EXTENDED REPORT

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)

Manjari Lahiri,[1,2] Robert N Luben,[3] Catharine Morgan,[1] Diane K Bunn,[1,4] Tarnya Marshall,[4] Mark Lunt,[1] Suzanne M M Verstappen,[1] Deborah P M Symmons,[1,5] Kay-Tee Khaw,[3] Nick Wareham,[3] Ian N Bruce[1,5]

Handling editor Tore K Kvien

▶ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2012-202481).

[1]Arthritis Research UK Epidemiology Unit, Manchester Academic Health Sciences Centre, University of Manchester, Manchester, UK
[2]University Medicine Cluster, National University Health System, Singapore, Singapore
[3]Department of Public Health and Primary Care, University of Cambridge School of Clinical Medicine, Cambridge, UK
[4]Department of Rheumatology, Norfolk Arthritis Register, Norfolk and Norwich University Hospital, Norwich, UK
[5]NIHR Manchester Musculoskeletal Biomedical Research Unit, Central Manchester NHS Foundation Trust, Manchester Academic Health Sciences Centre, Manchester, UK

Correspondence to
Professor Ian N Bruce, Arthritis Research UK Epidemiology Unit, Manchester Academic Health Sciences Centre, University of Manchester, Stopford Building, Oxford Road, Manchester M13 9PT, UK; ian.bruce@manchester.ac.uk

Received 9 August 2012
Revised 19 November 2012
Accepted 28 December 2012
Published Online First 16 March 2013

To cite: Lahiri M, Luben RN, Morgan C, *et al*. *Ann Rheum Dis* 2014;**73**:219–226.

ABSTRACT

Objectives To investigate the association of lifestyle factors with risk of inflammatory polyarthritis (IP) and rheumatoid arthritis (RA).

Methods The European Prospective Investigation of Cancer, Norfolk, UK (EPIC-Norfolk) gathered lifestyle data from participants aged 40–79 years from 1993 to 1997. Individuals who subsequently developed IP were identified by linkage with the Norfolk Arthritis Register. A Cox proportional hazard model was developed, and a score assigned to each risk factor to calculate the odds of developing IP.

Results 25 455 EPIC participants were followed for a median (IQR) of 14.2 (12.9, 15.3) years; 184 developed incident IP (138 cumulatively fulfilled criteria for RA; 107 were seropositive). Pack-years of smoking were associated with increased risk of IP and RA in men (HR 1.21 (95% CI 1.08 to 1.37) per 10-pack-years) and seropositive IP (HR 1.24 (95% CI 1.10 to 1.41)) for all. Diabetes mellitus was associated with increased risk of IP (HR 2.54 (95% CI 1.26 to 5.09)), while alcohol (HR 0.86 (95% CI 0.74 to 0.99) per unit/day) and higher social class (HR 0.36 (95% CI 0.15 to 0.89) for professionals vs manual workers) were associated with reduced risk. Body mass index was associated with seronegative IP (HR 2.75 (95% CI 1.39 to 5.46) for obese vs normal-weight participants). In women, parity (HR 2.81 (95% CI 1.37 to 5.76) for ≥2 vs no children) was associated with increased risk, and breast feeding (HR 0.66 (95% CI 0.46 to 0.94) for every 52 weeks of breast feeding) was inversely associated with risk. Risk factors from the model were used to generate a 'risk score'. A total of 1159 (8.4%) women had scores reflecting a >3-fold increased risk of IP over those with a score of 0.

Conclusions Several easily ascertained clinical and lifestyle factors can be used to stratify populations for risk of IP.

