the detection of associations. Nevertheless, the overall trend supports a beverage-specific effect driven by bioactive compounds.

5.5.3 Interpretation and implications

The inverse association between regular alcohol consumption and RA risk may be attributed to the significant immunomodulatory and anti-inflammatory properties of ethanol. Evidence suggests that ethanol suppresses the production of key pro-inflammatory cytokines, including TNF- α , IL-1, and IL-6, in a dose-dependent manner [40]. Experimental models further demonstrate that alcohol exposure reduces the expression of BCL6, PD-1, and IL-21 by T follicular helper (T_{fh}) cells, which are pivotal in the autoantibody production characteristic of RA

pathogenesis [41]. Another possible mechanism is that alcohol may exert influence through the “gut-joint axis”. Ethanol treatment has been shown to alter the gut microbiome, specifically increasing the abundance of *Muribaculaceae*, which subsequently elevates levels of the metabolite acetate [42]. This is of particular significance as acetate acts as a ligand for GPR43, a G protein-coupled receptor that maintains immune homeostasis by inhibiting excessive inflammatory responses. In our analysis, the more stable association observed in post-menopausal women suggests that as endogenous oestrogen, a key regulator of B-cell maturation and immune tolerance declines [43], the immune system may become more reliant on alternative regulatory pathways, such as the acetate–GPR43 axis, to suppress systemic inflammation.

The protective potential of alcohol could also be viewed through the lens of complex gene-environment interactions. A three-way interaction between alcohol consumption, smoking, and the HLA-DRB1 shared epitope (SE) has been identified, where the absence of alcohol consumption significantly compounds the risk of ACPA-positive RA in genetically predisposed smokers [44]. This suggests that ethanol may partially mitigate the pathogenic oxidative stress and cytokine surge induced by smoking.

However, the interpretation of these biological findings requires caution regarding inherent epidemiological biases. Within a competing risk framework, the observed inverse association may be partially attributed to “healthy subject bias”. Non-drinkers often exhibit poorer baseline health profiles and face higher premature mortality from cardiovascular disease or cancer. Consequently, regular drinkers who survive long enough to receive an RA diagnosis may represent a robust sub-population, potentially leading conventional models to overestimate alcohol’s biological protection. Furthermore, reverse causality, for example early, undiagnosed symptoms of RA or related comorbidities prompt individuals to reduce or cease alcohol intake, thereby inflating the risk associated with the non-drinking group.

5.5.4 Strengths and limitations

This study includes a large sample size with more than 20-years follow-up period; this has facilitated good statistical power to examine associations between

alcohol consumption and RA risk. We identified RA cases by linkage with the hospital records, which reduced reporting errors and loss to follow-up. Since RA is typically managed by specialist care, the completeness of diagnoses in hospital data is likely to be high. Finally, information on a wide range of factors was collected at baseline to adjust for potential confounders, and the competing risk models provide a more accurate estimate of their risk of RA, avoiding overestimate the incidence of an event if premature death is not considered.

However, we cannot exclude the potential for residual confounding or reverse causality in observational studies. For example, we could not adjust for family history of RA. Alcohol consumption for each participant was calculated based on the self-reported survey at recruitment, and while we collected data on the main beverage types (e.g., wine, beer, spirits), we could not disaggregate these into finer subtypes (e.g., types of beer or individual spirits). Future studies with more detailed drinking frequency categories (e.g., daily vs. weekly) would help refine this classification. Lifetime drinking trajectories over time were not captured, potentially resulting in attenuated estimates. Similarly, seronegative RA cases were identified, but the number was limited to explore the potential relationships. Our findings may not be generalizable to men or other ethnic groups, and UKWCS participants may be healthier on average than UK population due to a healthy participant bias, limiting the applicability of the results. Future research should utilise metabolomics and more granular data on drinking behaviours, such as binge patterns and long-term trends, to increase effect precision and further elucidate biological mechanisms like the gut-acetate-neutrophil axis.

5.6 Conclusions

This prospective study demonstrates that moderate-to-high alcohol intake is associated with a substantial reduction in RA risk among women. However, this association is partially influenced by healthy subject bias and potential reverse causality. Further studies are needed to understand the mechanisms behind the findings, which may contribute to define novel strategies for RA prevention.

5.7 References

1. Crowson CS, Liao KP, Davis JM, 3rd, Solomon DH, Matteson EL, Knutson KL, Hlatky MA, Gabriel SE: Rheumatoid arthritis and cardiovascular disease. *American heart journal* 2013, 166(4):622-628.e621.
2. Global, regional, and national burden of rheumatoid arthritis, 1990-2020, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. *The Lancet Rheumatology* 2023, 5(10):e594-e610.
3. Galloway J, Capron JP, De Leonardis F, Fakhouri W, Rose A, Kouris I, Burke T: The impact of disease severity and duration on cost, early retirement and ability to work in rheumatoid arthritis in Europe: an economic modelling study. *Rheumatology advances in practice* 2020, 4(2):rkaa041.
4. Scott IC, Tan R, Stahl D, Steer S, Lewis CM, Cope A: The protective effect of alcohol on developing rheumatoid arthritis: a systematic review and meta-analysis. 2013, 52(5):856-867.
5. SAMHSA C: 2024 National Survey on Drug Use and Health. Table 2.26A—Alcohol use in past year: among people aged 12 or older; by age group and demographic characteristics, numbers in thousands,. In.; 2024.
6. NHS Digital: Health Survey for England, 2024. In.; 2024.
7. Sornpaisarn B, Shield K, Manthey J, Limmade Y, Low WY, Van Thang V, Rehm J: Alcohol consumption and attributable harm in middle-income South-East Asian countries: Epidemiology and policy options. *The International journal on drug policy* 2020, 83:102856.
8. Azizov V, Zaiss MM: Alcohol Consumption in Rheumatoid Arthritis: A Path through the Immune System. *Nutrients* 2021, 13(4).
9. Barbería-Latasa M, Gea A, Martínez-González MA: Alcohol, Drinking Pattern, and Chronic Disease. *Nutrients* 2022, 14(9).
10. Le Daré B, Lagente V, Gicquel T: Ethanol and its metabolites: update on toxicity, benefits, and focus on immunomodulatory effects. *Drug metabolism reviews* 2019, 51(4):545-561.
11. Jin L, Zhang Z, Pan P, Zhao Y, Zhou M, Liu L, Zhai Y, Wang H, Xu L, Mei D, Zhang H, Yang Y, Hua J, Zhang X, Zhang L: Low-dose ethanol consumption inhibits neutrophil extracellular traps formation to alleviate rheumatoid arthritis. *Communications biology* 2023, 6(1):1088.

12. Martinsson K, Dürholz K, Schett G, Zaiss MM, Kastbom A: Higher serum levels of short-chain fatty acids are associated with non-progression to arthritis in individuals at increased risk of RA. *Annals of the rheumatic diseases* 2022, 81(3):445-447.
13. Qin B, Yang M, Fu H, Ma N, Wei T, Tang Q, Hu Z, Liang Y, Yang Z, Zhong R: Body mass index and the risk of rheumatoid arthritis: a systematic review and dose-response meta-analysis. *Arthritis research & therapy* 2015, 17(1):86.
14. Källberg H, Jacobsen S, Bengtsson C, Pedersen M, Padyukov L, Garred P, Frisch M, Karlson EW, Klareskog L, Alfredsson L: Alcohol consumption is associated with decreased risk of rheumatoid arthritis: results from two Scandinavian case-control studies. *Annals of the rheumatic diseases* 2009, 68(2):222-227.
15. Bergström U, Jacobsson LT, Nilsson J, Wirfält E, Turesson C: Smoking, low formal level of education, alcohol consumption, and the risk of rheumatoid arthritis. *Scandinavian journal of rheumatology* 2013, 42(2):123-130.
16. Sundström B, Johansson I, Rantapää-Dahlqvist S: Diet and alcohol as risk factors for rheumatoid arthritis: a nested case-control study. *Rheumatology international* 2015, 35(3):533-539.
17. Lu B, Solomon DH, Costenbader KH, Karlson EW: Alcohol consumption and risk of incident rheumatoid arthritis in women: a prospective study. *Arthritis & rheumatology (Hoboken, NJ)* 2014, 66(8):1998-2005.
18. Di Giuseppe D, Alfredsson L, Bottai M, Askling J, Wolk A: Long term alcohol intake and risk of rheumatoid arthritis in women: a population based cohort study. *BMJ (Clinical research ed)* 2012, 345:e4230.
19. Mazzucca CB, Scotti L, Cappellano G, Barone-Adesi F, Chiocchetti A: Nutrition and Rheumatoid Arthritis Onset: A Prospective Analysis Using the UK Biobank. *Nutrients* 2022, 14(8).
20. Bouajila N, Domenighetti C, Aubin HJ, Naassila M: Alcohol consumption and its association with cancer, cardiovascular, liver and brain diseases: a systematic review of Mendelian randomization studies. *Frontiers in epidemiology* 2024, 4:1385064.
21. Prevention. CfDCa: Facts About U.S. Deaths from Excessive Alcohol Use. August 2024.

22. Cade JE, Burley VJ, Alwan NA, Hutchinson J, Hancock N, Morris MA, Threapleton DE, Greenwood DC: Cohort Profile: The UK Women's Cohort Study (UKWCS). *International journal of epidemiology* 2017, 46(2):e11.
23. Rada-Fernandez de Jauregui D, Evans CEL, Jones P, Greenwood DC, Hancock N, Cade JE: Common dietary patterns and risk of cancers of the colon and rectum: Analysis from the United Kingdom Women's Cohort Study (UKWCS). *International journal of cancer* 2018, 143(4):773-781.
24. Lachat C, Hawwash D, Ocké MC, Berg C, Forsum E, Hörnell A, Larsson C, Sonestedt E, Wirfält E, Åkesson A, Kolsteren P, Byrnes G, De Keyzer W, Van Camp J, Cade JE, Slimani N, Cevallos M, Egger M, Huybrechts I: Strengthening the Reporting of Observational Studies in Epidemiology-Nutritional Epidemiology (STROBE-nut): An Extension of the STROBE Statement. *PLoS medicine* 2016, 13(6):e1002036.
25. Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, Charrondière UR, Hémon B, Casagrande C, Vignat J, Overvad K, Tjønneland A, Clavel-Chapelon F, Thiébaud A, Wahrendorf J, Boeing H, Trichopoulos D, Trichopoulou A, Vineis P, Palli D, Bueno-De-Mesquita HB, Peeters PH, Lund E, Engeset D, González CA, Barricarte A, Berglund G, Hallmans G, Day NE, Key TJ, Kaaks R, Saracci R: European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public health nutrition* 2002, 5(6b):1113-1124.
26. Cade JE, Burley VJ, Greenwood DC: The UK Women's Cohort Study: comparison of vegetarians, fish-eaters and meat-eaters. *Public health nutrition* 2004, 7(7):871-878.
27. NHS Digital: Alcohol units. [Available from: <https://www.nhs.uk/live-well/alcohol-advice/calculating-alcohol-units/>].
28. Tennant PWG, Murray EJ, Arnold KF, Berrie L, Fox MP, Gadd SC, Harrison WJ, Keeble C, Ranker LR, Textor J, Tomova GD, Gilthorpe MS, Ellison GTH: Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. *International journal of epidemiology* 2021, 50(2):620-632.
29. Wang T, Heianza Y, Sun D, Huang T, Ma W, Rimm EB, Manson JE, Hu FB, Willett WC, Qi L: Improving adherence to healthy dietary patterns, genetic

- risk, and long term weight gain: gene-diet interaction analysis in two prospective cohort studies. *BMJ (Clinical research ed)* 2018, 360:j5644.
30. Tomova GD, Arnold KF, Gilthorpe MS, Tennant PWG: Adjustment for energy intake in nutritional research: a causal inference perspective. *The American journal of clinical nutrition* 2022, 115(1):189-198.
 31. Austin PC, Fine JP: Practical recommendations for reporting Fine-Gray model analyses for competing risk data. *Statistics in medicine* 2017, 36(27):4391-4400.
 32. Hedenstierna L, Bellocco R, Ye W, Adami HO, Åkerstedt T, Trolle Lagerros Y, Hedström AK: Effects of alcohol consumption and smoking on risk for RA: results from a Swedish prospective cohort study. *RMD open* 2021, 7(1).
 33. Ascione S, Barde F, Artaud F, Nguyen Y, Macdonald C, Mariette X, Boutron-Ruault MC, Salliot C, Seror R: Association between beverage consumption and risk of rheumatoid arthritis: a prospective study from the French E3N Cohort. *Rheumatology (Oxford, England)* 2023, 62(5):1814-1823.
 34. Cerhan JR, Saag KG, Criswell LA, Merlino LA, Mikuls TR: Blood transfusion, alcohol use, and anthropometric risk factors for rheumatoid arthritis in older women. *The Journal of rheumatology* 2002, 29(2):246-254.
 35. VanEvery H, Yang W, Olsen N, Bao L, Lu B, Wu S, Cui L, Gao X: Alcohol Consumption and Risk of Rheumatoid Arthritis among Chinese Adults: A Prospective Study. 2021, 13(7):2231.
 36. Krok-Schoen JL, Brasky TM, Hunt RP, Rohan TE, Baker TA, Li W, Carbone L, Mackey RH, Snetselaar L, Lustberg MB, Neuhauser ML: Dietary Long-Chain n-3 Fatty Acid Intake and Arthritis Risk in the Women's Health Initiative. *Journal of the Academy of Nutrition and Dietetics* 2018, 118(11):2057-2069.
 37. Kovacs EJ, Messingham KA: Influence of alcohol and gender on immune response. *Alcohol research & health : the journal of the National Institute on Alcohol Abuse and Alcoholism* 2002, 26(4):257-263.
 38. Addissouky TA, El Sayed IET, Ali MMA, Wang Y, El Baz A, Elarabany N, Khalil AA: Oxidative stress and inflammation: elucidating mechanisms of smoking-attributable pathology for therapeutic targeting. *Bulletin of the National Research Centre* 2024, 48(1):16.

39. Dong Y, Greenwood DC, Webster J, Uzokwe C, Tao J, Hardie LJ, Cade JE: Dose-Response Associations Between Diet and Risk of Rheumatoid Arthritis: A Meta-Analysis of Prospective Cohort Studies. *Nutrients* 2024, 16(23).
40. Mandrekar P, Jeliaskova V, Catalano D, Szabo G: Acute Alcohol Exposure Exerts Anti-Inflammatory Effects by Inhibiting I κ B Kinase Activity and p65 Phosphorylation in Human Monocytes¹. *The Journal of Immunology* 2007, 178(12):7686-7693.
41. Azizov V, Dietel K, Steffen F, Dürholz K, Meidenbauer J, Lucas S, Frech M, Omata Y, Tajik N, Knipfer L, Kolenbrander A, Seubert S, Lapuente D, Sokolova MV, Hofmann J, Tenbusch M, Ramming A, Steffen U, Nimmerjahn F, Linker R, Wirtz S, Herrmann M, Temchura V, Sarter K, Schett G, Zaiss MM: Ethanol consumption inhibits TFH cell responses and the development of autoimmune arthritis. *Nature Communications* 2020, 11(1):1998.
42. Jin L, Zhang Z, Pan P, Zhao Y, Zhou M, Liu L, Zhai Y, Wang H, Xu L, Mei D, Zhang H, Yang Y, Hua J, Zhang X, Zhang L: Low-dose ethanol consumption inhibits neutrophil extracellular traps formation to alleviate rheumatoid arthritis. *Communications biology* 2023, 6(1):1088.
43. Santana-Sánchez P, Vaquero-García R, Legorreta-Haquet MV, Chávez-Sánchez L, Chávez-Rueda AK: Hormones and B-cell development in health and autoimmunity. *Frontiers in immunology* 2024, 15:1385501.
44. Hedström AK, Hössjer O, Klareskog L, Alfredsson L: Interplay between alcohol, smoking and HLA genes in RA aetiology. *RMD open* 2019, 5(1):e000893.

Chapter 6 Healthy lifestyle score, C-reactive protein levels, and incident rheumatoid arthritis: a prospective cohort study of 354,512 UK Biobank participants

Yuanyuan Dong¹, Darren C. Greenwood², Janet E. Cade¹

1 Nutritional Epidemiology Group, School of Food Science and Nutrition, University of Leeds, Leeds LS2 9JT, UK;

2 Leeds Institute for Data Analytics, Faculty of Medicine and Health, University of Leeds, Leeds, LS2 9JT, UK.

6.1 Abstract

Background: C-reactive protein (CRP) is a marker of systemic inflammation and has been associated with both lifestyle behaviours and rheumatoid arthritis (RA) risk. However, it is unknown whether CRP mediates the association between overall lifestyle and incident RA.

Methods: This prospective cohort study included 354,512 UK Biobank participants without RA at baseline. A healthy lifestyle score (range 0–4) was constructed from four components: never smoking, moderate alcohol consumption (≤ 1 drinks/day for women; ≤ 2 drinks/day for men), sufficient physical activity (≥ 600 MET-minutes/week), and a high-quality diet (meeting ≥ 5 out of 10 healthy food components). The cut-offs were based on established the WHO and UK guidelines. Cox proportional regression models were used to investigate the associations.

Results: Over a median follow-up of 15.6 years (5,341,886 person-years), 4,669 incident RA cases were documented. Compared with the lowest lifestyle score group (0–1 points), the medium (2 points) and highest (3–4 points) groups had significantly lower RA risks (HR (95% CI): 0.92 (0.86,0.99) and (0.81 (0.75, 0.88)), respectively. The inverse association was stronger in men ($P_{\text{interaction}} = 0.006$), participants with lower BMI (≤ 26.6 kg/m², $P_{\text{interaction}} = 0.03$), and those from England ($P_{\text{interaction}} = 0.04$). Mediation analyses showed that CRP explained 29% of the observed risk difference (95% CI: 21.7%,42.7%).

Conclusions: A higher healthy lifestyle score is associated with a lower risk of incident RA. This association is partly explained by CRP, suggesting that reducing systemic inflammation may be one pathway through which healthy lifestyles protect against RA. The findings support the potential for lifestyle-based prevention strategies for RA.

Keywords: Rheumatoid arthritis; healthy lifestyle; C-reactive protein; mediation analysis; UK Biobank

6.2 Introduction

Rheumatoid arthritis (RA) is the most common chronic inflammatory arthritis, affecting an estimated 17.6 million people worldwide [1], with women disproportionately affected [2]. It is characterised by synovial inflammation and progressive joint erosion [3]. Its prevalence is rising worldwide, driven largely by population aging and increased longevity [4]. Although the exact pathogenesis of RA remains unclear, it is believed to result from a complex interplay of genetic and environmental factors. This has shifted the focus from simply identifying risk factors to developing effective prevention strategies, where modifiable lifestyle behaviours play a central role.

The concept of a “healthy lifestyle”, typically encompassing never smoking, moderate alcohol consumption, adequate physical activity and a high-quality diet, has been associated with lower risks of cardiovascular disease [5], diabetes [6, 7] and certain cancers [8]. In contrast, data from the large prospective studies on whether a healthy lifestyle adherence is associated with RA risk remain sparse.

Our preceding chapters using the UK Women’s Cohort Study (UKWCS) have examined individual diet-related factors. Chapter 3 found that central adiposity measures were associated with increased RA risk [9]. Chapters 4 and 5 showed that higher adherence to the Alternative Mediterranean diet and low-level regular alcohol intake were associated with reduced RA risk. This chapter evaluates whether adherence to a comprehensive healthy lifestyle, within which diet is a core component, confers additional protection beyond these individual factors. To our knowledge, no studies have explicitly addressed whether long-term adherence to a healthy lifestyle facilitates RA prevention in the UK population.

Inflammation status is a strong determinant of both lifestyle behaviours and inflammatory arthritis outcomes. Higher levels of C-reactive protein (CRP), a well-known marker of systemic inflammation [10], are associated with increased diseases activity and progression of RA [11, 12]. Additionally, results of observational studies suggested that physical activity [13], non-smoking [14], better diet quality [15], and moderate alcohol drinking [16] are associated with lower CRP levels, respectively. Furthermore, systemic inflammation, measured

by CRP, could act as a mediator between lifestyle and RA risk, yet no studies have formally examined this pathway.

Using the UK Biobank, a large prospective cohort of middle-aged and older adults, we aimed to examine the association between a healthy lifestyle score and incident RA and determine the role of CRP as a potential mediators in this association.

6.3 Methods

6.3.1 Study population

The UK Biobank recruited more than 500,000 participants aged 37–73 years between 2006 and 2010 from 22 assessment centres across England, Scotland and Wales. Details of the study design and data collection have been described previously [17]. Participants were recruited via National Health Service patient registers, and attended one of 22 assessment centres across the UK, where participants completed a touchscreen questionnaire, verbal interview, physical measures, and a biosample collection. All participants provided written informed consent, and the study was approved by the North West Multicentre Research Ethics Committee (REC reference: 11/NW/0382).

For the present analysis, we excluded participants with genetic sex not matching self-reported sex ($n = 362$), Body Mass Index (BMI, <10 or >60 kg/m²) ($n = 3036$), and those recorded as lost to follow-up ($n = 1282$). We excluded those with a previous RA ($n = 6818$) on or before the date of recruitment. After exclusions, the analytical sample comprised 354,512 participants (**Appendix E: Figure E1**). We followed the Strengthening the Reporting of Observational Studies in Epidemiology – Nutritional Epidemiology (STROBE-nut) guidelines for the reporting of cohort studies (**Appendix E: Table E1**) [18].

6.3.2 Assessment of healthy lifestyle score

We constructed a healthy lifestyle score including smoking, alcohol drinking, diet quality and physical activity. All lifestyle factors were obtained through structured questionnaires and a touchscreen dietary questionnaire [19] that asked about their frequency of consumption of various meat, fish, eggs, and dairy products.

Never smoking was considered healthy, defined in the questionnaire as smoking fewer than 100 cigarettes in a lifetime. Current and former smokers (including those who smoked occasionally or only tried once or twice) were classified as unhealthy.

Alcohol intake was assessed using self-reported frequency and volume. A healthy level was defined daily consumption as ≤ 2 drinks for men and ≤ 1 drink for women (including non-drinkers), in accordance with the UK drinking guidelines [20]. One UK unit equals 8 g of ethanol [21].

Physical activity was assessed via self-reported frequency and duration of moderate and vigorous activities. Metabolic equivalent (MET) minutes per week were calculated for each participant. Following the World Health Organization physical activity guidelines [22] and standard UK Biobank practice [23], a healthy level was defined as ≥ 600 MET minutes/week (equivalent to ≥ 150 minutes of moderate-intensity or ≥ 75 minutes of vigorous-intensity activity per week).

We adapted a diet quality score previously used in UK Biobank for cardiovascular health [24]. This score reflects an anti-inflammatory dietary pattern through its emphasis on fruits, vegetables, whole grains, fish, and vegetable oils, and its limitation of refined grains, red/processed meats, and sugary drinks. A healthy diet was defined as meeting ≥ 5 of the 10 components (**Appendix E: Table E2**).

We assigned one point was assigned to a healthy level of each lifestyle factor and 0 points to an unhealthy level (**Appendix E: Table E3**). The healthy lifestyle score was the sum of these points, ranging from 0 to 4; higher scores indicated a healthier lifestyle. For analysis, the healthy lifestyle score was categorised into three groups: low (0–1 points), medium (2 points) and high (3–4 points), with the low group serving as the reference. This simple additive approach has been widely used [25, 26]. BMI was not included in the primary lifestyle score as it may act as a mediator between behavioural factors and health outcomes, and due to concerns about the obesity paradox [27, 28].

6.3.3 Outcome ascertainment

The primary outcome in this study was the incidence of RA. Baseline prevalent RA was defined using UK Biobank's first occurrence fields for International

Classification of Diseases (ICD)-10 codes M05 (Data-Field 131848) and M06 (Data-Field 131850). Participants with a first recorded diagnosis date on or before their baseline assessment centre date (Data-Field 53) were considered prevalent cases. Incident RA during follow-up was defined as a first recorded diagnosis date after the baseline assessment date. Participants were followed from baseline to the earliest of RA diagnosis, death, loss to follow-up, or the end of follow-up (25 October 2021).

6.3.4 Statistical analyses

6.3.4.1 Main analyses

All statistical methods were pre-registered on [ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT06670183) (NCT06670183).

Baseline characteristics of UK Biobank participants were summarised across healthy lifestyle score groups. Cox proportional hazards regression models were used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) for the association between lifestyle score groups and incident RA, with the lowest lifestyle score group (0–1 points) serving as the reference. The target estimand was the relative causal effect of a higher lifestyle score (medium or high) on RA risk compared with the lowest score group.

Age-adjusted and multivariable-adjusted models were applied, with calendar time as the timescale. Additional confounders included in the adjusted model were selected and comprised (all measured at recruitment): region (England, Scotland, Wales), ethnicity (white vs. Non-white), Townsend deprivation index (continuous), sex (male, female), BMI (continuous), menopausal status (premenopausal, postmenopausal), hormone replacement therapy (HRT; yes, no), diabetes (yes, no), and cardiovascular disease (CVD; yes, no). The proportional hazards assumption was tested using the Schoenfeld residuals, and no violations were observed.

6.3.4.2 Subgroup analyses

To evaluate potential effect modification by sex (male, female), region (England, others), and BMI (dichotomised at the median [26.6 kg/m²]), we used likelihood ratio tests comparing adjusted Cox regression models with and without an

interaction term between healthy lifestyle score groups and each potential effect modifier. In each analysis, the potential effect modifier itself was omitted from the adjustment set.

6.3.4.3 Mediation analyses

We explored the potential of CRP as a mediator because previous studies have consistently shown that CRP levels are associated with lifestyle behaviours and RA development [11, 29]. The mediation effect of CRP on the association between healthy lifestyle score and incident RA was examined using the product-of-coefficients method ($a \times b$) [30]. Linear regression estimated the a coefficient (lifestyle score $\rightarrow$ logCRP), and Cox proportional hazards regression estimated the b coefficient (logCRP $\rightarrow$ RA, adjusted for lifestyle score). Both models were adjusted for the same covariates as the main analysis. The indirect effect ($a \times b$), direct effect (c'), total effect (c), and the proportion mediated (indirect/total) were calculated with 95% confidence intervals derived from 1000 bootstrap resamples [31]. The proportional hazards assumption for logCRP was not satisfied ($P < 0.001$; **Appendix E: Figure E2**), indicating a time-varying effect, which was assessed using scaled Schoenfeld residuals. Therefore, the mediation estimates represent average mediation effects over the follow-up period. A detailed description of the mediation analysis is provided in **Appendix E: Table E4**.

6.3.4.4 Sensitivity analyses

We conducted several sensitivity analyses to test the robustness of our findings. These included excluding participants who developed RA within the first 3 years of follow-up, further adjusting for the initial assessment centre in the multivariable Cox models, and treating death during follow-up as a competing risk for RA using the Fine-Gray subdistribution hazard model. All analyses were performed using *R* (version 4.3.1).

6.4 Results

6.4.1 Participants

Of 490,438 participants potentially eligible at recruitment, those with missing covariate data for Townsend Deprivation Index ($n = 592$), CRP levels ($n = 32,036$),

smoking status (n = 345), alcohol intake (n = 1955), physical activity (n = 100,998) were excluded, leaving 354,512 participants for unadjusted and adjusted analyses. The study flow chart is given in **Appendix E: Figure E1**.

6.4.2 Descriptive data

Over a median follow-up of 15.6 years, 4,669 incident RA cases were observed (total person-years 5,341,886), corresponding to 1.3% of the cohort. Characteristics of the 354,512 UK Biobank participants at recruitment stratified by baseline CRP levels are summarised in **Table 6.1**. On average, participants with lower CRP levels were younger than those with higher CRP levels. The proportion of women was slightly higher in the lowest CRP group compared with the middle and high CRP groups. Socioeconomic indicators were more favourable among participants with lower CRP levels: a higher proportion had college or above education and a lower Townsend deprivation index. Participants with lower CRP levels also engaged in higher physical activity and were more likely to be never smokers. Conversely, BMI was inversely associated with CRP: mean BMI was 25.3 kg/m² in the >3 mg/L group and 30.2 kg/m² in the <1 mg/L group. The prevalence of CVD and diabetes at baseline was higher among participants with elevated CRP levels. Regarding female-specific covariates, the proportion of postmenopausal women and current HRT users was higher in the high CRP group compared with the low CRP group.

Sex- and lifestyle score-stratified characteristics (**Appendix E: Table E5**). Across both sexes, higher lifestyle scores were associated with younger age, higher physical activity levels, and a lower prevalence of current smoking. In men, the proportion of never smokers increased from 17.8% in the low score group to 86.9% in the high score group; in women, the corresponding increase was from 13.9% to 82.8%. Alcohol consumption was higher in men than in women, and within each sex, higher lifestyle scores were associated with slightly lower alcohol intake. BMI showed little variation across lifestyle score groups within each sex, though men had consistently higher BMI than women. Socioeconomic indicators were more favourable in higher lifestyle score groups: the proportion with college or above education was greater, and the Townsend deprivation index was lower. Among female participants, the proportion of postmenopausal women and HRT users was lower in the high lifestyle score group compared with the low score

group, whereas the proportion of premenopausal women was higher in the high score group.

Notably, the distribution of lifestyle scores differed by sex: a higher percentage of women (26.6%) were in the high score group compared with men (19.2%), while a higher percentage of men (25.9%) were in the low score group compared with women (17.1%). The number of incident RA cases was larger in women (1414 in the high score group) than in men (380 in the high score group).

6.4.3 Lifestyle score groups

The healthy lifestyle score was categorised as low (0–1 points), medium (2 points) and high (3–4 points). Compared with the low score group, there is no evidence of association with those in the medium score group, while a reduction in RA risk was observed after multivariable adjustment (HR 0.92 (95% CI 0.86,0.99)). Those in the highest score group had a significantly lower risk of RA (0.81 (0.75,0.88)) (**Figure 6.1**).

Table 6.1 Baseline characteristics of participants in the UK Biobank according to C-reactive protein levels.

Characteristics	Total	CRP levels		
		High (>3 mg/L)	Medium (1–3 mg/L)	Low (<1 mg/L)
Participants (%)	354,512	74,327(21)	134,478(37)	145,707(42)
Cases (%)	4669(1.3)	1680(2.3)	1726(1.3)	1263(0.9)
Socio-demographics				
Age, years (SD)	56.1(9.2)	59.5(9.2)	55.7(8.5)	52.6(9.4)
White ethnicity (%)	336,806(95)	70,345(95)	127,924(95)	138,537(95)
Sex (%)				
men	170,991(48)	32,792(44)	67,740(50)	70,459(48)
women	183,521(52)	41,535(56)	66,738(50)	75,248(52)
Region (%)				
England	313,672(88.5)	65,195(87.7)	119,054(88.5)	129,423(88.9)
Scotland	25,964(7.3)	55,91(7.5)	9578(7.2)	10,795(7.4)
Wales	14,876(4.2)	3541(4.8)	5846(4.3)	5489(3.7)
Household income (%)				
High (≥£52,000)	65,379(18.4)	14,088(19.0)	33,575(25.0)	17,716(12.2)
Medium (£18,000-£51,999)	114,959(32.4)	31,757(42.7)	62,382(46.4)	20,820(14.3)

Low (<£18,000)	135,992(38.4)	43,332(58.3)	67,146(49.9)	25,514(17.5)
Occupation (%)				
Employed	211,417(59.6)	39,428(53.0)	77,298(57.5)	94,691(65.0)
Retired	112,104(31.6)	26,243(35.3)	45,952(34.2)	39,909(27.4)
Unemployed or disabled	17,134(48.3)	5880(7.9)	6445(4.8)	4809(3.3)
Family/student/volunteer	11,344(32.0)	2216(3.0)	3880(2.9)	5248(3.6)
Education (%)				
College or above	178,738(50.4)	32,164(43.3)	65,243(48.5)	82,331(56.5)
High school or equivalent	123,916(35.0)	27,660(37.2)	48,140(35.8)	48,116(33.0)
Less than high school	50,858(14.3)	14,503(19.5)	21,095(15.7)	15,260(10.5)
Townsend deprivation index (SD)	-1.4(3.0)	-1.1(3.2)	-1.5(3.0)	-1.6(2.9)
Lifestyle				
Physical activity, MET.mins/week (SD)	1622(2016)	1417(1960)	1607(2016)	1742(2036)
Smoking status (%)				
Current	35,405(10)	10,504(15)	13,662(10)	11,239(8)
Former	122,036(35)	27,058(37)	47,636(36)	47,342(33)
Never	193,176(55)	36,031(48)	71,696(54)	85,449(59)

Alcohol consumption, units/week (SD)	2.2(2.6)	2.1(2.8)	2.3(2.7)	2.3(2.5)
BMI, kg/m ² (SD)	27.3(4.6)	30.2(5.6)	27.8(4.2)	25.3(3.5)
Comorbidities				
History of CVD (%)	17,837(5.0)	4566(6.1)	6904(5.1)	6367(4.4)
History of diabetes (%)	4577(1.3)	1152(1.5)	1741(1.3)	1684(1.2)
Female-specific covariates				
Menopausal status (%)				
Premenopausal	74,683(21.1)	16,477(22.2)	25,206(18.7)	33,000(22.6)
Postmenopausal	108,838(30.7)	25,058(33.7)	41,532(30.9)	42,248(29.0)
HRT use (%)				
Yes	67,148(18.9)	17,854(24.0)	26,508(19.7)	22,786(15.6)
No	116,373(32.8)	23,681(31.9)	40,230(29.9)	52,462(36.0)

BMI, body mass index; CRP, C-reaction protein; HRT, Hormone Replacement Therapy; MET, Metabolic equivalents; CVD, cardiovascular disease; SD, standard deviation; <£18 000, £18 000-£51 999, and ≥£52 000 of average total household income before tax represented the high, medium, and low family income level, respectively.

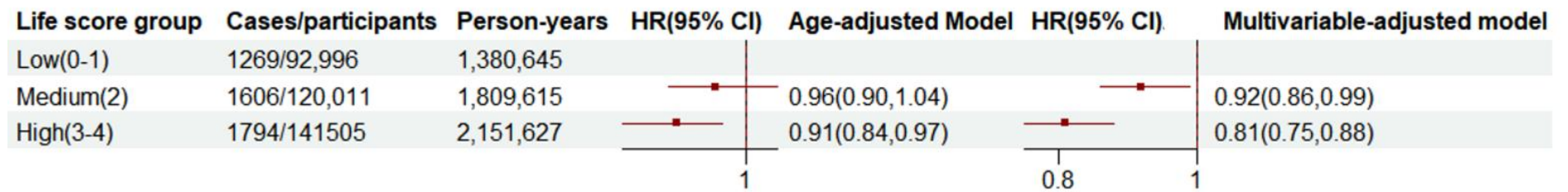


Figure 6.1 Risk of rheumatoid arthritis by healthy lifestyle score in the UK Biobank. Both models controlled for age, and the multivariable-adjusted model was adjusted for the following (all at recruitment): The multivariable-adjusted model was adjusted for the following covariates assessed at recruitment: age (continuous), region (England, Scotland, Wales), ethnicity (white vs. other), Townsend deprivation index (continuous), sex (male, female), body mass index (continuous), menopausal status (premenopausal, postmenopausal), hormone replacement therapy (yes, no), baseline diabetes (yes, no), and baseline cardiovascular disease (yes, no). HR (95% CI): hazard ratio (95% confidence interval).

6.4.4 Subgroup analyses

There was evidence of effect modification by sex ($P_{\text{interaction}} = 0.006$), region ($P_{\text{interaction}} = 0.04$), and BMI ($P_{\text{interaction}} = 0.03$) on the association between lifestyle score and RA risk. In male participants, a high lifestyle score was associated with a 30% lower RA risk (0.70 (0.62,0.79)), whereas in females the reduction was 15% (0.85 (0.77, 0.94)). The protective effect appeared similar in England (0.79 (0.73,0.85)) and other regions (0.76 (0.57,1.00)), although the confidence intervals for other regions were wider. For BMI, participants with a lower BMI ($\leq 26.6 \text{ kg/m}^2$) showed a more evident risk reduction (0.76 (0.67,0.85)) than those with higher BMI (0.84(0.76,0.93)) (**Table 6.2**).

Table 6.2 Risk of rheumatoid arthritis among life score groups in UK Biobank participants, stratified by sex, region, and body mass index.

Stratifying variable		n cases, adjusted HR (95% CI)			p interaction
Sex		Male		Female	
Low Lifescore(reference)	728/62,201	1	541/30,795	1	
Medium Lifescore	643/62,790	0.87(0.78,0.97)	963/57,221	0.96(0.87,1.07)	
High Lifescore	380/46,000	0.70(0.62,0.79)	1414/95,505	0.85(0.77,0.94)	0.006
Region		England		Others	
Low Lifescore(reference)	1176/82,093	1	93/10,903	1	
Medium Lifescore	1476/106,210	0.90(0.83,0.97)	130/13,801	1.00(0.77,1.31)	
High Lifescore	1666/125,369	0.79(0.73,0.85)	128/16,136	0.76(0.57,1.00)	0.04
BMI		≤ 26.6 kg/m ²		> 26.6 kg/m ²	
Low Lifescore(reference)	462/39,825	1	807/53,171	1	
Medium Lifescore	640/58,431	0.89(0.79,1.00)	966/61,580	0.94(0.85,1.03)	
High Lifescore	793/79,037	0.76(0.67,0.85)	1001/62,468	0.84(0.76,0.93)	0.03

The multivariable-adjusted model was adjusted for the following covariates assessed at recruitment: age (continuous), region (England, Scotland, Wales), ethnicity (White vs. other), Townsend deprivation index (continuous), sex (male, female), body mass index (continuous), menopausal status (premenopausal, postmenopausal), hormone replacement therapy (yes, no), baseline diabetes (yes, no), and baseline cardiovascular disease (yes, no). HR (95% CI): hazard ratio (95% confidence interval).

6.4.5 Mediation analyses

Potential mediation of CRP in the association between healthy lifestyle score and incident RA is shown in **Figure 6.2**. CRP levels were progressively lower with higher lifestyle scores ($\beta = -0.068$, $P < 0.001$). CRP was found to partially mediate the observed risk reduction, with an indirect effect ($a \times b$) of -0.0249 (95% CI: -0.0271 , -0.0227), showing that CRP explained 28.8% (95%CI: 21.7%, 42.7%) of the association between lifestyle score and RA risk.

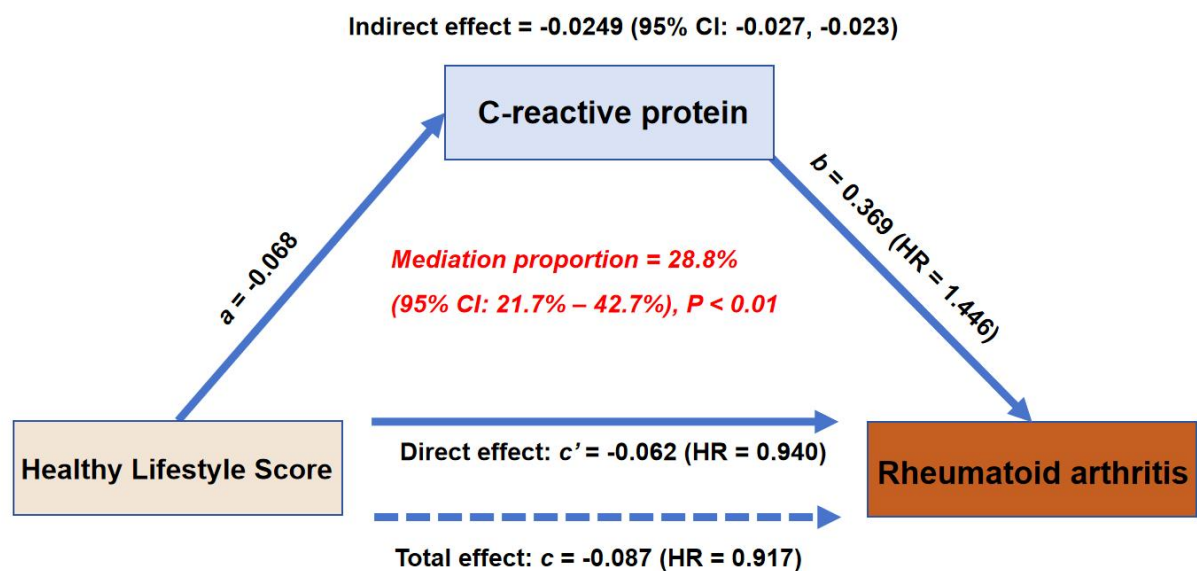


Figure 6.2 Mediation analysis of C-reactive protein in the association between healthy lifestyle score and incident rheumatoid arthritis in the UK Biobank. Path coefficients and hazard ratios (HR) were estimated from linear regression (for logCRP) and Cox proportional hazards models (for rheumatoid arthritis), adjusted for the covariates listed in the Methods. The a path (lifestyle score $\rightarrow$ logCRP) is shown as $a = -0.068$, a linear regression coefficient representing the change in logCRP per 1-unit increase in lifestyle score. The b path (logCRP $\rightarrow$ RA) is shown as $b = 0.369$, the log hazard ratio (logHR) for RA per unit increase in logCRP, corresponding to HR = 1.446. The indirect effect ($a \times b$) = -0.0249 is the product of these two coefficients (also on the log scale). The proportion mediated is 28.8% (95% CI: 21.7%–42.7%). Direct and total effects are also presented as logHRs (with corresponding HRs in parentheses). All coefficients are on the log scale; hazard ratios are obtained by exponentiation. The indirect effect ($a \times b$) and proportion mediated were calculated using the product-of-coefficients method with 1000 bootstrap resamples.

6.4.6 Sensitivity analyses

All sensitivity analyses are presented in **Appendix E: Table E6**. To address reverse causation, excluding participants with less than three years of follow-up slightly attenuated the association for the medium lifestyle score group (0.95 (0.88,1.03)), but the high score group remained protective (0.83 (0.76, 0.90)).

Additional adjustment for the assessment centre did not materially change the estimates (medium: 0.92 (0.85, 0.99); high: 0.81(0.75, 0.87)). In competing-risk regression treating death as a competing event, the subdistribution hazard ratios were virtually identical to the main Cox results (medium: 0.93 (0.86,1.00); high: 0.82 (0.75, 0.89)). Taken together, these sensitivity analyses support the robustness of our findings.

6.5 Discussion

6.5.1 Main findings

In this large prospective population- based cohort study, we observed that a higher combined healthy lifestyle score was associated with a lower risk of developing RA. The protective association was consistent across population subgroups, although it appeared stronger in men, in individuals with lower BMI, and in those recruited from England. Importantly, a substantial proportion of the association between healthy lifestyle and reduced RA risk was mediated by CRP, a well- established marker of systemic inflammation, suggesting that part of the beneficial effect of a healthy lifestyle may operate through lowering chronic low- grade inflammation.

6.5.2 Comparison with previous studies

Prospective evidence on the association between a combined healthy lifestyle score and incident RA is very limited. Only one directly comparable study comes from the Nurses' Health Study (NHS). This study found that a healthy lifestyle index score (including smoking, alcohol, BMI, physical activity, and diet) was associated with a lower risk of RA among US women [32]. Our findings are broadly consistent with those results, demonstrating a similar inverse relationship between higher lifestyle scores and RA risk in UK middle-age population. However, the confidence intervals around its effect estimates of the NHS analysis were relatively wide, reflecting lower statistical precision. In contrast, our study benefits from a five times larger number of incident RA cases more than in the NHS. Furthermore, our subgroup analyses revealed that the protective association was more obvious in men, suggesting potential sex- specific differences.

Another relevant investigation of mediating pathways was conducted within the UK Biobank population. Using a metabolic signature comprising 81 metabolites, that study found that metabolic factors explained a substantial proportion of the protective effect of a healthy lifestyle on RA risk, with Mendelian randomization analyses supporting a causal role [33]. Our study adds to this work by identifying CRP as an additional mediator, indicating that anti-inflammatory effects of healthy lifestyles may contribute independently of metabolic changes. Taken together, multiple biological pathways, including those involved in metabolic regulation and inflammation reduction, are related to RA risk in relation to lifestyle.

6.5.3 Interpretation and implications

The 19% lower relative risk of RA associated with the highest lifestyle score translates into a modest absolute risk reduction, given a relative low incidence of RA. This modest benefit should be weighed against the challenges of adopting multiple lifestyle changes. The mediating effect of CRP, which explained nearly 30% of the association, was not constant over time. The effect was strongest during the first few years and declined thereafter. This time-varying pattern could reflect early inflammatory drivers: baseline CRP might capture a pre-clinical inflammatory state that accelerates RA onset, while later cases are more likely driven by other factors (e.g., autoimmunity or genetic risk). Besides, survivor bias could play a role. Individuals with high CRP and high RA susceptibility may develop the disease early and leave the risk set, leaving a healthier residual population. Both interpretations align with the natural history of RA, where systemic inflammation often precedes clinical arthritis by several years [34]. Thus, lifestyle interventions that lower CRP may be most beneficial when started early, before any pre-clinical signs or symptoms emerge.

A previous UK Biobank study found that a metabolic signature comprising 81 metabolites that strongly correlated with a healthy lifestyle explained around two-thirds of the association with RA incidence [33]. Another study showed that a metabolic signature linked to the Mediterranean diet mediated 22.4% of the diet–RA association in a UK Biobank analysis [35]. Thus, inflammatory and metabolic pathways appear to operate partly independently. The remaining proportion points to other mechanisms, such as gut microbiota, oxidative stress,

and adipokine profiles. Besides, subgroup analyses showed a stronger protective effect in men, leaner individuals, and those living in England. The sex difference is biologically plausible, as women mount stronger pro-inflammatory responses than men, with increased lymphocyte proliferation, activation and pro-inflammatory cytokine production, whereas men display expanded regulatory cell subsets [36]. These differences are largely driven by sex hormones. Oestrogen tends to enhance immune responses, potentially increasing RA susceptibility in women, whereas testosterone exerts immunosuppressive effects that may protect men [37, 38]. Regarding BMI, the stronger effect observed in participants with lower BMI can be interpreted through the lens of obesity-driven inflammation [39]. Adipose tissue in obesity releases elevated levels of pro-inflammatory cytokines, including tumour necrosis factor-alpha (TNF- α), interleukin-6 (IL-6), and IL-1 β , which actively contribute to systemic low-grade inflammation [40]. The regional difference between England and other parts of the UK may be due to chance or residual confounding, given the substantially smaller sample size outside England.

6.5.4 Strengths and limitations

This study has several strengths. A major strength of this study is the large sample size. It allows us to perform the mediation and stratified analyses with sufficient statistical power. The prospective cohort design with long follow-up and linkage to hospital and death records ensured ascertainment of incident RA cases, minimising loss to follow-up and outcome misclassification. The comprehensive data available in the UK Biobank allowed us to construct a multidimensional healthy lifestyle score and to adjust for a wide range of potential confounders, including socio-demographic factors, lifestyle behaviours, comorbidities, and female-specific covariates (menopausal status and HRT use). The availability of baseline CRP measurements enabled mediation analyses using the product-of-coefficients method with bootstrap resampling for quantifying indirect effects in the presence of a mediator. Sensitivity analyses excluding early RA cases, adjusting for assessment centre, and treating death as a competing risk yielded results consistent with the main analysis, supporting the robustness of our findings. Finally, this study included both men and women, providing evidence across sexes.

Our study also has important limitations. First, the healthy lifestyle score was based on self-reported behaviours at a single time point, so we cannot rule out misclassification due to changes in lifestyle during follow-up. Second, some unmeasured factors such as genetic predisposition or early-life exposures could partly account for the observed associations. Third, CRP was measured only at baseline, repeated measurements would have allowed a better characterisation of chronic inflammation. Moreover, the proportional hazards assumption for logCRP was violated, indicating that the effect of CRP on RA risk diminished over time. Therefore, the reported indirect effect and mediation proportion represent average estimates over the follow-up period and should be interpreted with caution. Fourth, the diet score was originally developed for cardiovascular health, so if a score specifically designed for RA might perform differently. Fifth, as noted by Wang and Albert [31], our mediation analysis assumed sequential ignorability, for example, no unmeasured confounding between lifestyle and CRP, CRP and RA, or lifestyle and RA. Despite extensive covariate adjustment, residual confounding cannot be fully excluded. Sixth, the UK Biobank cohort is healthier and predominantly of white European ancestry compared with the general British population, which may limit the generalisability of our findings to other ethnic groups or less healthy populations. Finally, although we conducted a competing-risk analysis for death, the number of deaths during follow-up was limited, and the subdistribution hazard ratios were very similar to the main Cox estimates, suggesting that competing risk did not affect the results.

6.6 Conclusions

Higher healthy lifestyle scores were associated with a lower risk of incident RA, and nearly 30% of this association was explained by CRP levels. Targeting chronic low-grade inflammation through individual modifiable lifestyle factors may help reduce RA risk and inform evidence-based prevention strategies.

6.7 References

1. Gao Y, Zhang Y, Liu X: Rheumatoid arthritis: pathogenesis and therapeutic advances. *MedComm (2020) 2024*, 5(3):e509.
2. Tang K, Zhu L, Shan S, Luo Z, Zhou J, Ying J, Wu J, Shen G, Song P: Global, regional and national trends in the epidemiology of rheumatoid

arthritis from 1990 to 2021: an age-period-cohort effect analysis of the global burden of disease study 2021. *RMD Open* 2025, 11(2).