INTRODUCTION

Rheumatoid arthritis (RA) is a chronic autoimmune disease affecting 0.3–0.8% of the population[1–3]; however, its aetiology remains an area of intense interest. There have been major advances in our understanding of genetic risk from genome-wide association studies.[4] There has also been renewed interest in environmental factors, especially lifestyle factors that are potentially modifiable.[5] Smoking is the most consistent association and contributes up to 25% of population attributable risk of RA.[6 7] The risk appears to be dose-related, stronger in men and in carriers of the shared epitope (SE), and for anti-citrullinated peptide antibody positive (ACPA+) RA.[7–11] Some prospective studies also support a protective role of breast feeding in women.[12–15] There are data supporting an inverse association with alcohol intake[9 10 16–18] and higher education/social class[19–21] and an increased risk with obesity,[9 22 23] but not from prospective studies.[24–30] Pregnancy or parity have not been shown to be protective in recent studies,[12–15 26 31–33] although a time-varying decrease in risk from 1 to 5 years postpartum has been postulated.[34] We examined these lifestyle factors further in a prospective cohort study, the European Prospective Investigation of Cancer, Norfolk (EPIC-Norfolk), involving adults already being followed for other health outcomes. The EPIC-Norfolk Study has established that several easily ascertained lifestyle factors can predict survival in middle-aged adults.[35] We hypothesised that similar factors may also predict the onset of inflammatory polyarthritis (IP). In addition to contributing to a better understanding of the aetiology of IP, such lifestyle factors may also help to stratify subjects in the general population according to their level of risk of IP for further genetic and serological testing, with the aim of preventing the development of IP.

PATIENTS AND METHODS

The study population

The EPIC-Norfolk Study is a population-based, prospective cohort study based in Norfolk, UK. From 1993 to 1997, individuals (99.5% Caucasian) aged 40–79 years from 35 general practice registers were invited to participate. A total of 25 639men and women (33% of those invited) were recruited and provided complete information, details of which are provided elsewhere.[36] Briefly, all participants completed a self-administered questionnaire

regarding demographic, health and lifestyle factors such as occupation, education, current and lifetime smoking, and physical exercise (see online supplementary material). Hormonal exposures in women, including age of menarche and menopause (if applicable), parity, breast feeding and use of hormone replacement therapy and oral contraceptives, were recorded. Self-reported details of physician-confirmed medical conditions, including depression, cardiovascular disease (CVD), diabetes mellitus (DM) (type 1/2 not specified) and hypertension, were also ascertained. Alcohol consumption (units/week, 1 unit=8 g of alcohol) was derived from a semiquantitative food frequency questionnaire.[37] Social class was defined according to the Registrar General's occupation-based classification scheme.[38] A validated physical activity index was derived from two questions on past-year work and recreational activities.[39] All participants underwent a clinical examination, including measurement of weight, height, blood pressure, and waist and hip circumference, and provided a blood sample. The vital status of all EPIC-Norfolk participants for this study was ascertained up to 31 March 2010 through linkage with the UK Office for National Statistics. No follow-up questionnaires were used in this analysis.

Table 1 Baseline characteristics of inflammatory polyarthritis (IP) cases versus non-cases by gender

	Men		Women	
Characteristic	**IP cases (N=56)**	**Non-cases (N=11499)**	**IP cases (N=128)**	**Non-cases (N=13772)**
Age at enrolment (years)	62.6 (55.0, 69.9)	59.5 (51.2, 67.3)	58.3 (51.0, 63.4)	58.3 (50.5, 66.5)
Socioeconomic status				
Professional	1 (1.9)	872 (7.7)	4 (3.1)	871 (6.5)
Manager/technical/skilled non-manual worker	28 (52.8)	5727 (50.7)	58 (45.7)	7384 (55.0)
Manual worker	24 (45.3)	4698 (41.6)	65 (51.2)	5158 (38.5)
Education				
Degree	7 (12.5)	1763 (15.3)	1 (0.8)	1497 (10.9)
No degree	49 (87.5)	9727 (84.7)	127 (99.2)	12266 (89.1)
Smoking status				
Current	11 (19.6)	1381 (12.1)	20 (15.9)	1538 (11.3)
Former	32 (57.1)	6224 (54.5)	37 (29.4)	4391 (32.2)
Non-smoker	13 (23.2)	3813 (33.4)	69 (54.8)	7709 (56.5)
Among smokers, pack-years of smoking	26.6 (10.1, 50.0)	19 (9.5, 32)	9.3 (5.0, 23.1)	11.5 (4.3, 22.0)
Consumption of some alcohol	52 (92.9)	10443 (91.4)	109 (85.8)	11586 (84.8)
Among drinkers, units of alcohol/ week	7.5 (2.5, 12.3)	7.5 (2.5, 15.5)	2.5 (1.0, 6.0)	3.0 (1.0, 7.5)
Self-reported DM	3 (5.4)	359 (3.1)	8 (6.3)	209 (1.5)
Self-reported hypertension	6 (10.7)	1634 (14.2)	18 (14.1)	1957 (14.2)
Self-reported baseline CVD	4 (7.1)	1258 (10.9)	6 (4.7)	654 (4.8)
Depression, ever	3 (9.4)	914 (11.7)	15 (14.8)	1880 (19.2)
Physical activity				
Inactive	12 (21.4)	3557 (30.9)	36 (28.1)	4197 (30.5)
Moderately inactive	15 (26.8)	2829 (24.6)	45 (35.2)	4402 (32.0)
Moderately active	18 (32.1)	2632 (22.9)	27 (21.1)	3066 (22.3)
Active	11 (19.6)	2480 (21.6)	20 (15.6)	2107 (15.3)
BMI	26.1 (24.5, 28.3)	26.2 (24.3, 28.4)	26.2 (24.1, 30.1)	25.5 (23.2, 28.5)
Weight category				
Normal/underweight (BMI<25)	19 (33.9)	3817 (33.3)	47 (36.7)	6120 (44.5)
Overweight (BMI 25 to <30)	31 (55.4)	6122 (53.4)	48 (37.5)	5309 (38.6)
Obese (BMI≥30)	6 (10.7)	1534 (13.4)	33 (25.8)	2315 (16.8)
Birth weight (kg)	3.2 (2.9, 3.6)	3.5 (3.2, 3.9)	3.2 (2.7, 3.4)	3.3 (2.8, 3.7)
Age of menarche (years)			13 (11, 14)	13 (12, 14)
Postmenopausal			111 (86.7)	11701 (85.0)
Age of menopause (years)*			50 (50, 50)	50 (50, 50)
Parity				
0			9 (7.0)	1936 (14.1)
1			12 (9.4)	1980 (14.4)
≥2			107 (83.6)	9841 (71.5)
Breast feeding†			81 (68.1)	8912 (75.4)
Duration of breast feeding (weeks)†			5 (0, 28)	10 (1, 36)
Ever use of OC			59 (46.1)	6344 (47.2)
Ever use of HRT*			42 (38.5)	4307 (37.8)

Values are number (%) or median (IQR).
*Postmenopausal women only.
†Parous women only.
BMI, body mass index; CVD, cardiovascular disease; DM, diabetes mellitus; HRT, hormone replacement therapy; OC, oral contraceptive.

Ascertainment of cases of IP

New cases of IP in the EPIC population were ascertained by linkage with the Norfolk Arthritis Register (NOAR), which covers all the EPIC-Norfolk general practices. Linkage was undertaken in August 2010, and IP cases with symptom onset on or before 31 March 2010 were included. Details of NOAR have also been published elsewhere.[40] Briefly, all patients presenting to a general practitioner with IP, defined as inflammation of two or more peripheral joints persisting for at least 4 weeks and onset after 1989, were notified to NOAR. All patients were interviewed and examined by a research nurse to confirm the diagnosis of IP and ascertain fulfilment of the American College of Rheumatology (ACR) 1987 criteria for RA.[41] Subjects who fulfilled the above criteria and who were not subsequently given an alternative diagnosis (other than RA, psoriatic or post-viral arthritis) by a rheumatologist were followed. All patients had a baseline blood sample drawn for rheumatoid factor (RF) and ACPA analysis. RF was measured using a particle-enhanced immunoturbidimetric assay where ≥40 IU/ml was considered positive (Orion-Diagnostica). ACPAs were measured using the Axis-Shield CCP2 antigen-plate DIASTAT kit (Axis-Shield, Dundee, UK) where >5 U/ml was considered positive. Ethics approval for both studies was obtained from the Norwich Research Ethics Committee, and all participants gave informed consent.