3. Wu D, Huang Y, Zhao J, Long W, Wang B, Wang Y, Chen H, Wu R: Synovial macrophages drive severe joint destruction in established rheumatoid arthritis. *Sci Rep* 2025, 15(1):12111.
4. Zhang Z, Gao X, Liu S, Wang Q, Wang Y, Hou S, Wang J, Zhang Y: Global, regional, and national epidemiology of rheumatoid arthritis among people aged 20-54 years from 1990 to 2021. *Sci Rep* 2025, 15(1):10736.
5. Qiu Z, Liu G, Hu Y, Wang S, Liu B, Rimm EB, Wang M, Manson JE, Hu FB, Sun Q: Adherence to Healthy Lifestyle and Risk of Cardiometabolic Diseases in Individuals With Hypertension. *JAMA Network Open* 2026, 9(3):e260937-e260937.
6. Kuwahara K, Yamamoto S, Honda T, Nakagawa T, Ishikawa H, Hayashi T, Mizoue T: Improving and maintaining healthy lifestyles are associated with a lower risk of diabetes: A large cohort study. *J Diabetes Investig* 2022, 13(4):714-724.
7. Jacobs S, Klapp R, Shvetsov YB, Kristal BS, Setiawan VW, Le Marchand L, Maskarinec G: Prospective association of a Lifestyle Risk Factor Index with type 2 diabetes in the Multiethnic Cohort. *Eur J Nutr* 2025, 64(5):222.
8. Huang W, Zhang S, Chen W, Wang W, Lin Y, Zhang S, Huang F: The mediating role of biological age acceleration in the associations between lifestyle factors and incident digestive system cancers: A Prospective Cohort Study from the UK Biobank. *J Nutr Health Aging* 2026, 30(6):100846.
9. Dong Y, Greenwood DC, Hardie LJ, Cade JE: Measures of Adiposity and Risk of Rheumatoid Arthritis in Middle-Aged UK Women: A Prospective Cohort Study. *Nutrients* 2025, 17(9).
10. Ridker PMJJotACoC: A test in context: high-sensitivity C-reactive protein. 2016, 67(6):712-723.
11. Pope JE, Choy EH: C-reactive protein and implications in rheumatoid arthritis and associated comorbidities. *Semin Arthritis Rheum* 2021, 51(1):219-229.
12. Erre GL, Cacciapaglia F, Sakellariou G, Manfredi A, Bartoloni E, Viapiana O, Fornaro M, Cauli A, Mangoni AA, Woodman RJ, Palermo BL, Gremese

- E, Cafaro G, Nucera V, Vacchi C, Spinelli FR, Atzeni F, Piga M: C-reactive protein and 10-year cardiovascular risk in rheumatoid arthritis. *Eur J Intern Med* 2022, 104:49-54.
13. Albert MA, Glynn RJ, Ridker PM: Effect of physical activity on serum C-reactive protein. *Am J Cardiol* 2004, 93(2):221-225.
 14. Hastie CE, Haw S, Pell JP: Impact of Smoking Cessation and Lifetime Exposure on C-Reactive Protein. *Nicotine & Tobacco Research* 2008, 10(4):637-642.
 15. Dias JA, Wirfält E, Drake I, Gullberg B, Hedblad B, Persson M, Engström G, Nilsson J, Schiopu A, Fredrikson GNJA: A high quality diet is associated with reduced systemic inflammation in middle-aged individuals. 2015, 238(1):38-44.
 16. Pai JK, Hankinson SE, Thadhani R, Rifai N, Pischon T, Rimm EBJA: Moderate alcohol consumption and lower levels of inflammatory markers in US men and women. 2006, 186(1):113-120.
 17. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, Liu B, Matthews P, Ong G, Pell J, Silman A, Young A, Sprosen T, Peakman T, Collins R: UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015, 12(3):e1001779.
 18. Lachat C, Hawwash D, Ocké MC, Berg C, Forsum E, Hörnell A, Larsson C, Sonestedt E, Wirfält E, Åkesson A, Kolsteren P, Byrnes G, De Keyzer W, Van Camp J, Cade JE, Slimani N, Cevallos M, Egger M, Huybrechts I: Strengthening the Reporting of Observational Studies in Epidemiology-Nutritional Epidemiology (STROBE-nut): An Extension of the STROBE Statement. *PLoS Med* 2016, 13(6):e1002036.
 19. Bradbury KE, Young HJ, Guo W, Key TJ: Dietary assessment in UK Biobank: an evaluation of the performance of the touchscreen dietary questionnaire. *J Nutr Sci* 2018, 7:e6.
 20. NHS Digital: The risks of drinking too much. 2022, [Available from: <https://www.nhs.uk/live-well/alcohol-advice/the-risks-of-drinking-too-much/>].
 21. NHS Digital: Alcohol units. 2024, [Available from: <https://www.nhs.uk/live-well/alcohol-advice/calculating-alcohol-units/>].

22. Organization WH: WHO guidelines on physical activity and sedentary behaviour. 2020, [Available from: <https://www.who.int/publications/i/item/9789240015128>].
23. Chudasama YV, Khunti KK, Zaccardi F, Rowlands AV, Yates T, Gillies CL, Davies MJ, Dhalwani NN: Physical activity, multimorbidity, and life expectancy: a UK Biobank longitudinal study. *BMC Med* 2019, 17(1):108.
24. Zhang YB, Chen C, Pan XF, Guo J, Li Y, Franco OH, Liu G, Pan A: Associations of healthy lifestyle and socioeconomic status with mortality and incident cardiovascular disease: two prospective cohort studies. *Bmj* 2021, 373:n604.
25. Ding J, Schöttker B, Brenner H, Hoffmeister M: Thirteen simple lifestyle scores and risk of cancer, cardiovascular disease, diabetes, and mortality: Prospective cohort study in the UK Biobank. *Int J Cancer* 2025, 157(12):2495-2505.
26. Tsai MC, Yeh TL, Hsu HY, Hsu LY, Lee CC, Tseng PJ, Chien KL: Comparison of four healthy lifestyle scores for predicting cardiovascular events in a national cohort study. *Sci Rep* 2021, 11(1):22146.
27. Baker JF, Billig E, Michaud K, Ibrahim S, Caplan L, Cannon GW, Stokes A, Majithia V, Mikuls TR: Weight Loss, the Obesity Paradox, and the Risk of Death in Rheumatoid Arthritis. *Arthritis Rheumatol* 2015, 67(7):1711-1717.
28. Kishore BKJP: Reverse epidemiology of obesity paradox: Fact or fiction? 2024, 12(21):e70107.
29. Guo L, Huang Y, He J, Li D, Li W, Xiao H, Xu X, Zhang Y, Wang R: Associations of lifestyle characteristics with circulating immune markers in the general population based on NHANES 1999 to 2014. *Sci Rep* 2024, 14(1):13444.
30. Lange T, Hansen JV: Direct and indirect effects in a survival context. *Epidemiology* 2011, 22(4):575-581.
31. Wang W, Albert JM: Causal Mediation Analysis for the Cox Proportional Hazards Model with a Smooth Baseline Hazard Estimator. *J R Stat Soc Ser C Appl Stat* 2017, 66(4):741-757.
32. Hahn J, Malspeis S, Choi MY, Stevens E, Karlson EW, Lu B, Cui J, Yoshida K, Kubzansky L, Sparks JA, Costenbader KH: Association of Healthy

- Lifestyle Behaviors and the Risk of Developing Rheumatoid Arthritis Among Women. *Arthritis Care Res (Hoboken)* 2023, 75(2):272-276.
33. Zhang J, Fang XY, Leng R, Chen HF, Qian TT, Cai YY, Zhang XH, Wang YY, Mu M, Tao XR, Leng RX, Ye DQ: Metabolic Signature of Healthy Lifestyle and Risk of Rheumatoid Arthritis: Observational and Mendelian Randomization Study. *Am J Clin Nutr* 2023, 118(1):183-193.
 34. Arend WP, Firestein GSJNRR: Pre-rheumatoid arthritis: predisposition and transition to clinical synovitis. 2012, 8(10):573-586.
 35. Song X, Ma X, Yang B, Zhang D, Zou Y, Lei B, Xiang R, Zhao X, Qu Y, Zheng S, Yu T, Zhou J, Han T, Zhong Y, Xia M, Alfredsson L, Leander K, Fan M, Jiang X: Mediterranean diet, metabolic signature, genetic predisposition, and risk of rheumatoid arthritis: a large-scale population-based prospective cohort study. *Am J Clin Nutr* 2025, 122(6):1778-1787.
 36. Dodd KC, Menon M: Sex bias in lymphocytes: Implications for autoimmune diseases. 2022, Volume 13 - 2022.
 37. Sciarra F, Campolo F, Franceschini E, Carlomagno F, Venneri MA: Gender-Specific Impact of Sex Hormones on the Immune System. *Int J Mol Sci* 2023, 24(7).
 38. Nowak TJ, Muehlenbein MP: Toward understanding sexual immune dimorphism in humans. 2025, Volume 16 - 2025.
 39. Ellulu MS, Patimah I, Khaza'ai H, Rahmat A, Abed Y: Obesity and inflammation: the linking mechanism and the complications. *Arch Med Sci* 2017, 13(4):851-863.
 40. Uti DE AI, Omang WA, Alum EU, Obeten UN, Udeozor PA, Agada SA, Bawa I, Ogbu CO.: Cytokines as key players in obesity low grade inflammation and related complications. *Obesity Medicine* 2025, 54.

Chapter 7 Summary and Conclusions

Rheumatoid arthritis (RA) is the most common inflammatory arthritis resulting in progressive disability, and increasing morbidity and mortality [1]. Diet is one of the major modifiable factors for RA risk [2], with millions of musculoskeletal disorders more broadly estimated to be attributable to a low intake of healthy foods and a high intake of unhealthy foods [3]. However, evidence on many aspects of diet, including foods, nutrients, and eating patterns remains unclear, hindering the development of evidence-based dietary guidelines. More recently, multi-component and comprehensive healthy lifestyles have gained attention in relation to RA prevention. Nonetheless, epidemiological evidence exploring their combined impact on RA risk, particularly regarding the underlying circulatory inflammatory pathways, remains very limited.

The aim of this thesis:

To investigate associations between dietary factors and rheumatoid arthritis risk in British adults using data linkage.

The preceding chapters describe the approach taken to achieve this aim, and present the findings through five papers:

1. **Paper 1 (Chapter 2)** Dose–response associations between diet and risk of rheumatoid arthritis: A meta-analysis of prospective cohort studies
2. **Paper 2 (Chapter 3)** Measures of adiposity and risk of rheumatoid arthritis in middle-aged UK women: A prospective cohort study
3. **Paper 3 (Chapter 4)** Association of Mediterranean diet with rheumatoid arthritis risk: Evidence from the UK Women’s Cohort Study
4. **Paper 4 (Chapter 5)** Consumption of alcoholic beverages and risk of rheumatoid arthritis: findings from the UK Women’s Cohort Study
5. **Paper 5 (Chapter 6)** Healthy lifestyle score, C-reactive protein levels, and incident rheumatoid arthritis: a prospective cohort study of 354,512 UK Biobank participants

This final chapter brings together and synthesises the evidence generated from the preceding studies (**Chapter 2-6**). It summarises main findings, highlights novel contributions, and critically evaluates the results within the wider literature. Possible mechanisms are explored, followed by a balanced discussion of strengths and limitations. The chapter then turns to public health messages and future research directions, before concluding the thesis.

7.1 Summary of principal findings

Principal findings of this thesis are highlighted in **Table 7.1**. First, the dose-response meta-analysis of prospective cohort studies (**Chapter 2**) showed that higher intakes of total alcohol, fruit, and cereals were associated with a reduced risk of RA, while tea and caffeinated coffee were linked to an increased risk. However, the evidence was limited for most dietary exposures, with moderate-to-low certainty and considerable heterogeneity. Second, in **Chapter 3**, using the UK Women's Cohort Study (UKWCS), we found that central adiposity measures (waist circumference [WC], waist-to-hip ratio [WHR], waist-to-height ratio [WHtR], and skirt size) were stronger predictors of RA risk than BMI, particularly in never-smokers and premenopausal women. Third, in **Chapter 4**, higher adherence to the Mediterranean diet was associated with a lower risk of RA in the UKWCS, although the protective effect was only evident among women with highest adherence category and was attenuated after excluding current smokers. Fourth, in **Chapter 5**, moderate-to-high alcohol intake (>14 units/week) was associated with a decreased RA risk in the UKWCS, and this inverse association was modified by smoking. Finally, in **Chapter 6**, using the UK Biobank, a higher combined healthy lifestyle score (including never smoking, moderate alcohol consumption, sufficient physical activity, and a high-quality diet) was associated with a lower risk of incident RA. The protective effect was stronger in men, leaner individuals, and those recruited from England. CRP partially mediated this association, explaining approximately 30% of the risk reduction, and the mediation effect was strongest during the first few years of follow-up. These findings suggest that promoting multiple healthy lifestyle behaviours may reduce RA risk, partly through lowering systemic inflammation, with potential sex- and adiposity-specific differences.

Table 7.1 Summary of key findings, potential explanations, and novelty in this thesis.

Chapter	Key finding	Potential explanations	Novelty
2	(1) Moderate alcohol intake (up to 60 g/week), higher fruit intake (per 80 g/day), and higher cereal intake (per 30 g/day) were associated with reduced RA risk. (2) Tea and caffeinated coffee were linked to increased risk. (3) Non-linear inverse associations observed for oily fish and vitamin D supplements.	- Alcohol may reduce inflammation via polyphenols (e.g., resveratrol) and modulation of immune response. - Fruits and cereals provide fibre, antioxidants, and anti-inflammatory phytochemicals. - The effects of tea and coffee may be due to the pro-inflammatory effects of high-dose caffeine, or by confounding factors linked to unhealthy dietary habits (e.g., added sugar).	- First comprehensive dose-response meta-analysis covering 32 dietary factors and RA risk. - Included only prospective cohort studies, reducing recall bias. - Provided moderate-certainty evidence for fruit, cereals, and alcohol; also highlighted research gaps for dietary patterns and micronutrients.
3	(4) Higher BMI, waist circumference (WC), waist-to-hip ratio (WHR), waist-to-height ratio (WHtR), and larger clothing size (blouse/skirt) were associated with	- Central adiposity drives low-grade inflammation via pro-inflammatory cytokines (TNF-α, IL-6, CRP) and adipokines (leptin, adiponectin). - Smoking may obscure obesity-RA association.	First study to simultaneously compare BMI, WC, WHR, WHtR, and clothing size (blouse/skirt) for RA risk in a UK female cohort.

	increased RA risk in the UKWCS. (5) WHtR ≥ 0.6 showed the strongest effects. (6) Associations were more stronger in never-smokers, premenopausal women, and those with substantial weight gain.	• Weight gain may influence metabolic and inflammatory consequences of adiposity. • Hormonal changes in premenopausal women may amplify obesity-related inflammation.	• Demonstrated that WHtR and skirt size (self-reported, as practical proxy) are more effective than BMI at identifying women at higher RA risk.
4	(7) Higher adherence to a Mediterranean diet was associated with a 33% lower risk of RA compared to low adherence.	• Mediterranean diet may reduce RA risk through synergistic anti-inflammatory effects (e.g., fibre, antioxidants, polyphenols) and modulation of gut microbiota. • Composite score captures overall pattern better than individual components due to complex interactions.	• First large-scale UK prospective cohort study to examine Mediterranean diet and RA risk in women. • First to use restricted cubic splines to visualise dose-response in this population.
5	(8) Total alcohol intake >14 units/week was associated with lower RA risk	• Ethanol suppresses pro-inflammatory cytokines (TNF-α, IL-1,	• First UK female cohort to examine total alcohol and beverage types (wine, beer, sherry, spirits) with RA risk

	compared with non-drinkers.	IL-6) and inhibits Tfh cell responses. • The higher risk in non-drinkers was partially explained by competing mortality (healthy-survivor bias)	using competing risk models. • Showed that competing mortality explains part of the excess risk in non-drinkers. • Provided evidence for beverage-specific effects (wine, spirits) beyond ethanol.
6	(9) A higher healthy lifestyle score was associated with a lower risk of incident RA. (10) CRP mediated the association, explaining about 30% of the effect. (11) The protective effect was stronger in men, leaner individuals, and those from England.	• Healthy lifestyles reduce systemic inflammation, which is a known driver of RA. • CRP's mediating effect was time-varying (stronger early), suggesting early inflammatory drivers or survivor bias. • Sex differences may reflect hormonal modulation of inflammation; obesity may attenuate additional lifestyle benefit.	• First large-scale UK Biobank study to examine a combined healthy lifestyle score with RA risk and formally test CRP as a mediator.

7.2 Highlights of what is novel in this thesis

1. The first comprehensive meta-analysis of 32 dietary exposures and RA risk using only prospective cohort studies with dose-response analyses, including foods, beverages, nutrients, and supplements, assessing both linear and non-linear associations and evaluating the certainty of evidence.
2. The first prospective study in UK women to simultaneously investigate associations between multiple adiposity indicators and incident RA.
3. The first UK prospective study of Mediterranean diet adherence and RA risk, using restricted cubic splines.
4. The first prospective study in British women to examine total and beverage-specific alcohol intake in relation to rheumatoid arthritis risk, applying the competing risk models to account for deaths from CVD and cancer as competing events.
5. This is the first prospective study in UK Biobank to examine the association between a healthy lifestyle score and incident RA, as well as being the first to quantify CRP as a mediator with a time-varying effect and to investigate effect modification by sex, BMI and region.

7.3 Comparison with previous literature

This section integrates findings from all prior chapters under several topics, highlighting how this thesis advances the evidence base on diet, lifestyle and RA risk.

7.3.1 The meta-analysis

Prior to this thesis, several systematic reviews had examined the association between diet and RA risk, but no study had comprehensively synthesised prospective evidence across a wide range of dietary exposures using dose-response analyses. Two recent systematic reviews focused on dietary patterns and RA risk but did not include individual dietary components in their search strategies [4, 5]. A 2014 meta-analysis of alcohol consumption and RA risk reported a non-linear protective association, but included non-peer-reviewed

literature and did not distinguish between beverage types [6]. For most other dietary factors, including fruits, vegetables, cereals, meat, fish, tea, coffee, and micronutrients, the evidence remained fragmented or inconclusive, with no quantitative synthesis available.

A total of 3,525 potentially relevant reports were identified, from which 55 full-text articles were assessed for eligibility. The final selection included 30 prospective cohort studies, representing 32 distinct dietary exposures in relation to RA risk. Among these associations, total alcohol, beer, fruits, cereals, and vitamin D supplements were associated with a reduced risk of RA, while tea and caffeinated coffee were associated with an increased risk. Most risk reductions ranged from 3% to 10% per increment in intake. Most identified associations were protective ($n = 6$) and a smaller number were harmful ($n = 2$), with moderate-to-low certainty evidence.

Additionally, Chapter 2 informed the original research presented in subsequent chapters of this thesis. The meta-analysis highlighted the need for more robust prospective studies with adequate power, long follow-up durations, and appropriate adjustment for confounders, particularly for dietary patterns and for beverage-specific alcohol analysis. For example, although the meta-analysis indicated that beer consumption was associated with a 10% reduction in risk, this finding was based on evidence of very low certainty from only four studies.

The meta-analysis also identified the absence of mediation analyses examining potential biological pathways linking dietary exposure to RA risk, and the lack of studies investigating effect modification by smoking status, BMI, menopause, or other demographic factors. These identified gaps directly motivated the design of the UKWCS and UK Biobank analyses presented in Chapters 4, 5, and 6, which address the need for detailed dietary pattern analysis (Mediterranean diet, Chapter 4), beverage-specific alcohol assessment (Chapter 5), investigation of mediating pathways such as CRP (Chapter 6), and exploration of effect modification by sex, BMI, and smoking status. To provide a comprehensive picture of nutritional status and its impact on RA, Chapter 3 extends the investigation from specific dietary intakes to anthropometric measures of

adiposity. Thus, the meta-analysis serves as both a foundational synthesis of existing evidence and a roadmap for the original contributions of this thesis.

7.3.2 Dietary assessment

In the meta-analysis presented in Chapter 2, most included studies relied on a single baseline food frequency questionnaire or dietary record to assess dietary intake, which introduced random error and contributed to considerable heterogeneity across studies. Only a few cohorts implemented repeated measurements. The Nurses' Health Study (NHS) I and II repeated dietary measurements every 2–4 years in a subsample [7]. The Swedish Mammography Cohort (SMC) administered detailed FFQs at three time points (1987, 1997, and 2008–2009) and conducted a validation study comparing FFQ results from 129 women against four 7-day dietary records [8], improving the reliability of dietary assessment. These repeated measurement designs could capture dietary intake over up to 10 years and minimise the attenuation of effect estimates due to measurement error. For example, a 2014 study using the SMC found that higher intake of omega-3 fatty acids was associated with reduced RA risk, a finding that may have been attenuated in other cohorts relying on single dietary assessments [9].

Both UK Biobank and the UKWCS collected repeated dietary measurements (up to five 24-hour recalls in UK Biobank [10]; a 4-day food diary at phase 2 in the UKWCS [11]), enabling better characterisation of long-term intake than single-baseline assessments. Nevertheless, both cohorts rely on self-reported dietary data, a limitation that typically attenuates observed associations. Regarding alcohol consumption, the UKWCS used a 217-item FFQ at baseline to assess dietary intake over the preceding 12 months. This was supplemented by additional questions on the typical weekly frequency of consumption of each specific alcoholic beverage. Alcohol unit conversions and portion size definitions for alcohol intake followed UK standards [12]. Energy adjustment employed the all-components model, adjusting for all non-alcohol caloric sources [13]. This approach avoids over-adjustment because alcohol-derived calories may lie on the causal pathway between drinking behaviour and health outcomes; adjusting

for total energy (including alcohol) would therefore remove part of the effect of interest.

Although UK Biobank collected detailed data on different alcoholic beverages (red wine, white wine/champagne, beer/cider, spirits) at baseline, to our knowledge, no published RA study using UK Biobank has utilised these beverage-specific data. The only published UK Biobank analysis of alcohol and RA risk [14] used a simplified variable categorising drinking frequency (“daily or almost daily”, “3–4 times/week”, “1–2 times/week”, “occasionally”, or “never”), without distinguishing between beverage types. Similarly, although EPIC-Norfolk collected beverage-specific alcohol information, its published alcohol-RA analysis reported only total alcohol intake [15]. The UKWCS is the only UK female cohort to have systematically evaluated beverage-specific alcohol consumption in relation to RA risk, which enabled wine to be identified as the primary driver of the observed inverse association.

A further source of heterogeneity in Chapter 2 meta-analysis concerns food classification. Several dietary exposures showed substantial heterogeneity and wide confidence intervals, which may be partly attributable to crude food classifications across different FFQs and dietary assessment methods. Definitions of “fruit”, “vegetables”, or “cereal products” varied between cohorts. Some cohorts, such as the UK Biobank and the Kailuan Study, used broad categories, whereas others, including the NHS [16, 17], and the Women's Health Initiative (WHI) [18], collected more detailed data, enabling classification of specific sub-types such as citrus fruits, cruciferous vegetables, and green leafy vegetables. For fish intake, some studies assessed only “total fish”, whereas others distinguished between oily fish (e.g., salmon, mackerel, sardines) and non-oily fish (e.g., cod, haddock) [19, 20]. This distinction is important because our non-linear dose-response analysis in Chapter 2 revealed a protective association only with oily fish. This finding could only be detected in cohorts that collected disaggregated fish data, such as EPIC-Norfolk, NHS and the WHI. Cruder classification tends to attenuate true associations and increase between-study heterogeneity, as the same exposure category may include foods with divergent biological effects.

7.3.3 Cohort differences

The two British cohorts examined in this thesis, the UKWCS (recruited 1995–1998) and UK Biobank (recruited 2006–2010) differ in their design, population characteristics, and analytical capabilities. UK Biobank, recruited more recently, may reflect population-level shifts in lifestyle behaviours and health policies over time, as outlined in Section 1.2.4. These differences do not cancel each other out but rather contribute complementary evidence on the associations between lifestyle and RA risk.

Sex composition is the most difference between the two cohorts. The UKWCS includes women only, which is consistent with previous prospective studies of RA risk such as the NHS, E3N, SMC, Iowa Women's Health Study (IWHS). By contrast, UK Biobank includes both sexes (48% men and 52% women), enabling the investigation of the effect of sex on modification in Chapter 6. This chapter revealed that adherence to a healthy lifestyle was associated with a 30% lower RA risk in men, compared to a 15% reduction in women (Chapter 6: Table 6.2). Examination of baseline characteristics stratified by sex and lifestyle score (Chapter 6, Appendix E: Table E5) showed that a higher proportion of women (26.6%) achieved high lifestyle scores than men (19.2%), while a higher proportion of men (25.9%) were in the low-scoring group than women (17.1%). This implies that men may have greater scope for lifestyle improvement, potentially leading to more evident risk reductions when healthy behaviours are adopted.

Both cohorts exhibit healthy volunteer bias, though to different degrees. In the UKWCS, only 11% of participants were current smokers and 58% reported taking nutritional supplements at baseline. In UK Biobank, 10% were current smokers (Chapter 6, Table 6.1), and the proportion with college-level education (50.4%) was higher than in the UKWCS (27.1%), reflecting shifting demographics and selection patterns over time. Both cohorts are healthier and more health-conscious than the general UK population in the same era, which is approximately 25% current smokers [21], 35% supplement users [22]. On the one hand, reduced variability in exposure makes it harder to detect true associations because few participants have highly unhealthy lifestyles. Conversely, if health

consciousness is correlated with certain exposures, such as moderate alcohol consumption, and RA detection rates, for example through more frequent healthcare seeking, this could inflate some associations. Despite this uncertainty, the directionally consistent associations observed across both cohorts for alcohol, Mediterranean diet, and healthy lifestyle score argue for the robustness of these findings.

A fundamental strength shared by both cohorts is the method of RA case ascertainment. In both the UKWCS and UK Biobank, incident RA cases were identified through linkage to Hospital Episode Statistics (HES) rather than self-report. This represents a major methodological advance over several earlier prospective studies. For example, the NHS initially relied on self-reported RA confirmed by supplementary questionnaires and medical record review, but not all self-reported cases can be validated. The IWHS and the Kailuan cohort relied partly on self-reported arthritis [23, 24]. Self-report is prone to recall bias and misclassification, particularly for a condition like RA which may be confused with other forms of arthritis. By contrast, HES linkage provides high diagnostic accuracy, reduces selective loss to follow-up, and eliminates recall bias.

The UK Biobank offers unique analytical capabilities due to its large sample size and availability of biomarker data. With 354,512 participants and 4,669 incident RA cases (compared with 23,699 participants and 255 cases in the UKWCS), UK Biobank provided much greater statistical power for subgroup analyses. Neither UKWCS nor previous cohort studies (NHS, EPIC-Oxford, SMC, E3N) had the necessary, full-cohort biomarker data to perform mediation analyses.

The regional discrepancy observed in UK Biobank, where the protective effect of healthy lifestyle score was stronger among participants recruited from England, is likely driven by methodological factors rather than true biological differences. First, data linkage completeness varies slightly by region. UK Biobank's integration with England's HES is the most established, minimising outcome misclassification for incident RA. Second, it comes down to statistical power. England contributed most of the participants and cases, so its results are more precise (narrower confidence intervals) and reliable than those from the much smaller Scottish or Welsh cohorts. Yu et al. [25] also used UK Biobank but

focused on gene–lifestyle interactions, whereas this thesis examines lifestyle factors alone and includes CRP mediation analysis. Their lifestyle score excluded diet, a key component of Chapter 6.

Despite differences in design, population, and historical context across chapters, and although the findings vary, a certain degree of consistency remains across the cohorts.

7.3.4 Dietary exposure: from single food component to dietary pattern

A central theme of this thesis is the expansion of the analytical scope, from the examining individual nutrients and food components (Chapters 2 and 5) to analysing of the Mediterranean dietary pattern (Chapter 4) and developing a healthy lifestyle score (Chapter 6). This approach not only satisfies the research gap identified in Chapter 2, namely the rarity of dietary pattern and lifestyle score examinations, but also reflects the complementary nature of the analysis of individual components and overall patterns.

The single-component approach can identify which specific dietary factors drive an association. For example, Chapter 2 found that fruit and cereals intakes were protective against RA, while tea drinking was harmful. However, this method assumes that each component acts independently and additively [26], an assumption that may not hold in complex biological systems where interactions are common. In contrast, the pattern-based approach captures the properties that cannot be reduced to the sum of individual parts. As shown in Chapter 4, no individual component of the Mediterranean diet was found to be linked to RA risk in isolation, but the diet as a whole was protective, indicating a synergistic effect. This finding aligns with broader evidence that various dietary quality indexes (Healthy Eating Index, Mediterranean Diet Score, DASH) consistently showed that higher diet quality is associated with lower health outcomes across multiple large cohorts [27-29], despite differences in how individual components are scored [30].

The optimal analytical approach depends on the research question. Pattern-based analyses are well-suited to detecting the effects of complex, synergistic systems, such as the effect of diet on health conditions, and to informing

population-level public health messages. Furthermore, as Neuhouwer [31] points out, dietary exposure is inherently long-term and multidimensional. People consume food, not individual nutrients. Therefore, pattern-based methods are better suited to capturing the cumulative effect of a habitual diet over time, whereas single-component approaches may miss the synergistic effects that characterise real-world eating behaviours.

Chapter 6 reported for the first time that the protective effect of a healthy lifestyle score was stronger in men than in women. Biologically, oestrogen enhances pro-inflammatory responses, increasing RA susceptibility in women, while testosterone exerts immunosuppressive effects that may protect men [32].

The association between alcohol consumption and RA risk was significantly modified by smoking status (Chapter 5). Non-drinking current smokers had a 3.72-fold higher risk compared with regular drinkers. This is consistent with the Swedish EIRA study, which reported a three-way interaction between alcohol, smoking, and HLA-DRB1 shared epitope. Our study is the first to confirm this interaction in a UK prospective cohort while accounting for competing risks.

Obesity-related RA risk (assessed by BMI, WC, WHR, WHtR) was stronger in premenopausal women than in postmenopausal women (Chapter 3), supporting the hormonal hypothesis of RA pathogenesis. This may reflect the pro-inflammatory effects of oestrogen, which decline after menopause [33], potentially attenuating the obesity-inflammation-RA pathway in postmenopausal women. Additionally, the metabolic and inflammatory consequences of adiposity may be more pronounced when oestrogen levels are higher [34].

For the Mediterranean diet, no evidence of effect modification was observed by BMI, smoking, age, or physical activity (Chapter 4), suggesting that this dietary pattern confers protection consistently across subgroups.

In summary, hormonal status (sex, menopause) and smoking are important modifiers of RA risk, whereas the Mediterranean diet appears robust across subgroups. However, as the statistical power for the subgroup analyses was limited, the findings should be replicated in larger studies in the future.

7.4 Critique of potential mechanisms

7.4.1 Alcohol intake

As demonstrated in Chapter 2 (meta-analysis) and Chapter 5 (UKWCS), moderate alcohol consumption is associated with a lower risk of RA. The meta-analysis showed that each 2-unit/week increase in total alcohol intake was associated with a 4% reduction in RA risk, with a J-shaped relationship suggesting protective effects up to approximately 60 g/week. The UKWCS analysis further showed that those consuming >14 units/week had a 43% lower risk compared with non-drinkers, an effect that remained robust after accounting for competing risks from CVD and cancer mortality.

In the UKWCS, alcohol was consumed regularly by 67% of the participants (at least once a week), with a mean alcohol consumption of 9.0 units per week recorded among all participants and 13.0 units per week among regular drinkers. Wine was the most popular drink, with 83% of participants consuming it. In the UK Biobank (Chapter 6), mean alcohol consumption was 2.2 units/week. Despite this difference, both cohorts showed inverse associations between alcohol and RA risk.

The apparent decline in protection beyond 60 g/week in the meta-analysis requires cautious interpretation, as it likely reflects data sparsity rather than a true biological threshold. Because most primary studies included in meta-analyses have very limited numbers of high-level drinkers, leading to unstable estimates at the upper end of the distribution. The UKWCS, which included a substantial proportion of women consuming >14 units/week, observed continued protection at this level. The 14 units per week threshold is the guideline set by the UK Chief Medical Officers for low-risk drinking in order to minimise overall health harm [35]. Notably, similar levels of moderate alcohol intake have been associated with a lower risk of sarcopenia in a J-shaped manner [36]. This implies that there are protective effects for certain musculoskeletal outcomes. Conversely, even moderate consumption is consistently linked to an increased risk of breast cancer [37,38]. Therefore, the more accurate conclusion is that while protection for RA may extend into higher intake levels, with insufficient evidence regarding very

high consumption, any public health interpretation must weigh this against alcohol's well-established harms for other diseases.

Inspired by the evidence above, Chapter 6 treated alcohol as an independent component of a composite healthy lifestyle score, alongside smoking, physical activity, and diet. This approach avoids attributing alcohol's potential benefit to other health behaviours [39]. The mediation analysis in Chapter 6 showed that CRP explained 30% of the association between healthy lifestyle score and incident RA. In a cross-sectional study of 2,006 German adults, Imhof et al. observed a U-shaped association between alcohol intake and CRP levels, with moderate drinkers having the lowest concentrations [40]. Thus, while the finding of Chapter 6 is not specific to alcohol, it is consistent with the hypothesis that alcohol's protective effect on RA operates partly via reducing CRP-mediated inflammation.

The UKWCS analysis found that wine, but not beer or spirits, drove the protective association (Chapter 5). In the UKWCS, regular drinkers (the majority of whom consumed wine) had more favourable health profiles compared with non-drinkers: they were younger, had a lower BMI, higher educational attainment, and were less likely to smoke (Chapter 5, Table 5.1). While the models adjusted for these factors, residual confounding remains a concern. Therefore, wine drinking could indicate broader health consciousness, such as more frequent medical check-ups, a higher socioeconomic status and better access to healthcare, rather than being a causal agent itself. Without randomisation, this cannot be conclusively determined. Besides, while wine's polyphenols have documented anti-inflammatory properties [41], the UKWCS lacked data on polyphenol metabolites or biomarkers, preventing direct testing of this pathway.

Ethanol is rapidly absorbed after oral administration and primarily metabolised in the liver, exerting effects directly on immune cells and indirectly via its metabolite acetate. Directly, ethanol inhibits NF- κ B activation and reduces pro-inflammatory cytokines in human monocytes [42]. Indirectly, low-dose ethanol alters gut microbiota composition (increasing *Muribaculaceae*), elevates acetate production, which reduces neutrophil extracellular trap formation via the GPR43 receptor [43], a pathway relevant to citrullinated autoantigen exposure, and limits CD4⁺ T cell migration to inflamed joints [44]. Ethanol also reduces T follicular

helper cell differentiation and IL-21 expression, potentially limiting autoantibody production [45].

However, these pathways were beyond the scope of direct investigation in this thesis. Data on inflammatory biomarkers (CRP, IL-6, TNF- α), gut microbiota, serum acetate, or autoantibodies were not available in the UKWCS. The UK Biobank measured CRP at baseline (with repeat measurements available for a subset), but data on gut microbiota or autoantibodies were not available in the cohort. Consequently, while the observed associations are consistent with these mechanisms, the thesis provides no direct evidence for them.

7.4.2 Mediterranean diet

As demonstrated in Chapter 4, higher adherence to the Mediterranean diet is associated with a lower risk of RA. In the UKWCS, high adherence (score ≥ 7) was associated with a 33% lower RA risk compared with low adherence (score ≤ 3), yet no single dietary component showed an independent association with RA risk. As illustrated in directed acyclic graph (DAG) in Chapter 4 (Figure 4.1), the separation of the overall dietary pattern from its individual components is a defining feature of dietary synergy. The restricted cubic spline analysis revealed a steady inverse trend. The dose-response pattern mirrors findings for other chronic conditions, such as cardiovascular disease [46] and type 2 diabetes [47]. This consistency suggests that the Mediterranean diet may act on core biological pathways common to these diseases, such as inflammation and metabolic regulation, rather than through disease-specific mechanisms.

The protective effect can be underpinned by synergistic anti-inflammatory actions of multiple bioactive compounds. Fatty fish provide n-3 PUFAs (EPA and DHA), which are precursors to pro-resolving mediators that actively reduce inflammation [48]; fruits and vegetables supply polyphenols and carotenoids that reduce oxidative stress and modulate NF- κ B signalling [49]; whole grains and legumes deliver dietary fibre, fermented by gut microbiota into short-chain fatty acids e.g., acetate, propionate, butyrate, with systemic immunomodulatory effects [50,51]. Evidence from the PREDIMED clinical trial supports this synergy. Mediterranean diet interventions supplemented with olive oil or nuts were found to significantly

reduce levels of IL-6, IL-8, CRP and TNF- α [52], an effect that is attributed to cumulative dietary changes rather than any single component.

The Mediterranean diet also can modulate gut microbiota. High adherence to the diet is associated with an increased abundance of bacteria that produce short-chain fatty acids (e.g. *Faecalibacterium prausnitzii*, *Roseburia* and *Bifidobacterium*) and a reduction in pro-inflammatory taxa [53,54]. This enhances intestinal barrier integrity and dampens systemic inflammation, which may be relevant to the pathogenesis of RA, given that gut dysbiosis has been found to precede the onset of clinical symptoms [55].

7.4.3 Central obesity

Adipose tissue, particularly visceral fat, is not merely an energy store but an active endocrine organ. In obesity, hypertrophied adipocytes and infiltrating M1-polarised macrophages promote inflammation through three pathways: (1) secretion of pro-inflammatory cytokines (TNF- α , IL-6, CRP) [56]; (2) dysregulation of adipokines (increased leptin, decreased adiponectin) [57]; and (3) amplification of cytokine production via macrophage recruitment [58]. Mendelian randomisation studies by Vasileiadis et al. [59] found no direct causal effect of circulating adiponectin, leptin, or resistin on RA risk after BMI adjustment [59]. This suggests that these specific adipokines are not independent drivers of RA risk. One possibility is that the association is driven by the broader inflammatory milieu of obesity.

The finding that central obesity measures outperformed BMI aligns with evidence from other disease contexts. In the Osteoarthritis Initiative, central obesity was associated with knee cartilage damage independent of BMI [60], and a systematic review confirmed a consistent detrimental effect of obesity on knee cartilage [61]. A meta-analysis of over 300,000 adults demonstrated that WHtR has superior discriminatory power to BMI for detecting cardiometabolic risk factors [62], and a large cohort study found that central obesity was positively associated with mortality [63]. Visceral adiposity is specifically linked to elevated IL-6 and CRP levels, whereas BMI shows weaker associations with these inflammatory markers [64,65]. This suggests central obesity captures a distinct inflammatory burden not

fully reflected by BMI, which may explain why it is a more relevant risk factor for inflammatory conditions such as RA.

As shown in Chapter 3, the obesity-RA association was stronger in premenopausal than postmenopausal women, and linear associations were only evident in those who had weight gained (>10 kg) since age 20. These findings support that oestrogen enhances pro-inflammatory responses, while its sharp decline at menopause may alter cytokine balance [66]. Weight gain since young adulthood may represent cumulative exposure to obesity-related inflammation [67].

7.4.4 Lifestyle, inflammation, and RA risk

The preceding sections have detailed how individual lifestyle factors: alcohol consumption (Section 7.4.1), the Mediterranean diet (Section 7.4.2), and central obesity (Section 7.4.3) influence RA risk through inflammatory mechanisms. Each factor has been linked to circulating levels of pro-inflammatory cytokines.

Chapter 6 tested whether an overall healthy lifestyle reduces RA risk via this same inflammatory pathway. A healthy lifestyle score, encompassing never smoking, moderate alcohol consumption, sufficient physical activity, and a high-quality diet, was inversely associated with RA risk. This is the first prospective study to formally quantify the extent to which systemic inflammation mediates the lifestyle-RA association. This does not preclude other mechanisms, such as gut microbiota modulation or oxidative stress, but demonstrates that the anti-inflammatory effect is a major, quantifiable component of the overall protection.

Depending on which lifestyle domain offers the greatest room for improvement, different individuals may achieve risk reduction through different behavioural changes. For example, physical activity for the sedentary, smoking cessation for current smokers, and dietary improvement for those with a poor diet. Reduced systemic inflammation is a common endpoint and provides a unifying intervention target, even when the optimal behavioural pathway varies from person to person.

7.5 Critique of potential limitations

7.5.1 The meta-analysis design

As the meta-analysis in Chapter 2 only included prospective cohort studies published up to 2024 [68], recently emerging studies were outside our tracking scope and not identified. Additionally, limitations of the included studies were inherited, such as biases in dietary assessment methods and residual confounding. The majority of included studies were conducted in European and North American populations with predominantly female participants, so findings may not be generalisable to men, to low- or middle-income countries, or to populations with different dietary patterns. For example, Asian populations with higher tea consumption [69], and alcohol metabolism differs due to ALDH2 polymorphisms [70]. The NutriGrade tool was used to evaluate evidence quality [71]. However, the NutriGrade tool automatically downgrades all non-randomised studies to “low-quality” evidence, regardless of how well confounding factors are controlled [71]. This may result in the quality of the evidence being underestimated. Third, for dietary patterns and several nutrients (e.g., anti-inflammatory diet, fibre), fewer than three studies were available, precluding formal meta-analysis. Fourth, substantial heterogeneity was observed for some exposures (e.g., vegetables, n-3 PUFAs), which could not be fully explained by subgroup analyses. Fifth, publication bias cannot be entirely ruled out for coffee consumption, although trim-and-fill analyses yielded consistent results. These limitations are discussed in Chapter 2 (section 2.5.3).

7.5.2 UKWCS and UK Biobank designs

First, both cohorts are subject to healthy volunteer bias. This selection bias may limit generalisability to less healthy populations. However, if healthy volunteer bias reduces exposure variability, this would underestimate the true association, making it harder to detect a real effect. Second, many large prospective cohorts of RA (e.g., NHS, E3N, SMC) are limited to women, as is the UKWCS in this thesis. The women-only cohorts may miss the protective effect of diet or lifestyle interventions in men for RA prevention. Third, the UKWCS and UK Biobank were recruited approximately a decade apart (1995–1998 vs 2006–2010). Dietary

patterns, food availability, and lifestyle behaviours may have changed over this period, potentially affecting the comparability of findings.

7.5.3 Dietary assessment

Both the UKWCS and UK Biobank used FFQs for dietary assessment, but with different instruments, each of which had its own limitations. The UKWCS employed a more comprehensive FFQ (217 items), which is subject to recall bias, systematic under-reporting of energy and nutrient intakes, use of standard portion sizes that fail to capture inter- and intra-individual variation, reliance on food composition tables that do not account for brand or regional differences, and single baseline measurement that cannot capture dietary changes over time, leading to random misclassification and attenuation of effect estimates.

The UK Biobank used a short FFQ (29 items) administered at baseline. Beyond the limitations shared with comprehensive FFQs, its fewer items and coarser frequency categories may introduce additional measurement error. Moreover, the short FFQ was only administered once at baseline for the vast majority of participants, with a smaller subsample ($n = 20,346$) completing a repeat assessment four years later [72]. This limits the ability to adjust for attenuation of effect estimates due to dietary variation within individuals over time.

Despite these limitations, the UKWCS FFQ was validated against 4-day weighed food diaries ($n = 283$) [73], supporting its reasonable accuracy. The UK Biobank short FFQ, while less precise, captured dietary data from over 500,000 participants, maximising statistical power.

We could not directly apply AUDIT-C thresholds [74] as the FFQ did not capture typical drinking day or frequency of 6+ drinks/occasion. However, we were able to look at associations using alcohol intake as both a continuous variable and categorical groups.

7.5.4 Case ascertainment using hospital records

While linkage to HES for RA case ascertainment is a strength, several limitations should be acknowledged. In both the UKWCS and UK Biobank, hospital data were used to identify incident RA cases using ICD-10 codes. However, serological information (ACPA/RF status) was not available in HES, preventing

us from distinguishing seropositive from seronegative RA. This distinction could be important, as previous studies have shown that environmental risk factors have stronger associations with seropositive RA [75]. Additionally, mild RA cases managed entirely in primary care without hospital referral may not be captured in HES, potentially leading to under-ascertainment of less severe cases. However, RA is typically managed by specialist rheumatology care, so the completeness of hospital-recorded diagnoses is likely to be high.

7.5.5 Statistical approaches and residual confounding

Despite covariate adjustment informed by DAGs in analyses, residual confounding remains a possibility in observational studies. In the UKWCS (Chapters 3–5), we could not adjust for family history of RA or genetic predisposition (e.g., HLA-DRB1 status). In the UK Biobank (Chapter 6), while genetic data are available, they were not used in the present analyses. Additionally, information on different types of physical activity (e.g., resistance training, weight-bearing exercise, aerobic exercise) was not available; these may have differential effects on immune function and RA risk [76]. Additionally, relationships between diet and chronic disease risk (including RA) are highly individualised due to genetic and lifestyle factors [77]. Therefore, findings in this thesis are applicable to British adults, but not to individuals *per se*.

In the mediation analysis presented in Chapter 6, the potential for residual confounding is higher, since confounding is possible at the exposure-outcome, mediator-outcome, and exposure-mediator levels [78]. Furthermore, the proportional hazards assumption for logCRP was violated, indicating that the effect of CRP on RA risk diminished over time. Therefore, the reported mediation proportion represents an average estimate over the follow-up period and should be interpreted with caution. Causal mediation analysis also assumes sequential ignorability (no unmeasured confounding between lifestyle and CRP, CRP and RA, or lifestyle and RA), an assumption that cannot be empirically tested.

In complete-case analyses presented in both cohorts, participants with missing covariate data were excluded, potentially introducing selection bias. However, in the UKWCS, characteristics of included and excluded participants were similar.

In UK Biobank, the amount of missing data per covariate was broadly similar across lifestyle score groups, reducing concerns about selection bias.

7.5.6 Generalisability

The findings of this thesis may not be generalisable to all populations. First, both cohorts consist predominantly of white Caucasian participants, limiting generalisability to other ethnic groups. Second, both cohorts are subject to healthy volunteer bias. Participants are generally healthier than the general UK population, with lower smoking rates and higher educational attainment [79]. Third, UKWCS includes only women, so findings from Chapters 3–5 may not directly apply to men. This is relevant given finding of sex-specific differences from UK Biobank, where the protective association of healthy lifestyle was stronger in men. Fourth, UKWCS data (1995–1998) may not fully reflect contemporary diet-disease relationships, as the food environment and dietary patterns have changed substantially over the past two decades. However, consistency with the more recently recruited UK Biobank (2006–2010) suggests the observed associations are robust to temporal changes.

7.6 Critique of study strengths

7.6.1 The meta-analysis design

The meta-analysis in Chapter 2 provides the first comprehensive synthesis of 32 dietary exposures in relation to RA risk using only prospective cohort studies with dose-response analyses to quantify the magnitude and shape of associations. Both linear and non-linear associations were systematically assessed using restricted cubic splines, which improve precision and avoid arbitrary categorisation, and the certainty of evidence for each association was evaluated using the NutriGrade tool. By identifying gaps and limitations in previously published evidence, such as the scarcity of data on dietary patterns and beverage-specific alcohol consumption, this meta-analysis highlighted the need for further cohort studies, informing the original research presented in Chapters 3–6.

7.6.2 UKWCS and UK Biobank designs

A major strength of this thesis lies in its use of two large, prospective, population-based cohorts, which ensured that dietary assessment occurred before the onset of RA, reducing the risk of reverse causation. In both cohorts, effect estimates were robust to sensitivity analyses where RA that occurred within the first three years of follow-up were excluded, showing no evidence of reverse causation for all associations observed. The UKWCS offers detailed dietary assessment via a validated 217-item FFQ and long follow-up. The UK Biobank includes both sexes, enabling sex-specific analyses, and provides biomarker data allowing mediation analysis. Both cohorts have larger sample sizes than previous studies, providing greater statistical power to estimate RA risk.

The two cohorts complement each other. For example, the UK Biobank collected dietary information at recruitment via a touchscreen FFQ, which covered a limited number of food items. Repeated 24-hour recalls (Oxford WebQ) were available, but were not used in this analysis because they only included half of the study population. In contrast, the UKWCS collected detailed dietary information through an FFQ from all participants at recruitment, enabling the derivation of food and nutrient intakes and adjustment for energy intake in analyses of food and nutrient intake.

7.6.3 Dietary assessment

In the UKWCS, the FFQ was validated through comparison with 4-day weighed food diaries and a repeated FFQ administration among a subsample of women in 2001 (n=283) [73]. In the UK Biobank, the touchscreen short FFQ demonstrated good repeatability. Among 20,348 participants who repeated the assessment after 4 years, over 90% reported the same or adjacent category for meat and fish, and over 70% for fruit and vegetables [80].

7.6.4 Case ascertainment using hospital records

In both cohorts, RA cases were identified through linkage to HES, which is a methodological improvement on previous studies that relied on self-reported RA. This approach minimises recall bias and selective loss to follow-up, while also

permitting complete case ascertainment. It also provides the exact date of RA diagnosis, which is essential for accurate time-to-event analysis.

7.6.5 Statistical approaches

To ensure robust causal inference, all diet–RA associations were adjusted for potential confounders identified through DAGs, enabling clear and transparent specification of confounders, moderators, mediators, and competing risks. A stepwise adjustment strategy was applied: Chapter 3 presented three sequential models: unadjusted (Model 1), adjusted for demographic and reproductive factors (Model 2), and further adjusted for lifestyle factors (Model 3) to assess whether associations were independent of health behaviours. In Chapter 5, a mutually adjusted model was used for beverage-specific analyses, additionally adjusting for all other alcoholic beverage types. Subgroup analyses were conducted across both cohorts to examine effect modification by age, sex, BMI, menopausal status, smoking, physical activity, and weight change. Competing risks were addressed using Fine-Gray models, which attenuated the elevated risk among non- drinkers after accounting for competing mortality. Prevalent RA cases were excluded at baseline, and sensitivity analyses excluding early RA cases showed no evidence of reverse causation. Finally, in Chapter 6, mediation analysis using the product-of-coefficients method with bootstrap resampling quantified the proportion of the lifestyle–RA association explained by CRP.

7.6.6 Generalisability

This thesis adopted a population-based approach to examine the relationship between dietary factors and RA risk in British adults. The consistency of the findings across two cohorts that were recruited a decade apart using different dietary assessment methods and that had different population characteristics strengthens the confidence that the observed associations are not artefacts of a single cohort.

7.7 Recommendations for further research

The findings of this thesis highlight several important gaps that should be discussed with patient representative groups and addressed in future research.

7.7.1 Establish causality and underpinning mechanisms

Due to the observational nature of the studies presented in this thesis, causality cannot be inferred. Although extensive covariate adjustment, sensitivity analyses excluding early cases, and competing risk models were applied, residual confounding and reverse causation remain possible. Future studies should employ molecular genetics, such as Mendelian randomisation, to strengthen causal inference.