Outcome measures

The primary outcome measure was development of incident IP. Secondary outcome measures were development of RA and seropositive (RF+ and/or ACPA+) IP. Cases were followed annually in the NOAR study, and RA was ascertained cumulatively by applying 1987 ACR criteria at every visit for up to 5 years.[42] It has recently been shown in NOAR that case ascertainment from applying 2010 criteria at baseline is the same as from cumulatively applied 1987 criteria.[43]

Statistical analysis

Analyses were conducted in Stata V.10. Risk was estimated using a Cox proportional hazards model with the robust option to address heterogeneity of variance. Prevalent cases of IP were excluded. Cases were followed until IP symptom onset; all other participants were censored at time of death, loss to follow-up or 31 March 2010, whichever came first. Time-varying differences were assessed through the proportional hazards test and applied to gender. First, possible risk factors were tested univariately, with adjustment for age and gender. Gender differences were tested through inclusion of interaction terms, and retained only for pack-years of smoking. Pack-years of smoking (every 10 pack-years, adjusted for never being a smoker), units of alcohol (units consumed/day, adjusted for being a teetotaller) and number of years of breast feeding (for women) were modelled as continuous variables. DM, body mass index (BMI) (kg/m^2)

Table 2 Predictors of inflammatory polyarthritis

	Model 1 (All, n=184/25271)				Model 2 (Women, n=128/13772)			
	Age and gender adjusted		Multivariable*		Age adjusted		Multivariable*	
Risk factor	HR (95% CI)	p Value	HR (95% CI)	p Value	HR (95% CI)	p Value	HR (95% CI)	Value
Pack-years of smoking (every 10 pack-years)								
Men	1.23 (1.10 to 1.37)	<0.001	1.21 (1.08 to 1.37)	0.002				
Women	1.00 (0.83 to 1.20)	0.99	1.00 (0.82 to 1.21)	0.98	0.99 (0.81 to 1.20)	0.90	Not included	
Smoking status								
Never smokers	1 (ref)		Not included		1 (ref)		1 (ref)	
Men, current smokers	4.81(1.75 to 13.25)	0.01						
Men, ex-smokers	2.70 (1.12 to 6.49)							
Women, current smokers	1.49 (0.91 to 2.45)	0.22			1.49 (0.91 to 2.45)	0.22	1.57 (0.95 to 2.60)	0.14
Women, ex-smokers	0.96 (0.64 to 1.43)				0.96 (0.64 to 1.43)		0.93 (0.61 to 1.40)	
Alcohol (every 7 units/ week)	0.84 (0.73 to 0.98)	0.02	0.86 (0.74 to 0.99)	0.04	0.76 (0.58 to 0.99)	0.04	0.80 (0.62 to 1.02)	0.07
BMI								
Normal or underweight (BMI<25)	1 (ref)		1 (ref)		1 (ref)		1 (ref)	
Overweight (BMI 25 to < 30)	1.11 (0.79 to 1.56)		1.08 (0.76 to 1.54)		1.17 (0.77 to 1.76)		1.06 (0.70 to 1.62)	
Obese (BMI>30)	1.52 (1.01 to 2.27)	0.12	1.45 (0.95 to 2.21)	0.20	1.84 (1.17 to 2.91)	0.03	1.61 (1.00 to 2.58)	0.11
Social class								
Manual worker	1 (ref)		1 (ref)		1 (ref)		1 (ref)	
Technical/skilled/managerial	0.71 (0.53 to 0.95)		0.72 (0.53 to 0.99)		0.63 (0.44 to 0.89)		0.67 (0.46 to 0.97)	
Professional	0.32 (0.13 to 0.80)	0.01	0.36 (0.15 to 0.89)	0.02	0.37 (0.13 to 1.01)	0.01	0.44 (0.16 to 1.22)	0.05
Education								
No degree	1 (ref)		Not included		1 (ref)		Not included	
Degree	0.33 (0.16 to 0.68)	0.002			0.06 (0.01 to 0.46)	0.01		
Diabetes mellitus	3.26 (1.74 to 6.12)	<0.001	2.54 (1.26 to 5.09)	0.01	4.55 (2.18 to 9.50)	<0.001	4.28 (2.04 to 9.01)	<0.001
Parity								
None					1 (ref)		1 (ref)	
1					1.30 (0.54 to 3.08)		1.34 (0.54 to 3.32)	
≥2					2.31 (1.18 to 4.52)	0.01	2.81 (1.37 to 5.76)	0.003
Breast feeding (every 52 weeks)					0.78 (0.58 to 1.04)†	0.09	0.66 (0.46 to 0.94)†	0.02

*Adjusted for all other variables in model.
†For all women, including nulliparous.
BMI, body mass index.

(as per the WHO definition,[44] regrouped as BMI <25 (normal or underweight, referent), BMI 25 to <30 (overweight) and BMI ≥30 (obese or severely obese)), occupational class (regrouped as professionals, non-manual managerial/technical/skilled workers and manual workers), education (regrouped as 'degree or equivalent' and 'no degree') and parity (for women) were modelled as categorical variables. Variables found to have a trend to significance (p<0.25) on univariate analysis were included in the initial multivariate model. Occupational class rather than education was retained as a measure of socioeconomic status, as this has been shown to be a better discriminator of differentials in mortality in the UK population.[45]

Results from original data are shown. Results from a sensitivity analysis using multiple imputation for missing values (<5%) were similar (not shown). All results are expressed as HRs with 95% CIs.

Calculation of incidence rates and development of a 'risk score'

We calculated the cumulative 10-year incidence of IP from the number of incident cases per person-year of follow-up. We then calculated a risk score based on the β coefficients from our IP models (with negative scores for protective factors), and the odds of developing IP based on the score by logistic regression. We performed internal validation via bootstrapping (not shown), which did not materially affect the CIs of the estimates we have reported.

RESULTS

After exclusion of 180 prevalent cases, 25 455EPIC participants remained for analysis. The median (IQR) age was 58.9 (50.9, 66.9) years, 45.4% were men, and the median (IQR) duration of follow-up was 14.2 (12.9, 15.3) years. During 342 916 person-years of follow-up, 184 participants (128 (69.6%) women) developed incident IP, of which 138 cumulatively fulfilled criteria for RA; 57.6% and 35.9% were, respectively, RF and ACPA positive (60.4% seropositive). The median (IQR) time to onset of IP was 62.7 (27.8, 104.0) months, and the median (IQR) age at IP onset was 65.2 (57.6, 73.2) years. Table 1 shows the baseline characteristics of cases and the unaffected cohort. Patients with IP were more likely to be current smokers (19.6% vs 12.1% for men, 15.9% vs 11.3% for women). Among male ever smokers, patients were more likely to be heavy smokers (median (IQR) 26.6 (10.1, 50.0) vs 19.0 (9.5, 32) pack-years, p=0.03). Female smokers, on average, smoked less than male smokers (median (IQR) 9.3 (5.0, 23.1) pack-years for women with IP vs 26.6 (10.1, 50.0) for men with IP). Women who developed IP were more likely than those without IP to be obese (25.8% vs 16.8%, p=0.01), have DM (6.3% vs 1.5%, p<0.001), be of a non-manual

Table 3 Predictors of rheumatoid arthritis (ACR 1987 criteria)