Beyond causal inference, this thesis could not fully test the proposed biological pathways due to the absence of relevant biomarker data. In the UKWCS, inflammatory markers and gut microbiota composition were not measured. While Chapter 6 used baseline CRP in mediation analysis, repeated measurements and other biomarkers remain unavailable. Future prospective studies should integrate repeated measurements of these biomarkers to test these pathways such as the gut–joint axis [81]. Randomised controlled trials of moderate wine consumption assessing changes in gut microbiota, inflammatory markers, and autoantibody profiles would be particularly valuable.

7.7.2 Improve dietary measurement accuracy and precision

A limitation across nutritional epidemiology to date, including this thesis, is the reliance on single baseline dietary assessment, which cannot capture changes in dietary habits over time [82]. Future studies should implement repeated dietary measurements to better characterise long-term exposures. Novel technology-based approaches offer promising solutions. For example, the web-based dietary record tool “MyFood24” enables users to record food consumption in real-time and calculates energy and nutrient intakes from a comprehensive food composition database, with validation against biomarkers [83]. This approach reduces recall bias and allows for repeated measurements over time, enabling future studies to examine whether changes in diet are associated with changes in RA risk.

7.7.3 Investigate RA subtypes and other outcomes

This thesis examined RA as a single outcome, but RA has distinct subtypes (seropositive and seronegative) with potentially different aetiologies. In the

UKWCS and UK Biobank, serological information, e.g., ACPA and RF status, was not fully available, which prevented subtype-specific analyses. Given that smoking has stronger associations with seropositive RA, future studies should investigate whether dietary factors also differ in their associations. Such analyses would require cohorts with linked serological data or biobank samples, such as the UK Biobank ACPA/RF data for a subset of participants.

Furthermore, this thesis focused exclusively on RA incidence. However, diet may also influence RA disease activity, progression, and response to treatment. Future studies should examine whether the dietary patterns identified as protective for incidence also improve outcomes in patients with established RA, including disease activity scores, physical function, and quality of life.

7.7.4 Expand evidence to diverse populations

Previous cohort studies and cohorts used in this thesis consist predominantly of white Caucasian participants, limiting generalisability to other ethnic groups. Dietary composition differs across countries according to cultural, economic, and agricultural factors [84,85]. Future studies in diverse populations are needed to determine whether the associations observed in this thesis are universal or context-specific. Additionally, the UKWCS included only women, and while UK Biobank enabled sex-specific analyses, further studies with adequate power for sex-stratified analyses are needed to confirm the stronger protective effect of healthy lifestyles in men observed in Chapter 6.

7.7.5 Address dietary exposures with insufficient evidence

The meta-analysis in Chapter 2 identified several dietary exposures for which evidence remained insufficient to draw conclusions, including dietary fibre, micronutrients (e.g., vitamin D, vitamin E, carotenoids), saturated fat, and n-3 polyunsaturated fatty acids (PUFAs), and dietary patterns. For these exposures, data were either too limited to enable pooling or derived from only a small number of studies with low or very low certainty evidence. Future prospective studies should prioritise these under-investigated dietary factors, using standardised intake assessments and consistent outcome definitions to enable quantitative synthesis. In particular, given the biological plausibility linking n-3 PUFAs to anti-

inflammatory pathways, well-powered cohort studies with repeated dietary measurements and adjustment for key confounders are needed to determine whether higher intakes of oily fish or n-3 supplements are associated with lower RA risk.

7.7.6 Elucidate the role of diet across the preclinical stages of RA

RA development is a multi-stage process. It progresses from genetic susceptibility (Stage 1) and systemic autoimmunity (Stage 2: elevated ACPA/RF without symptoms) to clinically suspected arthralgia (Stage 3), unclassified arthritis (Stage 4) and, finally, classifiable RA (Stage 5). Subsequent disease progression occurs in Stage 6 [86,87]. The observational cohorts used in this thesis, including UKWCS and UK Biobank, captured RA at the point of clinical diagnosis (Stage 5), but could not examine whether dietary factors influence earlier, preclinical stages of disease development.

Different dietary components may exert effects at distinct stages of this continuum. For example, alcohol intake or Mediterranean diet adherence might reduce the initial breach of immune tolerance, whereas central adiposity might promote the transition from autoimmunity to inflammatory arthritis via adipose-derived cytokines. Elucidating these stage-specific effects would require prospective cohorts with serial measurements of autoantibodies, inflammatory markers, and joint symptoms in at-risk populations, such as first-degree relatives of RA patients. Such studies could determine whether dietary interventions are most effective when implemented before the onset of autoantibody positivity or after the emergence of clinical symptoms, thereby informing the optimal timing of prevention strategies.

7.8 Public health messages

While novel evidence across two large British cohorts, findings of this thesis are observational and cannot establish causality. Therefore, firm recommendations for policymakers, clinicians, or the public cannot be made. However, the above section (Chapter 7: section 7.7) outlines several potential considerations for practice, which require confirmation in future before they can be integrated into clinical guidelines.

7.8.1 Fruit, cereals, and tea

The meta-analysis in Chapter 2 provided moderate-certainty evidence that higher fruit intake (per 80 g/day, equivalent to one portion of the 5-A-Day recommendation) and higher whole grain intake (per 30 g/day) are associated with lower RA risk, while higher tea consumption (per cup/day) is associated with increased risk. Although fruit intake findings are based on moderate-certainty evidence and causality remains uncertain, they align with existing dietary guidelines [88]. This thesis suggest that meeting 5-A-Day recommendation may also confer benefits for RA prevention. For tea, moderate-certainty evidence suggests a positive association with RA risk, although findings from other studies remain inconsistent [89,90]. While further research is needed, if the association is causal, individuals with high tea consumption may consider moderating their intake, particularly if they have other RA risk factors.

7.8.2 Maintain a healthy central adiposity

Analysis for central obesity in Chapter 3 demonstrated that waist-to-height ratio was a stronger predictor of RA risk than BMI-defined obesity, with skirt size showing a similar magnitude of association. These findings suggest that simple, low-cost measures, measuring waist-to-height ratio or even clothing size, could help identify women at higher RA risk in primary care settings. Public health messages should emphasise maintaining a healthy waist circumference for RA prevention. The strongest associations were observed in premenopausal women and those with substantial weight gain since early adulthood, suggesting that weight management efforts may be particularly beneficial if initiated before menopause and that preventing weight gain across adulthood could be important for RA prevention.

7.8.3 Adherence to a Mediterranean-style diet

In the UKWCS, women with high versus low adherence to the Mediterranean diet had a 33% lower risk of RA. No single dietary component was independently associated with RA risk, suggesting that the protective effect derives from the overall dietary pattern rather than individual food component. Since the Mediterranean diet is already recommended for the prevention of cardiovascular

disease and type 2 diabetes, the findings of this thesis add RA to the list of conditions that may be prevented.

The protective effect of the Mediterranean diet was independent of BMI, smoking, and physical activity. However, adopting a Mediterranean dietary pattern may be challenging for individuals accustomed to Western-style diets. Public health messages should therefore focus on promoting incremental changes, such as increasing intakes of vegetables, legumes, whole grains, and fish, while reducing red and processed meat consumption. For individuals already following a Mediterranean diet, these findings provide reassurance of its potential benefits for RA prevention.

7.8.4 Moderate alcohol consumption

Chapter 5 showed that moderate alcohol consumption, particularly wine, was associated with a lower RA risk. However, this association was partially explained by healthy survivor bias and reverse causation, and sensitivity analyses excluding participants who had reduced alcohol intake for health reasons eliminated the protective effect for high alcohol intake (>14 units/week). Given the well-established harms of alcohol, including increased risks of cancer, liver disease, and cardiovascular disease at higher intakes, alcohol should not be recommended for RA prevention. The findings of this thesis do, however, suggest that for individuals who already consume alcohol moderately, wine may be a better choice than beer or spirits, but this observational finding requires confirmation in randomised trials before any recommendation can be made.

7.8.5 Adopt a healthy lifestyle

Chapter 6 demonstrated that adherence to a healthy lifestyle, encompassing never smoking, moderate alcohol consumption, sufficient physical activity, and a high-quality diet was associated with a 19% lower RA risk. Approximately 30% of this association was explained by lower systemic inflammation. This suggests that lifestyle modification may reduce RA risk, at least in part, through anti-inflammatory pathways. The stronger protective effect observed in men than in women suggests that targeted lifestyle interventions for men, who in this cohort had higher smoking rates and lower physical activity at baseline, could yield

substantial population-level benefits. Public health messages should encourage the adoption of multiple healthy behaviours simultaneously, because the protective effect of the combined lifestyle score exceeded the sum of individual components.

7.8.6 Balancing benefits and risks

As with all public health messages, the potential benefits of dietary modifications for RA prevention must be weighed up against other health outcomes. For example, while the Mediterranean diet is associated with lower RA risk, it is also associated with lower risks of cardiovascular disease [91,92], type 2 diabetes [93], and certain cancers [94]. Conversely, while moderate alcohol consumption was associated with lower RA risk in this thesis, alcohol is an established risk factor for breast cancer [37,95], liver disease [96], and other conditions [97]. Therefore, any public health message must be considered in the context of an individual's overall risk profile. For most individuals, adopting a Mediterranean-style diet, maintaining a healthy weight (particularly avoiding central obesity), and following a healthy lifestyle (never smoking, regular physical activity) are likely to confer broad health benefits with minimal risks. Alcohol should not be recommended for RA prevention, but for individuals who already drink moderately, wine may be preferred over beer or spirits, although moderation remains essential.

7.9 Conclusions

7.9.1 What was already known

RA is a common chronic autoimmune disease characterised by persistent joint inflammation, pain, and progressive functional decline. While genetic susceptibility and smoking are well-established risk factors, the role of modifiable lifestyle factors, particularly diet, has remained unclear. Prior to this thesis, no comprehensive synthesis of prospective evidence across multiple dietary exposures had been conducted, and evidence for specific factors was fragmented and inconsistent. Studies of dietary patterns, such as the Mediterranean diet, were limited to a few cohorts with conflicting results, and beverage-specific alcohol analyses were scarce. No study had formally tested whether systemic inflammation mediates the lifestyle-RA association.

Furthermore, the role of central adiposity had been under-investigated, and effect modification by sex, menopausal status, and smoking had not been systematically examined.

7.9.2 What this thesis adds

This thesis provides novel evidence from two large British cohorts supporting a role for diet-related lifestyle factors in RA prevention. For the first time in British women, associations for Mediterranean diet adherence, moderate wine consumption, and central adiposity were observed with RA risk. This thesis also strengthens the evidence that systemic inflammation mediates the protective effect of a healthy lifestyle score on RA risk, an effect that is stronger in men. This work highlights the need for further research into underlying mechanisms to confirm causality and inform evidence-based guidelines for RA prevention.

7.10 References

1. Smolen JS, Aletaha D, McInnes IB: Rheumatoid arthritis. *Lancet*. 2016;388(10055):2023–2038.
2. Nikiphorou E, Philippou E: Nutrition and its role in prevention and management of rheumatoid arthritis. *Autoimmun Rev*. 2023;22(7):103333.
3. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2019;393(10184):1958–1972.
4. Bäcklund R, Drake I, Bergström U, Compagno M, Sonestedt E, Turesson C: Diet and the risk of rheumatoid arthritis - A systematic literature review. *Semin Arthritis Rheum*. 2023;58:152118.
5. Guan CM, Beg S: Diet as a Risk Factor for Rheumatoid Arthritis. *Cureus*. 2023;15(5):e39273.
6. Jin Z, Xiang C, Cai Q, Wei X, He J: Alcohol consumption as a preventive factor for developing rheumatoid arthritis: a dose-response meta-analysis of prospective studies. *Ann Rheum Dis*. 2014;73(11):1962–1967.
7. Hu FB, Stampfer MJ, Rimm E, et al: Dietary fat and coronary heart disease: a comparison of approaches for adjusting for total energy intake and modeling repeated dietary measurements. *Am J Epidemiol*. 1999;149(6):531–540.

8. Khani BR, Ye W, Terry P, Wolk A: Reproducibility and validity of major dietary patterns among Swedish women assessed with a food-frequency questionnaire. *J Nutr*. 2004;134(6):1541–1545.
9. Di Giuseppe D, Wallin A, Bottai M, Askling J, Wolk A: Long-term intake of dietary long-chain n-3 polyunsaturated fatty acids and risk of rheumatoid arthritis: a prospective cohort study of women. *Ann Rheum Dis*. 2014;73(11):1949–1953.
10. Cole JB, Westerman KE, Manning AK, Florez JC, Hirschhorn JN: Genetic heritability as a tool to evaluate the precision of 24-hour recall dietary questionnaire variables in UK Biobank. *Front Genet*. 2022;13:1070511.
11. Cade JE, Burley VJ, Alwan NA, et al: Cohort Profile: The UK Women's Cohort Study (UKWCS). *Int J Epidemiol*. 2017;46(2):e11.
12. NHS Digital: Alcohol units. 2024;[Available from: <https://www.nhs.uk/live-well/alcohol-advice/calculating-alcohol-units/>]
13. Tomova GD, Arnold KF, Gilthorpe MS, Tennant PWG: Adjustment for energy intake in nutritional research: a causal inference perspective. *Am J Clin Nutr*. 2022;115(1):189–198.
14. Mazzucca CB, Scotti L, Cappellano G, Barone-Adesi F, Chiocchetti A: Nutrition and Rheumatoid Arthritis Onset: A Prospective Analysis Using the UK Biobank. *Nutrients*. 2022;14(8)
15. Lahiri M, Luben RN, Morgan C, et al: Using lifestyle factors to identify individuals at higher risk of inflammatory polyarthritis (results from the European Prospective Investigation of Cancer-Norfolk and the Norfolk Arthritis Register--the EPIC-2-NOAR Study). *Ann Rheum Dis*. 2014;73(1):219–226.
16. Bazzano LA, Li TY, Joshipura KJ, Hu FB: Intake of fruit, vegetables, and fruit juices and risk of diabetes in women. *Diabetes Care*. 2008;31(7):1311–1317.
17. Khandpur N, Rossato S, Drouin-Chartier JP, et al: Categorising ultra-processed foods in large-scale cohort studies: evidence from the Nurses' Health Studies, the Health Professionals Follow-up Study, and the Growing Up Today Study. *J Nutr Sci*. 2021;10:e77.
18. Prentice RL, Howard BV, Van Horn L, et al: Nutritional epidemiology and the Women's Health Initiative: a review. *Am J Clin Nutr*. 2021;113(5):1083–1092.
19. Patel PS, Sharp SJ, Luben RN, et al: Association between type of dietary fish and seafood intake and the risk of incident type 2 diabetes: the European

- prospective investigation of cancer (EPIC)-Norfolk cohort study. *Diabetes Care*. 2009;32(10):1857–1863.
20. Krok-Schoen JL, Brasky TM, Hunt RP, et al: Dietary Long-Chain n-3 Fatty Acid Intake and Arthritis Risk in the Women's Health Initiative. *J Acad Nutr Diet*. 2018;118(11):2057–2069.
21. NHS Digital: Statistics on Smoking - England, 2006. 2006;[Available from: <https://digital.nhs.uk/data-and-information/publications/statistical/statistics-on-smoking/statistics-on-smoking-england-2006>]
22. Harrison RA, Holt D, Pattison DJ, Elton PJ: Are those in need taking dietary supplements? A survey of 21 923 adults. *Br J Nutr*. 2004;91(4):617–623.
23. VanEvery H, Yang W, Olsen N, et al: Alcohol Consumption and Risk of Rheumatoid Arthritis among Chinese Adults: A Prospective Study. *Nutrients*. 2021;13(7)
24. Merlino LA, Curtis J, Mikuls TR, Cerhan JR, Criswell LA, Saag KG: Vitamin D intake is inversely associated with rheumatoid arthritis: results from the Iowa Women's Health Study. *Arthritis Rheum*. 2004;50(1):72–77.
25. Yu X-H, Bo L, Cao R-R, et al: Systematic Evaluation of Rheumatoid Arthritis Risk by Integrating Lifestyle Factors and Genetic Risk Scores. *Front Immunol*. 2022;Volume 13 - 2022
26. Jacques PF, Tucker KL: Are dietary patterns useful for understanding the role of diet in chronic disease? *Am J Clin Nutr*. 2001;73(1):1–2.
27. Brlek A, Gregorič M: Diet quality indices and their associations with all-cause mortality, CVD and type 2 diabetes mellitus: an umbrella review. *Br J Nutr*. 2023;130(4):709–718.
28. Morze J, Danielewicz A, Hoffmann G, Schwingshackl L: Diet Quality as Assessed by the Healthy Eating Index, Alternate Healthy Eating Index, Dietary Approaches to Stop Hypertension Score, and Health Outcomes: A Second Update of a Systematic Review and Meta-Analysis of Cohort Studies. *J Acad Nutr Diet*. 2020;120(12):1998–2031.e1915.
29. Taylor RM, Haslam RL, Herbert J, et al: Diet quality and cardiovascular outcomes: A systematic review and meta-analysis of cohort studies. *Nutr Diet*. 2024;81(1):35–50.
30. Lin J, Wang Q, Liu X, et al: Causal Inference Framework Reveals Mediterranean Diet Superiority and Inflammatory Mediation Pathways in Mortality

- Prevention: A Comparative Analysis of Nine Common Dietary Patterns. *Foods*. 2025;14(17)
31. Neuhouwer ML: The importance of healthy dietary patterns in chronic disease prevention. *Nutr Res*. 2019;70:3–6.
 32. Cutolo M, Serio B, Villaggio B, Pizzorni C, Cravetto C, Sulli A: Androgens and estrogens modulate the immune and inflammatory responses in rheumatoid arthritis. *Ann N Y Acad Sci*. 2002;966:131–142.
 33. Cutolo M, Gotelli E: Complex role of oestrogens in the risk and severity of rheumatoid arthritis in menopause. *RMD Open*. 2023;9
 34. Sophian A: Estrogen deficiency-induced lipid dysregulation in menopause: Mechanisms, metabolic consequences, and therapeutic strategies. *J Steroid Biochem Mol Biol*. 2026;265:107102.
 35. NHS Digital: Health Survey England Additional Analyses, Ethnicity and Health, 2011-2019 Experimental statistics: drinking alcohol. 2022;[Available from: <https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-england-additional-analyses/ethnicity-and-health-2011-2019-experimental-statistics/drinking-alcohol>]
 36. Bu YL, Wang C, Zhao C, Lu X, Gao W: The association of alcohol consumption with the risk of sarcopenia: a dose-response meta-analysis. *Am J Drug Alcohol Abuse*. 2024;50(3):305–320.
 37. Zhang SM, Lee I-M, Manson JE, Cook NR, Willett WC, Buring JEJAJoe: Alcohol consumption and breast cancer risk in the Women's Health Study. 2007;165(6):667–676.
 38. Chen WY, Rosner B, Hankinson SE, Colditz GA, Willett WCJJ: Moderate alcohol consumption during adult life, drinking patterns, and breast cancer risk. 2011;306(17):1884–1890.
 39. Hahn J, Malspeis S, Choi MY, et al: Association of healthy lifestyle behaviors and the risk of developing rheumatoid arthritis among women. 2023;75(2):272–276.
 40. Imhof A, Froehlich M, Brenner H, Boeing H, Pepys MB, Koenig W: Effect of alcohol consumption on systemic markers of inflammation. *Lancet*. 2001;357(9258):763–767.

41. Magrone T, Magrone M, Russo MA, Jirillo EJA: Recent advances on the anti-inflammatory and antioxidant properties of red grape polyphenols: in vitro and in vivo studies. 2019;9(1):35.
42. Azizov V, Zaiss MM: Alcohol Consumption in Rheumatoid Arthritis: A Path through the Immune System. *Nutrients*. 2021;13(4)
43. Jin L, Zhang Z, Pan P, et al: Low-dose ethanol consumption inhibits neutrophil extracellular traps formation to alleviate rheumatoid arthritis. 2023;6(1):1088.
44. Terracina S, Caronti B, Lucarelli M, et al: Alcohol consumption and autoimmune diseases. 2025;26(2):845.
45. Azizov V, Dietel K, Steffen F, et al: Ethanol consumption inhibits TFH cell responses and the development of autoimmune arthritis. 2020;11(1):1998.
46. Rosato V, Temple NJ, La Vecchia C, Castellan G, Tavani A, Guercio VJEjon: Mediterranean diet and cardiovascular disease: a systematic review and meta-analysis of observational studies. 2019;58(1):173–191.
47. Sarsangi P, Salehi-Abargouei A, Ebrahimpour-Koujan S, Esmailzadeh AJAiN: Association between adherence to the Mediterranean diet and risk of type 2 diabetes: an updated systematic review and dose–response meta-analysis of prospective cohort studies. 2022;13(5):1787–1798.
48. Harwood JLJIJoMS: Polyunsaturated fatty acids: conversion to lipid mediators, roles in inflammatory diseases and dietary sources. 2023;24(10):8838.
49. Zhang H, Tsao RJCOiFS: Dietary polyphenols, oxidative stress and antioxidant and anti-inflammatory effects. 2016;8:33–42.
50. Tannock GW, Liu YJJotRSoNZ: Guided dietary fibre intake as a means of directing short-chain fatty acid production by the gut microbiota. 2020;50(3):434–455.
51. Simpson HL, Campbell BJJAp, therapeutics: dietary fibre–microbiota interactions. 2015;42(2):158–179.
52. Urpi-Sarda M, Casas R, Sacanella E, et al: The 3-Year Effect of the Mediterranean Diet Intervention on Inflammatory Biomarkers Related to Cardiovascular Disease. *Biomedicines*. 2021;9(8)
53. Garcia-Mantrana I, Selma-Royo M, Alcantara C, Collado MCJFim: Shifts on gut microbiota associated to mediterranean diet adherence and specific dietary intakes on general adult population. 2018;9:890.

54. De Filippis F, Pellegrini N, Vannini L, et al: High-level adherence to a Mediterranean diet beneficially impacts the gut microbiota and associated metabolome. 2016;65(11):1812–1821.
55. Horta-Baas G, Romero-Figueroa MdS, Montiel-Jarquín AJ, Pizano-Zárate ML, García-Mena J, Ramírez-Durán NJJoir: Intestinal dysbiosis and rheumatoid arthritis: a link between gut microbiota and the pathogenesis of rheumatoid arthritis. 2017;2017(1):4835189.
56. Coppack SW: Pro-inflammatory cytokines and adipose tissue. *Proc Nutr Soc.* 2001;60(3):349–356.
57. Stofkova AJER: Leptin and adiponectin: from energy and metabolic dysbalance to inflammation and autoimmunity. 2009;43(4):157–168.
58. Udalova IA, Mantovani A, Feldmann MJNRR: Macrophage heterogeneity in the context of rheumatoid arthritis. 2016;12(8):472–485.
59. Vasileiadis GK, Sayols S, Zhao SS, Fatima T, Maglio C: Adipokines and risk of rheumatoid arthritis: A two-sample multivariable Mendelian randomisation study. *PLoS One.* 2023;18(6):e0286981.
60. Joseph GB, Firouzabadi FD, Lynch JA, et al: Evaluation of the Association between Central Obesity and MRI-based Biomarkers of Knee Osteoarthritis Independent of BMI: a Cross-Sectional Study from the OAI database. 2026:100810.
61. Mezhov V, Ciccutini F, Hanna F, et al: Does obesity affect knee cartilage? A systematic review of magnetic resonance imaging data. 2014;15(2):143–157.
62. Ashwell M, Gunn P, Gibson SJOr: Waist-to-height ratio is a better screening tool than waist circumference and BMI for adult cardiometabolic risk factors: systematic review and meta-analysis. 2012;13(3):275–286.
63. Xia X, Chen S, Tian X, et al: Roles of general and central adiposity in cardiometabolic multimorbidity: revisiting the obesity paradox using a multistate model. 2024;32(4):810–821.
64. Park HS, Park JY, Yu RJDr, practice c: Relationship of obesity and visceral adiposity with serum concentrations of CRP, TNF- α and IL-6. 2005;69(1):29–35.
65. Khaodhiar L, Ling PR, Blackburn GL, Bistran BRJJoP, Nutrition E: Serum levels of interleukin-6 and C-reactive protein correlate with body mass index across the broad range of obesity. 2004;28(6):410–415.

66. Motta F, Di Simone N, Selmi C: The Impact of Menopause on Autoimmune and Rheumatic Diseases. *Clin Rev Allergy Immunol*. 2025;68(1):32.
67. Nan K, Zhang M, Hu S, et al: Relationship of weight change patterns from young to middle adulthood with incident rheumatoid arthritis and osteoarthritis: a retrospective cohort study. *Front Endocrinol (Lausanne)*. 2023;14:1308254.
68. Dong Y, Greenwood DC, Webster J, et al: Dose-Response Associations Between Diet and Risk of Rheumatoid Arthritis: A Meta-Analysis of Prospective Cohort Studies. *Nutrients*. 2024;16(23)
69. Xie X, Fu J, Gou W, et al: Potential mechanism of tea for treating osteoporosis, osteoarthritis, and rheumatoid arthritis. *Front Med (Lausanne)*. 2024;11:1289777.
70. Chang TG, Yen TT, Wei CY, Hsiao TH, Chen IC: Impacts of ADH1B rs1229984 and ALDH2 rs671 polymorphisms on risks of alcohol-related disorder and cancer. *Cancer Med*. 2023;12(1):747–759.
71. Schwingshackl L, Knueppel S, Schwedhelm C, et al: Perspective: NutriGrade: a scoring system to assess and judge the meta-evidence of randomized controlled trials and cohort studies in nutrition research. 2016;7(6):994–1004.
72. Carter JL, Lewington S, Piernas C, et al: Reproducibility of dietary intakes of macronutrients, specific food groups, and dietary patterns in 211 050 adults in the UK Biobank study. *J Nutr Sci*. 2019;8:e34.
73. Cade JE, Burley VJ, Greenwood DC: The UK Women's Cohort Study: comparison of vegetarians, fish-eaters and meat-eaters. *Public Health Nutr*. 2004;7(7):871–878.
74. Frank D, DeBenedetti AF, Volk RJ, Williams EC, Kivlahan DR, Bradley KA: Effectiveness of the AUDIT-C as a screening test for alcohol misuse in three race/ethnic groups. *J Gen Intern Med*. 2008;23(6):781–787.
75. Wang C, Xie W, Wang C, Zhu Y, Zhong D: Causal Relationships Between Environmental Exposures, Iron Metabolism, Hematuria Markers, and Rheumatoid Arthritis: An Investigation Using Mendelian Randomization. *Biomedicines*. 2025;13(2)
76. Balchin C, Tan AL, Golding J, et al: Acute effects of exercise on pain symptoms, clinical inflammatory markers and inflammatory cytokines in people with rheumatoid arthritis: a systematic literature review. *Ther Adv Musculoskelet Dis*. 2022;14:1759720x221114104.

77. Ingegnoli F, Gandolfo S, Ciccia F, Caporali R: Integrating Nutrition into Precision Medicine for Controlling Systemic Inflammation in Rheumatoid Arthritis. *Inflammation*. 2025;49(1):6.
78. Rudolph KE, Williams N, Díaz I: Using instrumental variables to address unmeasured confounding in causal mediation analysis. *Biometrics*. 2024;80(1)
79. Fry A, Littlejohns TJ, Sudlow C, et al: Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *Am J Epidemiol*. 2017;186(9):1026–1034.
80. Bradbury KE, Young HJ, Guo W, Key TJ: Dietary assessment in UK Biobank: an evaluation of the performance of the touchscreen dietary questionnaire. *J Nutr Sci*. 2018;7:e6.
81. Jin L, Zhang Z, Pan P, et al: Low-dose ethanol consumption inhibits neutrophil extracellular traps formation to alleviate rheumatoid arthritis. *Commun Biol*. 2023;6(1):1088.
82. Naska A, Lagiou A, Lagiou PJF: Dietary assessment methods in epidemiological research: current state of the art and future prospects. 2017;6:926.
83. Bridger Staatz C, Bann D: Diet–Opportunities for Data Collection. 2021;
84. Bajan B, Piechocka MJ Ao-IPiE, Informatics: Economic Development and Diet Composition: Cross-Continental Insights into Bennett's Law. 2025;17(2):37–49.
85. Kearney J: Food consumption trends and drivers. *Philos Trans R Soc Lond B Biol Sci*. 2010;365(1554):2793–2807.
86. Deane KD: Preclinical rheumatoid arthritis (autoantibodies): an updated review. *Curr Rheumatol Rep*. 2014;16(5):419.
87. Gerlag DM, Raza K, van Baarsen LG, et al: EULAR recommendations for terminology and research in individuals at risk of rheumatoid arthritis: report from the Study Group for Risk Factors for Rheumatoid Arthritis. *Ann Rheum Dis*. 2012;71(5):638–641.
88. Wang DD, Li Y, Bhupathiraju SN, et al: Fruit and Vegetable Intake and Mortality: Results From 2 Prospective Cohort Studies of US Men and Women and a Meta-Analysis of 26 Cohort Studies. *Circulation*. 2021;143(17):1642–1654.
89. Mikuls TR, Cerhan JR, Criswell LA, et al: Coffee, tea, and caffeine consumption and risk of rheumatoid arthritis: results from the Iowa Women's Health Study. *Arthritis Rheum*. 2002;46(1):83–91.

90. Ascione S, Barde F, Artaud F, et al: Association between beverage consumption and risk of rheumatoid arthritis: a prospective study from the French E3N Cohort. *Rheumatology (Oxford)*. 2023;62(5):1814–1823.
91. Ahmad S, Moorthy MV, Demler OV, et al: Assessment of risk factors and biomarkers associated with risk of cardiovascular disease among women consuming a Mediterranean diet. 2018;1(8):e185708–e185708.
92. Li J, Guasch-Ferré M, Chung W, et al: The Mediterranean diet, plasma metabolome, and cardiovascular disease risk. 2020;41(28):2645–2656.
93. Martín-Peláez S, Fito M, Castaner OJN: Mediterranean diet effects on type 2 diabetes prevention, disease progression, and related mechanisms. A review. 2020;12(8):2236.
94. Schwingshackl L, Schwedhelm C, Galbete C, Hoffmann GJN: Adherence to Mediterranean diet and risk of cancer: an updated systematic review and meta-analysis. 2017;9(10):1063.
95. Scoccianti C, Lauby-Secretan B, Bello P-Y, Chajes V, Romieu IJAjopm: Female breast cancer and alcohol consumption: a review of the literature. 2014;46(3):S16–S25.
96. Roerecke M, Vafaei A, Hasan OS, et al: Alcohol consumption and risk of liver cirrhosis: a systematic review and meta-analysis. 2019;114(10):1574–1586.
97. Sterling SA, Palzes VA, Lu Y, et al: Associations between medical conditions and alcohol consumption levels in an adult primary care population. 2020;3(5):e204687.

Appendix A Chapter 2 Supplementary material

Table A1. Search strategy.

Ovid Medline strategy:

1	(food* or nutrition or nutrient* or diet* or meat or red meat or processed meat or meat consumption* or meat product* or fish consumption* or dietary fish or fish oil or fruit* or milk or dairy product* or soda consumption* or dietary sucrose or energy intake* or protein consumption* or dietary protein or calcium or coffee consumption* or Caffeine or Olive oil or food component* or Zinc or dietary pattern* or Mediterranean diet* or dietary approaches to stop hypertension or DASH diet* or Inflammatory dietary pattern* or eating pattern* or western diet* or long-chain n-3 polyunsaturated fatty acid* or PUFA* or omega-3 fatty acid* or fatty acid* or low-carbohydrate diet* or micronutrient*).tw.
2	(meat or "dairy product" or "dietary carbohydrates OR dietary proteins OR dietary supplements" or "fast food" or honey or nutrients or eggs or "diet, food ,and nutrition" or "food and beverages" or "DASH diet OR dietary approaches to stop hypertension" or "functional food" or micronutrients or "Mediterranean diet" or "energy intake" or "high-fat diet" or "western diet" or "Diabetic Diet" or "High-Protein Diet").tw.
3	(vegetable* or vegan or vegan diet* or vegetarian or vegetarian diet*).tw.
4	(vegetables or vegan or vegetarian or "vegan diet" or "vegetarian diet").tw.
5	(alcohol* or alcohol use* or alcohol intake* or alcohol exposure or alcohol consumption* or alcohol drinking).tw.
6	(alcohols or "alcoholic beverages" or "alcohol drinking").tw.
7	(antioxidant* or antioxidant intake*).tw.
8	antioxidants.tw.
9	(dietary fiber or dietary fibre or high-fiber dietary or fiber intake* or fibre intake* or probiotics).tw.
10	("dietary fiber" or probiotics).tw.
11	("vitamin K" or "vitamin E" or "vitamin D" or "vitamin B" or "vitamin B complex" or "vitamin A").tw.

12	(vitamin or vitamin intake* or vitamin B group or vitamin D or vitamin B6 or vitamin B12 or vitamin K or vitamin D deficiency or vitamin K2 or vitamin E).tw.
13	or/1-12
14	((("Rheumatoid arthritis" or "inflammatory arthritis").tw. or (Rheumatoid Arthritis.tw. or Arthritis, Rheumatoid/)) not treatment.tw.
15	randomized controlled trial.pt.
16	(random\$ or placebo\$ or single blind\$ or double blind\$ or triple blind\$).ti,ab.
17	(retraction of publication or retracted publication).pt.
18	or/15-17
19	(animals not humans).sh.
20	((comment or editorial or meta-analysis or practice-guideline or review or letter) not "randomized controlled trial").pt.
21	(random sampl\$ or random digit\$ or random effect\$ or random survey or random regression).ti,ab. not "randomized controlled trial".pt.
22	18 not (19 or 20 or 21)
23	exp cohort studies/
24	cohort\$.tw.
25	controlled clinical trial.pt.
26	epidemiologic methods/
27	limit 26 to yr=1971-1988
28	or/23-25,27
29	22 or 28
30	13 and 14 and 29
31	limit 30 to (english language and humans)

Ovid Embase strategy:

1	(food* or nutrition or nutrient* or diet* or meat or red meat or processed meat or meat consumption* or meat product* or fish consumption* or dietary fish or fish oil or fruit* or milk or dairy product* or soda consumption* or dietary sucrose or energy intake* or protein consumption* or dietary protein or calcium or coffee
---	--

	consumption* or Caffeine or Olive oil or food component* or Zinc or dietary pattern* or Mediterranean diet* or dietary approaches to stop hypertension or DASH diet* or Inflammatory dietary pattern* or eating pattern* or western diet* or long-chain n-3 polyunsaturated fatty acid* or PUFA* or omega-3 fatty acid* or fatty acid* or low-carbohydrate diet* or micronutrient*).tw.
2	(meat or "dairy product" or "dietary carbohydratesOR dietary proteinsOR dietary supplements" or "fast food" or honey or nutrients or eggs or "diet,food,and nutrition" or "food and beverages" or "DASH dietOR dietary approaches to stop hypertension" or "functional food" or micronutrients or "Mediterranean diet" or "energy intake" or "high-fat diet" or "western diet" or "Diabetic Diet" or "High-Protein Diet").tw.
3	(vegetable* or vegan or vegan diet* or vegetarian or vegetarian diet*).tw.
4	(vegetables or vegan or vegetarian or "vegan diet" or "vegetarian diet").tw.
5	(alcohol* or alcohol use* or alcohol intake* or alcohol exposure or alcohol consumption* or alcohol drinking).tw.
6	(alcohols or "alcoholic beverages" or "alcohol drinking").tw.
7	(antioxidant* or antioxidant intake*).tw.
8	antioxidants.tw.
9	(dietary fiber or dietary fibre or high-fiber dietary or fiber intake* or fibre intake* or probiotics).tw.
10	("dietary fiber" or probiotics).tw.
11	("vitamin K" or "vitamin E" or "vitamin D" or "vitamin B" or "vitamin B complex" or "vitamin A").tw.
12	(vitamin or vitamin intake* or vitamin B group or vitamin D or vitamin B6 or vitamin B12 or vitamin K or vitamin D deficiency or vitamin K2 or vitamin E).tw.
13	or/1-12
14	((("Rheumatoid arthritis" or "inflammatory arthritis").tw. or (Rheumatoid Arthritis.tw. or Arthritis, Rheumatoid/)) not treatment.tw.
15	(random\$ or placebo\$ or single blind\$ or double blind\$ or triple blind\$).ti,ab.
16	RETRACTED ARTICLE/
17	or/15-16

18	(animal\$ not human\$).sh,hw.
19	(book or conference paper or editorial or letter or review).pt. not exp randomized controlled trial/
20	(random sampl\$ or random digit\$ or random effect\$ or random survey or random regression).ti,ab. not exp randomized controlled trial/
21	17 not (18 or 19 or 20)
22	exp cohort analysis/
23	exp longitudinal study/
24	exp prospective study/
25	exp follow up/
26	cohort\$.tw.
27	or/22-26
28	21 or 27
29	13 and 14 and 28
30	limit 29 to (human and english language and yr="2000 -Current")

Web of Science strategy:

1	TS = (food* or nutrition or nutrient* or diet* or meat or red meat or processed meat or meat consumption* or meat product* or fish consumption* or dietary fish or fish oil or fruit* or milk or dairy product* or soda consumption* or dietary sucrose or energy intake* or protein consumption* or dietary protein or calcium or coffee consumption* or Caffeine or Olive oil or food component* or Zinc or dietary pattern* or Mediterranean diet* or dietary approaches to stop hypertension or DASH diet* or Inflammatory dietary pattern* or eating pattern* or western diet* or long-chain n-3 polyunsaturated fatty acid* or PUFA* or omega-3 fatty acid* or fatty acid* or low-carbohydrate diet* or micronutrient*)
2	TS= (meat or "dairy product" or "dietary carbohudratesOR dietary proteinsOR dietary supplements" or "fast food" or honey or nutrients or eggs or "diet,food,and nutrition" or "food and beverages" or "DASH dietOR dietary approaches to stop hypertension" or "functional food" or micronutrients or "Mediterranean diet" or "energy intake" or "high-fat diet" or "western diet" or "Diabetic Diet" or "High-Protein Diet")

3	TS= (Vegetable* or vegan or vegan diet* or vegetarian or vegetarian diet*)
4	TS= (vegetables or vegan or vegetarian or "vegan diet" or "vegetarian diet")
5	TS= (alcohol* or alcohol use* or alcohol intake* or alcohol exposure or alcohol consumption* or alcohol drinking)
6	TS= (alcohols or "alcoholic beverages" or "alcohol drinking")
7	TS= (antioxidant* or antioxidant intake*)
8	TS= antioxidants
9	TS= (dietary fiber or dietary fibre or high-fiber dietary or fiber intake* or fibre intake* or probiotics)
10	TS= ("dietary fiber" or probiotics)
11	TS= ("vitamin K" or "vitamin E" or "vitamin D" or "vitamin B" or "vitamin B complex" or "vitamin A")
12	TS= (vitamin or vitamin intake* or vitamin B group or vitamin D or vitamin B6 or vitamin B12 or vitamin K or vitamin D deficiency or vitamin K2 or vitamin E)
13	#12 OR #11 OR #10 OR #9 OR #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1
14	TS=((("Rheumatoid arthritis" or "inflammatory arthritis") or (Rheumatoid Arthritis or Arthritis, Rheumatoid/)) not treatment)
15	TS=(randomized controlled trial)
16	TS=(random\$ or placebo\$ or single blind\$ or double blind\$ or triple blind\$)
17	TS=(retraction of publication or retracted publication)
18	#17 OR #16 OR #15
19	TS=(animals not humans)
20	TS=((comment or editorial or meta-analysis or practice-guideline or review or letter) not "randomized controlled trial")
21	TS=((random sampl\$ or random digit\$ or random effect\$ or random survey or random regression) not "randomized controlled trial")
22	#18 NOT (#19 OR #20 OR #21)
23	TS=(cohort analysis OR longitudinal study OR prospective study OR follow up OR cohort\$)

24	#23 OR #22
25	#13 AND #14 AND #24
26	((#25) AND PY=(2000-2024)) AND LA=(English)

Table A2. PICOS criteria for inclusion and exclusion of studies.

	Inclusion criteria	Exclusion criteria
Population	General population of adults	Aged <18 years
Intervention / Exposure	Dietary intakes of Food component or Nutrient	No dietary exposure
Comparison	Relative Risk (RRs), or Hazard Ratios (HRs), or Incidence Risk Ratios (IRRs)	Other risk estimates
Outcomes	Risk of rheumatoid arthritis	Risk of other arthritis, other autoimmune disease
Study type	Prospective cohort studies	Cross-sectional, case-control, ecological, retrospective observational studies, clinical trials, and non-human studies

Table A3. Standard conversions for food groups and nutrients.

Food components	Standard amount
Total alcohol/beer/wine/liquor	1 unit = 8 grams 1 glass = 1 serving = 15 grams
Total coffee/caffeinated coffee/decaffeinated coffee/tea	1 cup = 1 serving
Sugar-sweetened soda	1 drink = 1 serving
Fruits/vegetables	1 portion = 1 serving = 80 grams
Total meat/red meat/processed meat/poultry	1 serving = 90 grams
Total fish/oily fish	1 serving = 100 grams
Total dairy/milk/cheese	1 serving = 1 cup
Legumes	1 serving = 80 grams
Cereal products	1 serving = 30 grams
Vitamin D	400 IU
Vitamin C	40 mg
Vitamin A	600 IU
Vitamin E	3 mg
N-3 PUFA	100 mg
α -carotenoid/ β -carotenoid/ β -cryptoxanthin/lycopene/Lutein/zeaxanthin	200 mg

Table A4. Studies reviewed in full text for eligibility (excluded reasons for meta-analysis).

No.	Results from full text review	Reasons for excluded
1	Samira Alizadeh, Amir Ghorbanihaghjo, Sousan Kolahi, et al. Effect of omega-3 fatty acids supplementation on circulating osteoprotegerin ligand to osteoprotegerin ratio in female patients with rheumatoid arthritis, Clinical Biochemistry, 2011; 44(13):S142.	conference abstract.
2	Sofia Charneca, Margarida Ferro, João Vasques, et al. Adherence to the mediterranean diet is associated with patient-reported functional status in rheumatoid arthritis. Clinical Nutrition, 2021; 9:356.	conference abstract.
3	Sigrun Chrubasik. Vitamin E for rheumatoid arthritis or osteoarthritis: low evidence of effectiveness. Zeitschrift für Rheumatologie.2003;62(5):491	review.
4	Cooles FAH, Pratt AG, Ng WF, et al. The effect of vitamin D on early rheumatoid arthritis: A retrospective cohort analysis. Arthritis & Rheumatism, 2013;131:578-579.	conference abstract.
5	Hahn, J., Malspeis, S., Choi, M.Y., et al. Association of Healthy Lifestyle Behaviors and the Risk of Developing Rheumatoid Arthritis Among Women. Arthritis Care Res, 2023;75: 272-276.	exposure is not dietary consumption.

- | | | |
|----|--|--------------------------------------|
| 6 | Schernhammer E S, Hu F B, Giovannucci E, et al. Sugar-sweetened soft drink consumption and risk of pancreatic cancer in two prospective cohorts. <i>Cancer Epidemiology Biomarkers & Prevention</i> , 2005, 14(9): 2098-2105. | conference abstract. |
| 7 | Lahiri M, Luben RN, Morgan C, et al. Using lifestyle factors to identify individuals at higher risk of inflammatory polyarthritis (results from the European Prospective Investigation of Cancer-Norfolk and the Norfolk Arthritis Register--the EPIC-2-NOAR Study). <i>Ann Rheum Dis</i> . 2014;73(1):219-26. | conference abstract. |
| 8 | Liu C, Meng X, Zhang H, et al. Early-life famine exposure and rheumatoid arthritis in Chinese adult populations: a retrospective cohort study. <i>BMJ Open</i> . 2021;11(7):e043416. | exposure is not dietary consumption. |
| 9 | Maria López Lasanta, et al. Assessing dietary habits in a large cohort of rheumatoid arthritis and psoriatic arthritis patients: Results of the Spanish imid consortium. <i>American college of rheumatology</i> , 2014. | conference abstract. |
| 10 | Sparks JA, Barbhaiya M, Tedeschi SK,et al. Inflammatory dietary pattern and risk of developing rheumatoid arthritis in women. <i>Clin Rheumatol</i> . 2019;38(1):243-250. | exposure is not dietary consumption. |
| 11 | Rosell M, Wesley AM, Rydin K, et al. Dietary fish and fish oil and the risk of rheumatoid arthritis. <i>Epidemiology</i> . 2009;20(6):896-901. | case-control study. |
| 12 | Sparks JA, Chen CY, Hiraki LT,et al. Contributions of familial rheumatoid arthritis or lupus and environmental factors to risk of rheumatoid arthritis in women: a prospective cohort study. <i>Arthritis Care Res (Hoboken)</i> . 2014;66(10):1438-46. | exposure is not dietary consumption. |

- | | | |
|----|--|--------------------------------------|
| 13 | Yang WM, Lv JF, Wang YY, et al. The Daily Intake Levels of Copper, Selenium, and Zinc Are Associated with Osteoarthritis but Not with Rheumatoid Arthritis in a Cross-sectional Study. <i>Biol Trace Elem Res</i> . 2023. | cross-sectional study. |
| 14 | Cai B, Zhou M, Xiao Q, et al. L-shaped association between serum 25-hydroxyvitamin D and all-cause mortality of individuals with rheumatoid arthritis. <i>Rheumatology (Oxford)</i> . 2023 Feb 1;62(2):575-582. | exposure is not dietary consumption. |
| 15 | De Vito R, Fiori F, Ferraroni M, et al. Olive Oil and Nuts in Rheumatoid Arthritis Disease Activity. <i>Nutrients</i> . 2023;15(4):963. | cross-sectional study. |
| 16 | Pengpid S, Peltzer K. Prevalence and associated factors of cross-sectional and incident self-reported arthritis or rheumatism among a national community sample of middle-aged and older adults in Thailand. <i>Front Public Health</i> . 2023;11:1064751. | cross-sectional study. |
| 17 | Carubbi F, Alunno A, Mai F, et al. Adherence to the Mediterranean diet and the impact on clinical features in primary Sjögren's syndrome. <i>Clin Exp Rheumatol</i> . 2021;39 Suppl 133(6):190-196. | outcome is not RA. |
| 18 | VanEvery H, Yang W, Su J, et al. Low-Density Lipoprotein Cholesterol and the Risk of Rheumatoid Arthritis: A Prospective Study in a Chinese Cohort. <i>Nutrients</i> . 2022;14(6):1240. | exposure is not dietary consumption. |
| 19 | Liang X, Chou OHI, Cheung CL, et al. Is hypertension associated with arthritis? The United States national health and nutrition examination survey 1999-2018. <i>Ann Med</i> . 2022;54(1):1767-1775. | exposure is not dietary consumption. |

- | | | |
|-----------|--|--------------------------------------|
| 20 | Crowson CS, Gunderson TM, Dykhoff HJ, et al. Comprehensive assessment of multimorbidity burden in a population-based cohort of patients with rheumatoid arthritis. RMD Open. 2022;8(1):e002022. | exposure is not dietary consumption. |
| 21 | Guan T, Wu Z, Xu C, et al. The association of trace elements with arthritis in US adults: NHANES 2013-2016. J Trace Elem Med Biol. 2023;76:127122. | exposure is not dietary consumption. |
| 22 | Moradi A, Nezamoleslami S, Nezamoleslami S, et al. The association between dietary total antioxidant capacity with risk of rheumatoid arthritis in adults: A case-control study. Clin Nutr ESPEN. 2022;51:391-396. | case-control study. |
| 23 | Liu L, Xie S. Dietary fiber intake associated with risk of rheumatoid arthritis among U.S. adults: NHANES 2010-2020. Medicine (Baltimore). 2023;102(12):e33357. | cross-sectional study. |
| 24 | Kronzer VL, Lennon RJ, Sparks JA, Myasoedova E, Davis JM, Crowson CS. Association between work physical activity, dietary factors, and risk of rheumatoid arthritis. Semin Arthritis Rheum. 2022;57:152100. | case-control study. |
| 25 | Hatami E, Aghajani M, Pourmasoumi M, Haeri F, Boozari B, Nezamoleslami S, Clark CCT, Nezamoleslami S, Ghiasvand R. The relationship between animal flesh foods consumption and rheumatoid arthritis: a case-control study. Nutr J. 2022;30;21(1):51. | case-control study. |

Table A5. Characteristics of the included studies on the consumption of food, and nutrient and risk of rheumatoid arthritis.

Author, year (cohort study)	Country, recruitment, sex, baseline age(yrs), follow-up(yrs)	Case ascertainment	Cohort size (No of cases), Dietary assessment (No of FFQ item)	Dietary exposure	Adjustment for confounding variables
Cerhan 2003(IWHS) ¹	US, female,55- 69,11	Diagnosis by one rheumatologist using ACR criteria	n=29 368(152), validated semi-quantitative FFQ	vitamin C, vitamin E, carotenoids,α-carotenoid,β- carotenoid, β-cryptoxanthin, lycopene, lutein + zeaxanthin, zinc,copper, selenium supplements, manganese, all fruits,citrus fruits, oranges, orange juice,grapefruit,grapefruit juice, all vegetables, green leafy vegetables, yellow vegetables, vegetables, cabbages, cauliflower,legumes, broccoli	1, 5, 9, 15, 16, 21, 36, 38
Costenbader 2008(NHS &	NHS:US,femal e,30-55,22	Diagnosis by two rheumatologists using	NHS: n=91 739(559), NHSII: n=94 650(163),	vitamin D	3, 4, 5, 8, 12, 15, 16, 17, 34,

NHSII) ²	NHSII:US,female,25-42,12	ACR criteria	validated semi-quantitative FFQ		43
Costenbader 2010(NHS & NHSII) ³	NHS:US,female,30-55,24 NHSII:US,female,25-42,12	Diagnosis by two rheumatologists using ACR criteria	NHS: n=90 721(619), NHSII: n=93 922(168), validated semi-quantitative FFQ	vitamin A, vitamin C,vitamin E, α-carotenoid,β-carotenoid, β-cryptoxanthin, lycopene, lutein/zeaxanthin	3, 4, 5, 12, 15, 16
Di Giuseppe 2012(SMC) ⁴	Sweden,female,54-89,6	Outpatient Register and the Swedish Rheumatology Register	n=34 141(197), validated FFQ	alcohol,beer,wine,liquor	1, 4, 5, 6, 15,17, 23, 47
Sundström 2019(SMC) ⁵	Sweden,female,54-8,12	National Patient Register and the Swedish Rheumatology Register	n=35 600(368), validated FFQ	Meat(overall), red meat, processed meat, poultry, total dairy, milk,cheese	1, 5, 14, 21, 23, 24
Di Giuseppe 2014(SMC) ⁶	Sweden,female,54-8,7.5	National Patient Register and the Swedish Rheumatology Register	n=32 232(205), validated FFQ	long-chain n-3 PUFAs, fish	1, 5, 14, 21, 22
Hedenstierna 2021(SNMC) ⁷	Sweden,both sexes,51.5±15.9,18	Swedish National Patient Registers	n=41 068(577), questionnaire for epidemiological research	alcohol	2, 4, 5, 6, 12, 27
Hiraki 2012(NHS &	NHS:US,female,30-55,30	Diagnosis by two rheumatologists using	NHS:n=73 629(652),	vitamin D	1, 3, 4, 5, 12, 14, 15, 21, 31,

NHSII) ⁸	NHSII:US,female,25-42,18	ACR criteria	NHSII:n=45 544(148), validated FFQ		32, 33, 35
Merlino 2004(IWHS) ⁹	US,female,55-69,11	Diagnosis by two rheumatologists using ACR criteria	n=29 368(152), validated semi-quantitative FFQ	total vitamin D, dietary vitamin D, supplemental vitamin D, milk products, butter, margarine, skim milk, whole milk	1, 5, 21, 30, 44
Lahiri 2012(EPIC- Norfolk Study) ¹⁰	UK, both sexes, 40-79, 4	1987 ACR criteria	n=25 271(138), validated semi-quantitative FFQ	alcohol	4, 5, 6, 17, 27, 31, 43
Nguyen 2020(E3N- EPIC Cohort Study) ¹¹	France, female, 52.5±6.5, 10	1987 ACR criteria	n=62 629(480), validated FFQ	raw vegetables, legumes,fruits, cereal products, fish, meat, dairy products, unsaturated fat, alcohol, MD	1, 4, 5, 6, 11, 12, 21
Hu 2017(NHS & NHSII) ¹²	NHS:US,female,30-55,26 NHSII:US,female,25-42,20	Diagnosis by two rheumatologists using ACR criteria	NHS:n=76 597(624), NHSII:n=91 393(383), validated semi-quantitative FFQ	red/processed meat, sugar- sweetened beverages,trans fat, sodium,nuts, vegetables, fruits, PUFA, alcohol, whole grain, long- chain n-3 PUFA	1, 4, 5,15, 16, 17, 20, 21, 43
Karlson 2003(NHS) ¹³	US,female,34-59, 8	Diagnosis by two rheumatologists using ACR criteria	n=83 124(731), validated semi-quantitative FFQ	caffeinated coffee, decaffeinated coffee, total coffee, tea,caffeine	1, 4, 5, 14, 15, 16, 17, 28, 35, 43

Krok-Schoen 2018(WHI- OS) ¹⁴	US,female,55- 79,8	Self-reported	n=80 551(3 348), validated semi-quantitative FFQ	alcohol, fruits,vegetables, red meat, long-chain n-3 PUFA, n-6 fatty acids	1, 3, 4, 6, 21, 25, 29, 46
Lamichhane 2019(WHI- OS) ¹⁵	US,female,55- 79,8	Self-reported & DMARD use	n=76 853(185), validated semi-quantitative FFQ	total coffee, caffeinated coffee, decaffeinated coffee, unfiltered coffee, filtered coffee, tea, caffeine	1,3, 4,5, 6, 9, 14, 16
Lu 2014(NHS & NHSII) ¹⁶	NHS:US,femal e,30-55,28 NHSII:US,fem ale,25-42,20	Diagnosis by two rheumatologists using ACR criteria	NHS:n=82 472(580), NHSII:n=110 737(323), validated semi-quantitative FFQ	alcohol, beer,wine,liquor	1, 4, 5, 12, 13, 15, 16, 17, 19, 21, 43
Mikuls 2002(IWHS) ¹⁷	US,female,55- 69,11	Diagnosis by two rheumatologists using ACR criteria	n=31 336(158), validated semi-quantitative FFQ	total coffee, caffeinated coffee, decaffeinated coffee,tea,caffeine	1, 5, 9, 14, 15, 16
Pedersen 2005(DCH cohort) ¹⁸	Denmark,both sexes,50- 64,5.3	Diagnosis by rheumatologists using ACR criteria	n=56 691(69), validated self- administered FFQ	fish, lean fish ,medium fat fish,fat fish,olive oil,all fruits and fruit juices ,citrus fruits,all vegetables/vegetable juices, all vegetables/fruits/juices, b- carotene, vitamin A, vitamin E, vitamin C, vitamin D, zinc, selenium, iron, red	1, 2, 5, 6

				meat/fish/poultry/processed meat, red meat, coffee	
Rubin 2020(DCH cohort) ¹⁹	Denmark, both sexes,50-64,5.3	Diagnosis by rheumatologists using ACR criteria	n=56 075(797), validated self-administered FFQ	dietary fibre, meat	1, 2, 5, 6, 11, 14, 21, 27
Sparks 2019(NHS & NHSII) ²⁰	NHS:US,female,30-55,20 NHSII:US,female,25-42,24	Diagnosis by rheumatologists using ACR criteria	n=166 013(1 080), validated self-administered FFQ	fish, marine n-3 PUFA	1, 4, 5, 14, 46, 20, 21, 37, 39, 40, 41, 42, 45, 46
VanEvery 2021(Kailuan Study) ²¹	China, both sexes, mean age:51.8,10	Diagnosis by two rheumatologists using ACR criteria	n=87 118(87), validated self-reported alcohol consumption questionnaire	alcohol	1, 2, 4, 5
Cerhan 2002(IWHS) ²²	US,female,55-69,11	Diagnosis by two rheumatologists using ACR criteria	n=31 336(158), validated semi-quantitative FFQ	alcohol, beer, red wine, white wine, liquor	1
Benito-Garcia 2007(NHS) ²³	US,female,30-55, 22	Diagnosis by two rheumatologists using ACR criteria	n=82 063(546), validated semi-quantitative FFQ	total protein, animal protein, vegetable protein,total iron intake (diet and supplements),dietary iron,iron from supplements,heme iron intake,meat,red meat,poultry meat,fish	1, 4, 5, 43

Hu 2014(NHS & NHSII) ²⁴	NHS:US,female,30-55,28 NHSII:US,female,25-42,20	Diagnosis by two rheumatologists using ACR criteria	NHS:n=79 570(559), NHSII:n=107 330(298), validated semi-quantitative FFQ	sugar-sweetened soda, diet soda	1, 4, 5, 12, 13, 14, 16, 17, 18, 21, 26, 27,43
Nguyen 2022(E3N-EPIC) ²⁵	France,female, 52.5±6.5,10	Self-reported & diagnosis by two rheumatologists using ACR criteria	n=62 629(480), validated French Dietary History Questionnaire	total fish, lean fish, oily fish	1, 4, 5, 6, 11, 12, 21
Heliövaara 2007(Mobile Clinic Health Examination Survey) ²⁶	Finland,both sexes,20-98,13	Social Insurance Institution's population register	n=18 981(126), validated self-filled questionnaires	coffee, alcohol	1, 2
Mazzucca 2022(UK Biobank) ²⁷	UK,both sexes,40-69,12	ICD-10	n=479 494(2 819),validated semi-quantitative FFQ	cooked vegetables, fruits,cheese,oily fish,processed meat,beef,lamb,pork,alcohol, tea,coffee,decaffeinated coffee, ground coffee, instant coffee,other coffee	1, 2, 4, 5, 6, 12, 13, 14, 27
Eun 2020(NHIS) ²⁸	South Korea,female, 61.4,9	N/A	n=1 357 736(6 056), self-report questionnaires	alcohol	1, 4, 5, 12, 13, 14, 16, 17, 26, 27, 28, 43

Ro 2022 (NHIS-NSC) ²⁹	South Korea, both sexes, 40- 79, 14	ACR criteria, positive for either rheumatoid factor or anti-citrullinated peptide antibody	n=517 053(1 948), self- report questionnaires	alcohol	1, 4, 5, 12, 13, 14, 27
Ascione 2023(E3N- EPIC Cohort Study) ³⁰	France, female, 52.5±6.5, 11.7	1987 ACR criteria	n=62 149(481), validated FFQ	total alcohol, beer, wine, liquor, total coffee, caffeinated coffee, decaffeinated coffee, tea	1, 4, 5, 6, 11, 12, 21

IWHS=Iowa Women's Health Study; NHS & NHSII=Nurses' Health Study and Nurses' Health Study II Cohorts; SMC=Swedish Mammography Cohort; SNMC=Swedish National March Cohort; HS-FFQ =high school-food frequency questionnaire; EPIC=European Prospective Investigation into Cancer study; E3N-EPIC=Etude Epidémiologique auprès des femmes de la Mutuelle générale de l' Education Nationale; WHI-OS=Women's Health Initiative Observational Study and Clinical Trials cohort; DCH=Diet, Cancer, and Health cohort; FFQ= food frequency questionnaire; F=female, SIICD= The Social Insurance Institution's Coronary Heart Disease Study; UKB = UK Biobank; ACR = the American College of Rheumatology; ICD-10=International Classification of Diseases, 10th Revision; aMed: HLIS=Healthy Lifestyle Index Score; BMI=body mass index; PUFA=polyunsaturated fatty acid; trans-FA=trans fatty acid; SFA=saturated fatty acid; MFA: monounsaturated fatty acids; Adjustment: age(1), sex(2), race/ethnicity(3), BMI(4), smoking status(5), education(6), civil status(7), occupation status(8), marital status(9), sugar-sweetened soda(9), total energy intake(10), gastrointestinal transit(11), physical activity(12), family income(13), alcohol use(14), age at menopause(15), use of hormone replacement therapy(16), parity(17), multivitamin use(18), beverages(19), cohort(20), total energy(21), use of aspirin (22), dairy consumption(23), fish consumption(24), fruits(25), age at menarche(26), co-morbidity(e.g., diabetes mellitus, cancer, cardiovascular disease)(27), oral contraceptive use(28), memory loss(29), decaffeinated coffee consumption(30), Socioeconomic Status(31), sunscreen use and sun sensitivity(32), birth weight(33), latitude of residence(34); age at first birth(35), tea consumption(36), α-linolenic acid(37), coffee consumption(38), trans fatty acids (g/d)(39), saturated fatty acids (g/d)(40), omega-6 polyunsaturated fatty acids (g/d)(41), protein (g/d)(42), breastfeeding(43),β-cryptoxanthin intake(44), monounsaturated fatty acids (g/d)(45); questionnaire circle(46); meat(47).

Table A6. Characteristics of 14 included cohort studies.

Cohort (country)	N of studies	Follow-up	Gender	Dietary assessment method and frequency	Dietary exposure
IWHS(US)	4	1987-1997	Female	FFQ; Single at baseline	vitamins, carotenoids, minerals, fruits, legumes, vegetables, milk, butter, coffee, tea, caffeine, alcohol
NHS(US)	11	1980-2002	Female	FFQ; Baseline and repeated every 4 years	vitamins, carotenoids, Mediterranean diet, red/processed meat, fish, vegetables, whole grain, nuts, MUFA:SFA ratio, beverages, trans fat, n-3 PUFA, coffee, tea, caffeine, alcohol, Empirical Dietary Inflammatory Pattern, fish, total protein, animal protein, vegetable protein, total meat
NHSII(US)	9	1991-2001	Female	FFQ; Baseline and repeated every 4 years	vitamins, carotenoids, Mediterranean diet, red/processed meat, fish, vegetables, whole grain, nuts, MUFA:SFA ratio, beverages, trans fat, n-3 PUFA, coffee, tea,

					caffeine, alcohol, Empirical Dietary Inflammatory Pattern, fish, total protein, animal protein, vegetable protein, total meat
SMC(Sweden)	3	2003-2014	Female	FFQ, Repeated in 1987 and 1997	alcohol, meat, milk, cheese, fish, n-3 PUFA
SNMC(Sweden)	1	1997-2016	Both sexes	Questionnaire for epidemiological research, Single at baseline	alcohol
EPIC-Nortfolk(UK)	1	1997-2010	Both sexes	FFQ, Single at baseline	alcohol
E3N-EPIC(France)	3	2004-2014	Female	FFQ, Single at baseline	vegetables, legumes, fruits, cereal products, fish, meat, dairy product, unsaturated fat, alcohol, Mediterranean diet, total coffee, caffeinated coffee, decaffeinated coffee, tea, total alcohol, beer, wine, liquor
WHI-OS(US)	2	8yrs	Female	FFQ; Single at baseline	alcohol, fruits,vegetables, red meat, energy, n-3 fatty acids, n-6 fatty acids, coffee, tea, caffeine
DCH(Denmark)	2	1997-2018	Both sexes	FFQ, Single at baseline	fish, olive oil, fruits, vegetables, vitamins, minerals, red meat, coffee, dietary fibre, meat

Kailuan(China)	1	2006-2018	Both sexes	Questionnaire, Single at baseline	alcohol
UK Biobank(UK)	1	2005-2017	Both sexes	FFQ, Single at baseline	vegetables,oily fish,fresh fruit, processed meat, poultry, beef, lamb, pork, cheese, bread, breakfast cereal, coffee, tea
NHIS-NSC(South Korea)	1	2002-2016	Both sexes	self-report questionnaires, Single at baseline	alcohol
NHIS(South Korea)	1	2009-2018	Female	self-report questionnaires; Single at baseline	alcohol
MCHES(Finland)	1	1977-1989	Both sexes	Mobile Clinic Health Examination Survey, Single at baseline	coffee, alcohol

*IWHS= Iowa Women's Health Study; NHS & NHSII=Nurses' Health Study and Nurses' Health Study II Cohorts; SMC= Swedish Mammography Cohort; SNMC= Swedish National March Cohort; EPIC=European Prospective Investigation into Cancer study;DCH = Diet, Cancer, and Health cohort; FFQ= food frequency questionnaire;WHI-OS= Women's Health Initiative Observational Study and Clinical Trials cohort; MCHES=Mobile Clinic Health Examination Survey; NHIS-NSC=Korean National Health Insurance Service (NHIS)-national sample cohort.

Table A7 (a). Relationship between food groups and beverages consumption and risk of rheumatoid arthritis: a pooled analysis comparing the highest with the lowest categories and dose-response analyses.

Exposure	Studies	Participants	Cases	RR (95%CI)	P	I²(%)
Total alcohol	12					
high v low	8	985,846	8,082	0.76(0.68,0.85)	<0.001	56
per 16 g/w	9	1,011,117	8,220	0.96(0.94,0.98)	<0.001	58
Total alcohol*	5					
high v low	4	5,444,640	8,735	0.84(0.73,0.96)	0.014	0
per 16 g/w	5	5,469,976	8,842	0.97(0.96,0.99)	<0.001	28
Wine	4					
high v low	4	320,832	1,669	0.95(0.79,1.14)	0.982	0.7
per 16 g/w	4	320,832	1,670	0.98(0.97,1.00)	0.074	0
Beer	3					
high v low	3	258,683	1,189	0.79(0.59,1.03)	0.083	0
per 16 g/w	3	258,683	1,189	0.90(0.84,0.97)	0.004	0
Liquor(spirit)	4					
high v low	4	284,832	1,670	0.93(0.77,1.11)	0.668	0
per 16 g/w	4	284,832	1,670	0.96(0.91,1.02)	0.163	0
Total coffee	7					
high v low	6	751,937	4,045	1.16(0.94,1.42)	0.164	63
per 1 cup/d	6	700,419	3,933	1.04(0.99,1.09)	0.133	80
Caffeinated coffee	4					
high v low	4	253,462	1,296	1.30(1.09,1.54)	0.003	0
per 1 cup/d	3	176,609	1,113	1.03(0.98,1.09)	0.277	50

Decaffeinated coffee	5					
high v low	4	732,956	1,642	1.34(0.95,1.88)	0.076	66
per 1 cup/d	3	593,954	980	1.07(0.96,1.19)	0.207	76
Decaffeinated coffee[#]	2					
high v low	2	1,419,351	299	1.81(0.72,4.58)	0.097	72
per 1 cup/d	2	1,865,843	299	1.12(0.98,1.28)	0.084	58
Caffeine	3					
high v low	3	191,313	805	1.24(0.93,1.64)	0.178	39
per 25 g/d	2	114,460	632	0.98(0.96,1.01)	0.116	0
Tea	5					
high v low	5	732,956	4,103	1.15(1.00,1.33)	<0.001	53
per 1 cup/d	4	656,103	3,992	1.04(1.02,1.05)	<0.001	0
Tea[#]	2					
high v low	2	1,875,138	377	0.64(0.15,2.80)	0.557	72
per 1 cup/d	2	1,875,138	377	0.96(0.87,1.07)	0.472	0
Sugar-sweetened soda	3					
high v low	2	249,049	1,338	1.13(0.80,1.58)	0.490	64
per 1 drink/w	1	186,900	857	1.20(0.89,1.62)	0.238	10
Total meat	5					
high v low	4	239994	2193	0.94(0.82,1.06)	0.285	0
per 70 g/d	2	101856	848	1.03(0.94,1.14)	0.505	0
Red meat	4					
high v low	3	201,841	4,139	1.06(0.91,1.24)	0.431	32
per 90 g/d	3	176,469	3,785	0.84(0.54,1.32)	0.053	84
Processed meat	2					
high v low	2	518,721	3,180	1.00(0.88,1.13)	0.953	0

per 90 g/d	2	518,721	3,180	1.00(0.98,1.01)	0.866	0
Poultry	3					
high v low	3	600,784	3728	1.11(0.95,1.28)	0.185	0
per 90 g/d	2	518,721	3180	0.89(0.37,2.13)	0.793	49
Total fish	7					
high v low	3	260,874	1,765	0.92(0.75,1.08)	0.313	9
per 1 serving/w	3	285,333	1,629	1.01(0.99,1.03)	0.333	0
Oily fish	3					
high v low	2	542,123	3,273	0.91(0.78,1.06)	0.229	33
per 1 serving/w	3	598,814	3,342	1.00(0.97,1.03)	0.983	12
Fruit	7					
high v low	5	728,639	7,225	0.88(0.79,0.97)	0.012	50
per 80 g/d	6	785,330	7,294	0.95(0.92,0.99)	0.016	57
Vegetables	8					
high v low	5	728,639	7,194	0.90(0.76,1.06)	0.002	81
per 80 g/d	6	785,330	7,263	0.97(0.93,1.01)	0.152	83
Legumes	3					
high v low	2	91,997	632	0.97(0.80,1.19)	0.794	0
per 1 serving/d	2	91,997	632	0.92(0.55,1.54)	0.740	0
Total dairy	3					
high v low	3	131,224	1,000	0.97(0.72,1.30)	0.850	56
per 1 serving/w	3	313,224	1,000	0.99(0.98,1.01)	0.312	21
Milk	2					
high v low	2	68,595	520	1.06(0.84,1.34)	0.631	0
per 1 serving/w	2	91,997	632	0.98(0.95,1.02)	0.277	0
Cheese	2					
high v low	2	518,721	3,187	0.94(0.64,1.39)	0.756	73
per 1 serving/w	2	518,721	3,179	0.99(0.92,1.07)	0.833	76

Cereal product	3					
high v low	3	618,720	4,306	0.87(0.80,0.96)	0.003	0
per 1 serving/w	3	618,720	4,295	0.97(0.96,0.99)	0.014	20

*seropositive RA; #Rheumatoid Factor-positive RA.

Table A7 (b). Relationship between nutrient consumption and risk of RA: a pooled analysis comparing the highest with the lowest categories and dose-response analyses.

Exposure	Studies	Person-Years	Cases	RR (95%CI)	<i>P</i>	<i>I</i> ² (%)
Vitamin C	3					
high v low	2	3,295,916	939	0.98(0.76,1.27)	0.882	38
per 40 mg/d	3	3,352,607	1,008	0.99(0.96,1.02)	0.415	32
Vitamin D	4					
high v low	2	3,256,089	874	0.86(0.63,1.18)	0.354	47
per 400 IU/d	3	3,313,780	943	0.98(0.95,1.02)	0.370	41
Vitamin D supplement	2					
high v low	2	3,383,325	874	1.08(0.55,2.10)	0.834	75
per 400 IU/d	2	3,383,325	874	1.01(0.78,1.31)	0.991	77
Vitamin A	2					
high v low	1	2,981,735	787	1.04(0.82,1.32)	0.737	0
per 600 IU/d	2	3,038,426	856	1.04(0.96,1.13)	0.315	0

Vitamin E	3					
high v low	2	3,295,916	939	0.84(0.69,1.03)	0.094	0
per 3 mg/d	3	3,352,607	1,008	0.99(0.98,1.00)	0.250	0
α-carotenoid	2					
high v low	2	2,152,395	939	1.02(0.78,1.35)	0.883	32
per 200 mg/d	1	1,838,214	787	1.01(0.96,1.07)	0.689	0
β-carotenoid	2					
high v low	2	3,295,917	939	1.06(0.79,1.43)	0.704	43
per 200 mg/d	1	2,981,736	787	1.00(0.99,1.01)	0.700	43
β-cryptoxanthin	2					
high v low	2	3,295,917	939	0.86(0.58,1.26)	0.433	68
per 200 mg/d	1	2,981,736	787	1.04(0.76,1.43)	0.804	0
Lycopene	2					
high v low	2	3,295,917	939	0.97(0.77,1.21)	0.763	0
per 200 mg/d	1	2,981,736	787	0.99(0.98,1.00)	0.019	0

Lutein/zeaxanthin 2

high v low	2	3,295,917	939	0.97(0.77,1.23)	0.815	8
------------	---	-----------	-----	-----------------	-------	---

per 200 mg/d	1	2,981,736	787	1.00(0.98,1.02)	1.000	0
--------------	---	-----------	-----	-----------------	-------	---

N-3 PUFA 4

high v low	3	2,414,739	3,997	0.90(0.73,1.10)	0.300	52
------------	---	-----------	-------	-----------------	-------	----

per 100 mg/d	3	2,414,739	3,997	0.96(0.90,1.03)	0.285	74
--------------	---	-----------	-------	-----------------	-------	----

Table A8. Scoring for the food groups, beverages and food components of NutriGrade for the risk of rheumatoid arthritis.

Exposure	Outcome	Risk of bias ¹	Precision ²	Heterogeneity ³	Directness ⁴	Publication bias ⁵	Funding bias ⁶	Effect size ⁷	Dose-response ⁸	Sum	NutriGrade
Total alcohol	RA	1.5	1	0	0	0.5	1	1	1	6	Moderate
Total alcohol	Seropositive RA	1.5	1	1	1	0.5	1	1	1	7	Moderate
Wine	RA	1.5	0	1	1	0	1	0	0	4.5	Low
Beer	RA	1.5	0	1	1	0	1	1	1	6.5	Very low
Liquor	RA	1.5	0	1	1	0	1	0	0	4.5	Low
Total coffee	RA	1.5	0	0	0	0	1	0	1	3.5	Very low
Caffeinated coffee	RA	1.5	0	1	0	0	1	1	0	4.5	Low
Decaffeinated coffee	RA	1.5	0	0	0	0.5	1	0	0	3	Very low
Decaffeinated coffee	Rheumatoid Factor-	1.5	0	0	1	0	1	0	0	3.5	Very low

		positive RA										
Caffeine	RA	1.5	0	0	0	0	1	0	0	2.5	Very low	
Tea	RA	1.5	0	1	0	0.5	1	1	1	6	Moderate	
Tea	Rheumatoid Factor- positive RA	1.5	0	0	1	0	1	0	0	3.5	Very low	
Sugar- sweetened soda	RA	2	0	0	1	0	1	0	0	4	Low	
Fruit	RA	1.5	1	0	1	0.5	1	1	1	7	Moderate	
Vegetable	RA	1.5	0	0	1	0.5	1	0	1	5	Low	
Total meat	RA	2	0	1	1	0	1	0	0	5	Low	
Red meat	RA	1.5	0	1	0	0	1	0	0	3.5	Very low	
Processed meat	RA	1.5	0	1	1	0	1	0	0	4.5	Low	
Poultry	RA	1.5	0	1	1	0	1	0	0	4.5	Low	
Total fish	RA	2	0	1	1	0	1	0	0	5	Low	
Oily fish	RA	1.5	0	1	1	0	1	0	1	5.5	Low	
Legumes	RA	2	0	1	1	0	1	0	0	5	Low	

Total dairy	RA	2	0	0	1	0	1	0	0	4	Low
Milk	RA	1.5	0	1	1	0	1	0	0	4.5	Low
Cheese	RA	1.5	0	0	1	0	1	0	0	3.5	Very low
Cereal products	RA	2	0	1	1	0	1	1	1	7	Moderate
Vitamin C	RA	1.5	0	1	1	0	1	0	0	4.5	Low
Vitamin D	RA	2	0	1	1	0	1	0	0	5	Low
Vitamin D supplement	RA	2	0	0	1	0	1	0	1	5	Low
Vitamin A	RA	1.5	0	1	1	0	1	0	0	4.5	Low
Vitamin E	RA	1.5	0	1	1	0	1	0	0	4.5	Low
α-carotenoid	RA	1.5	0	0	1	0	1	0	0	3.5	Very low
β-carotenoid	RA	1.5	0	0	1	0	1	0	0	3.5	Very low
β-cryptoxanthin	RA	1.5	0	1	1	0	1	0	0	4.5	Low
Lycopene	RA	1.5	0	1	1	0	1	0	0	4.5	Low
Lutein/zeaxanthin	RA	1.5	0	0	1	0	1	0	0	3.5	Very low
N-3 PUFA	RA	1.5	0	0	0	0	1	0	0	2.5	Very low

- ¹ Modifications of original NutriGrade-scoring system: Study quality Newcastle Ottawa Scale: ≥ 8 (2 points); ≥ 7 - <8 (1.5 points), ≥ 6 - <7 (1 point);
- ² >5000 cases and the 95%CI excludes the null value (1 point), <5000 cases or >5000 cases but 95%CI includes the null value (0 point);
- ³ 0 to 1 point on the basis of I^2 ;
- ⁴ No important differences in the population or hard clinical outcomes (1 point), important differences in the population (0 point);
- ⁵ <5 studies or severe evidence of bias or publication bias not assessed (0 point), no evidence of bias for 5-9 studies or moderate evidence of bias for ≥ 10 studies (0.5 point), no evidence of bias for ≥ 10 studies (1 point);
- ⁶ Funded by academic institutions or research institutions (1 point), funded by private institutions, foundations, or nongovernmental organizations (0.5 point), Industry funding or conflict of interest (0 point);
- ⁷ No effect (RR: 0.80-1.20) when comparing the highest vs. lowest category (0 point), moderate effect size (RR: <0.80 -0.50 and >1.20 -2, and corresponding test is statistically significant) when comparing the highest vs. lowest category (1 point), large effect size (RR: <0.50 and >2.00 , and corresponding test is statistically significant) when comparing the highest vs. lowest category (2 point);
- ⁸ No dose-response analysis or dose-response analysis with corresponding statistical test no significant (0 point), significant linear or nonlinear dose-response relationship (1 point).

Table A9. Newcastle Ottawa Scale scores of the included prospective cohort studies.

Study (Supplemental reference)	Selection			Comparability		Exposure			Total score
	Exposed	Non-exposed	Exposure	Demonstration ⁴	Confounding ⁵	Outcome	Follow-up	Non-response	
	cohort ¹	cohort ²	assessment ³			assessment ⁶	duration ⁷	rate ⁸	

Cerhan et al. 2003	1	1	1	1	2	1	0	1	8
Costenbader et al. 2008	1	1	1	1	1	1	1	0	7
Costenbader et al. 2010	1	1	1	1	1	1	1	0	7
Di Giuseppe et al. 2012	1	1	1	1	0	1	1	0	6
Sundström et al. 2019	1	1	1	1	2	1	1	0	7
Di Giuseppe et al. 2013	1	1	1	1	2	1	1	0	8
Hedenstierna et al. 2021	1	1	0	1	2	1	1	0	7
Hiraki et al. 2012	1	1	1	1	2	1	1	1	9
Merlino et al. 2004	1	1	1	1	2	1	1	0	8
Lahiri et al. 2012	1	1	1	1	2	1	0	1	7
Nguyen et al. 2020	1	1	1	1	2	1	1	1	9
Hu et al. 2015	1	1	1	1	2	1	1	0	8
Hu et al. 2017	1	1	1	1	2	1	1	0	8
Karlson et al. 2003	1	1	1	1	1	1	1	0	7
Krok-Schoen et al. 2018	0	1	1	1	2	1	0	0	6
Lamichhane et al. 2019	0	1	1	1	2	1	1	0	7
Lu et al. 2014	1	1	1	1	2	1	1	1	9
Sparks et al. 2019	1	1	1	1	2	1	1	1	9

Mikuls et al. 2002	1	1	1	1	1	1	1	1	8
Pedersen et al. 2005	1	1	1	1	1	1	0	0	6
Rubin et al. 2020	1	1	1	1	2	1	1	0	8
Sparks et al. 2019	1	1	1	1	2	1	1	0	8
VanEvery et al. 2021	1	1	0	1	1	1	1	0	6
Cerhan et al. 2002	1	1	1	1	0	1	1	0	6
Benito-Garcia et al. 2007	1	1	1	1	2	1	1	0	8
Hu et al. 2014	1	1	1	1	2	1	1	0	8
Nguyen et al. 2022	1	1	1	1	2	1	1	0	8
Heliövaara et al. 2000	1	1	1	1	0	1	1	0	6
Mazzucca et al. 2022	1	1	1	1	2	1	0	0	7
Eun et al. 2020	0	1	0	1	1	1	0	0	5
Ro et al. 2022	1	1	0	1	1	1	1	0	6
Ascione et al. 2023	1	1	1	1	2	1	1	1	9

¹ For the representativeness of the exposed cohort, if the participants were derived from general population, then this item would be given 1 point; if the participants were derived from occupational population, disease population, or population with higher risk factor (smoking population), then this item would be given 0 point;

² For the selection of the non-exposed cohort, if drawn from the same original population as the exposed cohort, then this item would be given 1 point; otherwise 0 point;

- ³ For the ascertainment of exposure, if intakes of food and nutrient were assessed by structured dietary assessment tools (e.g., 24h-recall, and FFQ), then this item would be given 1 point; if not, this item would be given 0 point;
- ⁴ For the demonstration that RA was not present at start of study, this item would be given 1 point; otherwise 0 points;
- ⁵ For comparability of cohorts, if the model adjusted for age, sex, and another socioeconomic factor, then this item would be given 1 point; if the model adjusted for smoking status, alcohol intake, physical activity and other dietary information, then this study would be given additional 1 point;
- ⁶ For the assessment of outcome, if the outcome was identified through medical records, or registered data, then this item would be given 1 point; otherwise this item would be given 0 point;
- ⁷ For follow-up duration, if it was 10 years or above, then this item would be given 1point; otherwise 0 point;
- ⁸ For the adequacy of follow up of cohorts, if the rate of loss to follow-up was less than 10%, then this item would be given 1 point; if $\geq 10\%$ or no information on the rate of loss to follow-up, then 0 point.

Table A10 (a). Subgroup analyses for total alcohol consumption and risk of rheumatoid arthritis.

	n	RR	95% CI	<i>I</i> ² (%)	<i>Q</i> within ¹	<i>Q</i> between ²	<i>P</i> ³
All studies	8	0.76	[0.68,0.85]	56	15.57		
Gender							
Both sexes	3	0.72	[0.65,0.80]	82	10.91	2.22	0.14
Female	5	0.80	[0.74,0.87]	0	2.44		
Geographic location							
Europe	4	0.72	[0.66, 0.80]	56	6.80	8.27	0.02
America	3	0.80	[0.73, 0.88]	0	0.50		
Asia	1	1.98	[0.93, 4.22]	-	0		
Follow-up duration							
<10 years	3	0.77	[0.72, 0.83]	15	2.35	0.12	0.73
≥10 years	5	0.75	[0.65, 0.87]	70	13.10		
Number of cases							
<500	4	0.86	[0.70,1.06]	57	6.97	1.33	0.25
≥500	4	0.76	[0.71,0.81]	59	7.27		
Outcome assessment							
Including self-reported data	4	0.81	[0.74,0.88]	0	1.07	3.18	0.07
Register data	4	0.72	[0.65,0.79]	74	11.31		

¹ *Q* within each subgroup;

² *Q* between subgroups;

³ *P* values for comparisons between subgroups.

Table A10 (b). Subgroup analyses for total alcohol consumption and risk of seropositive rheumatoid arthritis.

	n	RR	95% CI	I ² (%)	Q _{within} ¹	Q _{between} ²	P ³
All cohorts	5	0.84	[0.73,0.96]	0	3.30		
Gender							
Both sexes	1	0.75	[0.61,0.92]	0	1.39	1.92	0.17
Female	4	0.91	[0.76,1.07]	-	0		
Geographic location							
Europe	1	0.66	[0.41, 1.07]	-	0	1.32	0.51
America	2	0.81	[0.61, 1.07]	0	0.80		
Asia	2	0.88	[0.75, 1.02]	15.8	1.19		
Number of cases							
<500	3	0.77	[0.60,0.98]	0	1.31	0.80	0.37
≥500	2	0.87	[0.75,1.02]	15.8	1.19		

¹ Q within each subgroup;
² Q between subgroups;
³ P values for comparisons between subgroups.

Table A10 (c). Subgroup analyses for total coffee consumption and risk of rheumatoid arthritis.

	n	RR	95% CI	I ² (%)	Q _{within} ¹	Q _{between} ²	P ³
All studies	6	1.16	[0.94,1.42]	63	13.52		
Age (mean/median)							
<50 years	2	1.73	[0.59,5.07]	57.4	2.35	0.59	0.44

≥50 years	4	1.12	[0.88,1.43]	64.1	8.35		
Gender							
Both sexes	2	0.92	[0.83,1.00]	73.1	3.72	0.09	0.76
Female	4	1.26	[1.05,1.51]	0	0.50		
Geographic location							
America	3	1.27	[1.00,1.61]	0	0.49	0.33	0.56
Europe	3	0.92	[0.83,1.00]	74.5	7.83		
Follow-up duration							
<10 years	2	1.01	[0.74,1.39]	58.9	2.43	1.43	0.23
≥10 years	4	1.31	[0.98,1.74]	0	2.74		
N of cases							
<500	5	1.28	[1.07,1.53]	0	2.74	10.78	<0.01
≥500	1	0.91	[0.83,1.00]	-	0		
Outcome assessment							
Including self-reported data	4	1.26	[1.05,1.51]	0	0.50	9.30	<0.01
Register data	2	1.56	[0.39,6.23]	73.1	3.72		
Dietary assessment tool							
FFQ	5	0.97	[0.89,1.06]	60.5	10.14	3.39	0.06
Not mentioned	1	3.95	[0.89,17.52]	-	0		

¹Q within each subgroup;

²Q between subgroups;

³P values for comparisons between subgroups.

Table A10 (d). Subgroup analyses for decaffeinated coffee consumption and risk of rheumatoid arthritis.

	n	RR	95% CI	I ² (%)	Q ^{within} 1	Q ^{between} 2	P ³
All studies	5	1.34	[0.95,1.88]	66	11.90		
Age (mean/median)							
<50 years	1	1.00	[0.59,1.68]	-	0	0.83	0.36
≥50 years	4	1.29	[1.08,1.54]	72.9	11.08		
Gender							
Both sexes	1	0.96	[0.68,1.35]	-	0	3.11	0.08
Female	4	1.37	[1.13,1.65]	65.9	8.79		
Geographic location							
America	3	1.66	[1.22,2.26]	68	6.25	4.48	0.03
Europe	2	1.12	[0.92,1.37]	14.5	1.17		
Follow-up duration							
<10 years	2	1.10	[0.81,1.48]	62	2.63	1.13	0.29
≥10 years	3	1.33	[1.09,1.63]	75	8.15		
N of cases							
<500	4	1.36	[1.13,1.65]	65.9	8,79	3.11	0.08
≥500	1	0.96	[0.68,1.35]	-	0		
Outcome assessment							
Including self-reported data	4	1.37	[1.13,1.65]	65.9	8.79	3.11	0.08
Register data	1	0.96	[0.68,1.35]	-	0		

¹Q within each subgroup;

²Q between subgroups;

³P values for comparisons between subgroups.

Table A10 (e). Subgroup analyses for tea consumption and risk of rheumatoid arthritis.

	n	RR	95% CI	I^2 (%)	Q_{within}^1	$Q_{between}^2$	P^3
All studies	5	0.88	[0.79,0.97]	50	8.01		
Age (mean/median)							
<50 years	1	1.10	[0.73,1.66]	-	0	0.16	0.69
≥50 years	4	1.20	[1.10,1.31]	64.0	8.33		
Gender							
Both sexes	1	1.23	[1.12,1.36]	-	0	2.00	0.16
Female	4	1.04	[0.85,1.28]	53.8	6.49		
Geographic location							
America	3	1.06	[0.75,1.49]	69.1	6.47	0.49	0.48
Europe	2	1.20	[1.10,1.32]	34.5	1.53		
Follow-up duration							
<10 years	3	1.00	[0.80,1.24]	55	4.44	3.16	0.08
≥10 years	2	1.24	[1.12,1.36]	0	0.89		
N of cases							
<500	4	1.04	[0.85,1.28]	53.8	6.49	2.00	0.16
≥500	1	1.23	[1.11,1.36]	-	0		
Outcome assessment							
Including self-reported data	4	1.04	[0.85,1.28]	53.8	6.49	2.00	0.16
Register data	1	1.23	[1.12,1.36]	-	0		

¹ Q within each subgroup;

² Q between subgroups;

³ P values for comparisons between subgroups.

Table A10 (f). Subgroup analyses for fruit consumption and risk of rheumatoid arthritis.

	n	RR	95% CI	I^2 (%)	Q_{within} 1	$Q_{between}$ 2	P^3
All studies	5	0.88	[0.79,0.97]	50	8.01		
Age (mean/median)							
<50 years	1	1.03	[0.78,1.37]	-	0	1.51	0.2 2
≥50 years	4	0.86	[0.81,0.91]	54	6.50		
Gender							
Both sexes	1	0.92	[0.84,1.01]	-	0	3.01	0.0 8
Female	4	0.82	[0.76,0.90]	40	4.99		
Geographic location							
America	3	0.81	[0.74,0.88]	39	3.27	4.67	0.0 3
Europe	2	0.92	[0.85,1.00]	0	0.07		
Follow-up duration							
<10 years	3	0.85	[0.80,0.91]	65	5.67	2.15	0.1 4
≥10 years	2	0.98	[0.82,1.17]	0	0.19		
N of cases							
<500	3	0.94	[0.80,1.11]	0	1.78	1.11	0.2 9
≥500	2	0.85	[0.80,0.91]	80	5.11		
Outcome assessment							
Including self-reported data	4	0.82	[0.76,0.90]	40	5.00	3.01	0.0 8
Register data	1	0.92	[0.84,1.01]	0	0		

¹Q within each subgroup;

²Q between subgroups;

³P values for comparisons between subgroups.

Table A10 (g). Subgroup analyses for vegetables consumption and risk of rheumatoid arthritis.

	n	RR	95% CI	<i>I</i> ² (%)	<i>Q</i> _{within} ¹	<i>Q</i> _{between} ²	<i>P</i> ³
All studies	5	0.90	[0.76,1.06]	81	20.88		
Age (mean/median)							
<50 years	1	0.84	[0.63,1.12]	-	0	0.2	0.65
≥50 years	4	0.90	[0.84,0.97]	86	20.68		
Gender							
Both sexes	1	1.13	[0.99,1.28]	-	0	18.26	<0.001
Female	4	0.81	[0.74,0.88]	0	2.62		
Geographic location							
America	3	0.78	[0.71,0.86]	0	0.30	18.9	<0.001
Europe	2	1.08	[0.97,1.21]	41	1.68		
Follow-up duration							
<10 years	3	0.89	[0.83,0.97]	90	20.43	0.02	0.89
≥10 years	2	0.91	[0.76,1.08]	0	0.43		
N of cases							
<500	3	0.88	[0.75,1.04]	0	1.15	0.06	0.81
≥500	2	0.89	[0.73,0.97]	95	19.68		
Outcome assessment							

Including self-reported data	4	0.81	[0.74,0.88]	0	2.62	18.26	<0.001
Register data	1	1.13	[0.99,1.28]	-	0		

¹Q within each subgroup;

²Q between subgroups;

³P values for comparisons between subgroups.

Table A11 (a). Sensitivity analyses for food group consumption and risk of rheumatoid arthritis: excluding one study at a time¹.

Study (Supplemental reference)	Estimate	95% Confidence Interval	
Total alcohol			
Included all studies	0.76	0.68	0.85
Omitting Di Giuseppe 2012	0.77	0.68	0.87
Omitting Hedenstierna 2021	0.79	0.73	0.84
Omitting Ascione 2023	0.74	0.66	0.84
Omitting Krok-Schoen 2018	0.75	0.64	0.88
Omitting Hu 2017	0.77	0.66	0.89
Omitting VanEvery 2021	0.75	0.68	0.83
Omitting Cerhan 2002	0.76	0.67	0.86
Omitting Mazzucca 2022	0.77	0.65	0.92
Total alcohol (seropositive RA)			
Included all studies	0.84	0.73	0.96
Omitting Lu 2014 (NHS)	0.80	0.65	0.98
Omitting Lu 2014(NHSII)	0.86	0.75	0.98
Omitting Eun 2020	0.86	0.75	0.99
Omitting Ro 2022	0.75	0.61	0.93
Omitting Ascione 2023	0.86	0.75	0.99

Wine

Included all studies	0.95	0.79	1.14
Omitting Di Giuseppe 2012	0.96	0.78	1.18
Omitting Lu 2014(NHS)	0.86	0.70	1.06
Omitting Lu 2014(NHSII)	0.98	0.81	1.18
Omitting Cerhan 2002	0.95	0.77	1.17
Omitting Ascione 2023	0.97	0.74	1.28

Beer

Included all studies	0.87	0.74	1.02
Omitting Di Giuseppe 2012	0.87	0.73	1.03
Omitting Lu 2014(NHS)	0.90	0.76	1.06
Omitting Lu 2014(NHSII)	0.86	0.73	1.02
Omitting Cerhan 2002	0.86	0.73	1.02
Omitting Ascione 2023	0.78	0.59	1.03

Liquor

Included all studies	0.93	0.77	1.11
Omitting Di Giuseppe 2012	0.93	0.78	1.13
Omitting Lu 2014(NHS)	0.90	0.73	1.13
Omitting Lu 2014(NHSII)	0.93	0.78	1.12
Omitting Cerhan 2002	0.91	0.74	1.10
Omitting Ascione 2023	0.95	0.75	1.20

Total coffee

Included all studies	1.16	0.94	1.42
Omitting Lamichhane 2019	1.15	0.91	1.45
Omitting Karlson 2003	1.17	0.90	1.52
Omitting Mikuls 2002	1.12	0.91	1.39
Omitting Mazzucca 2022	1.12	0.92	1.37
Omitting Heliovaara 2000	1.10	0.87	1.40

Omitting Ascione 2023	1.16	0.89	1.50
Caffeinated coffee			
Included all studies	1.30	1.09	1.54
Omitting Mikuls 2002	1.35	1.12	1.62
Omitting Lamichhane 2019	1.29	1.07	1.55
Omitting Karlson 2003	1.25	1.00	1.54
Omitting Ascione 2023	1.29	1.03	1.62
Decaffeinated coffee			
Included all studies	1.34	0.95	1.88
Omitting Mazzucca 2022	1.48	1.00	2.18
Omitting Lamichhane 2019	1.28	0.86	1.90
Omitting Karlson 2003	1.43	0.95	2.15
Omitting Mikuls 2002	1.14	0.96	1.37
Omitting Ascione 2023	1.40	0.88	2.22
Decaffeinated coffee (RF-positive RA)			
Included all studies	1.80	0.90	3.58
Omitting Karlson 2003	2.64	1.46	4.78
Omitting Mikuls 2002	1.30	0.85	1.99
Caffeine			
Included all studies	1.24	0.93	1.64
Omitting Karlson 2003	1.30	0.68	2.49
Omitting Lamichhane 2019	1.13	0.88	1.44
Omitting Mikuls 2002	1.39	0.94	2.06
Tea			
Included all studies	1.15	1.00	1.33
Omitting Mikuls 2002	1.20	1.10	1.32
Omitting Lamichhane 2019	1.11	0.94	1.33
Omitting Karlson 2003	1.13	0.92	1.39

Omitting Mazzucca 2022	1.04	0.85	1.28
Omitting Ascione 2023	1.22	1.11	1.34
Tea (RF-positive RA)			
Included all studies	0.64	0.15	2.80
Omitting Karlson 2003	0.26	0.06	1.11
Omitting Mikuls 2002	1.20	0.61	2.35
Sugar-sweetened soda			
Included all studies	1.13	0.80	1.58
Omitting Hu 2014(NHS)	0.94	0.75	1.18
Omitting Hu 2014(NHSII)	1.19	0.71	2.00
Omitting Ascione 2023	1.29	0.82	2.01
Fruit			
Included all studies	0.88	0.79	0.97
Omitting Cerhan 2013	0.89	0.79	0.99
Omitting Nguyen 2020	0.87	0.76	0.96
Omitting Hu 2017(NHS,NHSII)	0.86	0.76	0.96
Omitting Krok-Schoen 2018	0.92	0.85	1.00
Omitting Mazzucca 2022	0.86	0.74	0.99
Vegetable			
Included all studies	0.90	0.76	1.06
Omitting Cerhan 2013	0.92	0.77	1.10
Omitting Nguyen 2020	0.88	0.72	1.09
Omitting Hu 2017(NHS, NHSII)	0.91	0.74	1.11
Omitting Krok-Schoen 2018	0.96	0.80	1.15
Omitting Mazzucca 2022	0.81	0.73	0.91
Legumes			
Included all studies	0.97	0.80	1.19
Omitting Cerhan 2013	0.98	0.79	1.22

Omitting Nguyen 2020	0.95	0.61	1.48
Total meat			
Included all studies	0.94	0.82	1.06
Omitting Sundstrom 2019	0.91	0.80	1.05
Omitting Nguyen 2020	0.90	0.77	1.04
Omitting Rubin 2020	1.00	0.85	1.18
Omitting Benito-Garcia 2007	0.95	0.90	1.12
Red meat			
Included all studies	1.06	0.91	1.24
Omitting Krok-Schoen 2018	0.95	0.76	1.19
Omitting Benito-Garcia 2007	1.13	1.03	1.24
Omitting Sundstrom 2019	1.02	0.79	1.32
Processed meat			
Included all studies	1.00	0.88	1.13
Omitting Mazzucca 2022	0.84	0.58	1.21
Omitting Sundstrom 2019	1.02	0.89	1.17
Poultry			
Included all studies	1.11	0.95	1.28
Omitting Sundstrom 2019	1.15	0.98	1.35
Omitting Mazzucca 2022	1.05	0.79	1.38
Omitting Benito-Garcia 2007	1.06	0.85	1.33
Total fish			
Included all studies	0.92	0.78	1.08
Omitting Di Giuseppe 2013	0.97	0.81	1.16
Omitting Nguyen 2020	0.83	0.64	1.08
Omitting Sparks 2019	0.87	0.64	1.20
Oily fish			
Included all studies	0.91	0.78	1.06

Omitting Mazzucca 2022	1.01	0.81	1.25
Omitting Nguyen 2020	0.86	0.75	0.99
Milk			
Included all studies	1.06	0.84	1.34
Omitting Sundstrom 2019	1.04	0.69	1.57
Omitting Merlino 2004	1.07	0.80	1.43
Cheese			
Included all studies	0.94	0.64	1.39
Omitting Mazzucca 2022	1.20	0.81	1.77
Omitting Sundstrom 2019	0.80	0.69	0.93
Total dairy			
Included all studies	0.97	0.72	1.30
Omitting Nguyen 2020	0.86	0.53	1.41
Omitting Merlino 2004	1.11	0.92	1.34
Omitting Sundstrom 2019	0.89	0.53	1.49
Cereal product			
Included all studies	0.87	0.80	0.96
Omitting Nguyen 2020	0.87	0.79	0.95
Omitting Mazzucca 2022	0.89	0.73	1.07
Omitting Hu 2017(NHS, NHSII)	0.88	0.80	0.97

¹Effect sizes were pooled using random-effects model.

Table A11 (b). Sensitivity analyses for nutrient intake and risk of rheumatoid arthritis: excluding one study at a time¹.