	Model 1 (All, n=138/25271)				Model 2 (Women, n=102/13772)			
	Age and gender adjusted		Multivariable*		Age adjusted		Multivariable*	
Risk factor	HR (95% CI)	p Value	HR (95% CI)	p Value	HR (95% CI)	p Value	HR (95% CI)	p Value
Pack-years of smoking (every 10 pack-years)								
Men	1.30 (1.16 to 1.45)	<0.001	1.26 (1.10 to 1.44)	0.001				
Women	1.01 (0.85 to 1.19)	0.94	0.93 (0.74 to 1.17)	0.53	0.97 (0.78 to 1.22)	0.81	Not included	
Smoking status								
Never smokers	1 (ref)		Not included		1 (ref)		1 (ref)	
Men, current smokers	4.67(1.53 to 14.24)	0.03						
Men, ex-smokers	2.76 (1.06 to 7.17)							
Women, current Smokers	1.33 (0.74 to 2.39)	0.56			1.33 (0.74 to 2.39)	0.56	1.45 (0.81 to 2.60)	0.41
Women, ex-smokers	0.95 (0.61 to 1.48)				0.95 (0.61 to 1.48)		0.97 (0.62 to 1.52)	
Alcohol (every 7 units/ week)	0.82 (0.68 to 1.00)	0.05	0.86 (0.72 to 1.04)	0.11	0.67 (0.49 to 0.93)	0.02	0.75 (0.55 to 1.03)	0.07
BMI								
Normal or underweight (BMI<25)	1 (ref)		1 (ref)		1 (ref)		1 (ref)	
Overweight (BMI 25 to < 30)	1.17 (0.79 to 1.72)		1.16 (0.78 to 1.74)		1.18 (0.75 to 1.87)		1.09 (0.68 to 1.75)	
Obese (BMI>30)	1.57 (0.99 to 2.50)	0.16	1.49 (0.91 to 2.42)	0.28	1.89 (1.13 to 3.15)	0.04	1.62 (0.95 to 2.75)	0.17
Social class								
Manual worker	1 (ref)		1 (ref)		1 (ref)		1 (ref)	
Technical/skilled/managerial	0.62 (0.44 to 0.88)		0.65 (0.45 to 0.93)		0.62 (0.42 to 0.92)		0.65 (0.43 to 0.99)	
Professional	0.33 (0.12 to 0.90)	0.005	0.37 (0.14 to 1.03)	0.02	0.46 (0.17 to 1.28)	0.03	0.56 (0.20 to 1.53)	0.10
Education								
No degree	1 (ref)		Not included		1 (ref)		Not included	
Degree	0.17 (0.05 to 0.53)	0.002			0.09 (0.01 to 0.61)	0.01		
Diabetes mellitus	3.14 (1.50 to 6.60)	0.002	2.16 (0.92 to 5.07)	0.07	4.68(2.13 to 10.31)	<0.001	4.23 (1.92 to 9.33)	<0.001
Parity								
None					1 (ref)		1 (ref)	
1					0.97 (0.36 to 2.59)		1.10 (0.40 to 3.01)	
≥2					2.14 (1.05 to 4.37)	0.02	2.55 (1.19 to 5.48)	0.01
Breast feeding (every 52 weeks)					0.84 (0.62 to 1.14)†	0.27	0.71 (0.49 to 1.03)†	0.07

*Adjusted for all other variables in model.
†For all women, including nulliparous.
ACR, American College of Rheumatology; BMI, body mass index.

occupational class (48.8% vs 61.5%, p=0.003) and have at least two children (83.6% vs 71.5%, p=0.01), but breast fed for a shorter time (median (IQR) 5 (0, 28) vs 10 (1, 36) weeks, p=0.02).

The risk of developing IP, RA and seropositive IP associated with each risk factor are presented in tables 2–4, respectively. Two models were developed: one for the whole cohort, and one for women only. Smoking was associated with a dose-dependent linear 20% increase in IP risk for every 10 pack-years smoked in men (adjusted HR 1.21 (1.08 to 1.37)), but not in women (adjusted HR 1.00 (0.82 to 1.21)). A similar trend was seen for RA; however, risk of seropositive IP was increased in both genders (adjusted HR 1.24 (1.10 to 1.41)). When analysed by smoking status, female current versus non-smokers were at about a 50% increased risk of IP, RA and seropositive IP. Alcohol appeared protective for the development of IP (adjusted HR 0.86 (0.74 to 0.99)), RA and seropositive IP, with a 14% risk reduction per unit consumed per day. Higher BMI showed a trend towards association with risk of IP (adjusted HR 1.45 (0.95 to 2.21) for BMI ≥ 30 vs normal weight) and RA, but not seropositive IP (HR 1.05 (0.61 to 1.79), age- and gender-adjusted). A post hoc analysis looking at seronegative IP revealed a nearly threefold increase in risk (HR 2.75 (1.39 to 5.46) for BMI ≥ 30 vs normal weight, age- and gender-adjusted). Self-reported DM was associated with an increased risk of IP, especially in women (adjusted HR 4.28 (2.04 to 9.01)), which was independent of BMI. Higher occupational class or degree education was associated with reduced risk of IP. In women, having two or more children was associated with a doubling of risk (adjusted HR 2.81 (1.37 to 5.76) vs nulliparous), while breast feeding (adjusted HR 0.66 (0.46 to 0.94) per year of breast feeding) showed a dose-dependent inverse association with IP, RA and seropositive IP.