Study (Supplemental reference)	Estimate	95% Confidence Interval	
N-3 PUFA			
Included all studies	0.90	0.73	1.10
Omitting Di Giuseppe et al	0.98	0.86	1.11
Omitting Hu et al (NHS, NHSII)	0.87	0.59	1.28
Omitting Krok-Schoen et al (WHI)	0.80	0.64	1.01
Vitamin C			
Included all studies	0.98	0.76	1.27
Omitting Cerhan et al (IWHS)	1.10	0.89	1.36
Omitting Costenbader et al (NHS)	0.88	0.58	1.34
Omitting Costenbader et al (NHSII)	0.92	0.61	1.39
Vitamin D			
Included all studies	0.79	0.62	1.00
Omitting Hiraki et al (NHS)	0.80	0.52	1.24
Omitting Hiraki et al (NHSII)	0.74	0.57	0.97
Omitting Merlino et al (IWHS)	0.86	0.64	1.16
Vitamin D supplement			
Included all studies	1.08	0.55	2.10
Omitting Costenbader et al (NHS)	1.18	0.35	4.00
Omitting Costenbader et al (NHSII)	0.80	0.54	1.21
Omitting Merlino et al (IWHS)	1.43	0.64	3.21
Vitamin A			
Included all studies	1.04	0.82	1.32
Omitting Costenbader et al (NHS)	0.90	0.57	1.42

Omitting Costenbader et al (NHSII)	1.10	0.83	1.46
------------------------------------	------	------	------

Vitamin E

Included all studies	0.84	0.69	1.03
Omitting Cerhan et al (IWHS)	0.84	0.67	1.05
Omitting Costenbader et al (NHS)	0.92	0.66	1.28
Omitting Costenbader et al (NHSII)	0.81	0.65	1.01

α-carotenoid

Included all studies	1.02	0.78	1.35
Omitting Costenbader et al (NHS)	0.88	0.61	1.26
Omitting Costenbader et al (NHSII)	1.14	0.90	1.45
Omitting Cerhan et al (IWHS)	0.97	0.58	1.62

β-carotenoid

Included all studies	1.06	0.79	1.43
Omitting Costenbader et al (NHS)	0.88	0.64	1.23
Omitting Costenbader et al (NHSII)	1.18	0.91	1.55
Omitting Cerhan et al (IWHS)	1.06	0.66	1.69

β-cryptoxanthin

Included all studies	0.86	0.58	1.26
Omitting Costenbader et al (NHS)	0.71	0.47	1.07
Omitting Costenbader et al (NHSII)	0.82	0.45	1.52
Omitting Cerhan et al (IWHS)	1.06	0.84	1.33

Lycopene

Included all studies	0.97	0.77	1.21
Omitting Costenbader et al (NHS)	0.94	0.61	1.65

Omitting Costenbader et al (NHSII)	0.91	0.71	1.18
Omitting Cerhan et al (IWHS)	1.05	0.81	1.37
Lutein/zeaxanthin			
Included all studies	0.97	0.77	1.23
Omitting Costenbader et al (NHS)	0.83	0.59	1.16
Omitting Costenbader et al (NHSII)	1.04	0.83	1.32
Omitting Cerhan et al (IWHS)	0.94	0.61	1.43

¹Effect sizes were pooled using random-effects model.

Table A12 (a). Trim and fill tests for food group consumption and risk of rheumatoid arthritis.

Food exposure	n to trim	Estimated effect after adding hypothesized studies
Total alcohol	0	0.76(0.68-0.85)
Total alcohol*	2	0.88(0.77,0.99)
Wine	2	0.99(0.84-1.17)
Beer	0	0.87(0.74-1.02)
Liquor	2	0.94(0.80-1.12)
Total coffee	3	1.03(0.84,1.25)
Caffeinated coffee	0	1.30(1.09,1.54)
Decaffeinated coffee	1	1.16(0.77,1.75)
Caffeine	0	1.24(0.93,1.64)
Tea	1	1.21(1.10,1.32)

Fruit	0	0.88(0.79,0.97)
Vegetable	0	0.90(0.76,1.06)
Total meat	1	0.91(0.81,1.03)
Red meat	2	1.13(0.98,1.30)
Poultry	0	1.11(0.95,1.28)
Total fish	0	0.92(0.78,1.08)
Total dairy	0	0.97(0.72,1.30)
Cereal product	0	0.87(0.80,0.96)

*seropositive RA

Table A12 (b). Trim and fill tests for nutrient intake and risk of rheumatoid arthritis.

Nutrients exposure	n to trim	Estimated effect after adding hypothesized studies
N-3 PUFA	2	1.01(0.82,1.24)
Vitamin C	0	0.98(0.76,1.27)
Vitamin D	1	0.74(0.59,0.93)
Vitamin D supplement	2	0.66(0.31,1.42)
Vitamin E	2	0.80(0.67,0.95)

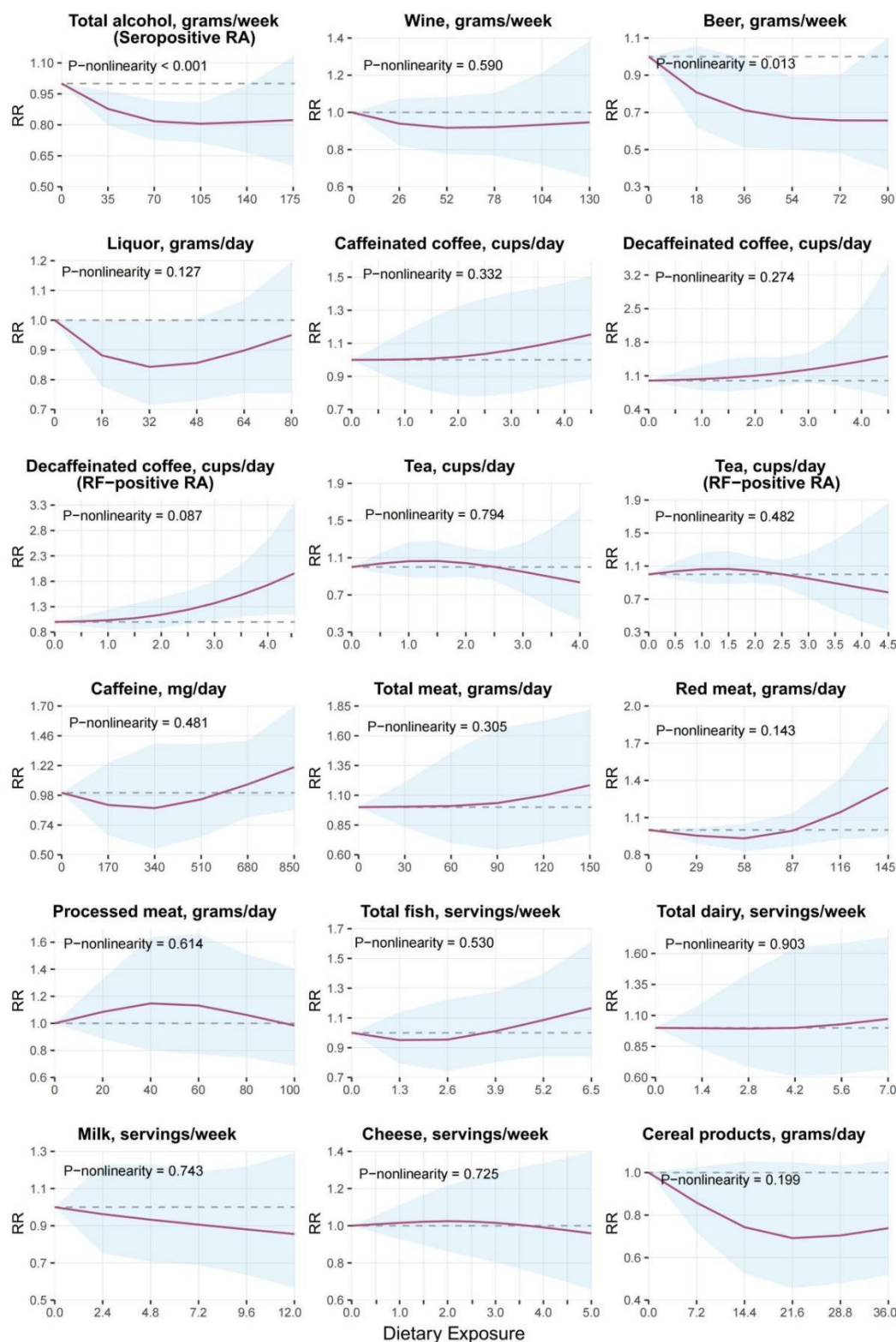


Figure A1. Non-linear dose-response meta-analyses for consumption of food group, and beverages and risk of rheumatoid arthritis using restricted cubic splines.

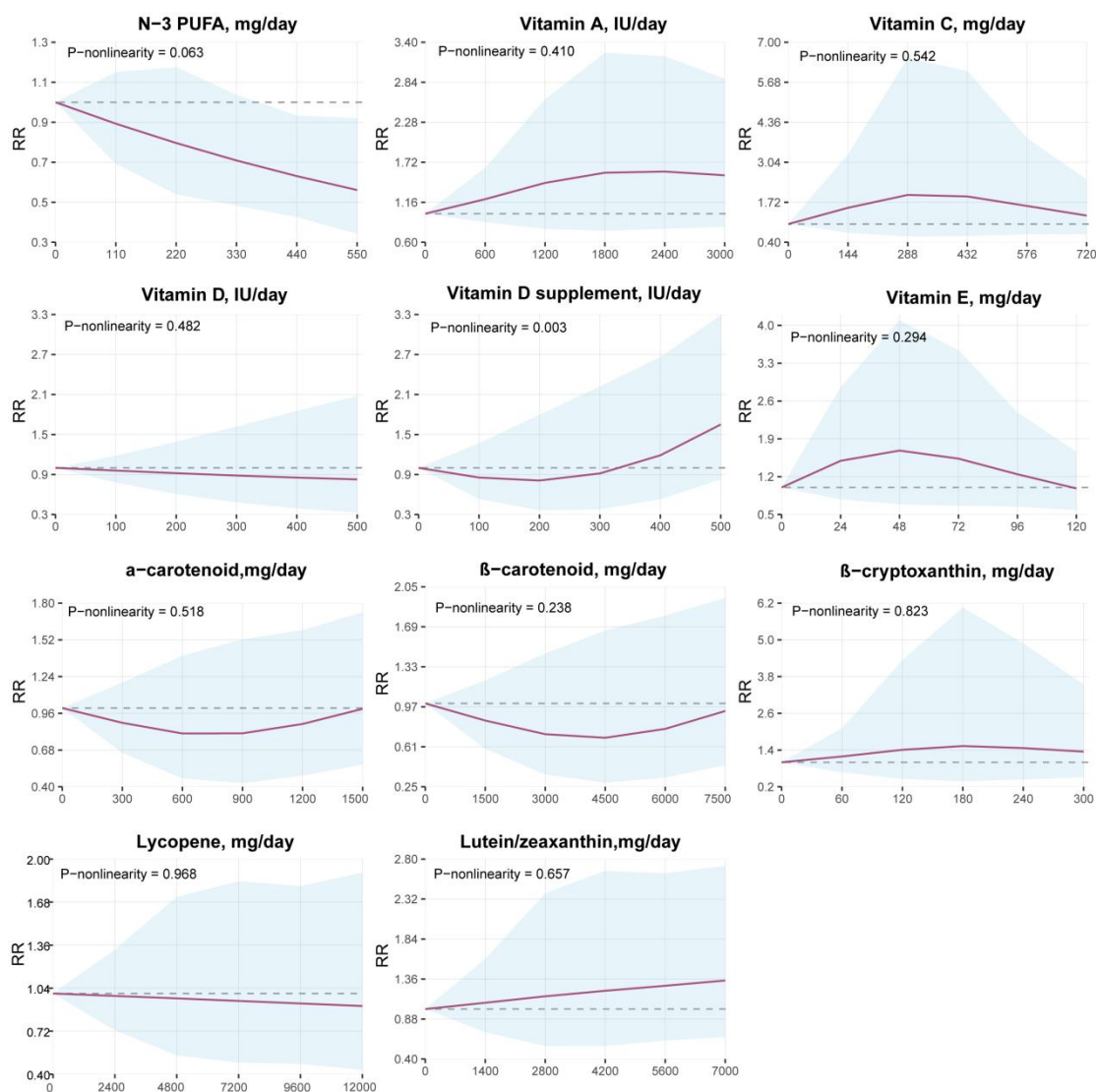


Figure A2. Non-linear dose-response meta-analyses for nutrient intake and risk of rheumatoid arthritis using restricted cubic splines.

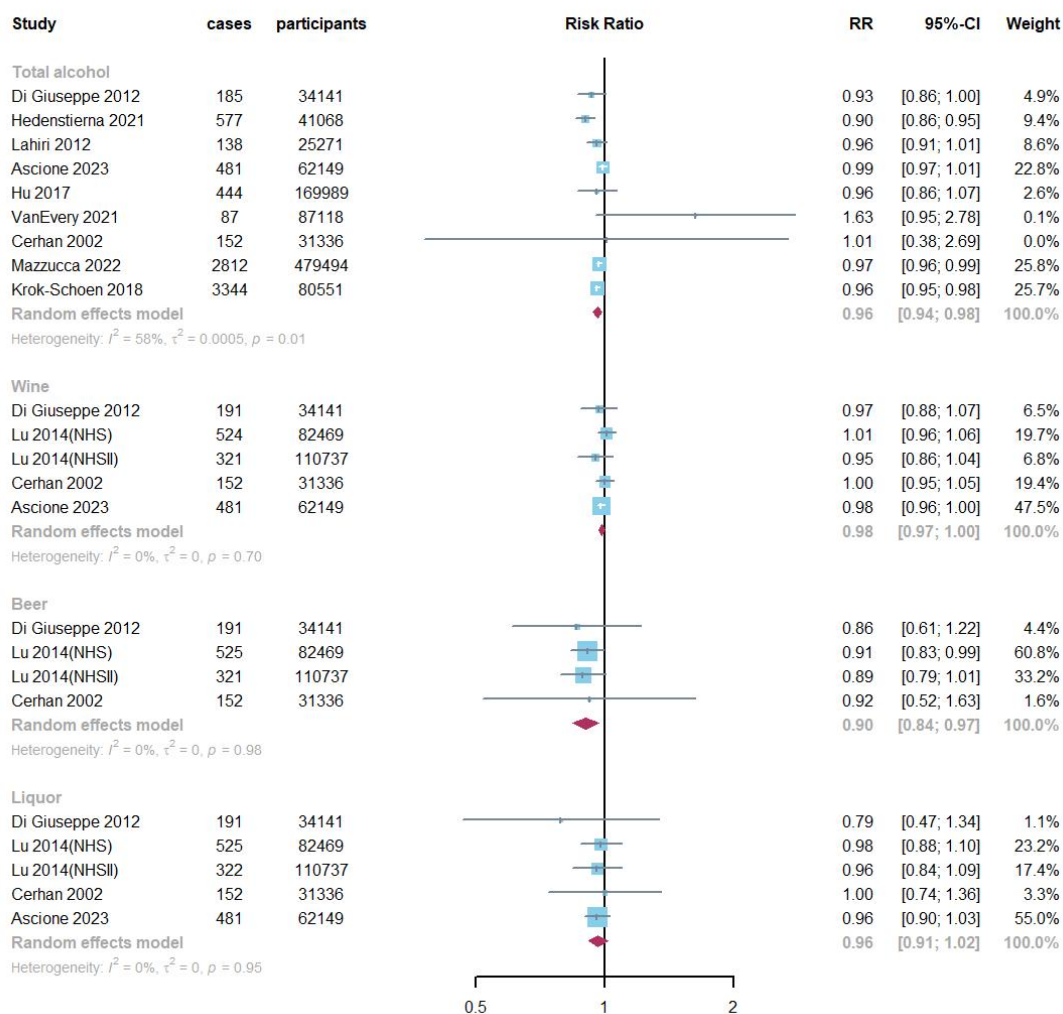


Figure A3. Linear dose-response analyses for alcohol consumption and risk of rheumatoid arthritis.

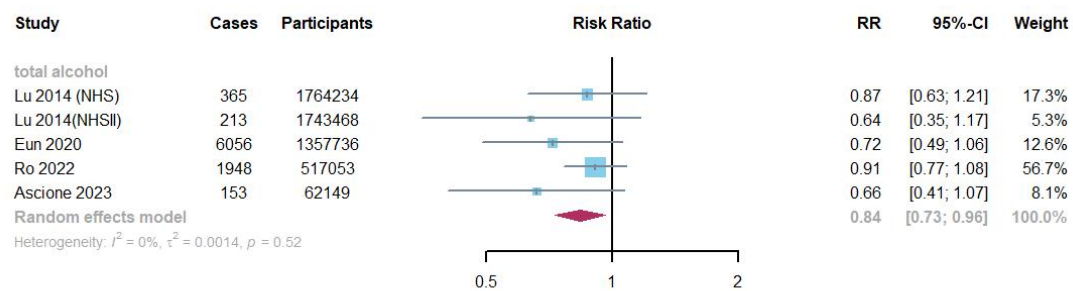


Figure A4. Meta-analysis of total alcohol consumption and risk of seropositive rheumatoid arthritis comparing highest with the lowest categories.

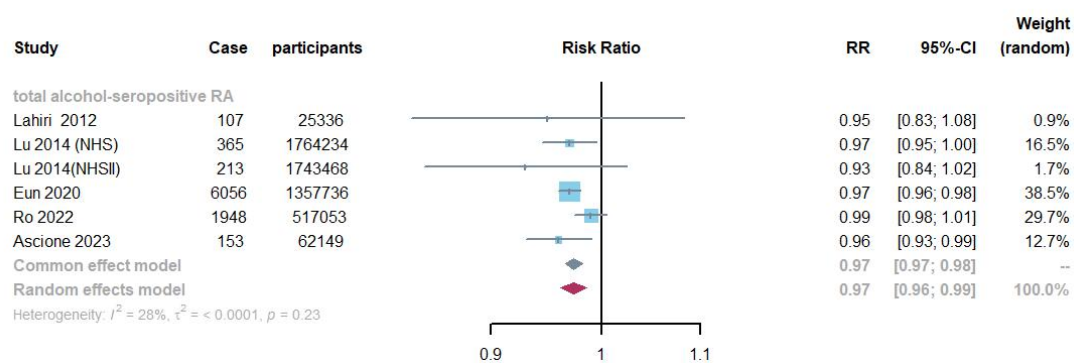


Figure A5. Linear dose-response analyses for total alcohol consumption and risk of seropositive rheumatoid arthritis.

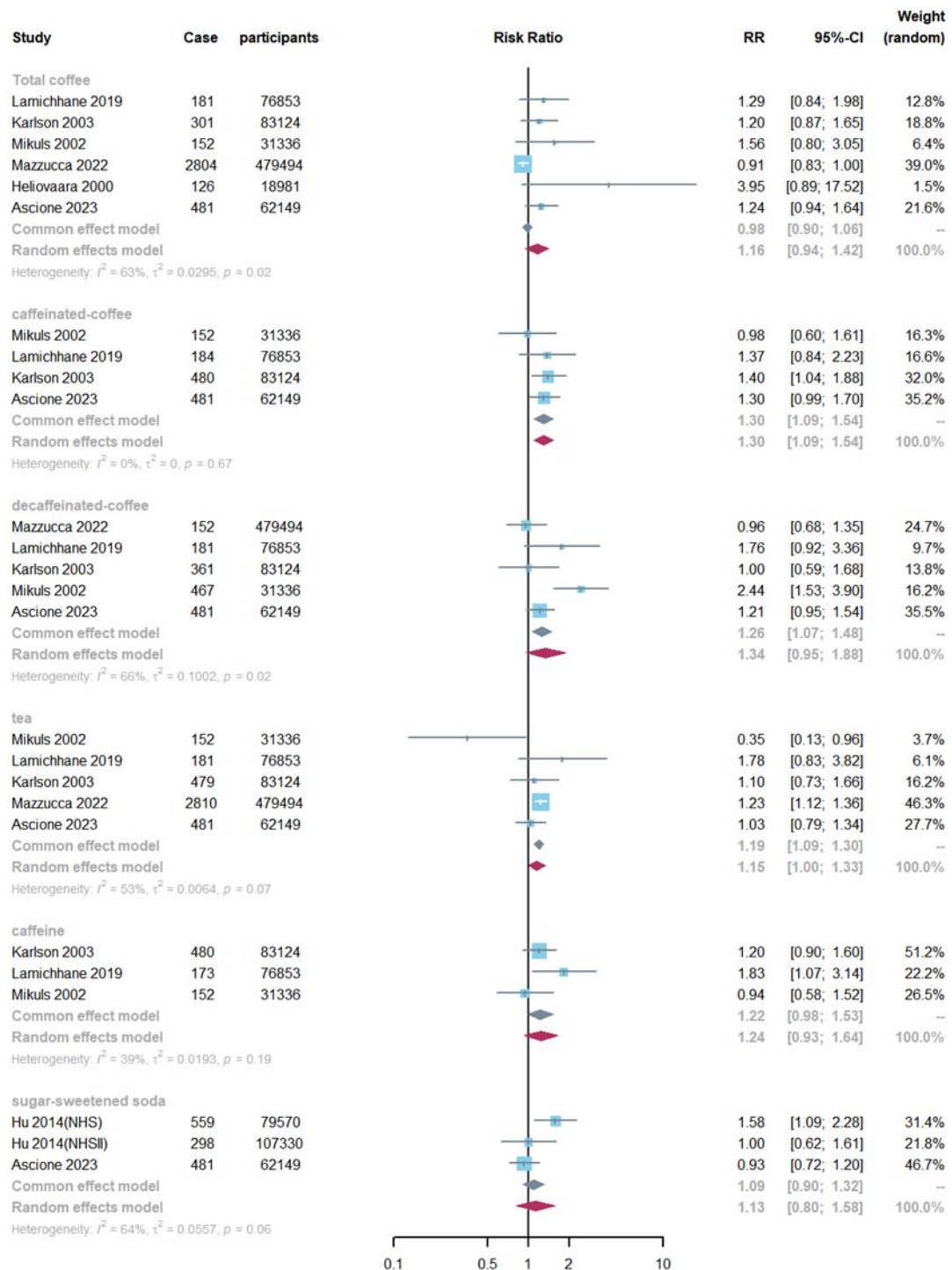


Figure A6. Meta-analysis of non-alcoholic beverages (tea and coffee, soda) consumption and risk of rheumatoid arthritis comparing highest with the lowest categories.

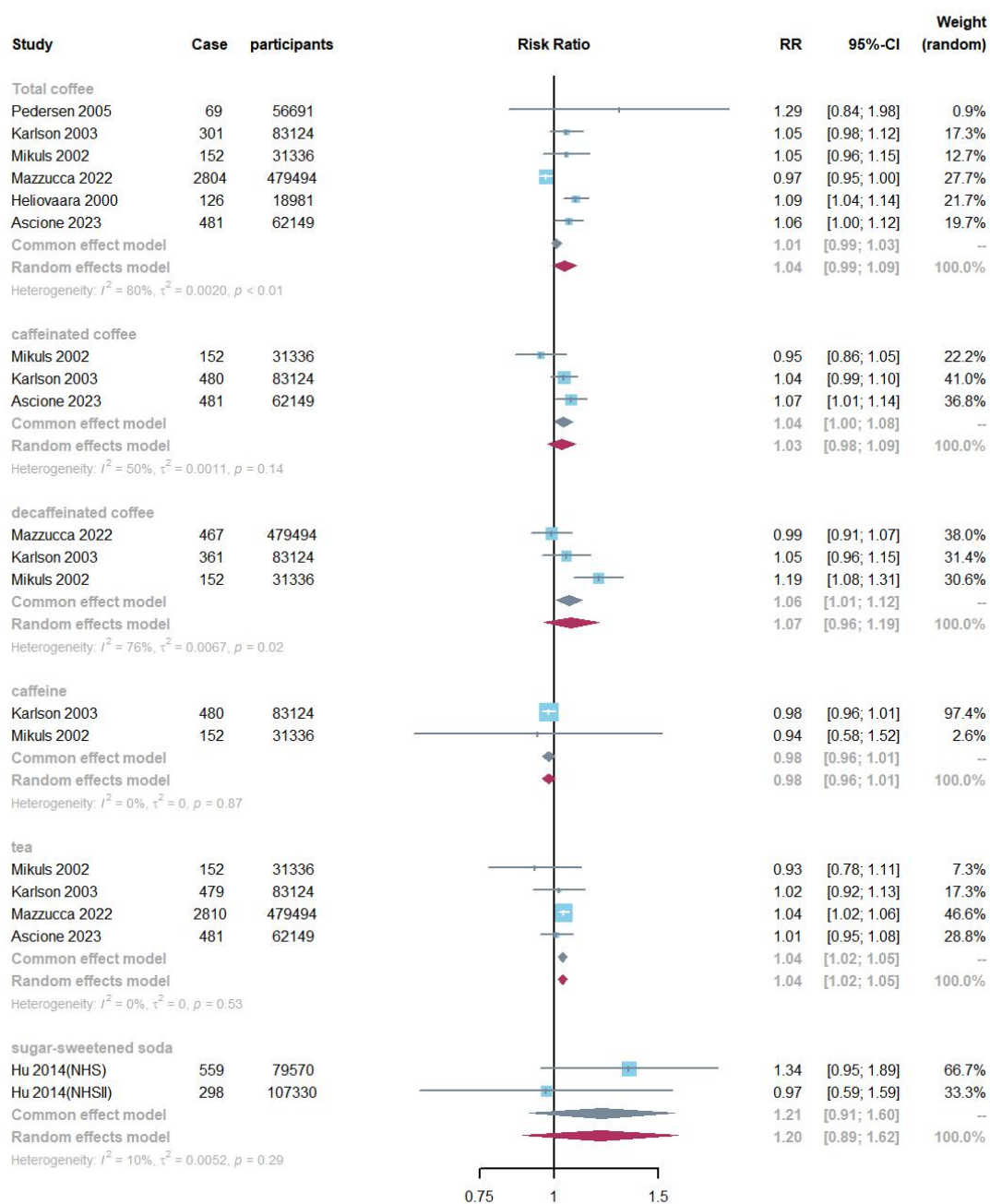


Figure A7. Linear dose-response analyses for non-alcoholic beverages (tea, coffee, and soda) consumption and risk of rheumatoid arthritis.

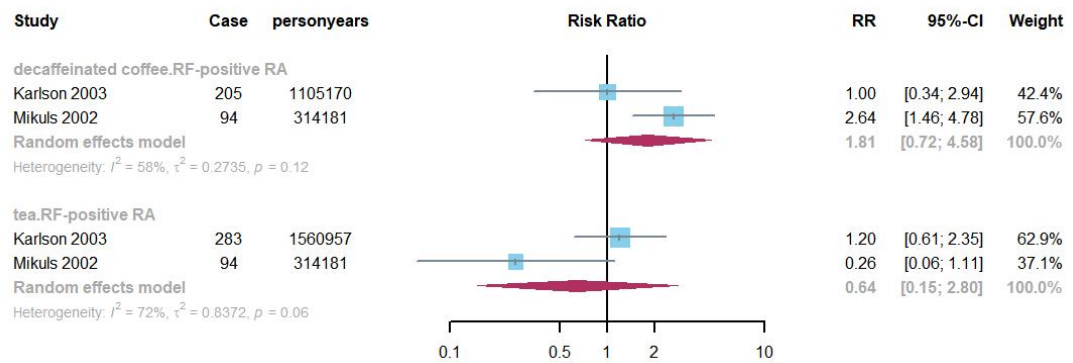


Figure A8. Meta-analysis of decaffeinated coffee, and tea consumption and risk of RF-positive rheumatoid arthritis comparing highest with the lowest categories.

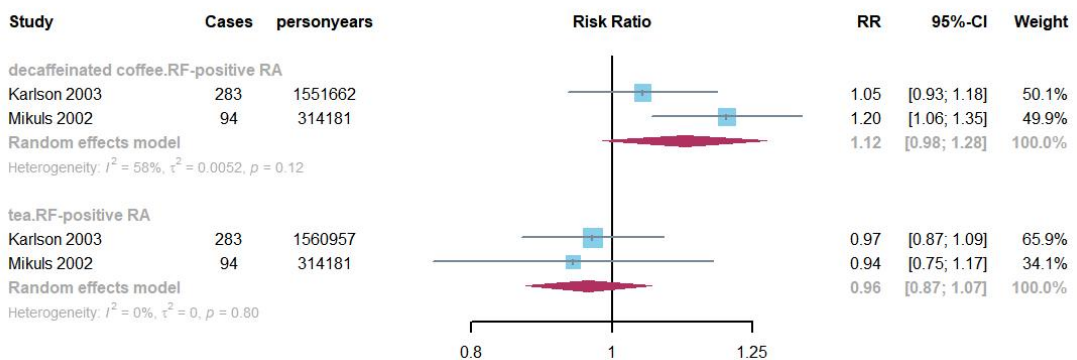


Figure A9. Linear dose-response analyses for decaffeinated coffee, and tea consumption and risk of RF-positive rheumatoid arthritis.

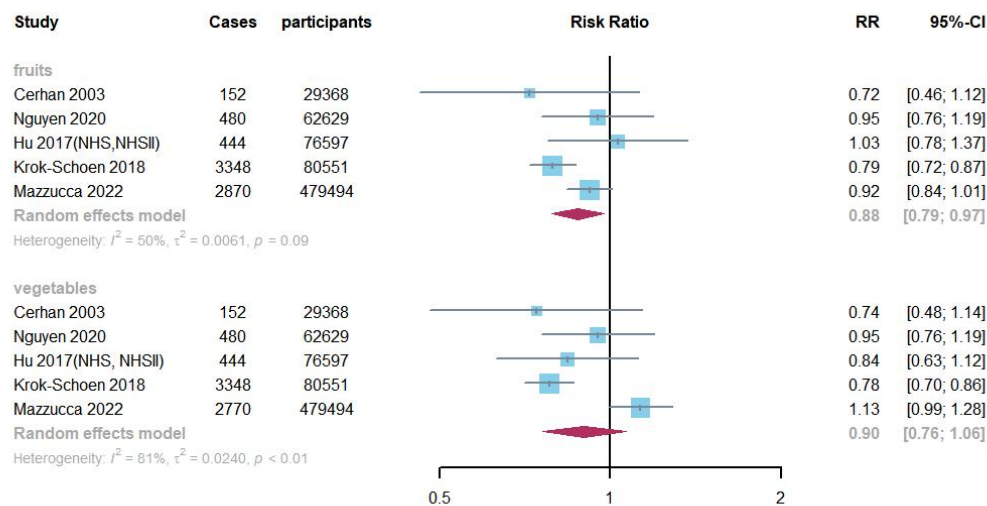


Figure A10. Meta-analysis of fruit, and vegetable consumption and risk of rheumatoid arthritis comparing highest with the lowest categories.

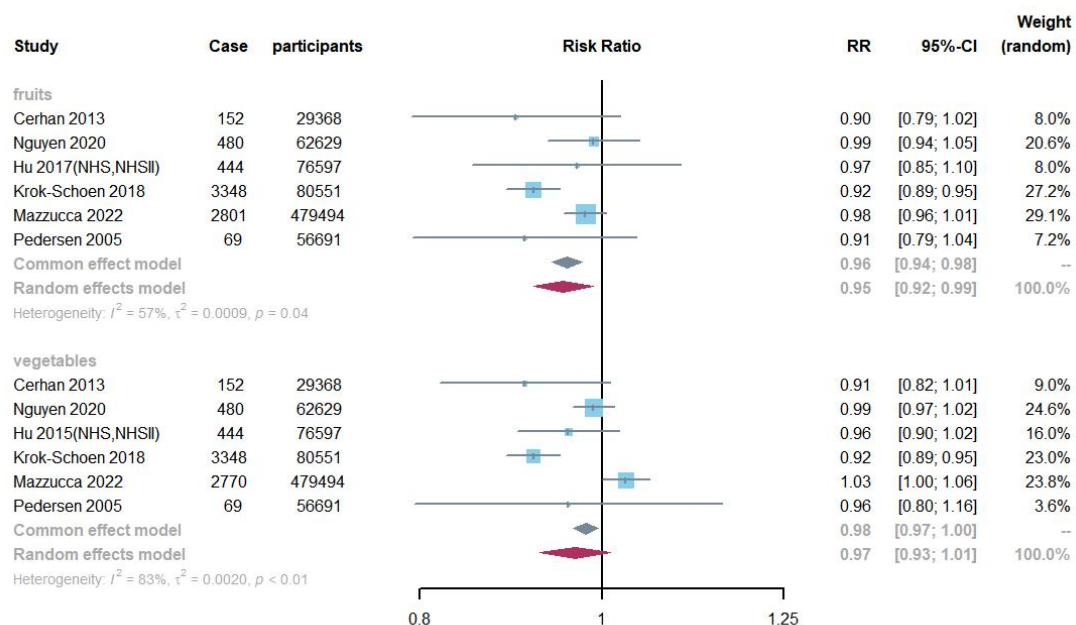


Figure A11. Linear dose-response analyses for fruit, and vegetable consumption and risk of rheumatoid arthritis.

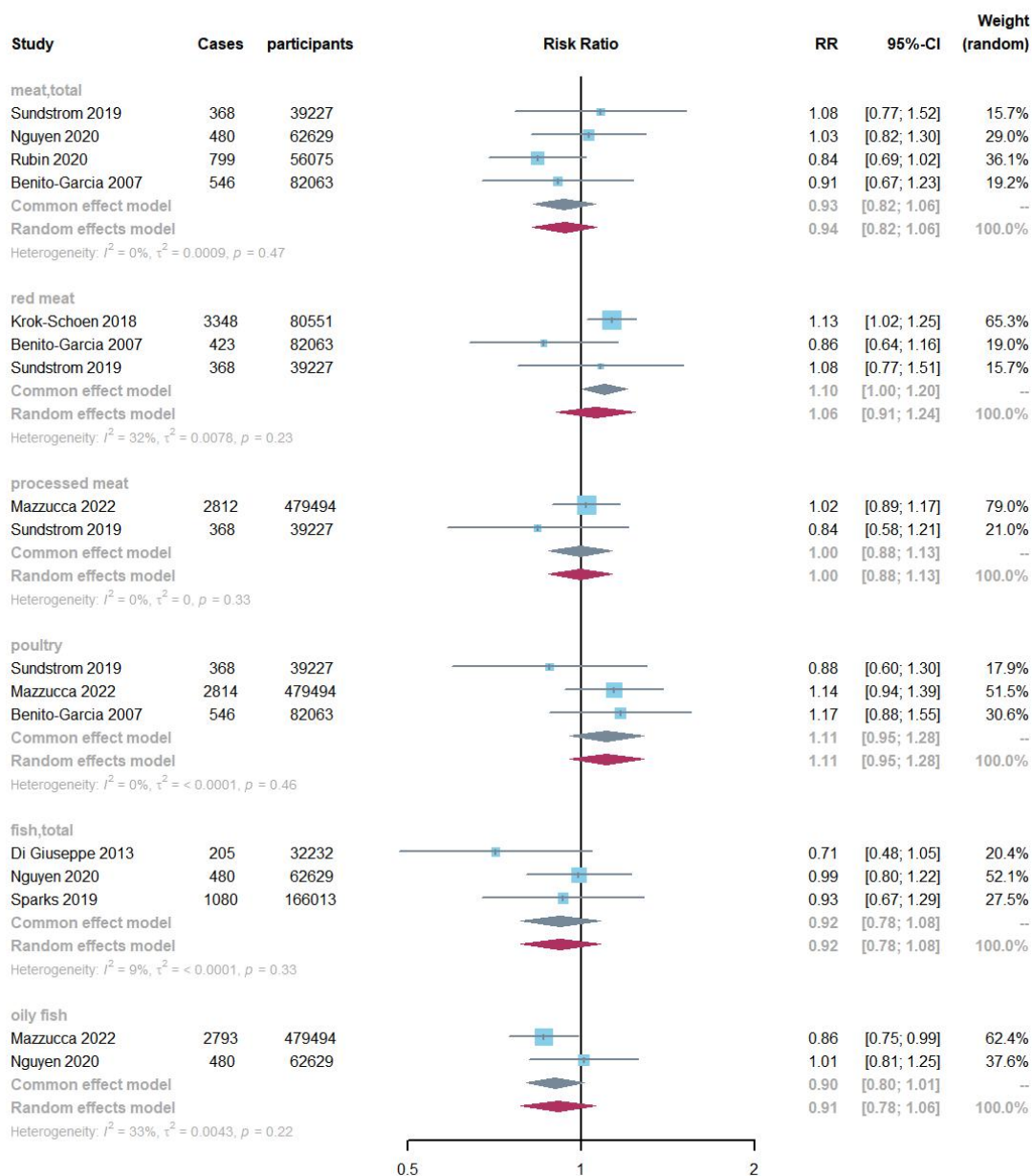


Figure A12. Meta-analysis of meat, and fish consumption and risk of rheumatoid arthritis comparing highest with the lowest categories.

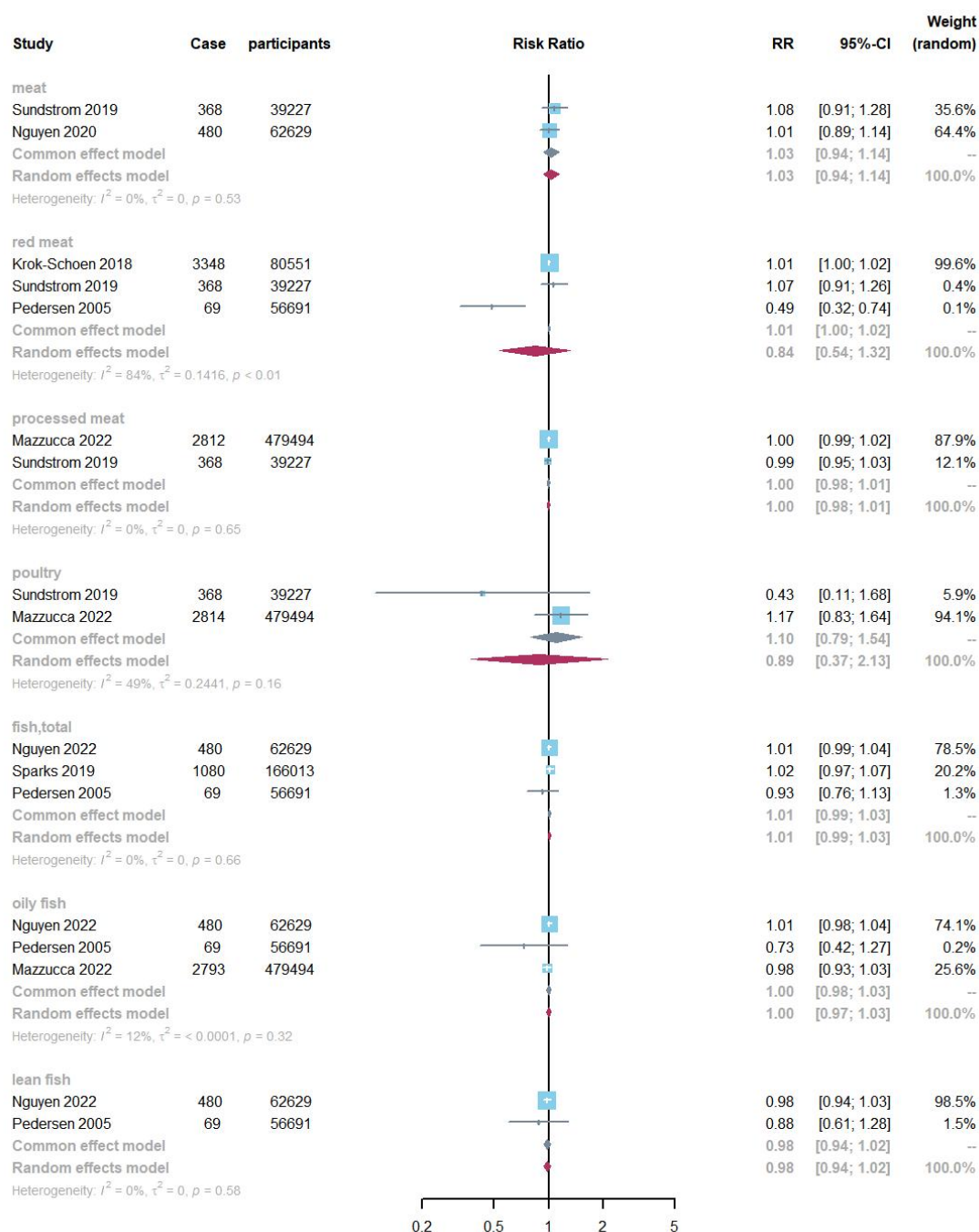


Figure A13. Linear dose-response analyses for meat, and fish consumption and risk of rheumatoid arthritis.

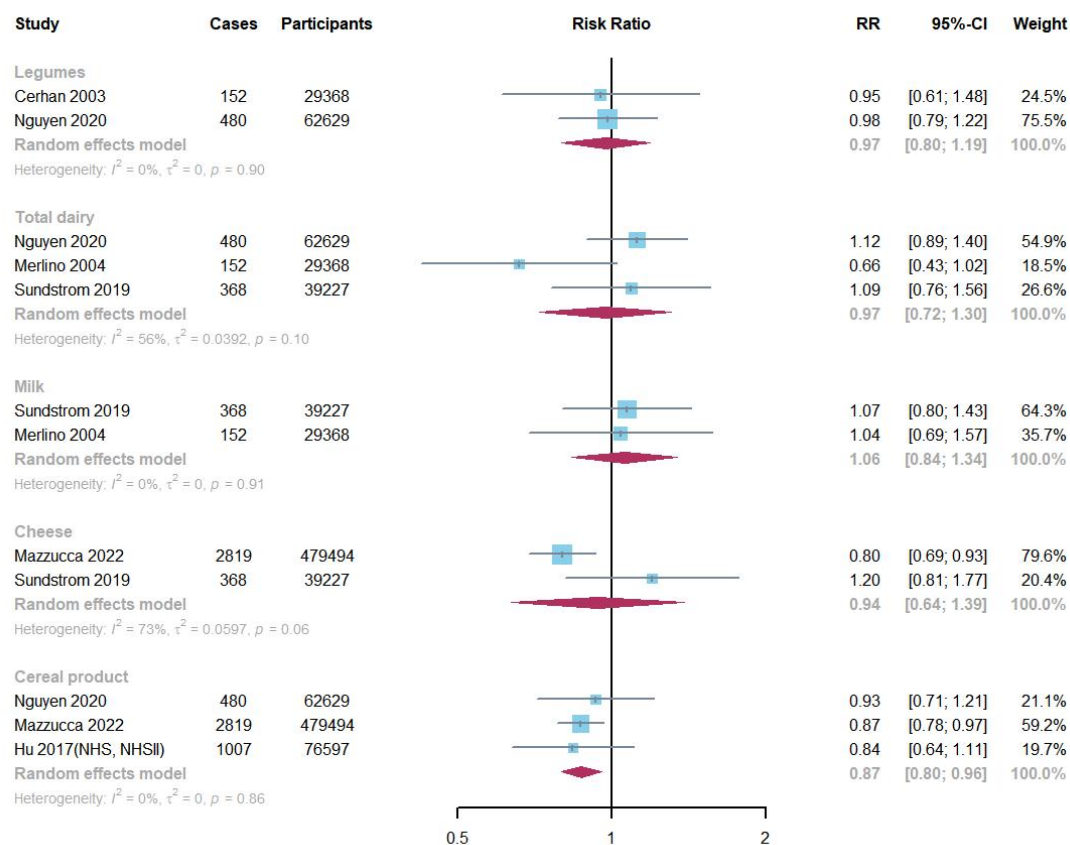


Figure A14. Meta-analysis of other food group consumption and risk of rheumatoid arthritis comparing highest with the lowest categories.

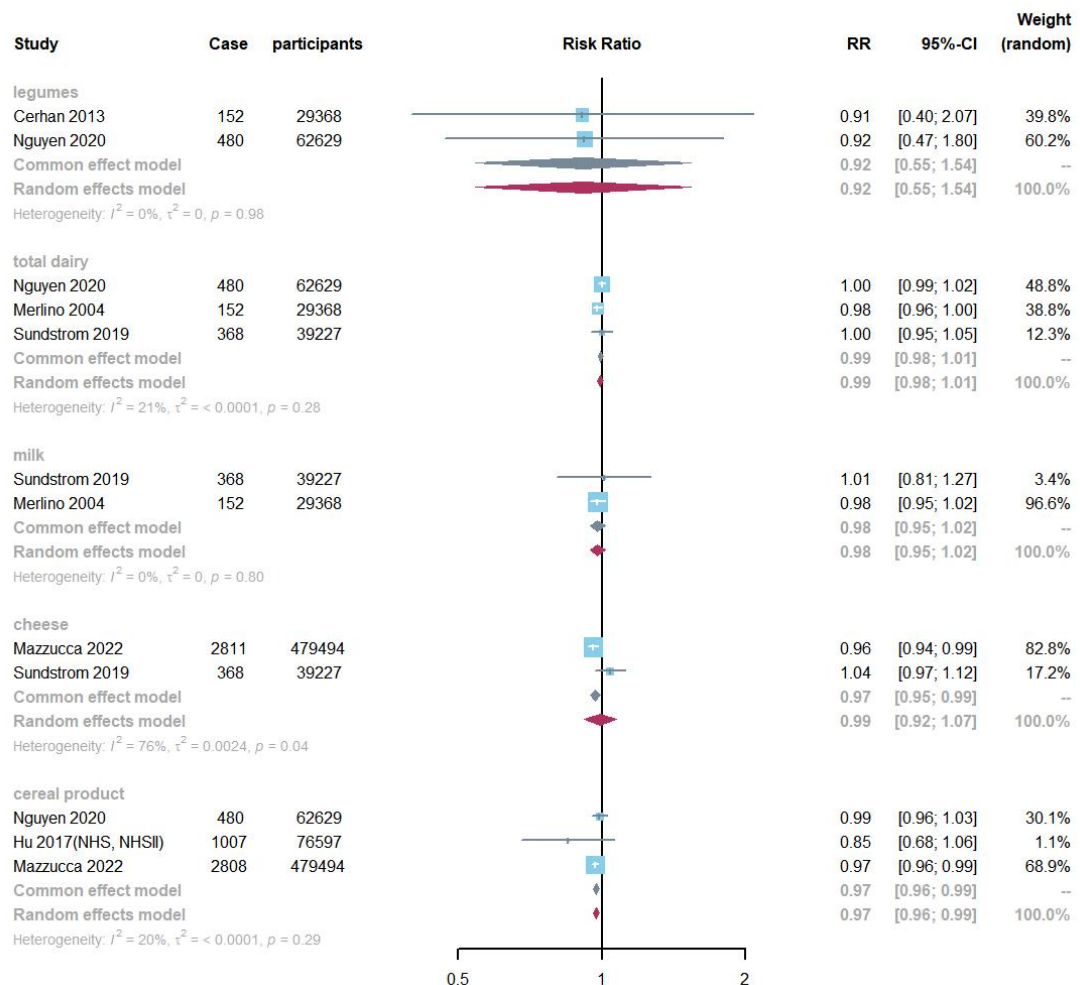


Figure A15. Linear dose-response analyses for other food group consumption and risk of rheumatoid arthritis.

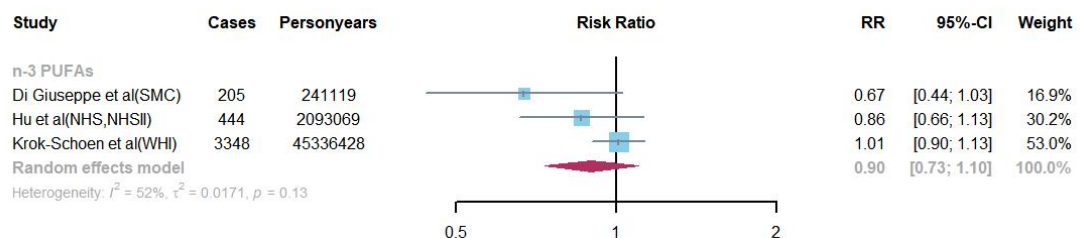


Figure A16. Meta-analysis of dietary n-3 PUFA and risk of rheumatoid arthritis comparing highest with the lowest categories.

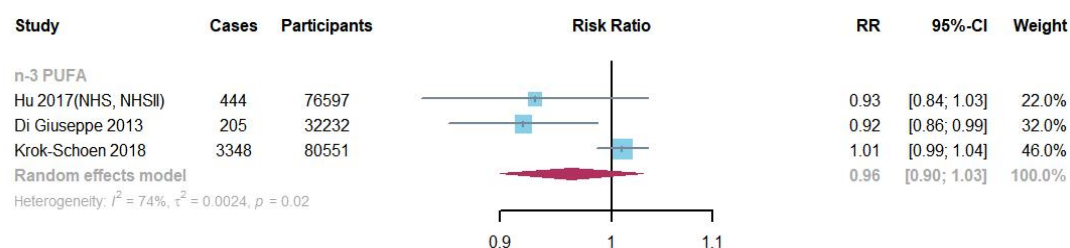


Figure A17. Linear dose-response analyses for dietary n-3 PUFA and risk of rheumatoid arthritis.

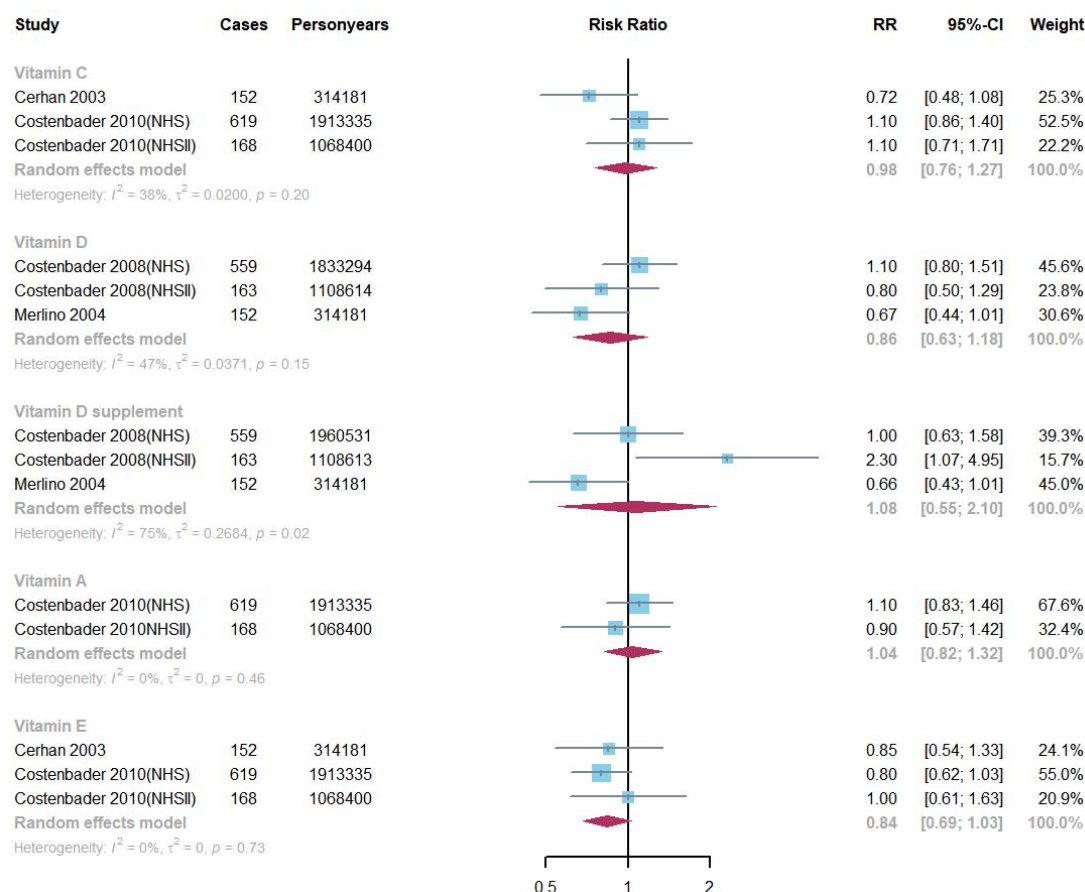


Figure A18. Meta-analysis of vitamin intake and risk of rheumatoid arthritis comparing highest with the lowest categories.

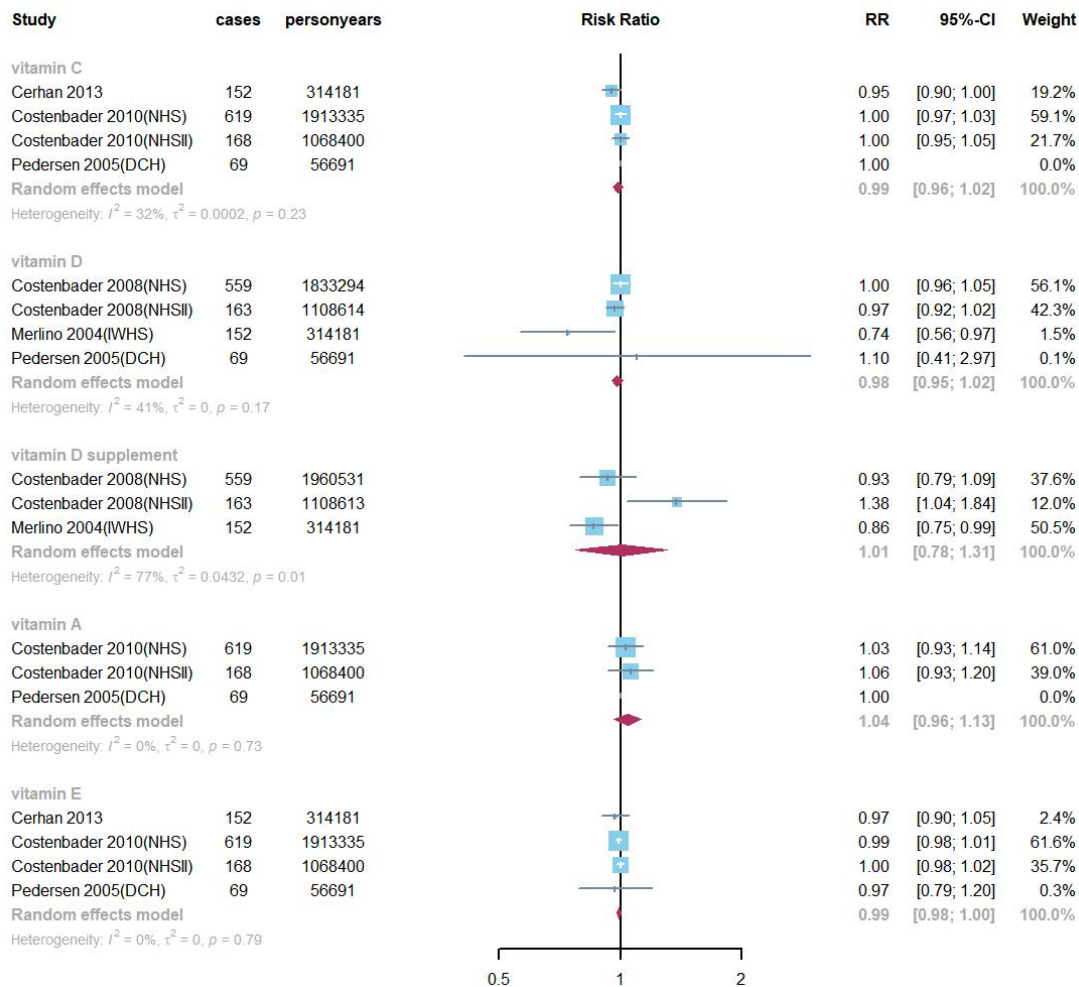


Figure A19. Linear dose-response analyses for vitamins intake and risk of rheumatoid arthritis.

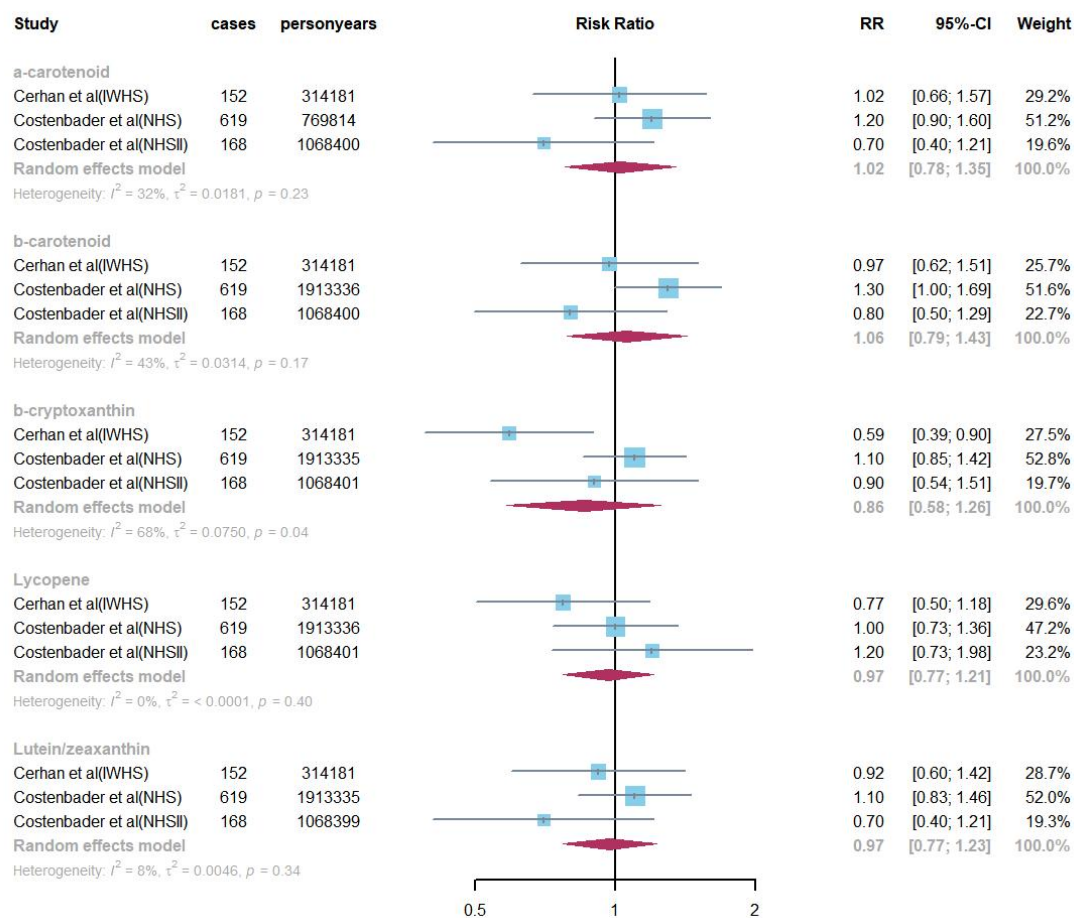


Figure A20. Meta-analysis of dietary carotenoids intake and the risk of rheumatoid arthritis comparing highest with the lowest categories.

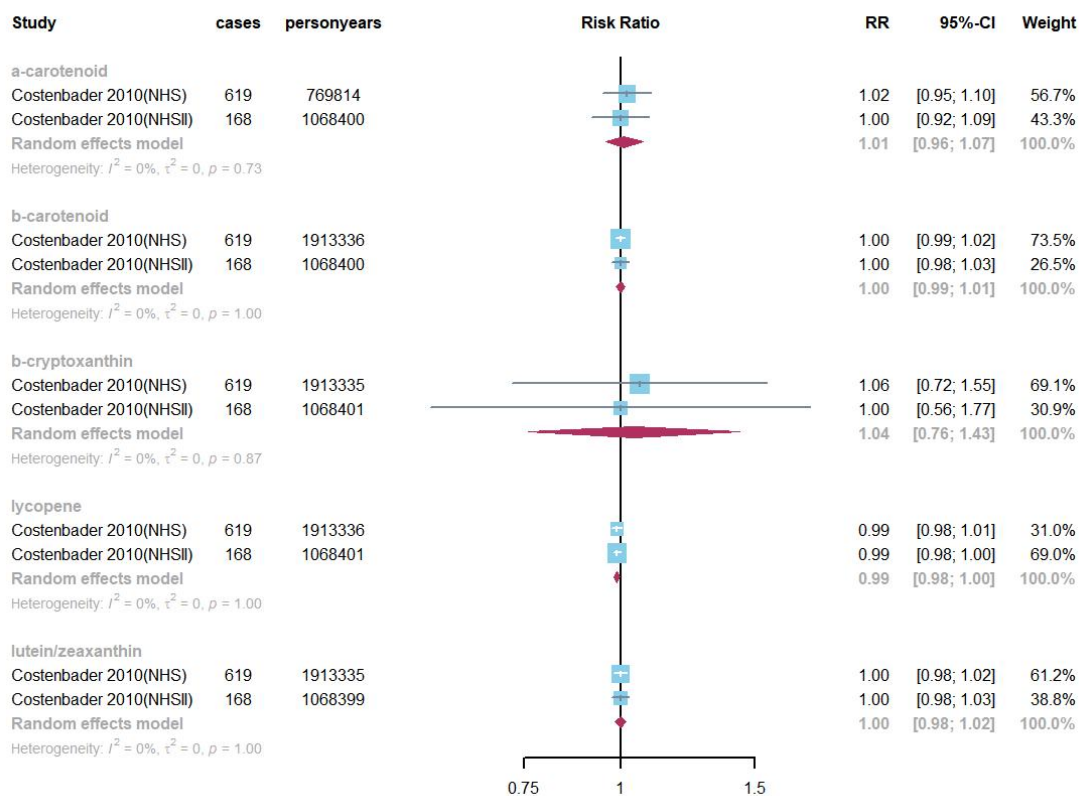


Figure A21. Linear dose-response analyses for dietary carotenoids intake and risk of rheumatoid arthritis.

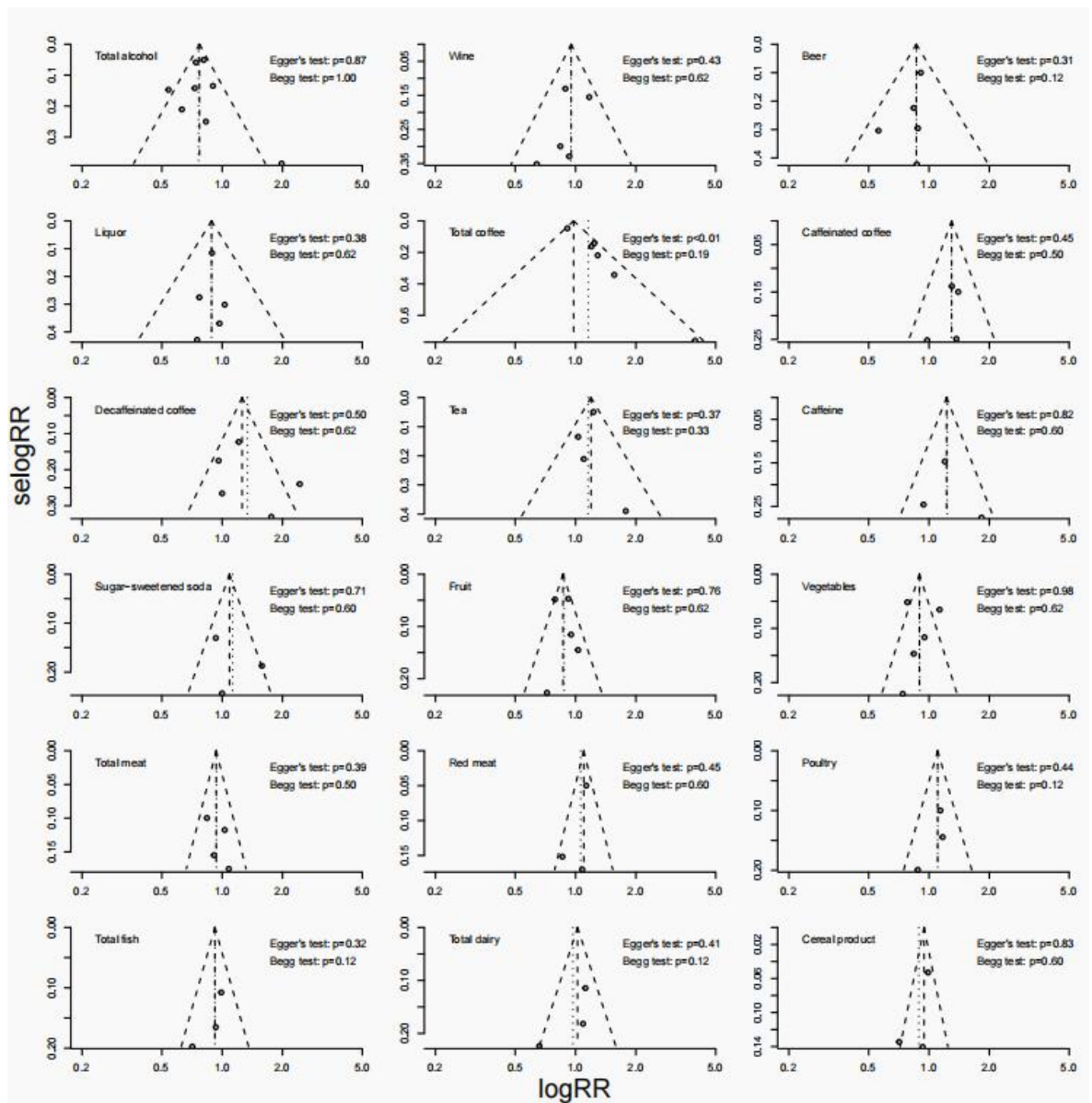


Figure A22. Funnel plots for food group, and beverages consumption and risk of rheumatoid arthritis.

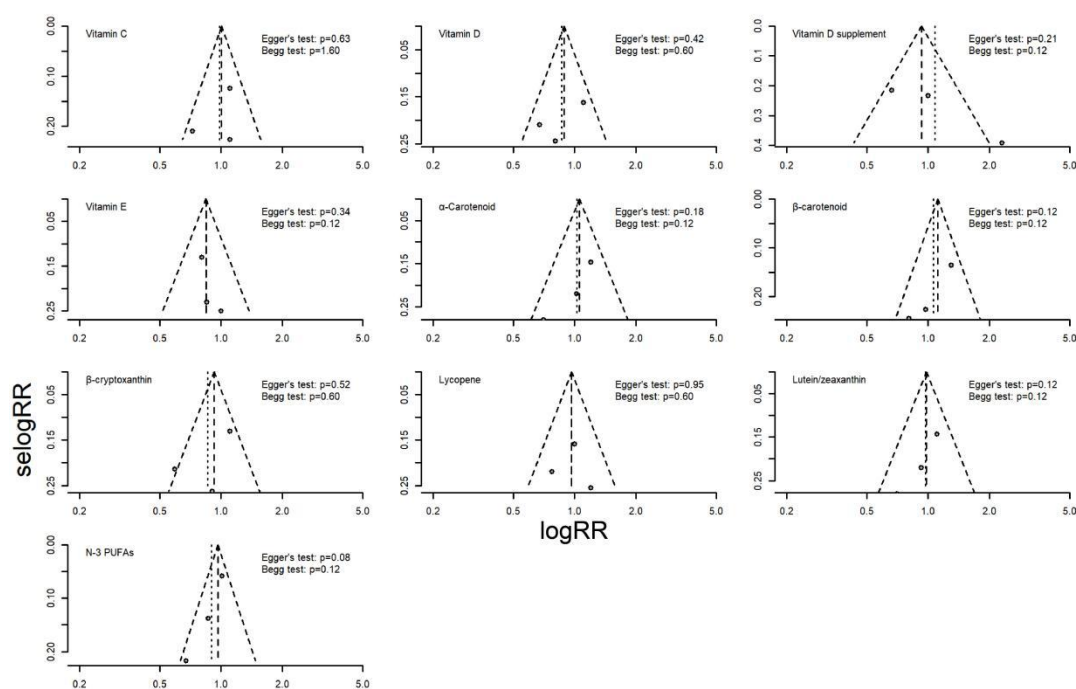


Figure A23. Funnel plots for nutrients intake and risk of rheumatoid arthritis.

References

1. Cerhan JR, Saag KG, Merlino LA, et al. Antioxidant micronutrients and risk of rheumatoid arthritis in a cohort of older women. *American journal of epidemiology* 2003;157(4):345-54.
2. Costenbader KH, Feskanich D, Holmes M, et al. Vitamin D intake and risks of systemic lupus erythematosus and rheumatoid arthritis in women. *Annals of the rheumatic diseases* 2008;67(4):530-5.
3. Costenbader KH, Kang JH, Karlson EW. Antioxidant intake and risks of rheumatoid arthritis and systemic lupus erythematosus in women. *American journal of epidemiology* 2010;172(2):205-16.
4. Di Giuseppe D, Alfredsson L, Bottai M, et al. Long term alcohol intake and risk of rheumatoid arthritis in women: a population based cohort study. *BMJ (Clinical research ed)* 2012;345:e4230.
5. Sundström B, Ljung L, Di Giuseppe D. Consumption of Meat and Dairy Products Is Not Associated with the Risk for Rheumatoid Arthritis among Women: A Population-Based Cohort Study. *Nutrients* 2019;11(11).
6. Di Giuseppe, D.; Wallin, A.; Bottai, M.; Askling, J.; Wolk, A. Long-term intake of dietary long-chain n-3 polyunsaturated fatty acids and risk of

- rheumatoid arthritis: A prospective cohort study of women. *Ann. Rheum. Dis.* 2014, 73, 1949–1953.
7. Hedenstierna L, Bellocco R, Ye W, et al. Effects of alcohol consumption and smoking on risk for RA: results from a Swedish prospective cohort study. *RMD open* 2021;7(1).
 8. Hiraki LT, Munger KL, Costenbader KH, et al. Dietary intake of vitamin D during adolescence and risk of adult-onset systemic lupus erythematosus and rheumatoid arthritis. *Arthritis care & research* 2012;64(12):1829-36.
 9. Merlino LA, Curtis J, Mikuls TR, et al. Vitamin D intake is inversely associated with rheumatoid arthritis: results from the Iowa Women's Health Study. *Arthritis and rheumatism* 2004;50(1):72-7.
 10. Lahiri M, Luben RN, Morgan C, et al. Using lifestyle factors to identify individuals at higher risk of inflammatory polyarthritis (results from the European Prospective Investigation of Cancer-Norfolk and the Norfolk Arthritis Register--the EPIC-2-NOAR Study). *Annals of the rheumatic diseases* 2014;73(1):219-26.
 11. Nguyen Y, Salliot C, Gelot A, et al. Mediterranean Diet and Risk of Rheumatoid Arthritis: Findings From the French E3N-EPIC Cohort Study. *Arthritis & rheumatology (Hoboken, NJ)* 2021;73(1):69-77.
 12. Hu Y, Sparks JA, Malspeis S, et al. Long-term dietary quality and risk of developing rheumatoid arthritis in women. *Annals of the rheumatic diseases* 2017;76(8):1357-64.
 13. Karlson EW, Mandl LA, Aweh GN, et al. Coffee consumption and risk of rheumatoid arthritis. *Arthritis and rheumatism* 2003;48(11):3055-60.
 14. Krok-Schoen JL, Brasky TM, Hunt RP, et al. Dietary Long-Chain n-3 Fatty Acid Intake and Arthritis Risk in the Women's Health Initiative. *Journal of the Academy of Nutrition and Dietetics* 2018;118(11):2057-69.
 15. Lamichhane D, Collins C, Constantinescu F, et al. Coffee and Tea Consumption in Relation to Risk of Rheumatoid Arthritis in the Women's Health Initiative Observational Cohort. *Journal of clinical rheumatology : practical reports on rheumatic & musculoskeletal diseases* 2019;25(3):127-32.

16. Lu B, Solomon DH, Costenbader KH, et al. Alcohol consumption and risk of incident rheumatoid arthritis in women: a prospective study. *Arthritis & rheumatology (Hoboken, NJ)* 2014;66(8):1998-2005.
17. Mikuls TR, Cerhan JR, Criswell LA, et al. Coffee, tea, and caffeine consumption and risk of rheumatoid arthritis: results from the Iowa Women's Health Study. *Arthritis and rheumatism* 2002;46(1):83-91.
18. Pedersen M, Stripp C, Klarlund M, et al. Diet and risk of rheumatoid arthritis in a prospective cohort. *The Journal of rheumatology* 2005;32(7):1249-52.
19. Rubin KH, Rasmussen NF, Petersen I, et al. Intake of dietary fibre, red and processed meat and risk of late-onset Chronic Inflammatory Diseases: A prospective Danish study on the "diet, cancer and health" cohort. *International journal of medical sciences* 2020;17(16):2487-95.
20. Sparks JA, O'Reilly É J, Barbhaiya M, et al. Association of fish intake and smoking with risk of rheumatoid arthritis and age of onset: a prospective cohort study. *BMC musculoskeletal disorders* 2019;20(1):2.
21. VanEvery H, Yang W, Olsen N, et al. Alcohol Consumption and Risk of Rheumatoid Arthritis among Chinese Adults: A Prospective Study. *Nutrients* 2021;13(7).
22. Cerhan JR, Saag KG, Criswell LA, et al. Blood transfusion, alcohol use, and anthropometric risk factors for rheumatoid arthritis in older women. *The Journal of rheumatology* 2002;29(2):246-54.
23. Benito-Garcia E, Feskanich D, Hu FB, et al. Protein, iron, and meat consumption and risk for rheumatoid arthritis: a prospective cohort study. *Arthritis Res Ther* 2007;9(1):R16.
24. Hu Y, Costenbader KH, Gao X, et al. Sugar-sweetened soda consumption and risk of developing rheumatoid arthritis in women. *The American journal of clinical nutrition* 2014;100(3):959-67.
25. Nguyen Y, Salliot C, Mariette X, et al. Fish Consumption and Risk of Rheumatoid Arthritis: Findings from the E3N Cohort Study. *Nutrients* 2022;14(4).
26. Heliövaara M, Aho K, Knekt P, et al. Coffee consumption, rheumatoid factor, and the risk of rheumatoid arthritis. *Annals of the rheumatic diseases* 2000;59(8):631-5.

27. Mazzucca CB, Scotti L, Cappellano G, et al. Nutrition and Rheumatoid Arthritis Onset: A Prospective Analysis Using the UK Biobank. *Nutrients* 2022;14(8).
28. Eun Y, Jeon KH, Han K, et al. Menopausal factors and risk of seropositive rheumatoid arthritis in postmenopausal women: a nationwide cohort study of 1.36 million women. *Scientific reports* 2020;10(1):20793.
29. Ro J, Kim SH, Kim HR, et al. Impact of lifestyle and comorbidities on seropositive rheumatoid arthritis risk from Korean health insurance data. *Scientific reports* 2022;12(1):2201.
30. Ascione S, Barde F, Artaud F, et al. Association between beverage consumption and risk of rheumatoid arthritis: a prospective study from the French E3N Cohort. *Rheumatology (Oxford, England)* 2023;62(5):1814-23.

Appendix B Chapter 3 Supplementary material

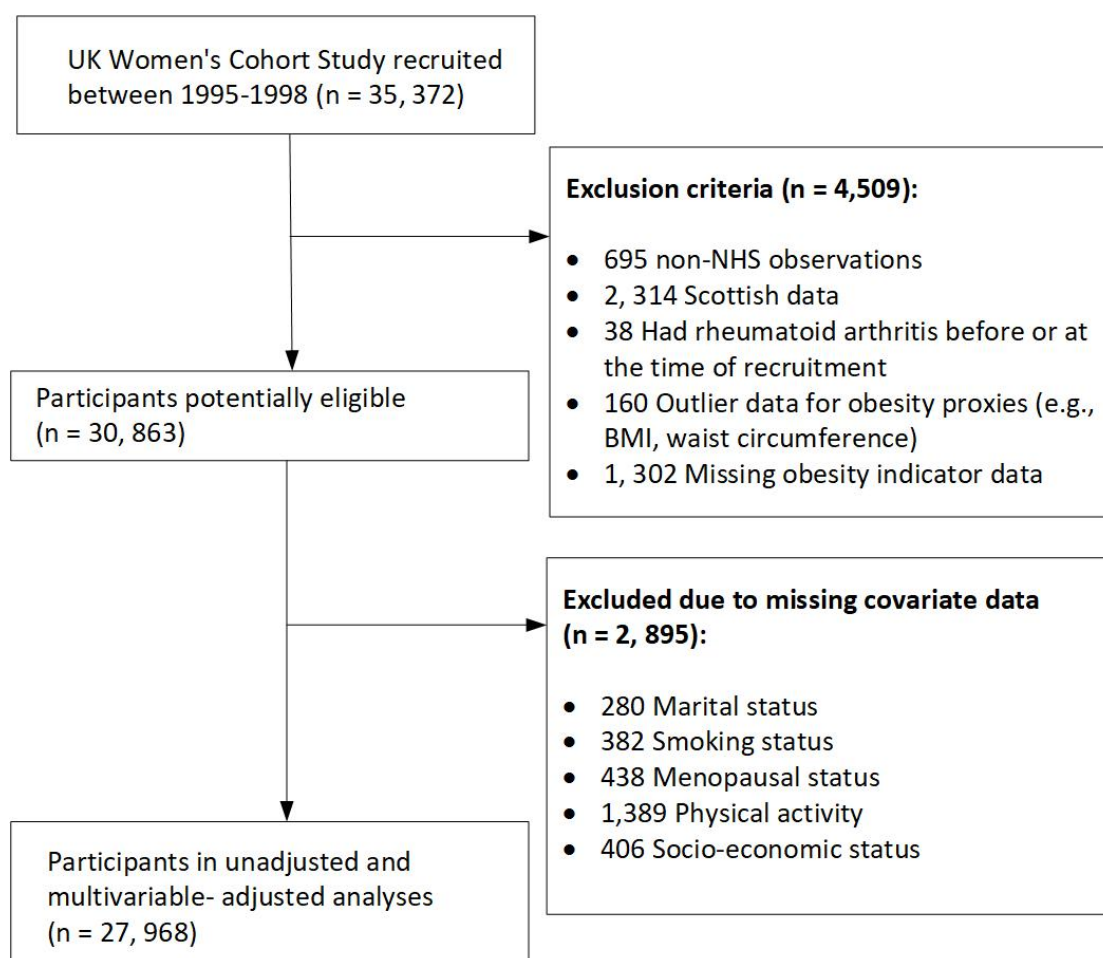


Figure B1. Flow chart of UKWCS participants.

Table B1. Strengthening the reporting of observational studies in nutritional epidemiology (STROBE-Nut) checklist.

Topic	Item number	Recommendation	Page (line number)
Title and abstract	1	Indicate the study's design with a commonly used term in the title or the abstract	1(1-4)
		Provide in the abstract an informative and balanced summary of what was done and what was found	1(12-36)
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	1(38-42), 2(43-77)
Objectives	3	State specific objectives, including any prespecified hypotheses	2(68-77)
Methods			
Study design	4	Present key elements of study design early in the manuscript	3(84-100)
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	3(84-100)

Participants	6	Cohort study - give the eligibility criteria, and the sources and methods of selection of participants; describe methods of follow-up	4(139-142), Supplementary Table A1.
		Cohort study - for matched studies, give matching criteria and number of exposed and unexposed Case-control study - for matched studies, give matching criteria and the number of controls per case	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers; give diagnostic criteria, if applicable	3(101-125), 4(125-147) Table A1, Table A5, A6
Data sources/measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement); describe comparability of assessment methods if there is more than one group	3(80-125), 4(126-147) Table A1, Figure A1, Table A5 and A6
Bias	9	Describe any efforts to address potential sources of bias	4(151-174), 5(175-188)
Study size	10	Explain how the study size was arrived at	5(190-193), Figure A1
Quantitative variables	11	Explain how quantitative variables were handled in the analyses; if applicable, describe which groupings were chosen and why	4(151-162)

Statistical methods	12	Describe all statistical methods, including those used to control for confounding	4(143-170) 5(171-183)
		Describe any methods used to examine subgroups and interactions	4(172-174), 5(175-180)
		Explain how missing data were addressed	5(185-188)
		Cohort study - if applicable, explain how loss to follow-up was addressed	N/A
		Describe any sensitivity analyses	5(182-188)
Results			
Participants	13*	Report numbers of individuals at each stage of study - e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analyzed	5(190-193)
		Give reasons for nonparticipation at each stage Consider use of a flow diagram	Figure A1
Descriptive data	14*	Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders	5(190-221)

		Indicate number of participants with missing data for each variable of interest Cohort study - summarize follow-up time (e.g., average and total amount)	5(190-221)
Outcome data	15*	Cohort study - report numbers of outcome events or summary measures over time	5(184-214) Table 1, Table A5 and A6
Main results	16	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval); make clear which confounders were adjusted for and why they were included Report category boundaries when continuous variables were categorized	Figure 1, Figure 2
		If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done - e.g., analyses of subgroups and interactions, and sensitivity analyses	6(239-264), Table A7-A19
Discussion			
Key results	18	Summarize key results with reference to study objectives	10(292-302)
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision; discuss both direction and	12(394-414)

		magnitude of any potential bias	
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	11(303-337),12(338-383)
Generalizability	21	Discuss the generalizability (external validity) of the study results	13(416-421)
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	14(457-458)

Table B2. Adiposity group categorisation and definitions.

Classification Systems	Category	Definition
BMI	Underweight	BMI < 18.5 kg/m ²
	Normal weight	BMI 18.5 - < 25 kg/m ²
	Overweight	BMI 25.0 - < 30 kg/m ²
	Obese	BMI ≥30.0 kg/m ²
WC	No abdominal obesity	WC ≤88 cm
	Abdominal obesity	WC > 88 cm
WHR	Normal	WHR: < 0.8
		WHR: 0.8-0.85
	Elevated	WHR ≥ 0.85
WHtR	Normal	WHtR < 0.5
	Overweight	WHtR 0.5- < 0.6
	Obese	WHtR ≥ 0.6

BMI, WC, and WHR categories are defined by World Health Organization (WHO) criteria[1]. WHtR categories are defined by NICE criteria[2]. BMI: body mass index. WC: waist circumference. WHR: waist-to-hip ratio. WHtR: waist-to-height ratio.

[2] Ashwell M, Mayhew L, Richardson J, Rickayzen B. Waist-to-height ratio is more predictive of years of life lost than body mass index. PLoS One. 2014 Sep 8;9(9):e103483. doi: 10.1371/journal.pone.0103483.

Table B3. Covariates at recruitment and their derivation.

Covariate	How the variable was derived
Socio-demographic variables	
Age	Calculated as year differences between date of birth and date of recruitment and was considered a continuous variable in adjustment sets.
Socio-economic status	Participants were asked about their occupation. Options were 'never had paid job', 'managers and administrators', 'professional', 'technical and associate professional', 'clerical and secretarial', 'craft and skilled', 'personal and protective', 'sales', 'plant and machine operatives', or 'other'. We condensed these options into 'routine/manual', 'intermediate', or 'managerial/professional'.
Education level	Participants were asked what their highest educational qualification was. Options were 'no qualifications', 'O level', 'A level', 'degree', or 'missing'.
Marriage status	Participants were asked 'what is your marital status?' with options of 'married or living as married', 'divorced', 'widowed', 'single', or 'separated'. We combined 'divorced' and 'separated' together, and 'widowed' and 'single' together.
Weight change	Participants were asked about weight at age 20 and current weight from the baseline survey. We categorised weight change as a categorical variable: 'weight loss' (< -2.5 kg), 'weight gain' (moderate: 2.5 to 10 kg; large: > 10 kg), and 'stable weight' (± 2.5 kg).
Lifestyle and other variables	
Physical activity	Participants reported the type (e.g., walking, cycling, housework, gardening) and duration of physical activity per week. Physical activity levels were calculated using METs and categorized as low activity (< 600 MET-min/week), moderate activity (600-1500 MET-min/week), and high activity (> 1500 MET-min/week).

Smoking status	Participants were asked to describe their smoking habit as 'smoke daily', 'smoke occasionally', 'ex-smoker', or 'never'. We combined daily and occasional smokers into 'smokers', and kept 'ex-smoker' and 'never smoked' the same.
Alcohol consumption	Participants were asked how often they drink alcohol. Options were '> 1/week', '1/week', '< 1/week', or 'never'. This was computed as grams per week.
Energy intake	Participants were asked how often they eat the specific food item and completed a 217-item FFQs. Portion sizes were assigned based on McCance and Widdowson's Composition of Foods database, which provides standard UK portion sizes. Each food item in the FFQ was linked to nutrient composition tables to calculate total energy intake. This was computed as kcal per day.
AHEI-2010	AHEI-2010 was used to assess diet quality based on adherence to dietary recommendations, consisting of 11 food components. It was calculated using food intake data from a validated 217-item FFQs. The total AHEI-2010 score was obtained by summing across 11 food components (vegetables, fruits, whole grains, sugar-sweetened beverages and fruit juice, nuts and legumes, red and processed meat, trans fat, polyunsaturated fatty acids, long-chain omega-3 fats, sodium, alcohol), and participants were categorised into tertiles based on their AHEI-2010 scores.
Menopausal status	Categorised participants as pre-menopausal or post-menopausal. Criteria for postmenopausal was: age > 55 years, both ovaries removed, currently on hormone replacement therapy, or no periods in the last 12 months.
Hormone replacement therapy	Participants were asked 'have you ever used hormone replacement therapy?' and 'are you using HRT now?' – based on these yes or no answers, we categorised hormone replacement therapy use as 'current', 'ex-user', and 'never'.
Prevalence of CVD, cancer, or diabetes	Participants were asked 'Have you ever had a XX' including cancer, stroke, and diabetes; - based on these yes or no answers, we categorised prevalence of CVD, cancer, or diabetes as 'Yes' and 'No'.

METS: Metabolic Equivalent of Task, FFQ: food frequency questionnaire, AHEI-2010: Alternate Healthy Eating Index-2010, CVD: cardiovascular disease.

Table B4. The proportional hazards assumption test for obesity indicators and rheumatoid arthritis risk (in the fully-adjusted model).

Obesity Indicator	PH Test (P-Value)
BMI	
Continuous (per 1 kg/m ²)	0.18
Categorical (normal, overweight, obese)	0.16
WC	
Continuous (per 5 cm increment)	0.21
Categorical (≤ 88 cm vs. > 88 cm)	0.28
WHR categories (≤ 0.85 vs. >0.85)	0.18
WHtR categories (< 0.5, $0.5-0.6$, ≥ 0.6)	0.20
Blouse sizes	
Continuous (per 2-sizes increment)	0.26
Categorical (8-10, 12-14, ≥ 16)	0.38
Skirt sizes	
Continuous (per 2-sizes increment)	0.24
Categorical (8-10, 12-14, ≥ 16)	0.46

The PH assumption was tested using Schoenfeld residuals. A P-value < 0.05 indicates a violation of the PH assumption. PH: proportional hazards; BMI: body mass index; WC: waist circumference; WHR: waist-to-hip ratio; WHtR: waist-to-height ratio.

Table B5. Baseline characteristics of 27,968 UKWCS participants stratified by body mass index, waist circumference, waist-to-hip ratio, and waist-to-height ratio.

Characteristics	BMI (kg/m ²)			WC (cm)		WHR			WHtR		
	18.5 -< 25.0	25.0-< 30.0	≥ 30.0	≤ 88	> 88	≤ 0.80	0.80-0.85	> 0.85	< 0.5	0.5 -< 0.6	≥ 0.6
Participants	17558	7035	2679	25307	2406	22848	3186	1934	22212	4764	737
Cases(%)	141()	74()	36()	218()	37()	189()	38()	28()	182()	57()	16()
Age, years (SD)	51.0(9.1)	53.7(9.1)	53.7(9.1)	51.5(9.0)	55.8(9.2)	51.4(9.0)	53.6(9.7)	55.0(10.0)	51.0(8.9)	55.4(9.5)	56.2(9.3)
Degree-level education(%)	4957(28.2)	1423(20.2)	486(18.1)	6639(26.2)	418(17.4)	5957(26.1)	738(23.2)	362(18.7)	6045(27.2)	906(19.0)	106(14.4)
Socio-economic status (%)											
Professional or managerial	11449(65.2)	4321(61.4)	1618(60.4)	16291(64.4)	1462(60.8)	14550(63.7)	2022(63.5)	1181(61.1)	14409(64.9)	2911(61.1)	433(58.8)
Intermediate	4797(27.3)	2055(29.2)	739(27.6)	7016(27.7)	719(29.9)	6335(27.7)	849(26.6)	551(28.5)	6114(27.5)	1407(29.5)	214(29.0)
Routine or manual	1453(12.7)	659(9.4)	322(12.0)	2218(8.8)	262(10.9)	1963(8.6)	315(9.9)	202(10.4)	1871(8.4)	503(10.6)	106(14.4)
Married (%)	13489(76.8)	5357(76.1)	1989(74.2)	19454(76.9)	1753(72.9)	17480(76.5)	2349(73.7)	1378(71.3)	17127(77.1)	3555(74.6)	525(71.2)
Physical activity (METs), h/week (SD)	16.8(11.4)	16.1(11.0)	15.8(12.9)	16.6(11.3)	16.1(13.1)	16.6(11.3)	16.6(12.0)	16.5(12.7)	16.7(11.3)	16.2(12.1)	15.8(12.6)
Smoking status (%)											
Current	2022(11.5)	678(9.6)	275(10.3)	2849(11.3)	232(9.6)	2501(10.9)	367(11.5)	213(11.0)	2550(11.5)	459(9.6)	72(9.8)
Former	5223(29.7)	2371(33.7)	1007(37.6)	7788(30.8)	934(38.8)	6886(30.1)	1105(34.7)	731(37.8)	6673(30.0)	1755(36.9)	294(39.9)
Never	10454(59.5)	3986(56.7)	1397(52.1)	14888(58.8)	1277(53.1)	13461(58.9)	1714(53.8)	990(51.2)	13171(59.3)	2607(54.7)	387(52.5)

Alcohol consumption, g/week (SD)	9.4(10.4)	8.7(10.5)	6.8(10.4)	9.1(10.4)	7.5(11.2)	9.0(10.2)	8.8(11.2)	8.7(12.1)	9.2(10.3)	8.1(10.7)	6.3(12.2)
Total energy intake, kcal/day (SD)	2348(696.6)	2308(689.2)	2314(746.9)	2332(698.7)	2350(730.0)	2331(699.7)	2355(697.8)	2334(728.3)	2333(698.3)	2336(703.3)	2348(781.5)
AHEI-2010 score (SD)	66.4(11.9)	66.4(10.8)	63.3(10.8)	65.8(11.0)	63.4(10.8)	65.9(10.9)	64.9(10.9)	63.3(1.2)	66.1(11.0)	64.0(10.9)	62.9(10.8)
Menopausal status (%)											
Pre-menopausal	9346(53.2)	2862(40.7)	1055(39.4)	12782(50.5)	789(32.8)	11511(50.4)	1358(42.6)	702(36.3)	11690(52.6)	1652(34.7)	229(31.1)
Postmenopausal	8353(47.6)	4173(59.3)	1624(60.6)	12743(50.4)	1654(68.7)	11337(49.6)	1828(57.4)	1232(63.7)	10704(48.2)	3169(66.5)	524(71.1)
Hormone replacement therapy (%)	4516(25.7)	2258(32.1)	817(30.5)	6996(27.6)	702(29.2)	6285(27.5)	873(27.4)	540(27.9)	5995(27.0)	1496(31.4)	207(28.1)
Prevalence of CVD, cancer, or diabetes (%)	1027(5.8)	527(7.5)	309(11.5)	1619(6.4)	289(12.0)	1421(6.2)	295(9.3)	192(9.9)	1358(6.1)	435(9.1)	115(15.6)

SD standard deviation, BMI body mass index, WC waist circumference, WHR waist-to-hip ratio, WHtR waist-to-height ratio, MET Metabolic Equivalent of Task, AHEI-2010 Alternative Healthy Eating Index-2010, CVD cardiovascular diseases.

Table B6. Baseline characteristics of 27,968 UKWCS participants stratified by clothing size.

Characteristics	Blouse size						Skirt size					
	≤ 10	12	14	16	18	≥ 20	≤ 10	12	14	16	18	≥ 20
Participants	3719	8222	7670	4667	2063	1467	2809	6774	7901	5476	2750	2018
Cases(%)	29(0.8)	66(0.8)	57(0.7)	53(1.1)	26(1.3)	24(1.6)	15(0.5)	51(0.8)	68(0.9)	63(1.2)	24(0.9)	24(1.2)
Age, years (SD)	48.2(7.9)	49.8(8.6)	52.4(9.1)	54.6(9.3)	56.1(9.2)	55.9(9.2)	47.1(7.5)	48.9(8.0)	51.7(9.0)	54.4(9.2)	56.3(9.3)	56.9(9.3)
Degree-level education(%)	1173(31.5)	2375(28.9)	1868(24.4)	1017(21.8)	377(18.3)	247(16.8)	922(32.8)	2030(30.0)	1989(25.2)	1231(22.5)	533(19.4)	352(17.4)

Socio-economic status (%)														
Professional or managerial	2475(66.6)	5348(65.0)	4828(62.9)	2931(62.8)	1286(62.3)	885(60.3)	1914(68.1)	4441(65.6)	5027(63.6)	3412(62.3)	1742(63.3)	1217(60.3)		
Intermediate	981(26.4)	2220(27.0)	2196(28.6)	1348(28.9)	569(27.6)	421(28.7)	699(24.9)	1850(27.3)	2228(28.2)	1607(29.3)	768(27.9)	583(28.9)		
Routine or manual	263(7.1)	654(79.5)	703(91.7)	441(9.4)	234(11.3)	185(12.6)	196(7.0)	534(7.9)	714(9.0)	520(9.5)	264(9.6)	252(12.5)		
Married (%)	2846(76.5)	6339(77.1)	5934(77.4)	3511(75.2)	1538(74.6)	1039(70.8)	2112(75.2)	5261(77.7)	6166(78.0)	4204(76.8)	2041(74.2)	1423(70.5)		
Physical activity (METs), h/week (SD)	17.1(11.5)	17.0(11.4)	16.4(11.2)	16.2(11.0)	16.1(12.6)	15.8(12.6)	17.7(12.0)	17.1(11.6)	16.7(11.2)	16.0(10.9)	15.9(12.0)	15.4(12.0)		
Smoking status (%)														
Current	483(13.0)	928(11.3)	810(10.6)	497(10.6)	201(9.7)	162(11.0)	424(15.1)	801(11.8)	859(10.9)	558(10.2)	231(8.4)	208(10.3)		
Former	942(25.3)	2419(29.4)	2448(31.9)	1588(34.0)	767(37.2)	558(38.0)	714(25.4)	2013(29.7)	2471(31.3)	1813(33.1)	952(34.6)	759(37.6)		
Never	2294(61.7)	4875(59.3)	4469(58.3)	2635(56.5)	1121(54.3)	771(52.6)	1671(59.5)	4011(59.2)	4639(58.7)	3168(57.9)	1591(57.9)	1085(53.8)		
Alcohol consumption, g/week (SD)	9.2(9.9)	9.5(10.3)	9.3(10.7)	8.7(10.7)	7.9(10.7)	5.9(9.3)	9.6(10.3)	9.7(10.5)	9.3(10.5)	8.8(10.7)	7.8(10.2)	6.2(9.1)		
Total energy intake, kcal/day (SD)	2288(700.0)	2335(693.0)	2345(689.1)	2349(704.6)	2321(739.9)	2356(744.7)	2287(714.8)	2328(697.8)	2339(687.5)	2350(691.1)	2334(723.2)	2353(744.0)		
AHEI-2010 score (SD)	66.8(11.0)	66.4(11.0)	65.7(10.9)	64.8(10.9)	63.8(10.9)	62.8(10.9)	67.0(11.2)	66.8(11.0)	65.7(10.9)	64.9(10.8)	64.4(10.8)	63.2(10.8)		
Menopausal status (%)														
Premenopausal	2412(64.9)	4752(57.8)	3619(47.2)	1703(36.5)	638(30.9)	447(30.5)	1925(68.5)	4190(61.9)	3979(50.4)	2076(37.9)	845(30.7)	556(27.6)		
Postmenopausal	1307(35.1)	3470(42.2)	4108(53.6)	3017(64.6)	1451(70.3)	1044(71.2)	884(31.5)	2635(38.9)	3990(50.5)	3463(63.2)	1920(69.8)	1496(74.1)		
Hormone replacement therapy (%)	791(21.3)	2087(25.4)	2221(29.0)	1513(32.4)	671(32.5)	415(28.3)	564(20.1)	1627(24.0)	2308(29.2)	1747(31.9)	856(31.1)	596(29.5)		
Prevalence of CVD, cancer, or diabetes (%)	184(4.9)	458(5.6)	499(6.5)	340(9.7)	205(9.9)	222(15.1)	145(5.2)	331(4.9)	498(6.3)	396(7.2)	251(9.1)	287(14.2)		

SD standard deviation, MET Metabolic Equivalent of Task, AHEI-2010 Alternative Healthy Eating Index-2010, CVD cardiovascular diseases.

Table B7. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by age.

Obesity Indicator	Age (years)		<i>P</i> -interaction
	< 55	≥ 55	
	(142/18252)	(113/9716)	
BMI (per 2.5 kg/m² increment)	1.16(1.08, 1.24)	1.04(0.92, 1.17)	0.1
WC (per 5 cm increment)	1.14(1.05, 1.23)	1.02(0.92, 1.12)	0.1
WHR (per 0.1 increment)	1.32(1.05, 1.64)	1.12(0.86, 1.47)	0.4
WHtR (per 0.1 increment)	1.54(1.26, 1.88)	1.11(0.81, 1.51)	0.1
Blouse size (per one-size increase)	1.21(1.09, 1.34)	1.09(0.95, 1.25)	0.2
Skirt size (per one-size increase)	1.21(1.10, 1.33)	1.10(0.98, 1.24)	0.2

Models were adjusted for socio-economic status, marital status, menopausal status, hormone replacement therapy, and prevalence of CVD, cancer, or diabetes at recruitment, physical activity, smoking status, alcohol consumption, total energy intake, and Alternate Healthy Eating Index -2010; HR (95% CI), hazard ratio (95% confidence interval).

Table B8. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by weight change.

Adiposity Indicator	Weight change (kg)				<i>P</i> -interaction
	Weight loss	Stable weight	Weight gain, moderate	Weight gain, large	
	(< -2.5)	(±2.5)	(2.5-< 10)	(≥ 10)	
	(31/4147)	(26/2909)	(98/10657)	(108/10255)	
BMI (per 2.5 kg/m² increment)	1.06(0.73, 1.56)	1.15(0.83, 1.59)	1.02(0.88, 1.19)	1.14(1.03, 1.26)	0.8

WC (per 5 cm increment)	0.70(0.50, 0.97)	1.11(0.89, 1.38)	0.95(0.81, 1.11)	1.15(1.06, 1.24)	0.03
WHR (per 0.1 increment)	0.60(0.28, 1.28)	0.98(0.57, 1.70)	0.83(0.56, 1.24)	1.46(1.24, 1.74)	0.01
WHtR (per 0.1 increment)	0.49(0.21, 1.19)	1.39(0.67, 2.87)	0.92(0.59, 1.44)	1.57(1.27, 1.94)	0.05
Blouse size (per one-size increase)	1.02(0.58, 1.81)	1.43(1.06, 1.93)	1.21(0.97, 1.52)	1.11(0.98, 1.26)	0.4
Skirt size (per one-size increase)	1.12(0.79, 1.58)	1.31(1.07, 1.61)	1.23(1.04, 1.44)	1.11(0.99, 1.26)	0.4

Models were adjusted for age, socio-economic status, marital status, menopausal status, hormone replacement therapy, and prevalence of CVD, cancer, or diabetes at recruitment, physical activity, smoking status, alcohol consumption, total energy intake, and Alternate Healthy Eating Index -2010; HR (95% CI), hazard ratio (95% confidence interval).

Table B9. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women’s Cohort Study participants, stratified by AHEI-2010.

Adiposity Indicator	AHEI-2010 tertiles			<i>P</i> -interaction
	Low	Moderate	High	
	(83/9267)	(86/9356)	(86/9345)	
BMI (per 2.5 kg/m² increment)	1.05(0.93, 1.19)	1.11(1.00, 1.22)	1.18(1.05, 1.32)	0.3
WC (per 5 cm increment)	1.07(0.96, 1.19)	1.10(0.99, 1.22)	1.09(0.97, 1.23)	0.7
WHR (per 0.1 increment)	1.17(0.91, 1.51)	1.29(0.87, 1.90)	1.22(0.93, 1.59)	0.8
WHtR (per 0.1 increment)	1.24(0.97, 1.68)	1.36(0.99, 1.87)	1.48(1.06, 2.05)	0.5
Blouse size (per one-size increase)	1.14(0.98, 1.33)	1.17(1.03, 1.34)	1.15(0.98, 1.36)	0.8
Skirt size (per one-size increase)	1.10(0.95, 1.27)	1.17(1.05, 1.31)	1.21(1.06, 1.39)	0.3

Models were adjusted for age, socio-economic status, marital status, menopausal status, hormone replacement therapy, and prevalence of CVD, cancer, or diabetes at recruitment, physical activity, smoking status, alcohol consumption, total energy intake; HR (95% CI), hazard ratio (95% confidence interval).

Table B10. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by physical activity.

Adiposity Indicator	METs (min/week)			<i>P</i> -interaction
	Low(< 600) (71/8665)	Moderate(600-1500) (148/14352)	High(> 1500) (36/4951)	
BMI (per 2.5 kg/m² increment)	1.08(0.97, 1.21)	1.11(1.01, 1.22)	1.19(1.03, 1.37)	0.5
WC (per 5 cm increment)	1.07(0.94, 1.20)	1.08(1.00, 1.19)	1.13(0.96, 1.33)	0.7
WHR (per 0.1 increment)	1.08(0.74, 1.58)	1.26(1.04, 1.53)	1.30(0.86, 1.96)	0.8
WHtR (per 0.1 increment)	1.22(0.85, 1.74)	1.37(1.08, 1.74)	1.56(1.00, 2.43)	0.6
Blouse size (per one-size increase)	1.14(0.98, 1.33)	1.17(1.04, 1.31)	1.13(0.88, 1.45)	0.9
Skirt size (per one-size increase)	1.16(1.02, 1.32)	1.15(1.04, 1.28)	1.18(0.96, 1.44)	0.8

Models were adjusted for age, socio-economic status, marital status, menopausal status, hormone replacement therapy, and prevalence of CVD, cancer, or diabetes at recruitment, smoking status, alcohol consumption, total energy intake, and Alternate Healthy Eating Index -2010; HR (95% CI), hazard ratio (95% confidence interval).

Table B11. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by menopausal status.

Adiposity Indicator	Menopausal status		<i>P</i> -interaction
	Pre-menopausal (93/13571)	Post-menopausal (162/14397)	
BMI (per 2.5 kg/m² increment)	1.22(1.13, 1.33)	1.02(0.94, 1.11)	0.003
WC (per 5 cm increment)	1.17(1.06, 1.29)	1.03(0.95, 1.12)	0.04

WHR (per 0.1 increment)	1.37(0.98, 1.91)	1.16(0.94, 1.43)	0.4
WHtR (per 0.1 increment)	1.70(1.35, 2.14)	1.13(0.88, 1.45)	0.02
Blouse size (per one-size increase)	1.19(1.02, 1.38)	1.12(1.01, 1.25)	0.5
Skirt size (per one-size increase)	1.24(1.10, 1.40)	1.10(1.01, 1.21)	0.1

Models were adjusted for age, socio-economic status, marital status, hormone replacement therapy, and prevalence of CVD, cancer, or diabetes at recruitment, physical activity, smoking status, alcohol consumption, total energy intake, and Alternate Healthy Eating Index -2010; HR (95% CI), hazard ratio (95% confidence interval).

Table B12. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women’s Cohort Study participants, stratified by smoking status.

Adiposity Indicator	Smoking status		<i>P</i> -interaction
	Smoker	Non-smoker	
	(130/11803)	(125/16165)	
BMI (per 2.5 kg/m² increment)	1.10(1.02, 1.20)	1.11(1.00, 1.23)	0.9
WC (per 5 cm increment)	1.07(0.98, 1.16)	1.09(0.99, 1.21)	0.7
WHR (per 0.1 increment)	1.17(0.95, 1.44)	1.24(0.93, 1.67)	0.7
WHtR (per 0.1 increment)	1.29(1.03, 1.63)	1.37(1.02, 1.83)	0.8
Blouse size (per one-size increase)	1.12(1.00,1.25)	1.17(1.03, 1.35)	0.6
Skirt size (per one-size increase)	1.15(1.05, 1.27)	1.15(1.02, 1.29)	0.9

Models were adjusted for age, socio-economic status, marital status, hormone replacement therapy, and prevalence of CVD, cancer, or diabetes at recruitment, physical activity, smoking status, alcohol consumption, total energy intake, and Alternate Healthy Eating Index - 2010; HR (95% CI), hazard ratio (95% confidence interval).

Table B13. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by age.

Adiposity Indicator	Category	Age(years)		P-interaction
		< 55	≥ 55	
		(142/18252)	(113/9716)	
BMI (kg/m²)	Normal (18.5-< 25)	1(ref)	1(ref)	0.3
	Overweight (25.0-< 30.0)	1.44(0.98, 2.12)	1.09(0.72, 1.65)	
	Obese (≥ 30.0)	2.11(1.31, 3.42)	1.19(0.68, 2.10)	
WC (cm)	Normal (≤ 88)	1(ref)	1(ref)	0.3
	Abdominal obese (> 88)	2.06(1.24, 3.41)	1.41(0.87, 2.28)	
WHR	≤ 0.8	1(ref)	1(ref)	0.1
	0.8-0.85	1.89(1.20, 2.97)	0.96(0.55, 1.66)	
	> 0.85	2.12(1.22, 3.72)	1.30(0.74, 2.28)	
WHtR	Normal (< 0.5)	1(ref)	1(ref)	0.1
	Overweight (0.5-< 0.6)	1.84(1.22, 2.78)	1.01(0.66, 1.56)	
	Obese (≥ 0.6)	3.52(1.71, 7.22)	1.76(0.85, 3.64)	
Blouse size	≤ 10	1(ref)	1(ref)	0.2
	12	1.12(0.64, 1.95)	0.76(0.37, 1.54)	
	14	1.20(0.68, 2.13)	0.49(0.24, 1.01)	
	16	1.93(1.07, 3.50)	0.69(0.34, 1.41)	
	18	1.60(0.72, 3.55)	0.95(0.44, 2.03)	
	≥ 20	2.81(1.32, 5.96)	1.00(0.45, 2.23)	

Skirt size	≤ 10	1(ref)	1(ref)	0.3
	12	1.22(0.63, 2.36)	1.78(0.52, 6.04)	
	14	1.43(0.75, 2.74)	1.59(0.49, 5.22)	
	16	2.43(1.26, 4.67)	1.49(0.45, 4.90)	
	18	1.61(0.69, 3.72)	1.29(0.37, 4.51)	
	≥ 20	3.13(1.42, 6.88)	2.45(0.73, 8.21)	

Models were adjusted for socio-economic status, marital status, hormone replacement therapy, menopausal status, and prevalence of CVD, cancer, or diabetes at recruitment, physical activity, smoking status, alcohol consumption, total energy intake, and Alternate Healthy Eating Index -2010; HR (95% CI), hazard ratio (95% confidence interval).

Table B14. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants stratified by weight change.

Adiposity Indicator	Category	Weight change (kg)				P-interaction
		Weight loss (< -2.5) (31/4147)	Stable weight (± 2.5) (26/2909)	Weight gain, moderate (2.5-< 10) (98/10657)	Weight gain, large (≥ 10) (108/10255)	
BMI (kg/m²)	Normal (18.5-< 25)	1(ref)	1(ref)	1(ref)	1(ref)	0.3
	Overweight (25.0-< 30.0)	1.87(0.43, 8.18)	0.89(0.21, 3.80)	1.47(0.92, 2.35)	0.92(0.58, 1.46)	
	Obese (≥ 30.0)	N/A	3.60(0.48, 26.9)	N/A	1.43(0.88, 2.33)	
WC (cm)	Normal (≤ 88)	1(ref)	1(ref)	1(ref)	1(ref)	<0.001
	Abdominal obese (> 88)	N/A	N/A	0.72(0.18, 2.94)	1.87(1.24, 2.82)	
WHR	≤ 0.8	1(ref)	1(ref)	1(ref)	1(ref)	<0.001
	0.8-0.85	N/A	1.10(0.25, 4.83)	0.85(0.41, 1.78)	2.12(1.34, 3.35)	
	> 0.85	N/A	N/A	0.93(0.34, 2.55)	2.32(1.44, 3.78)	
WHtR	Normal (< 0.5)	1(ref)	1(ref)	1(ref)	1(ref)	0.4
	Overweight (0.5-< 0.6)	N/A	0.96(0.13, 7.38)	0.88(0.44, 1.78)	1.77(1.17, 2.69)	

	Obese (≥ 0.6)	N/A	N/A	N/A	3.17(1.78, 5.67)	
Blouse size	≤ 10	1(ref)	1(ref)	1(ref)	1(ref)	<0.001
	12	0.37(0.15, 0.90)	0.51(0.18, 1.45)	1.71(0.72, 4.10)	N/A	
	14	0.35(0.11, 1.17)	0.70(0.21, 2.27)	1.73(0.71, 4.21)	N/A	
	16	0.87(0.23, 3.36)	2.91(0.88, 9.59)	1.68(0.63, 4.47)	N/A	
	18	1.45(0.16, 12.80)	3.25(0.43, 24.66)	2.62(0.79, 8.67)	N/A	
	≥ 20	3.96(0.46, 33.50)	4.41(0.59, 33.23)	3.42(0.70, 16.84)	N/A	
Skirt size	≤ 10	1(ref)	1(ref)	1(ref)	1(ref)	<0.001
	12	0.86(0.32, 2.24)	1.80(0.56, 5.77)	1.39(0.48, 4.00)	N/A	
	14	1.16(0.40, 3.35)	1.66(0.46, 5.99)	1.53(0.54, 4.34)	N/A	
	16	1.12(0.26, 4.78)	3.99(0.92, 17.28)	2.10(0.72, 6.09)	N/A	
	18	1.57(0.17, 14.68)	3.23(0.37, 28.55)	1.18(0.30, 4.66)	N/A	
	≥ 20	2.77(0.30, 25.80)	4.82(0.55, 42.02)	4.23(1.15, 15.57)	N/A	

Models were adjusted for age, socio-economic status, marital status, hormone replacement therapy, menopausal status, and prevalence of CVD, cancer, or diabetes at recruitment, physical activity, smoking status, alcohol consumption, total energy intake, and Alternate Healthy Eating Index -2010; HR (95% CI), hazard ratio (95% confidence interval). N/A indicates an insufficient sample size for a reliable estimate.

Table B15. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by AHEI-2010.

Adiposity Indicator	Category	AHEI-2010 tertiles			<i>P</i> -interaction
		Low (83/9267)	Moderate (86/9356)	High (86/9345)	
BMI (kg/m²)	Normal (18.5-< 25)	1(ref)	1(ref)	1(ref)	0.7
	Overweight (25.0-< 30.0)	1.14(0.69, 1.91)	1.09(0.67, 1.79)	1.63(1.01, 2.63)	
	Obese (≥ 30.0)	1.46(0.78, 2.73)	1.94(1.08, 3.50)	1.45(0.68, 3.07)	
WC (cm)	Normal (≤ 88)	1(ref)	1(ref)	1(ref)	0.5
	Abdominal obese (> 88)	1.42(0.76, 2.64)	1.64(0.90, 2.98)	2.09(1.11, 3.91)	
WHR	≤ 0.8	1(ref)	1(ref)	1(ref)	0.4
	0.8-0.85	1.25(0.67, 2.34)	1.74(1.00, 3.04)	1.16(0.59, 2.27)	

	> 0.85	1.32(0.65, 2.67)	1.45(0.69, 3.05)	2.45(1.28, 4.71)	
WHtR	Normal (< 0.5)	1(ref)	1(ref)	1(ref)	0.9
	Overweight (0.5-< 0.6)	1.27(0.74, 2.19)	1.47(0.87, 2.46)	1.38(0.80, 2.39)	
	Obese ($\geq$ 0.6)	2.71(1.21, 6.08)	1.78(0.64, 4.92)	3.04(1.22, 7.64)	
Blouse size	$\leq$ 10	1(ref)	1(ref)	1(ref)	0.8
	12	1.21(0.57, 2.59)	1.19(0.50, 2.82)	0.75(0.38, 1.51)	
	14	0.78(0.35, 1.79)	1.38(0.59, 3.22)	0.68(0.32, 1.45)	
	16	1.09(0.47, 2.54)	1.84(0.76, 4.48)	1.21(0.57, 2.55)	
	18	1.03(0.38, 2.79)	2.20(0.84, 5.79)	1.41(0.57, 3.48)	
	$\geq$ 20	2.13(0.86, 5.26)	2.38(0.85, 6.68)	1.29(0.44, 3.78)	
Skirt size	$\leq$ 10	1(ref)	1(ref)	1(ref)	0.4
	12	1.42(0.57, 3.54)	1.27(0.42, 3.88)	1.43(0.53, 3.87)	
	14	1.21(0.49, 3.00)	1.88(0.66, 5.38)	1.56(0.58, 4.20)	
	16	1.02(0.38, 2.70)	2.30(0.79, 6.75)	2.99(0.40, 4.62)	

18	1.10(0.37, 3.26)	2.04(0.65, 6.49)	1.36(0.40, 4.62)
≥ 20	2.15(0.80, 5.81)	3.33(1.06, 10.42)	3.21(1.04, 9.96)

Models were adjusted for age, socio-economic status, marital status, hormone replacement therapy, menopausal status, and prevalence of CVD, cancer, or diabetes at recruitment, physical activity, smoking status, alcohol consumption, total energy intake; HR (95% CI), hazard ratio (95% confidence interval).

Table B16. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women’s Cohort Study participants, stratified by physical activity.

Adiposity Indicator	Category	METs (min/week)			<i>P</i> -interaction
		Low(< 600)	Moderate(600-1500)	High(> 1500)	
		(71/8665)	(148/14352)	(36/4951)	
BMI (kg/m²)	Normal (18.5-< 25)	1 (ref)	1 (ref)	1 (ref)	0.9
	Overweight (25.0-< 30.0)	1.44 (0.85, 2.44)	1.16 (0.79, 1.70)	1.43 (0.21,12.02)	

	Obese (≥ 30.0)	1.37 (0.68, 2.79)	1.61 (0.97, 2.66)	2.59 (1.08, 6.26)	
WC (cm)	Normal (≤ 88)	1 (ref)	1 (ref)	1(ref)	0.9
	Abdominal obese (> 88)	1.49 (0.77, 2.87)	1.82 (1.13, 2.92)	1.56(0.61, 4.00)	
WHR	≤ 0.8	1 (ref)	1 (ref)	1(ref)	0.9
	0.8-0.85	1.63 (0.86, 3.08)	1.26 (0.78, 2.05)	1.51(0.61, 3.73)	
	> 0.85	2.01 (1.01, 4.01)	1.65 (0.95, 2.84)	1.14(0.34, 3.80)	
WHtR	Normal (< 0.5)	1 (ref)	1 (ref)	1(ref)	0.8
	Overweight ($0.5-< 0.6$)	1.64 (0.95, 2.81)	1.35 (0.89, 2.05)	0.95 (0.38, 2.33)	
	Obese (≥ 0.6)	1.81 (0.64, 5.13)	2.76 (1.38, 5.51)	2.94 (0.88, 9.84)	
Blouse size	≤ 10	1 (ref)	1 (ref)	1 (ref)	0.9
	12	1.56 (0.58, 4.23)	0.94 (0.54, 1.65)	0.76 (0.28, 2.10)	
	14	1.39 (0.49, 3.88)	0.87 (0.49, 1.56)	0.55 (0.18, 1.66)	
	16	1.75 (0.60, 5.15)	1.37 (0.75, 2.49)	0.83 (0.27,2.61)	

	18	2.61 (0.84, 8.08)	1.39 (0.68, 2.83)	0.54 (0.12, 2.68)	
	≥ 20	2.25 (0.65, 7.77)	1.81 (0.86, 3.82)	1.87 (0.56, 6.27)	
Skirt size	≤ 10	1 (ref)	1 (ref)	1 (ref)	0.9
	12	2.34 (0.52, 10.48)	1.16 (0.59, 2.30)	1.61 (0.33, 7.74)	
	14	2.62 (0.59, 11.68)	1.16 (0.59, 2.27)	2.45 (0.56, 10.82)	
	16	3.88 (0.86, 17.45)	1.53 (0.76, 3.09)	2.37 (0.52, 10.91)	
	18	3.59 (0.73, 17.54)	1.25 (0.56, 2.79)	2.34 (0.51, 10.87)	
	≥ 20	4.33 (0.86, 21.72)	2.30 (1.08, 4.91)	4.44 (0.91, 21.54)	

Models were adjusted for age, socio-economic status, marital status, hormone replacement therapy, menopausal status, and prevalence of CVD, cancer, or diabetes at recruitment, smoking status, alcohol consumption, total energy intake, and Alternate Healthy Eating Index -2010; HR (95% CI), hazard ratio (95% confidence interval).

Table B17. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women's Cohort Study participants, stratified by smoking status.

Adiposity Indicator	Category	Smoking status		<i>P</i> -interaction
		Smoker (71/8665)	Non-smoker (148/14352)	
BMI (kg/m²)	Normal (18.5-< 25)	1(ref)	1(ref)	0.7
	Overweight (25.0-< 30.0)	1.30(0.87, 1.94)	1.22(0.82, 1.84)	
	Obese (≥ 30.0)	1.78(1.10, 2.91)	1.37(0.77, 2.44)	
WC (cm)	Normal (≤ 88)	1(ref)	1(ref)	0.5
	Abdominal obese (> 88)	1.45(0.89, 2.38)	1.84(1.10, 3.09)	
WHR	≤ 0.8	1(ref)	1(ref)	0.9
	0.8-0.85	1.32(0.81, 2.1)	1.42(0.85, 2.38)	
	> 0.85	1.58(0.92, 2.71)	1.64(0.89, 3.01)	
WHtR	Normal (< 0.5)	1(ref)	1(ref)	0.3
	Overweight (0.5-< 0.6)	1.56(1.04, 2.34)	1.10(0.68, 1.79)	
	Obese (≥ 0.6)	1.85(0.85, 4.04)	3.08(1.53, 6.18)	
Blouse size	≤ 10	1(ref)	1(ref)	0.7
	12	0.98(0.52, 1.83)	1.02(0.55, 1.87)	

	14	0.71(0.36, 1.39)	1.06(0.57, 1.98)	
	16	1.44(0.75, 2.75)	1.10(0.55, 2.20)	
	18	1.32(0.62, 2.85)	1.46(0.67, 3.17)	
	≥ 20	1.64(0.75, 3.58)	1.97(0.87, 4.43)	
Skirt size	≤ 10	1(ref)	1(ref)	0.8
	12	0.93(0.43, 2.03)	2.05(0.85, 4.92)	
	14	1.24(0.59, 2.60)	1.92(0.81, 4.58)	
	16	1.69(0.80, 3.59)	2.31(0.94, 5.66)	
	18	1.26(0.53, 3.02)	1.70(0.62, 4.66)	
	≥ 20	2.24(0.99, 5.07)	3.40(1.29, 8.91)	

Models were adjusted for age, socio-economic status, marital status, hormone replacement therapy, menopausal status, and prevalence of CVD, cancer, or diabetes at recruitment, physical activity, alcohol consumption, total energy intake, and Alternate Healthy Eating Index -2010; HR (95% CI), hazard ratio (95% confidence interval).

Table B18. Associations between adiposity indicators and rheumatoid arthritis incidence in UK Women’s Cohort Study participants, stratified by menopausal status.

Adiposity Indicator	Category	Menopausal status		<i>P</i> -interaction
		Pre-menopausal	Post-menopausal	
		(93/13571)	(162/14397)	
BMI (kg/m²)	Normal (18.5-< 25)	1(ref)	1(ref)	0.2
	Overweight (25.0-< 30.0)	1.57(0.97, 2.54)	1.10(0.78, 1.56)	

	Obese (≥ 30.0)	2.58(1.45, 4.58)	1.22(0.76, 1.96)	
WC (cm)	Normal (≤ 88)	1(ref)	1(ref)	0.2
	Abdominal obese (> 88)	1.41(0.92, 2.16)	2.40(1.31, 4.39)	
WHR	≤ 0.8	1(ref)	1(ref)	0.1
	0.8-0.85	1.11(0.70, 1.76)	2.02(1.18, 3.49)	
	> 0.85	1.35(0.82, 2.23)	2.46(1.27, 4.77)	
WHtR	Normal (< 0.5)	1(ref)	1(ref)	0.03
	Overweight (0.5- < 0.6)	1.08(0.74, 1.57)	2.12(1.28, 3.51)	
	Obese (≥ 0.6)	1.78(0.93, 3.41)	4.59(2.01, 10.47)	
Blouse size	≤ 10	1(ref)	1(ref)	0.5
	12	1.07(0.58, 1.98)	0.88(0.47, 1.65)	
	14	0.70(0.37, 1.33)	1.16(0.61, 2.19)	
	16	1.20(0.64, 2.26)	1.31(0.63, 2.73)	
	18	1.28(0.64, 2.26)	1.51(0.58, 3.91)	
	≥ 20	1.60(0.79, 3.29)	2.17(0.84, 5.63)	
Skirt size	≤ 10	1(ref)	1(ref)	0.4
	12	1.68(0.70, 4.03)	1.07(0.49, 2.35)	

14	1.55(0.66, 3.65)	1.41(0.66, 3.00)
16	1.66(0.70, 3.95)	2.48(1.15, 5.35)
18	1.45(0.57, 3.66)	1.27(0.42, 3.80)
≥ 20	2.57(1.04, 6.30)	3.08(1.19, 7.95)

Models were adjusted for age, socio-economic status, marital status, hormone replacement therapy, and prevalence of CVD, cancer, or diabetes at recruitment, smoking status, physical activity, alcohol consumption, total energy intake, and Alternate Healthy Eating Index -2010; HR (95% CI), hazard ratio (95% confidence interval).

Table B19. Risk of rheumatoid arthritis by adiposity indicators with varying restrictions in the UKWCS.

Adiposity indicator	Category	HR (95% CI)	<i>P</i>
Fully-adjusted models (255 cases/ 27,968 participants)^a			
BMI (per 2.5 kg/m ² increment)		1.09 (1.02, 1.16)	< 0.01
WC (per 5 cm increment)		1.07 (1.00, 1.13)	0.04
WHR (per 0.1 increment)		1.20 (1.01, 1.43)	0.04
WHtR (per 0.1 increment)		1.27 (1.05, 1.53)	0.01
Blouse size (per one-size increase)		1.13 (1.04, 1.22)	< 0.01
Skirt size (per one-size increase)		1.13 (1.05, 1.22)	< 0.01
BMI (kg/m ²)	Normal (18.5-< 25)	1 (ref)	
	Overweight (25.0-< 30.0)	1.27 (0.95, 1.70)	0.11
	Obese (≥ 30.0)	1.49 (1.02, 2.18)	0.04

WC (cm)	Normal (≤ 88)	1 (ref)	
	Abdominal obese (> 88)	1.59 (1.10, 2.28)	0.01
WHR	≤ 0.8	1 (ref)	
	0.8-0.85	1.38 (0.97, 1.97)	0.08
	> 0.85	1.56 (1.03, 2.36)	0.04
WHtR	Normal (< 0.5)	1 (ref)	
	Overweight ($0.5-< 0.6$)	1.26 (0.92, 1.73)	0.15
	Obese (≥ 0.6)	2.26 (1.34, 3.81)	< 0.01
Blouse size	≤ 10	1 (ref)	
	12	0.97 (0.62, 1.52)	0.89
	14	0.87 (0.55, 1.37)	0.54
	16	1.24 (0.77, 1.99)	0.37
	18	1.28 (0.73, 2.23)	0.39
	≥ 20	1.64 (0.93, 2.89)	0.09
Skirt size	≤ 10	1 (ref)	
	12	1.35 (0.76, 2.40)	0.31
	14	1.42 (0.81, 2.50)	0.23
	16	1.73 (0.97, 3.08)	0.06
	18	1.33 (0.68, 2.58)	0.41
	≥ 20	2.36 (1.25, 4.45)	< 0.01
Excluding participants with survival time < 5 years (255 cases/ 27,968 participants)			
BMI (per 2.5 kg/m ² increment)		1.08 (1.01, 1.16)	0.02
WC (per 5 cm increment)		1.08 (1.00, 1.16)	0.04

WHR (per 0.1 increment)		1.21 (1.00, 1.48)	0.05
WHtR (per 0.1 increment)		1.30 (1.07, 1.59)	< 0.01
Blouse size (per one-size increase)		1.14 (1.04, 1.26)	< 0.01
Skirt size (per one-size increase)		1.14 (1.05, 1.24)	< 0.01
BMI (kg/m ²)	Normal (18.5-< 25)	1 (ref)	
	Overweight (25.0-< 30.0)	1.30 (0.95, 1.78)	0.09
	Obese (≥ 30.0)	1.49 (0.98, 2.25)	0.06
WC (cm)	Normal (≤ 88)	1 (ref)	
	Abdominal obese (> 88)	1.55 (1.03, 2.32)	0.04
WHR	≤ 0.8	1 (ref)	
	0.8-0.85	1.38 (0.93, 2.03)	0.11
	> 0.85	1.51 (0.94, 2.40)	0.09
WHtR	Normal (< 0.5)	1 (ref)	
	Overweight (0.5-< 0.6)	1.31 (0.92, 1.86)	0.14
	Obese (≥ 0.6)	2.29 (1.27, 4.13)	< 0.01
Blouse size	≤ 10	1 (ref)	
	12	1.13 (0.69, 1.86)	0.63
	14	0.94 (0.56, 1.58)	0.82
	16	1.41 (0.83, 2.40)	0.20
	18	1.54 (0.83, 2.86)	0.18
	≥ 20	1.90 (1.01, 3.56)	0.05
Skirt size	≤ 10	1 (ref)	

	12	1.43 (0.75, 2.73)	0.28
	14	1.71 (0.91, 3.20)	0.09
	16	1.75 (0.91, 3.38)	0.09
	18	1.73 (0.84, 3.57)	0.14
	≥ 20	2.68 (1.33, 5.43)	< 0.01
Excluding participants who are current smoker (255 cases/ 27,968 participants)			
BMI (per 2.5 kg/m ² increment)		1.13 (1.06, 1.21)	< 0.01
WC (per 5 cm increment)		1.11 (1.04, 1.19)	< 0.01
WHR (per 0.1 increment)		1.30 (1.11, 1.53)	< 0.01
WHtR (per 0.1 increment)		1.42 (1.17, 1.74)	< 0.01
Blouse size (per one-size increase)		1.17 (1.07, 1.29)	< 0.01
Skirt size (per one-size increase)		1.17 (1.08, 1.28)	< 0.001
BMI (kg/m ²)	Normal (18.5-< 25)	1 (ref)	
	Overweight (25.0-< 30.0)	1.45 (1.05, 1.98)	0.02
	Obese (≥ 30.0)	1.77 (1.19, 2.66)	< 0.01
WC (cm)	Normal (≤ 88)	1 (ref)	
	Abdominal obese (> 88)	1.84 (1.26, 2.70)	< 0.01
WHR	≤ 0.8	1 (ref)	
	0.8-0.85	1.48 (1.00, 2.18)	0.05
	> 0.85	1.80 (1.15, 2.80)	< 0.01
WHtR	Normal (< 0.5)	1(ref)	
	Overweight (0.5-< 0.6)	1.44 (1.01, 2.03)	0.04
	Obese (≥ 0.6)	2.50 (1.42, 4.40)	< 0.01

Blouse size	≤ 10	1 (ref)	
	12	0.90 (0.55, 1.48)	0.67
	14	0.85 (0.51, 1.42)	0.54
	16	1.43 (0.85, 2.39)	0.17
	18	1.32 (0.71, 2.42)	0.38
	≥ 20	1.96 (1.07, 3.59)	0.03
Skirt size	≤ 10	1 (ref)	
	12	1.07 (0.57, 2.01)	0.84
	14	1.30 (0.71, 2.39)	0.71
	16	1.76 (0.94, 3.28)	0.08
	18	1.21 (0.59, 2.49)	0.60
	≥ 20	2.60 (1.33, 5.10)	< 0.01
Excluding participants who developed chronic diseases at recruitment (255 cases/ 27,968 participants)			
BMI (per 2.5 kg/m ² increment)		1.09 (1.01, 1.17)	0.02
WC (per 5 cm increment)		1.07 (1.00, 1.15)	0.05
WHR (per 0.1 increment)		1.20 (0.99, 1.45)	0.07
WHtR (per 0.1 increment)		1.31 (1.07, 1.62)	0.01
Blouse size (per one-size increase)		1.15 (1.05, 1.27)	< 0.01
Skirt size (per one-size increase)		1.15 (1.05, 1.25)	< 0.01
BMI (kg/m ²)	Normal (18.5-< 25)	1 (ref)	
	Overweight (25.0-< 30.0)	1.28 (0.94, 1.75)	0.11
	Obese (≥ 30.0)	1.57 (1.04, 2.36)	0.03

WC (cm)	Normal (≤ 88)	1 (ref)	
	Abdominal obese (> 88)	1.60 (1.07, 2.39)	0.02
WHR	≤ 0.8	1 (ref)	
	0.8-0.85	1.30 (0.88, 1.94)	0.19
	> 0.85	1.63 (1.04, 2.55)	0.03
WHtR	Normal (< 0.5)	1 (ref)	
	Overweight ($0.5-< 0.6$)	1.19 (0.83, 1.69)	0.35
	Obese (≥ 0.6)	2.69 (1.56, 4.65)	< 0.001
Blouse size	≤ 10	1 (ref)	
	12	0.91 (0.56, 1.48)	0.71
	14	0.97 (0.59, 1.58)	0.90
	16	1.20 (0.72, 2.01)	0.49
	18	1.40 (0.77, 2.55)	0.27
	≥ 20	1.67 (0.90, 3.12)	0.11
Skirt size	≤ 10	1 (ref)	
	12	1.34 (0.72, 2.50)	0.36
	14	1.50 (0.81, 2.76)	0.19
	16	1.75 (0.93, 3.30)	0.08
	18	1.54 (0.75, 3.13)	0.24
	≥ 20	2.50 (1.25, 4.99)	< 0.01
Excluding participants with weight change of more than 10 kg (255 cases/ 27, 968 participants)			
BMI (per 2.5 kg/m ² increment)		0.94 (0.83, 1.06)	0.30

WC (per 5 cm increment)		1.03 (0.91, 1.17)	0.60
WHR (per 0.1 increment)		0.81 (0.59, 1.12)	0.20
WHtR (per 0.1 increment)		0.88 (0.61, 1.27)	0.50
Blouse size (per one-size increase)		1.23 (1.03, 1.46)	0.02
Skirt size (per one-size increase)		1.19 (1.04, 1.36)	0.01
BMI (kg/m ²)	Normal (18.5-< 25)	1 (ref)	
	Overweight (25.0-< 30.0)	1.58 (1.05, 2.38)	0.03
	Obese ($\geq$ 30.0)	0.55 (0.08, 3.95)	0.55
WC (cm)	Normal ($\leq$ 88)	1 (ref)	
	Abdominal obese (> 88)	0.56 (0.14, 2.25)	0.41
WHR	$\leq$ 0.8	1 (ref)	
	0.8-0.85	0.83 (0.44, 1.55)	0.56
	> 0.85	0.72 (0.27, 1.95)	0.27
WHtR	Normal (< 0.5)	1 (ref)	
	Overweight (0.5-< 0.6)	0.82 (0.43, 1.59)	0.56
	Obese ($\geq$ 0.6)	N/A	N/A
Blouse size	$\leq$ 10	1 (ref)	
	12	0.79 (0.49, 1.28)	0.33
	14	0.87 (0.52, 1.45)	0.59
	16	1.25 (0.69, 2.26)	0.47
	18	1.55 (0.63, 3.79)	0.34
	$\geq$ 20	2.97 (1.08, 8.12)	0.03

Skirt size	≤ 10	1 (ref)	
	12	1.19 (0.65, 2.16)	0.57
	14	1.29 (0.71, 2.36)	0.41
	16	1.70 (0.88, 3.30)	0.12
	18	1.30 (0.51, 3.35)	0.59
	≥ 20	3.18 (1.22, 8.31)	0.02
Further adjusted for dietary vitamin D (255 cases/ 27,968 participants)			
BMI (per 2.5 kg/m ² increment)		1.09 (1.02, 1.17)	0.01
WC (per 5 cm increment)		1.07 (1.00, 1.14)	0.05
WHR (per 0.1 increment)		1.20 (1.00, 1.44)	0.05
WHtR (per 0.1 increment)		1.27 (1.05, 1.54)	0.02
Blouse size (per one-size increase)		1.13 (1.03, 1.24)	< 0.01
Skirt size (per one-size increase)		1.13 (1.04, 1.22)	< 0.01
BMI (kg/m ²)	Normal (18.5-< 25)	1 (ref)	
	Overweight (25.0-< 30.0)	1.27 (0.95, 1.69)	0.11
	Obese (≥ 30.0)	1.48 (1.02, 2.17)	0.04
WC (cm)	Normal (≤ 88)	1 (ref)	
	Abdominal obese (> 88)	1.58 (1.10, 2.28)	0.01
WHR	≤ 0.8	1 (ref)	
	0.8-0.85	1.38 (0.96, 1.97)	0.08
	> 0.85	1.56 (1.02, 2.37)	0.04
WHtR	Normal (< 0.5)	1 (ref)	
	Overweight (0.5-< 0.6)	1.26 (0.91, 1.74)	0.16

	Obese (≥ 0.6)	1.25 (1.32, 3.84)	< 0.01
Blouse size	≤ 10	1 (ref)	
	12	0.97 (0.62, 1.52)	0.89
	14	0.87 (0.54, 1.38)	0.55
	16	1.24 (0.77, 2.01)	0.38
	18	1.27 (0.72, 2.25)	0.41
	≥ 20	1.64 (0.93, 2.90)	0.09
Skirt size	≤ 10	1 (ref)	
	12	1.35 (0.75, 2.40)	0.32
	14	1.42 (0.80, 2.51)	0.23
	16	1.73 (0.96, 3.10)	0.07
	18	1.32 (0.67, 2.60)	0.41
	≥ 20	2.36 (1.25, 4.46)	< 0.01
Used competing risk regression model (Fine-Gray method, 255 cases/27,968 participants)			
BMI (per 2.5 kg/m ² increment)		1.09 (1.02, 1.17)	0.01
WC (per 5 cm increment)		1.06 (1.00, 1.14)	0.05
WHR (per 0.1 increment)		1.20 (1.00, 1.44)	0.05
WHtR (per 0.1 increment)		1.27 (1.05, 1.55)	0.02
Blouse size (per one-size increase)		1.13 (1.03, 1.24)	< 0.01
Skirt size (per one-size increase)		1.13 (1.04, 1.22)	< 0.01
BMI (kg/m ²)	Normal (18.5-< 25)	1 (ref)	
	Overweight (25.0-< 30.0)	1.27 (0.95, 1.69)	0.11

WC (cm)	Obese (≥ 30.0)	1.49 (1.02, 2.17)	0.04
	Normal (≤ 88)	1 (ref)	
	Abdominal obese (> 88)	1.58 (1.10, 2.28)	0.01
WHR	≤ 0.8	1 (ref)	
	0.8-0.85	1.38 (0.96, 1.97)	0.08
	> 0.85	1.56 (1.02, 2.37)	0.04
WHtR	Normal (< 0.5)	1 (ref)	
	Overweight ($0.5-< 0.6$)	1.26 (0.91, 1.74)	0.16
	Obese (≥ 0.6)	2.25 (1.32, 3.84)	< 0.01
Blouse size	≤ 10	1 (ref)	
	12	0.97 (0.62, 1.52)	0.89
	14	0.87 (0.54, 1.38)	0.55
	16	1.24 (0.77, 2.01)	0.38
	18	1.27 (0.72, 2.25)	0.40
	≥ 20	1.64 (0.92, 2.91)	0.09
Skirt size	≤ 10	1 (ref)	
	12	1.35 (0.75, 2.41)	0.32
	14	1.42 (0.80, 2.51)	0.23
	16	1.73(0.96, 3.11)	0.07
	18	1.32 (0.67, 2.60)	0.42
	≥ 20	2.36 (1.24, 4.47)	< 0.01

^a Models were adjusted for age, socio-economic status, marital status, physical activity, smoking status, alcohol consumption, total energy intake, the Alternate Healthy Eating Index -2010, menopausal status, hormone replacement therapy, and the

prevalence of CVD, cancer, or diabetes. All sensitivity analyses were based on the adjusted model. HR (95% CI): hazard ratio (95% confidence interval).

Supplementary methods

Cox proportional hazards regression models stratified by age group

In this study, age was used as the time scale. Due to a violation of the proportional hazards assumption (assessed using Schoenfeld residuals), we stratified the Cox model by age group (≤ 55 and > 55 years). This approach adjusts for non-proportionality related to age, allowing for a single summary HR for each adiposity indicator while accounting for age-related differences in baseline risk. The updated proportional hazards test results are presented in Table B4.

Appendix C Chapter 4 Supplementary material

Figure C1. Flow chart of UKWCS participants.

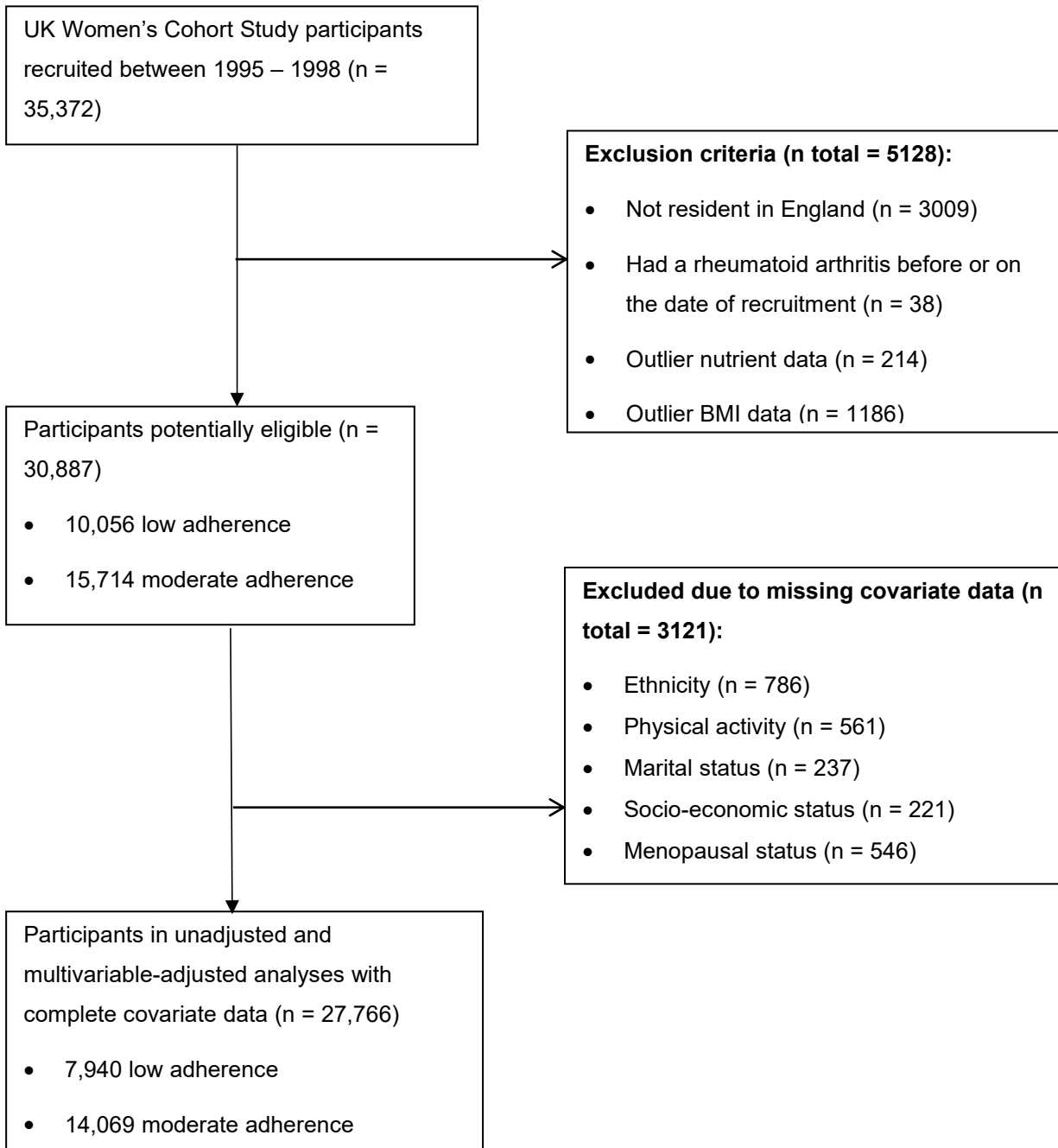


Table C1. Strengthening the reporting of observational studies in nutritional epidemiology (STROBE-Nut) checklist.

Topic	Item number	Recommendation	Page (line number)
Title and abstract	1	Indicate the study's design with a commonly used term in the title or the abstract	1
		Provide in the abstract an informative and balanced summary of what was done and what was found	1
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	2
Objectives	3	State specific objectives, including any prespecified hypotheses	2
Methods			
Study design	4	Present key elements of study design early in the manuscript	1,2
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	2,3
Participants	6	Cohort study - give the eligibility criteria, and the sources and methods of selection of participants; describe methods of follow-up	2,3
		Cohort study - for matched studies, give matching criteria and number of exposed and unexposed Case-control study - for matched studies, give matching criteria and the number of controls per case	N/A

Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers; give diagnostic criteria, if applicable	3, Figure 1
Data sources/measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement); describe comparability of assessment methods if there is more than one group	3,4
Bias	9	Describe any efforts to address potential sources of bias	3,4
Study size	10	Explain how the study size was arrived at	3,4
Quantitative variables	11	Explain how quantitative variables were handled in the analyses; if applicable, describe which groupings were chosen and why	3,4
Statistical methods	12	Describe all statistical methods, including those used to control for confounding	3,4
		Describe any methods used to examine subgroups and interactions	4
		Explain how missing data were addressed	2,4
		Cohort study - if applicable, explain how loss to follow-up was addressed	2
		Describe any sensitivity analyses	
Results			
Participants	13*	Report numbers of individuals at each stage of study - e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analyzed	5
		Give reasons for nonparticipation at each stage Consider use of a flow diagram	5, Supplementary Figure S1

Descriptive data	14*	Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders	5,6
		Indicate number of participants with missing data for each variable of interest Cohort study - summarize follow-up time (e.g., average and total amount)	5
Outcome data	15*	Cohort study - report numbers of outcome events or summary measures over time	5,6
Main results	16	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval); make clear which confounders were adjusted for and why they were included Report category boundaries when continuous variables were categorized	6,7, Figure 2
		If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done - e.g., analyses of subgroups and interactions, and sensitivity analyses	7,8, Supplementary Figure S5, S6, S7
Discussion			
Key results	18	Summarize key results with reference to study objectives	8
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision; discuss both direction and magnitude of any potential bias	8,9,10
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	9, 10
Generalizability	21	Discuss the generalizability (external validity) of the study results	10, 11

Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	12

Table C2. Covariates at recruitment and their derivation.

Covariate	How the variable was derived
Socio-demographic variables	
Age	Calculated as year differences between date of birth and date of recruitment and was considered a continuous variable in adjustment sets.
Socio-economic status	Participants were asked about their occupation. Options were 'never had paid job', 'managers and administrators', 'professional', 'technical and associate professional', 'clerical and secretarial', 'craft and skilled', 'personal and protective', 'sales', 'plant and machine operatives', or 'other'. We condensed these options into 'routine/manual', 'intermediate', or 'managerial/professional'.
Education	Participants were asked what their highest educational qualification was. Options were 'no qualifications', 'O level', 'A level', 'degree', or 'missing'.
Marriage	Participants were asked 'what is your marital status?' with options of 'married or living as married', 'divorced', 'widowed', 'single', or 'separated'. We combined 'divorced' and 'separated' together, and 'widowed' and 'single' together.
Lifestyle and other variables	
Physical activity	Participants were asked "In a typical week during the last 12 months, how many hours did you spend on the following activities: housework, do-it-yourself, gardening, walking, cycling, other physical activity". Physical activity was calculated as metabolic equivalent tasks (METs) by multiplying the duration of activities by their respective MET values.
Smoking status	Participants were asked to describe their smoking habit as 'smoke daily', 'smoke occasionally', 'ex-smoker', or 'never'. We combined daily and

	occasional smokers into 'current smokers', and kept 'ex-smoker' and 'never smoked' the same.
Alcohol consumption	Participants were asked how often they drink alcohol. Options were '> 1/week', '1/week', '< 1/week', or 'never'. This was regrouped as '≥ 1/week', '< 1/week', or 'never'. Participants reported typical weekly consumption of beer, wine, sherry, and spirits. This was computed as grams per week.
Body weight	Self-reported continuous variable
Height	Self-reported continuous variable
Body mass index	Calculated as self-reported body weight (kg) divided by the square of self-reported height (m), considered as a continuous variable (kg/m ²).
Menopausal status	Categorised participants as pre-menopausal or post-menopausal. Criteria for postmenopausal was: age > 55 years, both ovaries removed, currently on hormone replacement therapy, or no periods in the last 12 months.
Hormone replacement therapy use	Participants were asked 'have you ever used hormone replacement therapy?' and 'are you using HRT now?' – based on these yes or no answers, we categorised hormone replacement therapy use as 'user' and 'non-user'.
Energy intake	Participants were asked how often they eat the specific food item and completed a 217-item FFQs. Portion sizes were assigned based on McCance and Widdowson's Composition of Foods database, which provides standard UK portion sizes. Each food item in the FFQ was linked to nutrient composition tables to calculate total energy intake. This was computed as kcal per day.
Dietary nutrient supplements	Participants were asked 'Do you take vitamins, minerals, or food supplements?' - based on these yes

	or no answers, we categorised nutrient supplements as 'Yes' and 'No'.
Prevalence of CVD, cancer, or diabetes	Participants were asked 'Have you ever had a XX' including cancer, stroke, and diabetes; - based on these yes or no answers, we categorised prevalence of CVD, cancer, or diabetes as 'Yes' and 'No'.

Table C3. Grading of the Mediterranean diet score.

Components	Median Cut-off (g/day or ratio)	Points (0 or 1)
Vegetables (g/day)	281.8	1
Legumes (g/day)	31.0	1
Fruit and nuts (g/day)	269.7	1
Cereals (g/day)	224.7	1
Fish (g/day)	29.0	1
MUFA + PUFA : SFA*	0.9	1
Meat (g/day)	80.5	1
Poultry (g/day)	10.2	1
Dairy products (g/day)	396.7	1
Alcohol (g/day)	5–25	1
Total score range		0–10

Ratio of the sum of monounsaturated fatty acids and polyunsaturated fatty acids to saturated fatty acids. MUFA = monounsaturated fatty acids; PUFA = polyunsaturated fatty acids; SFA = saturated fatty acids. Each component receives 1 point if intake is above the median for beneficial components (e.g., vegetables, legumes, fruit and nut, cereals, fish, MUFA + PUFA : SFA), or below for potentially adverse components (e.g., meat, poultry, dairy). Alcohol scoring is based on moderate intake: 5–25 g/day for women.

Table C4. Dietary characteristics of the UKWCS by adherence to Mediterranean diet at recruitment.

Characteristics , mean (SD) or n (%)	Total	Low adherence	Moderate adherence	High adherence
Participants	27,766	7940	14,069	5757
Foods and beverages intake				
Fruit (g/day)	398.3 (275.9)	283.2 (209.6)	403.7 (264.1)	543.8 (310.6)
Vegetables (g/day)	314.2 (175.6)	212.9 (105.5)	319.3 (161.5)	441.2 (198.4)
Total meat (g/day)	86.2 (81.5)	118.8 (66.6)	89.5 (83.9)	33.3 (66.6)
Red meat (g/day)	39.3 (47.1)	55.5 (44.2)	40.5 (48.9)	13.9 (34.1)
Poultry (g/day)	16.8 (20.2)	22.0 (18.7)	17.7 (20.7)	7.3 (17.3)
Processed meat (g/day)	28.2 (30.9)	38.9 (28.4)	29.2 (31.7)	11.2 (24.4)
Total fish (g/day)	34.0 (30.7)	28.1 (20.4)	35.7 (30.9)	37.9 (39.3)
Oily fish (g/day)	9.1 (12.5)	6.3 (7.7)	9.6 (12.1)	11.7 (17.3)
Lean fish (g/day)	24.9 (23.7)	21.8 (16.9)	26.1 (24.5)	26.2 (29.0)
Total dairy (g/day)	414.2 (215.8)	471.3 (201.2)	415.0 (214.8)	333.4 (212.5)
Milk (g/day)	304.4 (191.5)	369.8 (183.4)	301.7 (186.5)	220.8 (180.3)
Yoghurt (g/day)	60.2 (69.4)	53.9 (66.5)	62.5 (71.2)	63.1 (68.2)
Cheese (g/day)	27.8 (27.8)	22.2 (22.7)	28.8 (29.3)	33.1 (29.0)

Cream (g/day)	1.7 (3.4)	2.1 (4.4)	1.6 (3.0)	1.3 (2.7)
Dairy desserts (g/day)	20.1 (26.4)	23.4 (29.2)	20.3 (26.9)	15.1 (19.6)
Eggs (number/day)	22.4 (18.9)	21.4 (17.9)	22.8 (19.1)	22.6 (19.6)
Tea (cups/day)	3.0 (2.0)	2.9 (2.0)	3.0 (2.0)	3.1 (2.0)
Total coffee (cups/day)	2.0 (1.8)	2.1 (1.9)	2.0 (1.8)	1.8 (1.8)
Caffeinated coffee (cups/day)	1.5 (1.7)	1.6 (1.7)	1.5 (1.7)	1.3 (1.6)
Decaffeinated coffee (cups/day)	0.5 (1.1)	0.5 (1.1)	0.5 (1.1)	0.5 (1.1)
Nutrients				
Protein (g/day)	88.8 (26.8)	84.94 (19.6)	90.12 (28.6)	90.97 (18.8)
Protein (g/kg BW/day)	1.39 (0.5)	1.30 (0.4)	1.40 (0.5)	1.47 (0.5)
Carbohydrate intake (g/day)	308.6 (49.5)	271.1 (84.6)	310.5 (49.5)	355.5 (97.9)
Carbohydrate (% energy)	53.1 (6.7)	51.1 (6.7)	53.3 (24.8)	55.5 (6.3)
Fibre intake (g/day)	25.3 (9.9)	18.8 (6.5)	25.6 (8.7)	33.4 (10.4)
Fibre (% energy)	2.2 (0.5)	1.8 (0.5)	2.2 (0.6)	2.6 (0.6)
Fat intake (g/day)	83.8 (30.9)	79.2 (28.7)	83.8 (31.4)	90.2 (31.3)
Fat (% energy)	32.2 (5.6)	33.2 (5.6)	31.9 (5.6)	31.3 (5.6)
Saturated fat intake (g/day)	29.0 (12.9)	30.5 (13.1)	28.9 (13.1)	26.9 (11.6)
Saturated fat (% energy)	11.1 (3.1)	12.7 (3.2)	10.9 (2.9)	9.3 (2.6)
MUFA (g/day)	27.5 (10.6)	25.9 (9.6)	27.5 (10.8)	29.7 (11.3)
MUFA (% energy)	10.54 (2.2)	10.9 (2.0)	10.5 (2.2)	10.3 (2.4)

PUFA (g/day)	16.1 (6.6)	13.0 (4.8)	16.1 (6.1)	20.4 (7.5)
PUFA (% energy)	6.2 (1.6)	5.5 (1.3)	6.2 (1.6)	7.1 (1.8)
Vitamin C (mg/day)	171.3 (83.6)	131.0 (59.6)	174.3 (79.3)	219.7 (93.9)
Vitamin D (µg/day)	3.1 (1.7)	2.9 (1.3)	3.2 (1.7)	3.2 (2.0)
Zinc (mg/day)	11.3 (3.6)	10.6 (3.1)	11.5 (3.8)	12.0 (3.6)
Calcium (mg/day)	1145.6 (374.4)	1097.6 (329.8)	1154.5 (389.5)	1189.9 (387.4)
Folate (µg/day)	400.1 (130.6)	336.2 (96.8)	405.0 (125.1)	476.4 (140.2)

SD standard deviation. BW body weight (kg).

Table C5. Associations between individual Mediterranean diet components and risk of rheumatoid arthritis in the UK Women's Cohort Study.

Food components	Median intake (g/day or ratio)	Cases/participants	HR(95% CI)	
			Age-adjusted model	*multivariable-adjusted model
Vegetables				
Low	166.3	94/9256	1	1
Moderate	281.8	76/9179	0.81 (0.60, 1.10)	0.82 (0.60, 1.11)
High	448.9	85/9170	0.90 (0.67, 1.20)	0.87 (0.64, 1.20)
Legumes				
Low	12.8	77/9242	1	1
Moderate	31.2	98/9125	1.34 (0.99, 1.81)	1.31 (0.97, 1.77)
High	72.5	80/9144	1.14 (0.83, 1.56)	1.08 (0.78, 1.50)
Fruit and nuts				

Low	133.5	97/9256	1	1
Moderate	269.7	74/9255	0.73 (0.54, 0.99)	0.74 (0.54, 1.01)
High	481.7	84/9255	0.81 (0.61, 1.09)	0.78 (0.57, 1.07)
Cereals				
Low	131.1	92/9164	1	1
Moderate	224.7	74/9181	0.83 (0.61, 1.12)	0.81 (0.59, 1.11)
High	350.0	89/9166	1.02 (0.76, 1.36)	0.96 (0.68, 1.35)
Fish				
Low	4.2	79/9196	1	1
Moderate	29.0	80/9171	0.97 (0.71, 1.32)	0.98 (0.71, 1.34)
High	57.5	96/9144	1.14 (0.84, 1.54)	1.15 (0.84, 1.57)
MUFA + PUFA: SFA				
Low	0.8	90/9166	1	1
Moderate	1.0	93/9162	1.07 (0.80, 1.43)	1.10 (0.82, 1.47)
High	1.2	72/9183	0.83 (0.61, 1.13)	0.87 (0.63, 1.19)
Meat				
Low	0	84/9172	1	1
Moderate	80.5	87/9169	0.96 (0.71, 1.30)	0.96 (0.70, 1.30)
High	162.3	84/9170	0.96 (0.70, 1.30)	0.87 (0.63, 1.21)
Poultry				
Low	0	77/9185	1	1
Moderate	14	104/10,730	1.09 (0.81, 1.47)	1.08 (0.79, 1.46)
High	40	74/7596	1.13 (0.82, 1.55)	1.10 (0.79, 1.53)
Dairy				
Low	193.7	77/9179	1	1
Moderate	396.8	87/9169	1.12 (0.83, 1.52)	1.11 (0.81, 1.51)
High	637.8	91/9163	1.14 (0.84, 1.55)	1.05 (0.76, 1.45)
Alcohol				

Low	0	98/9182	1	1
Moderate	50.4	89/9161	0.95 (0.71, 1.26)	1.04 (0.76, 1.43)
High	152.8	68/9168	0.74 (0.54, 1.01)	1.03 (0.58, 1.82)

Adjusted for age, education, socio-economic status, marital status, BMI, smoking status, energy intake, alcohol consumption, prevalence of cardiovascular disease, cancer, or diabetes, menopausal status, and hormone replacement therapy use. HR (95% CI), hazard ratio (95% confidence interval).

Table C6. Association between adherence to the Mediterranean diet and risk of rheumatoid arthritis by potential effect modifiers in the UKWCS.

Stratifying variable		n cases, *adjusted HR (95% CI)		P- interaction
Age	≤ 55 years	> 55 years		
Low adherence	4 1	3 1		
(reference)	4	9		
Moderate adherence	7 0.87 (0.60, 1.27)	5 0.97 (0.65, 1.45)		
	5	9		
High adherence	2 0.57 (0.34, 0.95)	1 0.78 (0.43, 1.42)		0.7
	3	5		
Menopausal status	Pre-menopausal	Post-menopausal		
Low adherence	5 1	2 1		
(reference)	5	8		
Moderate adherence	8 0.83 (0.52, 1.32)	4 0.98 (0.70, 1.37)		
	7	7		
High adherence	1 0.68 (0.38, 1.22)	1 0.62 (0.37, 1.05)		0.7
	9	9		
Physical activity	Inadequate physical activity	Adequate physical activity		

	≤ 10 MET-hrs/week		> 10 MET-hrs/week		
Low adherence	2	1	5	1	
(reference)	4		9		
Moderate adherence	3	1.04 (0.62, 1.73)	9	0.84 (0.61, 1.17)	
	9		5		
High adherence	9	0.67 (0.31, 1.46)	2	0.58 (0.37, 0.91)	0.8
			9		
Body mass index	≤ 25.0 kg/m²		> 25.0 kg/m²		
Low adherence	4	1	3	1	
(reference)	5		8		
Moderate adherence	7	0.88 (0.61, 1.27)	5	0.96 (0.64, 1.46)	
	8		6		
High adherence	2	0.62 (0.38, 1.00)	1	0.67 (0.35, 1.28)	0.9
	6		2		
Smoking status	No-smoker		Ex-smoker		Current smoker
Low adherence	4	1	2	1	1
(reference)	5		2		6
Moderate adherence	6	0.78 (0.53, 1.15)	4	1.05 (0.64, 1.75)	2
	2		9		3

High adherence	1	0.54 (0.31, 0.95)	1	0.88 (0.48, 1.63)	2	0.25(0.06,1.10)	0.4
	7		9				

Adjusted for age, education, socio-economic status, marital status, BMI, smoking status, energy intake, alcohol consumption, prevalence of cardiovascular disease, cancer, or diabetes, menopausal status, and hormone replacement therapy use. HR (95% CI), hazard ratio (95% confidence interval). The respective stratification variable was excluded from the adjustment model in each subgroup analyses.

Table C7. Sensitivity analysis for the association between risk of rheumatoid arthritis and the Mediterranean diet adherence with varying restrictions in the UKWCS.

Model adjustments	± further	HR (95% CI)		
		Low adherence (reference)	Moderate adherence	High adherence
Model 1 ^a		1.0	0.90(0.68,1.18)	0.61(0.42,0.90)
Model 2 ^b		1.0	0.94(0.71,1.25)	0.67(0.45,1.00)
Sensitivity analyses (based on Model 2):				
Excluding subjects with < 5 years of follow-up		1.0	0.95(0.70,1.29)	0.61(0.39,0.96)
Excluding subjects with chronic diseases		1.0	0.84(0.62,1.13)	0.63(0.42,0.96)
Excluding subjects as current smokers		1.0	0.82(0.55,1.21)	0.61(0.34,1.09)
Model 2 + vitamin D intake		1.0	0.94(0.71,1.24)	0.67(0.45,1.00)
Model 2 + nutrient supplements		1.0	0.98(0.73,1.31)	0.61(0.40,0.95)
Competing risk model (Fine and Gray method)		1.0	0.94(0.72,1.24)	0.67(0.44,0.99)

^aModel 1: adjusted for age (continuous) only.

^bModel 2: adjusted for age, education, socio-economic status, marital status, BMI, smoking status, energy intake, alcohol consumption, prevalence of cardiovascular disease, cancer, or diabetes, menopausal status, and hormone replacement therapy use.

All other models were based on the 27,766 participants in model 2. HR (95% CI), hazard ratio (95% confidence interval).

Supplementary Methods

Categorisation of Mediterranean diet adherence

Adherence to the Mediterranean diet was evaluated using a modified 10-point scoring system adapted from Trichopoulou et al.[1-2], based on dietary data collected via a Food Frequency Questionnaire administered at recruitment. The FFQ assessed consumption frequencies of foods and beverages using 10 response options: 0 'never', 1 '< once/month', 2 '1–3/month', 3 'once/week', 4 '2–4/week', 5 '5–6/week', 6 'once/day', 7 '2–3/day', 8 '4–5/day', or 9 '6+/day*'. Responses were converted to daily frequencies (e.g., 0, 0.02, 0.07, 0.14, 0.4, 0.8, 1, 2.5, 4.5, 6 times/day) to estimate intake of key Mediterranean diet components.

For this non-Mediterranean British cohort, binary scores (0 or 1) were assigned to ten food components based on cohort-specific median intake values. Participants received 1 point for:

- Beneficial components (vegetables, legumes, fruit/nuts, cereals, fish, MUFA+PUFA:SFA ratio) if intake was $\geq$ median.
- Potentially detrimental components (meat, poultry, dairy) if intake was $<$ median.
- Alcohol if intake was 5–25 g/day.

Total Mediterranean diet scores (range: 0–10) classified participants into: low (0–3), moderate (4–6), or high (7–10) adherence.

References

- [1] Trichopoulou A, Kouris-Blazos A, Wahlqvist ML, Gnardellis C, Lagiou P, Polychronopoulos E, Vassilakou T, Lipworth L, Trichopoulos D: Diet and overall survival in elderly people. *Bmj* 1995, 311(7018):1457-1460.
- [2] Trichopoulou A, Bamia C, Norat T, Overvad K, Schmidt EB, Tjønneland A, Halkjaer J, Clavel-Chapelon F, Vercambre MN, Boutron-Ruault MC, Linseisen J, Rohrmann S, Boeing H, Weikert C, Benetou V, Psaltopoulou T, Orfanos P, Boffetta P, Masala G, Pala V, Panico S, Tumino R, Sacerdote C, Bueno-de-Mesquita HB, Ocke MC, Peeters PH, Van der Schouw YT, González C, Sanchez MJ, Chirlaque MD, Moreno C, Larrañaga N, Van Guelpen B, Jansson JH, Bingham S, Khaw KT,

Spencer EA, Key T, Riboli E, Trichopoulos D: Modified Mediterranean diet and survival after myocardial infarction: the EPIC-Elderly study. *Eur J Epidemiol* 2007, 22(12):871-881.

Appendix D Chapter 5 Supplementary material

Table D1. Strengthening the reporting of observational studies in nutritional epidemiology (STROBE-Nut) checklist.

Section/topic	Item number	Recommendation	Page (line number)
Title and abstract	1	Indicate the study's design with a commonly used term in the title or the abstract	1
		Provide in the abstract an informative and balanced summary of what was done and what was found	1
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	2,3
Objectives	3	State specific objectives, including any prespecified hypotheses	1-3
Methods			
Study design	4	Present key elements of study design early in the manuscript	3,4, Figure 2
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	3,4, Figure 1
Participants	6	Cohort study - give the eligibility criteria, and the sources and methods of selection of participants; describe methods of follow-up	3,4

		Cohort study - for matched studies, give matching criteria and number of exposed and unexposed Case-control study - for matched studies, give matching criteria and the number of controls per case	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers; give diagnostic criteria, if applicable	4-6
Data sources/measurements	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement); describe comparability of assessment methods if there is more than one group	4,5, Table S2,S3
Bias	9	Describe any efforts to address potential sources of bias	7, Table S4
Study size	10	Explain how the study size was arrived at	8,Figure 1
Quantitative variables	11	Explain how quantitative variables were handled in the analyses; if applicable, describe which groupings were chosen and why	6, Table S5
Statistical methods	12	Describe all statistical methods, including those used to control for confounding	6,7
		Describe any methods used to examine subgroups and interactions	7
		Explain how missing data were addressed	3,4
		Cohort study - if applicable, explain how loss to follow-up was addressed	3,4
		Describe any sensitivity analyses	7
Results			
Participants	13*	Report numbers of individuals at each stage of study - e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and	Figure 1

		analyzed	
		Give reasons for nonparticipation at each stage Consider use of a flow diagram	Figure 1
Descriptive data	14*	Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders	Table 1
		Indicate number of participants with missing data for each variable of interest Cohort study - summarize follow-up time (e.g., average and total amount)	8, Figure 1
Outcome data	15*	Cohort study - report numbers of outcome events or summary measures over time	8
Main results	16	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval); make clear which confounders were adjusted for and why they were included Report category boundaries when continuous variables were categorized	11-13, Table 2,3
		If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done - e.g., analyses of subgroups and interactions, and sensitivity analyses	16, Table 4, S6-S7
Discussion			
Key results	18	Summarize key results with reference to study objectives	17-20
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision; discuss both direction and magnitude of any potential bias	20,21

Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	19-21
Generalizability	21	Discuss the generalizability (external validity) of the study results	21
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	22

Table D2. Conversion of alcohol nutrient intake ^[1].

Alcohol	Grams pint/glass	per Alcohol nutrient/100g	Alcohol units	nutrient
Wine	125	9.25	1.5	
Beer	287	3.08	1.1	
Cider	287	5.98	2.1	
Spirits	23	31.70	0.9	
Sherry	40	16.65	0.8	

1 unit equals 8g of pure alcohol, from “Calculating alcohol units”, by NHS. (<https://www.nhs.uk/live-well/alcohol-advice/calculating-alcohol-units/>).

[1] Nur U, Longford NT, Cade JE, Greenwood DC. The impact of handling missing data on alcohol consumption estimates in the UK women cohort study. *Eur J Epidemiol.* 2009;24(10):589-95. doi: 10.1007/s10654-009-9384-1. PMID: 19728116.

The pure alcohol of each drink was calculated by multiplying the total weight of a drink (in grams) by quantities of its nutrient intake (per 100g) and dividing the result by 100.

Table D3. Components of death causes.

Components	ICD-9/ICD-10 (from HES)	Number of events
CVD Death	390-448 (ICD-9)	42
Cancer Death	140-208 (ICD-9)	122
CVD Death	I00-I99 (ICD-10)	903
Cancer Death	C00-C97(ICD-10)	824
Total		1891

Table D4. Covariates at recruitment and their derivation.

Covariate	How the variable was derived
Socio-demographic variables	
Age	Calculated as year differences between date of birth and date of recruitment and was considered a continuous variable in adjustment sets.
Socio-economic status	Participants were asked about their occupation. Options were 'never had paid job', 'managers and administrators', 'professional', 'technical and associate professional', 'clerical and secretarial', 'craft and skilled', 'personal and protective', 'sales', 'plant and machine operatives', or 'other'. We condensed these options into 'routine/manual', 'intermediate', or 'managerial/professional'.
Marital status	Participants were asked 'what is your marital status?' with options of 'married or living as married', 'divorced', 'widowed', 'single', or 'separated'. We combined 'divorced' and 'separated' together, and 'widowed' and 'single' together.
Lifestyle factors	
Physical activity	Participants were asked how long they perform exercises that makes them sweat per week (in hours and minutes per week). This was computed as hours per day.
Smoking	Participants were asked to describe their smoking habit as 'smoke daily', 'smoke occasionally', 'ex-smoker', or 'never'. We combined daily and occasional smokers into 'smokers', and kept 'ex-smoker' and 'never smoked' the same.
Body mass index	Calculated as self-reported weight divided by the square of self-reported height, considered as a continuous variable
Hormones/reproductive factors	

Menopausal status	Categorised participants as pre-menopausal or post-menopausal. Criteria for postmenopausal was: age > 55 years, both ovaries removed, currently on hormone replacement therapy, or no periods in the last 12 months.
Hormone replacement therapy use	Participants were asked 'have you ever used hormone replacement therapy?' and 'are you using HRT now?' – based on these yes or no answers, we categorised hormone replacement therapy use as 'current', 'former user', and 'never'.
Prevalence of stroke, diabetes, and cancer	Participants were asked 'Have you ever had a XX' including cancer, stroke, and diabetes; - based on these yes or no answers, we categorised prevalence of CVD, cancer, or diabetes as 'Yes' and 'No'.
Diet quality	AHEI-2010
Total energy intake except alcohol	Participants were asked how often they eat the specific food item and completed a 217-item FFQs. Portion sizes were assigned based on McCance and Widdowson's Composition of Foods database, which provides standard UK portion sizes. Each food item in the FFQ was linked to nutrient composition tables to calculate total energy intake. This was computed as kcal per day.

Table D5. Baseline characteristics of 23,699 UKWCS participants stratified by drinking frequency.

Characteristics	Alcohol consumption, units per week			
	None	< 7	7-14	>14
Participants (%)	2151 (9.1)	10884 (45.9)	5246 (22.1)	5418 (22.9)
Age, years (SD)	53.7 (9.5)	52.0 (9.2)	50.8 (9.0)	50.2 (8.5)
Degree-level education (%)	406 (20.9)	2451 (24.6)	1592 (32.5)	1713 (33.7)

Socio-economic status (%)				
Professional or managerial	1291 (60.0)	6655 (61.1)	3438 (65.5)	3784 (69.8)
Intermediate	597 (27.8)	3144 (28.9)	1403 (26.7)	1289 (23.8)
Routine or manual	263 (12.2)	1085 (10.0)	405 (7.7)	345 (6.4)
Married (%)	1471 (68.4)	7966 (73.2)	4025 (76.7)	4209 (77.7)
White ethnicity (%)	2015 (96.5)	10506 (98.8)	5094 (99.2)	5259 (99.3)
Anthropometric indices				
Height, cm (SD)	163.3 (7.3)	163.6 (6.6)	164.0 (6.5)	164.4 (6.7)
Weight, kg (SD)	66.3 (13.7)	66.1 (12.4)	64.5 (10.5)	64.7 (10.7)
BMI at recruitment, kg/m2 (SD)	24.9 (5.0)	24.7 (4.4)	24.0 (3.7)	23.9 (3.8)
Physical activity (METs), hours/week (SD)	15.1 (10.4)	14.9 (10.1)	14.5 (9.6)	14.4 (9.7)
Smoking status (%)				
Current	199 (9.3)	1024 (9.4)	575 (11.0)	890 (16.4)
Former	463 (21.5)	2949 (27.1)	1755 (33.5)	2156 (39.8)
Never	1489 (69.2)	6911 (63.5)	2916 (55.6)	2372 (43.8)
Alcohol consumption, units/week (SD)	0(0)	2.6 (2.1)	10.2 (2.0)	24.4 (10.9)
Total energy intake except alcohol, kcal/day (SD)	2291 (752)	2295 (693)	2311 (699)	2208 (688)
AHEI-2010 score (SD)	60.1 (10.9)	70.3 (10.5)	69.2 (10.4)	61.0 (10.3)
Menopausal status (%)				
Premenopausal	863 (40.1)	5285 (48.6)	2831 (54.0)	3015 (55.7)
Postmenopausal	1288 (59.9)	5599 (51.4)	2415 (46.0)	2403 (44.3)
HRT (%)	565 (26.3)	2849 (26.2)	1400 (26.7)	1506 (27.8)
Prevalence of CVD, cancer, or diabetes (%)	239 (11.1)	967 (8.9)	425 (8.1)	423 (7.8)
CVD, cancer deaths(%)	232 (10.8)	899 (8.3)	366 (7.0)	394 (7.3)

SD standard deviation, BMI body mass index, MET Metabolic Equivalent of Task, AHEI-2010 Alternative Healthy Eating Index-2010, HRT hormone replacement therapy, CVD cardiovascular disease.

Table D6. Risk of rheumatoid arthritis according to drinking status stratified by potential effect modifiers in the UKWCS.

Stratifying variable		n cases, adjusted HR (95% CI)				p interaction
Age		< 55 years		≥ 55 years		0.9
Regular drinkers (reference)	91 1			57 1		
Occasional drinkers	37 1.27 (0.85, 1.88)			35 1.45 (0.89, 2.35)		
Non-drinkers	14 1.23 (0.69, 2.20)			21 1.22 (0.66, 2.27)		
Menopausal status		Pre-menopausal		Post-menopausal		0.4
Regular drinkers (reference)	85 1			63 1		
Occasional drinkers	52 1.09 (0.66, 1.80)			20 1.49 (1.01, 2.20)		
Non-drinkers	23 1.48 (0.76, 2.88)			12 1.08 (0.63, 1.87)		
Physical activity		Low (tertile 1)		Moderate (tertile 2)		High (tertile 3) 0.7
Regular drinkers (reference)	33 1			67 1	48 1	
Occasional drinkers	20 1.88 (1.06, 3.38)			29 1.37 (0.87, 2.18)	23 1.01 (0.57, 1.80)	
Non-drinkers	6 1.37 (0.57, 3.29)			18 1.25 (0.66, 2.39)	11 1.21 (0.59, 2.48)	
SES		Professional/managerial		Intermediate		Routine/manual 0.9
Regular drinkers (reference)	14 1			47 1	87 1	

Occasional drinkers	10	1.17 (0.66, 2.06)	19	1.41 (0.95, 2.10)	43	1.43 (0.58, 3.50)
Non-drinkers	5	1.49 (0.75, 2.97)	14	1.08 (0.59, 1.98)	16	1.28 (0.41, 3.98)
BMI	< 25		≥ 25		0.9	
Regular drinkers (reference)	85	1	63	1		
Occasional drinkers	35	1.41(0.93, 2.14)	37	1.24(0.79, 1.94)		
Non-drinkers	18	1.26 (0.70,2.27)	17	1.21(0.67, 2.21)		

Models were adjusted for the following: age, socio-economic status (SES, professional/managerial, intermediate, routine/manual), physical activity (METs, continuous), total energy intake (continuous), marriage status (married/living as married, separated/divorced, single/widowed), body mass index (BMI, continuous), diet quality (AHEI-2010 score, continuous), hormone replacement therapy, and menopausal status (premenopausal, postmenopausal). HR (95% CI) hazard ratio (95% confidence interval).

Table D7. Risk of rheumatoid arthritis according to amount of alcohol drinking stratified by potential effect modifiers in the UKWCS.

Stratifying variable		n cases, adjusted HR (95% CI)		p interaction
Age		< 55 years	≥ 55 years	0.5
None (reference)	14	1	19	1
< 7	61	0.68 (0.38,1.22)	60	0.94 (0.50,1.77)

7-14	37	0.88 (0.47,1.62)	18	0.75 (0.35,1.60)		
> 14	30	0.64 (0.34,1.21)	16	0.60 (0.26,1.35)		
Menopausal status		Pre-menopausal		Post-menopausal		0.3
None (reference)	13	1	20	1		
< 7	39	0.50 (0.26,0.97)	82	1.06 (0.59,1.88)		
7-14	26	0.66 (0.33,1.31)	29	0.99 (0.52,1.90)		
> 14	17	0.41 (0.19,0.85)	29	0.89 (0.46,1.72)		
Physical activity		Low (tertile 1)		Moderate (tertile 2)	High (tertile 3)	0.02
None (reference)	5	1	14	1	14	1
< 7	33	1.16 (0.45,3.02)	51	0.83 (0.42,1.65)	37	0.52 (0.26,1.05)
7-14	12	0.95 (0.34,2.71)	25	0.85 (0.40,1.79)	18	0.74 (0.35,1.57)
> 14	9	0.52 (0.17,1.65)	24	0.84 (0.40,1.79)	13	0.44 (0.19,1.02)
SES		Professional/managerial		Intermediate		Routine/manual 0.9
None (reference)	14	1	14	1	5	1
< 7	71	0.97 (0.52,1.81)	36	0.50 (0.25,1.00)	14	0.96 (0.27,3.43)
7-14	33	0.97 (0.50,1.90)	16	0.58 (0.27,1.26)	6	1.29 (0.32,5.17)
> 14	28	0.72 (0.36,1.42)	14	0.51 (0.22,1.15)	4	0.76 (0.15,3.76)

Smoking status	Current		Former		Never		0.2
None (reference)	10	1	9	1	14	1	
< 7	15	0.23 (0.10,0.56)	38	0.94 (0.40,2.24)	68	1.06 (0.56,2.03)	
7-14	8	0.30 (0.12,0.78)	24	1.05 (0.43,2.58)	23	0.93 (0.45,1.95)	
> 14	11	0.24 (0.10,0.60)	17	0.57 (0.22,1.46)	18	0.81 (0.37,1.78)	
BMI	< 25 kg/m²		≥ 25 kg/m²				0.4
None (reference)	18	1	15	1			
< 7	58	0.68 (0.38,1.24)	63	0.87 (0.47,1.65)			
7-14	31	0.72 (0.38,1.37)	24	1.00 (0.50,1.99)			
> 14	31	0.69 (0.36,1.30)	15	0.50 (0.23,1.11)			

Models were adjusted for the following: age, socio-economic status (SES, professional/managerial, intermediate, routine/manual), physical activity (METs, continuous), total energy intake (continuous), marriage status (married/living as married, separated/divorced, single/widowed), body mass index (BMI, continuous), diet quality (AHEI-2010 score, continuous), hormone replacement therapy, and menopausal status (premenopausal, postmenopausal). HR (95% CI) hazard ratio (95% confidence interval).

Table D8. Sensitivity analyses of alcohol consumption and rheumatoid arthritis incidence in the UKWCS.

Alcohol consumption	Original model		Exclusion of events in first 2 years of follow-up		Excluded those who had reduced alcohol intake at baseline		Excluded those who had long-term treatment for illness	
	Cases/ participants	HR (95% CI)	Cases/ participants	HR (95% CI)	Cases/ participants	HR (95% CI)	Cases/ participants	HR (95% CI)
Alcohol status								
Regular drinkers (reference)	148/15932	1.00	147/15931	1.00	139/14654	1.00	67/10917	1.00
Occasional drinkers	72/5306	1.43(1.07,1.90)	68/5302	1.17 (0.85, 1.63)	63/4849	1.34(0.99,1.82)	21/3291	0.98(0.60,1.61)
Non-drinkers	35/2461	1.45(1.00,2.12)	33/2459	1.13 (0.73, 1.74)	34/2460	1.38(0.94, 2.03)	12/1359	1.32(0.70,2.47)
Total alcohol (continuous, per 1 unit/week)	255/23699	0.98(0.96,0.99)	248/23692	0.97 (0.96, 0.99)	236/21963	1.00(0.67,1.49)	100/15567	0.98(0.96,1.01)
Total alcohol intake (units/week):								
None	33/2118	1.00	31/2149	1.00	31/2147	1.00	13/1205	1.00
< 7	121/10763	0.76(0.50,1.14)	116/10879	0.87 (0.57, 1.33)	110/10019	0.79(0.53,1.20)	43/6965	0.60(0.32,1.13)
7-14	55/5191	0.73(0.46,1.14)	55/5246	1.19 (0.67, 2.10)	49/4818	0.75(0.47,1.18)	24/3610	0.67(0.34,1.31)
> 14	46/5372	0.57(0.36,0.89)	46/5418	1.54 (0.67, 3.56)	46/4979	0.65(0.41,1.05)	20/3787	0.54(0.27,1.09)
Wine only (servings/week):								

0	42/3983	1.00	40/4023	1.00	35/3729	1.00	15/2352	1.00
1-2	61/6112	1.05(0.67,1.63)	60/6172	1.00(0.66,1.49)	55/5595	1.06(0.69,1.65)	26/4088	1.02(0.54,1.92)
3-4	41/3262	1.34(0.81,2.20)	40/3302	1.25(0.80,1.96)	38/3012	1.40(0.87,2.24)	23/2299	1.66(0.86,3.20)
> 4	55/6437	1.03(0.65,1.64)	55/6492	0.87(0.57,1.33)	52/5998	0.97(0.62,1.50)	21/4601	0.79(0.41,1.54)
Wine(continuous, per 1 unit/week):	255/23669	0.99(0.97,1.00)	248/23692	0.98(0.96,1.00)	236/21963	0.98(0.96,1.03)	100/15567	0.98(0.96,1.02)
Beer only (servings/week):								
0	69/6936	1.00	66/7002	1.00		1.00	28/4352	1.00
< 1	10/1151	0.96(0.47,1.96)	9/1160	0.84(0.42,1.70)	54/6116	0.99(0.49,2.00)	4/819	0.73(0.25,2.10)
1-2	20/2116	1.01(0.58,1.77)	20/2136	0.99(0.59,1.65)	9/1030	1.13(0.66,1.94)	9/1513	0.91(0.43,1.93)
≥2	18/2001	1.05(0.59,1.86)	18/2019	0.92(0.52,1.61)	19/1845	0.97(0.53,1.76)	6/1453	0.62(0.24,1.57)
Beer (continuous, per 1 unit/week):	255/23669	0.98(0.93,1.04)	248/23692	0.98(0.94,1.02)	16/1784	0.97(0.94,1.03)	100/15567	0.98(0.92,1.04)
Sherry only(servings/week):								
0	65/6942	1.00	63/7005	1.00	52/6026	1.00	28/4484	1.00
1	21/2577	0.89(0.51,1.55)	20/2597	0.80(0.48,1.34)	17/2281	0.80(0.46,1.40)	8/1683	0.74(0.34,1.62)
2-3	10/1504	0.68(0.30,1.53)	10/1514	0.66(0.34,1.29)	10/1377	0.76(0.39,1.49)	4/993	0.61(0.22,1.69)
>3	13/1016	1.45(0.74,2.86)	13/1029	1.16(0.63,2.13)	13/972	1.27(0.69,2.35)	2/666	0.47(0.11,1.94)
Sherry (continuous, per 1 unit/week):	255/23699	0.98(0.90,1.07)	248/23692	0.98(0.91,1.06)	236/23699	0.99(0.91,1.06)	100/15567	0.90(0.74,1.09)
Spirit only (servings/week):								
0	64/6447	1.00	62/6509	1.00	64/6447	1.00	24/4089	1.00

1	30/2573	1.32(0.83,2.10)	30/2603	1.14(0.74,1.78)	30/2573	1.26(0.79,2.01)	13/1716	1.26(0.64,2.46)
2	22/1726	1.35(0.78,2.32)	20/1746	1.12(0.67,1.87)	22/1726	1.24(0.73,2.12)	10/1203	1.41(0.67,2.95)
≥3	30/3098	1.02(0.62,1.69)	30/3128	0.90(0.56,1.43)	30/3098	1.00(0.61,1.62)	13/2071	1.10(0.56,2.21)
Spirit (continuous, per 1 unit/week):	255/23699	0.95(0.90,1.01)	248/23692	0.96(0.92,1.01)	236/23699	0.96(0.92,1.01)	100/15567	1.00(0.95,1.06)

Models were adjusted for the following: age, socio-economic status (SES, professional/managerial, intermediate, routine/manual), physical activity (METs, continuous), total energy intake (continuous), marriage status (married/living as married, separated/divorced, single/widowed), body mass index (BMI, continuous), diet quality (AHEI-2010 score, continuous), hormone replacement therapy, and menopausal status (premenopausal, postmenopausal). HR (95% CI) hazard ratio (95% confidence interval).

Appendix E Chapter 6 Supplementary material

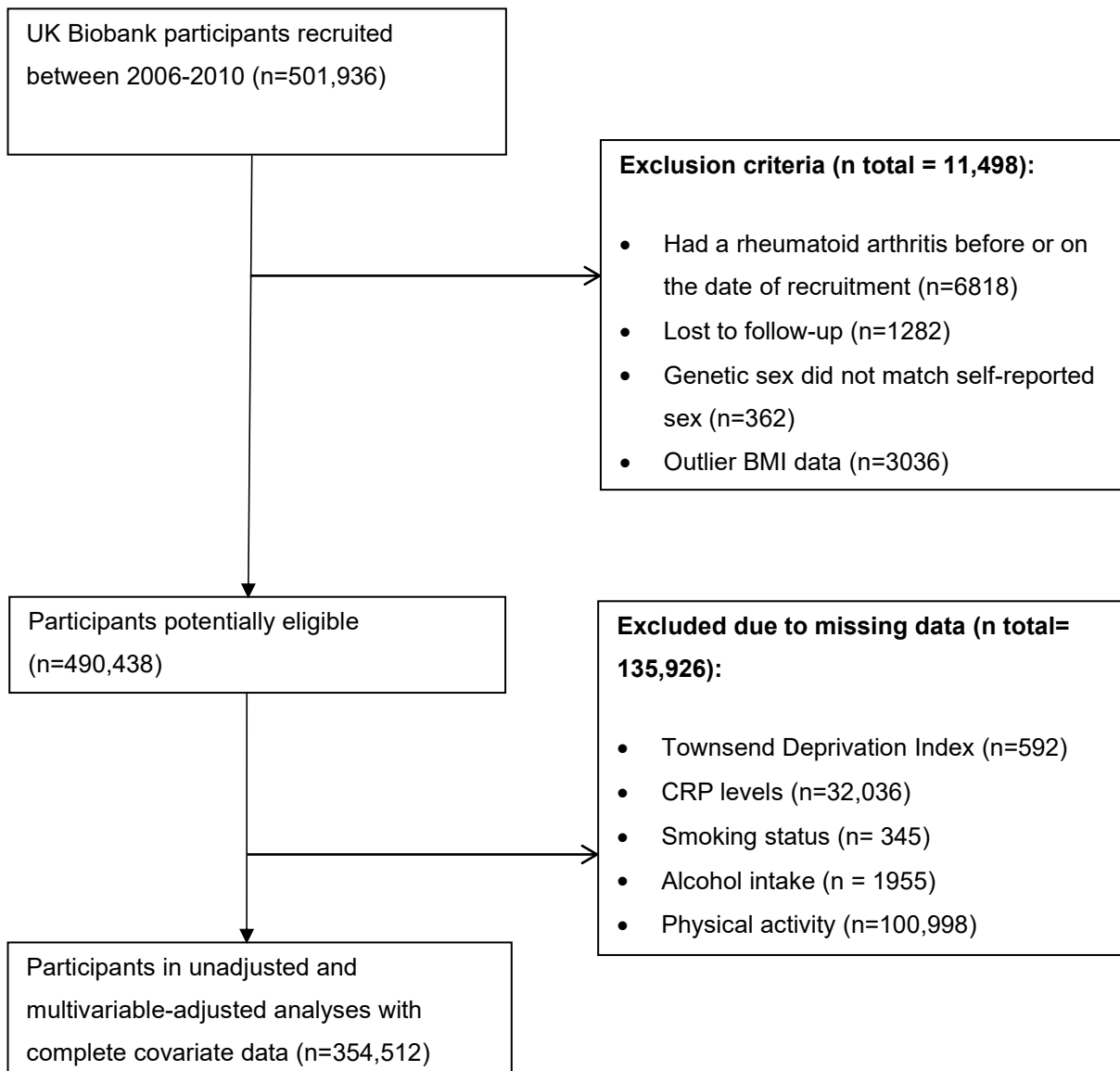


Figure E1. Flow chart of UK Biobank participants.

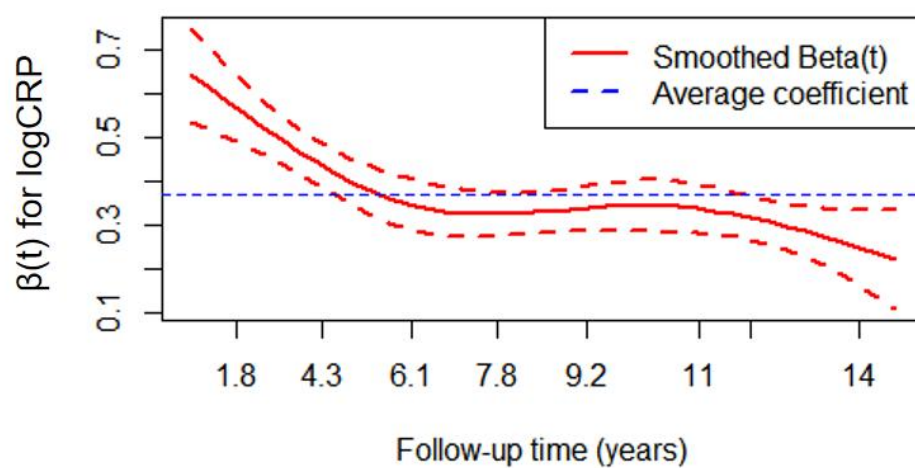


Figure E2. Time-varying coefficient ($\beta(t)$) of log-transformed C-reactive protein (logCRP) in the Cox model.

Table E1. Strengthening the reporting of observational studies in nutritional epidemiology (STROBE-Nut) checklist.

Topic	Item number	Recommendation	Page (line number)
Title and abstract	1	Indicate the study's design with a commonly used term in the title or the abstract	1
		Provide in the abstract an informative and balanced summary of what was done and what was found	1
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	1, 2
Objectives	3	State specific objectives, including any prespecified hypotheses	2
Methods			
Study design	4	Present key elements of study design early in the manuscript	3, Table 1
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	3-5, Figure S1
Participants	6	Cohort study - give the eligibility criteria, and the sources and methods of selection of participants; describe methods of follow-up	3-6, Table S1.

		Cohort study - for matched studies, give matching criteria and number of exposed and unexposed Case-control study - for matched studies, give matching criteria and the number of controls per case	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers; give diagnostic criteria, if applicable	3-5; Table 1
Data sources/measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement); describe comparability of assessment methods if there is more than one group	3-5; Table S2-S3
Bias	9	Describe any efforts to address potential sources of bias	7-8
Study size	10	Explain how the study size was arrived at	8; Figure S1
Quantitative variables	11	Explain how quantitative variables were handled in the analyses; if applicable, describe which groupings were chosen and why	Table S2
Statistical methods	12	Describe all statistical methods, including those used to control for confounding	6-8
		Describe any methods used to examine subgroups and interactions	7
		Explain how missing data were addressed	Figure S1
		Cohort study - if applicable, explain how loss to follow-up was addressed	N/A

		Describe any sensitivity analyses	8
Results			
Participants	13*	Report numbers of individuals at each stage of study - e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analyzed	8-10; Table 1
		Give reasons for nonparticipation at each stage Consider use of a flow diagram	Figure S1
Descriptive data	14*	Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders	Table 1
		Indicate number of participants with missing data for each variable of interest Cohort study - summarize follow-up time (e.g., average and total amount)	8-9
Outcome data	15*	Cohort study - report numbers of outcome events or summary measures over time	8-9; Table 1, Table S5
Main results	16	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval); make clear which confounders were adjusted for and why they were included Report category boundaries when continuous variables were categorized	Figure 1, Table 2
		If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A

Other analyses	17	Report other analyses done - e.g., analyses of subgroups and interactions, and sensitivity analyses	Table 2, Figure 2, Table S6
Discussion			
Key results	18	Summarize key results with reference to study objectives	15-16
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision; discuss both direction and magnitude of any potential bias	15-16
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	16-18
Generalizability	21	Discuss the generalizability (external validity) of the study results	18-19
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	20

Table E2. Diet group categorisation and definitions.

Food Component	Healthy level (score = 1	Original UKB field(s)	Notes on derivation
Fruit	≥3 pieces/day	p1309_i0 (fresh/dried fruit, pieces/day) p1319_i0 (tinned fruit, pieces/week)	Weekly intake divided by 5 to convert to daily; sum of both sources.
Vegetables	≥3 tablespoons/day	p1289_i0 (cooked vegetable, tablespoons/day) p1299_i0 (raw vegetable/salad, tablespoons/day)	Sum of both sources; each divided by 3 to convert tablespoons to standard servings.
Whole grain	≥1 serving/day	p1448_i0 (type of bread) p1438_i0 (bread slices/week) p1468_i0 (type of breakfast cereal) p1458_i0 (cereal bowls/week)	Wholemeal/wholegrain bread or wholegrain cereal (Muesli, oat, bran, biscuit). Weekly intake divided by 7 to daily.
Fish	Oily fish ≥1 time/week	p1329_i0 (oily fish frequency)	Frequency categories mapped to weekly times.
Dairy	Both cheese healthy AND milk healthy	p1408_i0 (cheese frequency) p1418_i0 (type of milk)	Cheese healthy: ≥1 time/week; milk healthy: semi- skimmed, skimmed, or soya.
Oil	Healthy spread AND ≥2 slices bread/day	p1428_i0, p2654_i0 (type of spread) p1438_i0 (bread slices/week)	Healthy spreads: polyunsaturated/olive oil-based, low-fat, or plant stanol-enriched.
Refined grains	≤2 servings/day	p1448_i0, p1438_i0 (white/other bread) p1468_i0, p1458_i0 (non-wholegrain cereals)	White/other bread and non- wholegrain breakfast cereals. Weekly intake divided by 7

			to daily.
Processed meat	<2 times/week	p1349_i0 (processed meat frequency)	Frequency mapped to weekly times.
Unprocessed meat	≤2 meat types with intake ≥2 times/week	p1359_i0, p1369_i0, p1379_i0, p1389_i0 (poultry, beef, lamb, pork frequencies) p3680_i0 (total unprocessed meat, g/week)	Count of meat types each consumed ≥2 times/week; healthy if count ≤2. Also considered zero if total intake = 0.
Sugar- sweetened beverages	<1 serving/week	p6144_i0 (sugary drink intake, times/week)	Healthy if weekly intake <1.

Table E3. Healthy lifestyle Score definition.

Health behaviour	How scored
Smoking	Non-smoker (never smoked) = 1 point
Alcohol intake	Healthy drinking defined as weekly alcohol units ≤ 2 for men, ≤ 1 for women = 1 point
Diet	Healthy diet score: based on 10 food components (fruits, vegetables, whole grains, refined grains, dairy, fish, oil, processed meat, unprocessed meat, sugary drinks). Each component meeting recommended intake = 1 point. Sum ≥ 7 points = diet healthy = 1 point
Physical activity	Sufficient physical activity: ≥ 600 MET-min/week (or equivalent, e.g., at least 150 min moderate or 75 min vigorous activity per week) = 1 point

Table E4. Summary of mediation analyses using the product-of-coefficients method with bootstrap resampling (UK Biobank).

Step	Description	Model
1. a path	Linear regression: lifestyle score → logCRP; Adjusted for same covariates as main model.	Linear regression, adjusted for covariates
2. b path	Cox regression: logCRP → RA (adjusted for lifestyle score); adjusted for same covariates.	Cox proportional hazards regression, adjusted for covariates
3. Total effect (c)	Cox regression: lifestyle score → RA (without mediator).	Cox proportional hazards regression, adjusted for covariates
4. Direct effect (c')	Cox regression: lifestyle score + logCRP → RA.	Cox proportional hazards regression, adjusted for covariates
5. Indirect effect (a × b)	$a \times b$	Product of coefficients
6. Proportion mediated	$(a \times b) / c \times 100\%$	Percentage of total effect explained.

Table E5. Characteristics of UK Biobank participants by lifestyle score group at recruitment, stratified by sex.

Characteristics, mean (SD), or n (%)	Male			Female		
	High score group	Medium score group	Low score group	High score group	Medium score group	Low score group
Participants (%)	45,620 (12.9)	62,147 (17.5)	61,473 (17.4)	94,091 (26.5)	56,258 (15.9)	30,254 (8.5)
Cases (%)	380 (8.1)	643 (13.7)	728 (15.5)	1414 (30.1)	963 (20.5)	541 (11.5)
Socio-demographics						
Age, years (SD)	55.1(9.0)	56.5 (8.9)	58.7 (8.1)	52.6 (9.4)	56.5 (8.9)	57.7 (8.0)
White ethnicity (%)	42,735 (93)	59,690 (95)	60,039 (97)	89,806 (94)	54,565 (95)	29,971 (97)
Region (%)						
England	40,815 (88.7)	55,648 (88.6)	54,952 (88.3)	84,554 (88.5)	50,562 (88.4)	27,141 (88.1)
Scotland	3219 (7.0)	4467 (7.1)	4576 (7.4)	7047 (7.4)	4281 (7.5)	2374 (7.7)
Wales	1966 (4.3)	2675 (4.3)	2673 (4.3)	3904 (4.1)	2378 (4.2)	1280 (4.2)
Household income (%)						
High (≥£52,000)	12,523 (27.2)	17,900 (28.5)	17,691 (28.4)	20,284 (14.4)	13,108 (22.9)	7671 (24.9)
Medium (£18,000-£51,999)	21,964 (47.7)	29,602 (47.1)	28,481 (45.8)	42,803 (44.8)	26,160 (45.7)	14,093 (45.8)
Low (<£18,000)	7558 (16.4)	10,062 (16.0)	11,202 (18.0)	18,659 (19.5)	10,797 (18.9)	5772 (18.7)
Occupation (%)						

Employed	29,138 (63.3)	39,019 (62.1)	37,517 (60.3)	52,109 (54.6)	34,177 (59.7)	19,457 (63.2)
Retired	13,856 (30.1)	19,322 (30.8)	18,954 (30.5)	34,470 (36.1)	17,271 (30.2)	8231 (26.7)
Unemployed or disabled	2128 (4.6)	3467 (5.5)	4841 (7.8)	2725 (2.9)	2403 (4.0)	1570 (5.1)
Family/student/volunteer	488 (1.1)	570 (0.9)	475 (0.8)	5496 (5.8)	2958 (5.2)	1357 (4.4)
Education (%)						
College or above	25,400 (55)	32,566 (52)	29,498 (47)	49,767 (52)	27,921 (49)	14586 (47)
High school or equivalent	14,619 (32)	21,160 (34)	22,250 (36)	32,850 (34)	21,185 (37)	11852 (39)
Less than high school	5981 (13)	9064 (14)	10453 (17)	12,888 (14)	8115 (14)	4357 (14)
Townsend deprivation index (SD)	-1.6 (3.0)	-1.5 (3.0)	-1.3 (3.1)	-1.7 (2.9)	-1.3 (3.0)	-1.1 (3.1)
Lifestyle	27.3 (4.0)	25.8 (3.6)	25.8 (3.9)	27.5 (5.2)	26.4 (4.8)	25.0 (4.5)
Physical activity, MET.mins/week (SD)	2452 (2305)	1928 (2223)	1041 (1862)	1971 (1925)	1204 (1628)	633 (1209)
Smoking status (%)						
Current	1138 (2.5)	5868 (9.4)	12,894 (20.7)	2598 (2.7)	6011 (10.5)	6896 (22.4)
Former	4780 (10.4)	24,613 (39.2)	37,437 (60.2)	13,168 (13.8)	23,134 (40.4)	18,904 (61.4)
Never	39,956 (86.9)	31,748 (50.6)	11,084 (17.8)	79,101 (82.8)	27,019 (47.2)	4268 (13.9)
Alcohol consumption, units/week (SD)	1.5 (2.1)	2.9 (2.8)	4.1 (3.4)	0.9 (1.2)	1.8 (2.1)	3.1 (2.6)

BMI, kg/m ² (SD)	27.2 (4.1)	27.7 (4.1)	28.2 (4.3)	26.5 (4.9)	27.1 (5.2)	27.2 (5.2)
Comorbidities						
History of CVD (%)	2909 (6.3)	4694 (7.5)	5381 (8.7)	2374 (2.5)	1579 (2.8)	900 (2.9)
History of diabetes (%)	822 (1.8)	1048 (1.7)	985 (1.6)	935 (1.0)	554 (1.0)	233 (0.8)
Female-specific covariates						
Menopausal status (%)						
Premenopausal	N/A	N/A	N/A	58,249 (61.0)	33,086 (57.8)	17,503 (56.8)
Postmenopausal	N/A	N/A	N/A	23,281 (24.4)	15,361 (26.8)	8321 (27.0)
HRT use (%)						
Yes	N/A	N/A	N/A	34,352 (36.0)	20,991 (36.7)	11,805 (38.3)
No	N/A	N/A	N/A	61,153 (64.0)	36,230 (63.3)	18,990 (61.7)

BMI, body mass index; CRP, C-reaction protein; HRT, Hormone Replacement Therapy; MET, Metabolic equivalents; CVD, cardiovascular disease; SD, standard deviation; <£18 000, £18 000-£51 999, and ≥£52 000 of average total household income before tax represented the high, medium, and low family income level, respectively.

Table E6. Risk of rheumatoid arthritis by lifestyle score group in the UK Biobank with varying restrictions.

Lifestyle score	HR (95% CI)
Adjusted model	
Low Lifescore(reference)	1
Medium Lifescore	0.92(0.86,0.99)
High Lifescore	0.81(0.75,0.88)
Excluding participants with < 3 years of follow-up	
Low Lifescore(reference)	1
Medium Lifescore	0.95 (0.88, 1.03)
High Lifescore	0.83 (0.76, 0.90)
Additionally adjust for the assessment centre	
Low Lifescore(reference)	1
Medium Lifescore	0.92(0.85,0.99)
High Lifescore	0.81(0.75,0.87)
Accounting for death as a competing risk	
Low Lifescore(reference)	1
Medium Lifescore	0.93(0.86,1.00)
High Lifescore	0.82(0.75,0.89)

The multivariable-adjusted model was adjusted for the following covariates assessed at recruitment: age (continuous), region (England, Scotland, Wales), ethnicity (White vs. other), Townsend deprivation index (continuous), sex (male, female), body mass index (continuous), menopausal status (premenopausal, postmenopausal), hormone replacement therapy (yes, no), baseline diabetes (yes, no), and baseline cardiovascular disease (yes, no). HR (95% CI): hazard ratio (95% confidence interval).