Calculation of incidence and development of risk score

The 10-year cumulative incidence of IP was 0.37% in men and 0.67% in women. We devised a risk score based on the β coefficients from our IP models (table 5). Continuous variables were assigned points for each unit of measurement up to a maximum which was within the range for our cohort. For men, every additional 10 pack-years of smoking (up to a maximum of 4 points for >30 pack-years), being obese and having DM were scored positively, whereas drinking up to 3 units of alcohol per day (1 point for each whole unit, up to a maximum of 3 points) and being of a higher occupational class were scored negatively. For women, 'current' smoking replaced pack-years smoked, and additional scores were assigned for having ≥2 children (positive) and every additional 6 months of breast feeding (negative, up to a maximum of 4 points for ≥2 years). The total score is the sum of all individual scores, and a higher total score implies a higher risk of IP. We derived relative estimates of risk at each score by logistic regression (figure 1, table S1). For example, the risk for

Table 4 Predictors of seropositive (RF+ or ACPA+) inflammatory polyarthritis

	Model 1 (All, n=107/25336)				Model 2 (Women, n=74/13821)			
	Age and gender adjusted		Multivariable*		Age adjusted		Multivariable*	
Risk factor	HR (95% CI)	p Value	HR (95% CI)	p Value	HR (95% CI)	p Value	HR (95% CI)	p Value
Pack-years of smoking (every 10 pack-years)	1.24 (1.11 to 1.38)	<0.001	1.24 (1.10 to 1.41)	0.001	1.17 (0.98 to 1.40)	0.08	Not included	
Smoking status								
Never smokers	1 (ref)		Not included		1 (ref)		1 (ref)	
Men, current smokers	4.64(1.53 to 14.06)	0.02						
Men, ex-smokers	2.20 (0.80 to 6.03)							
Women, current smokers	1.61 (0.87 to 3.01)	0.12			1.61 (0.86 to 3.00)		1.66 (0.88 to 3.13)	
Women, ex-smokers	0.77 (0.44 to 1.33)				0.77 (0.44 to 1.33)	0.12	0.78 (0.44 to 1.39)	0.12
Alcohol (every 7 units/ week)	0.83 (0.69 to 0.98)	0.03	0.86 (0.73 to 1.01)	0.07	0.67 (0.48 to 0.95)	0.02	0.78 (0.57 to 1.08)	0.13
BMI†								
Normal or underweight (BMI<25)	1 (ref)		Not included		1 (ref)		Not included	
Overweight (BMI 25 to < 30)	0.73 (0.48 to 1.13)				0.74 (0.43 to 1.25)			
Obese (BMI>30)	1.05 (0.61 to 1.79)	0.27			1.18 (0.64 to 2.17)	0.31		
Social class								
Manual worker	1 (ref)		1 (ref)		1 (ref)		1 (ref)	
Technical/skilled/managerial	0.63 (0.42 to 0.92)		0.61 (0.41 to 0.92)		0.48 (0.30 to 0.76)		0.47 (0.29 to 0.77)	
Professional	0.21 (0.05 to 0.87)	0.01	0.24 (0.06 to 0.97)	0.02	0.28 (0.07 to 1.15)	0.003	0.32 (0.08 to 1.35)	0.01
Education								
No degree	1 (ref)		Not included		1 (ref)		Not included	
Degree	0.36 (0.15 to 0.89)	0.03			0.11 (0.01 to 0.81)	0.03		
Diabetes mellitus	2.99 (1.26 to 7.08)	0.01	1.99 (0.69 to 5.74)	0.20	4.95(1.93 to 12.72)	0.001	5.20 (2.00 to 13.52)	0.01
Parity								
None					1 (ref)		1 (ref)	
1					1.56 (0.51 to 4.78)		1.50 (0.49 to 4.56)	
≥2					2.36 (0.96 to 5.81)	0.11	2.43 (0.99 to 6.00)	0.09
Breast feeding (every 52 weeks)					0.77 (0.53 to 1.14)‡	0.19	0.66 (0.41 to 1.05)‡	0.08

*Adjusted for all other variables in model.
†BMI was a significant predictor of seronegative inflammatory polyarthritis. Age- and gender-adjusted HR (95% CI) for being overweight=2.23 (1.22 to 4.08) and for being obese=2.75 (1.39 to 5.46) (whole cohort). For women, age-adjusted HR (95% CI) for being overweight=2.65 (1.29 to 5.46) and for being obese=3.66 (1.68 to 7.98).
‡For all women, including nulliparous.
ACPA+, anti-citrullinated peptide antibody positive; BMI, body mass index; RF+, Rheumatoid factor positive.

Table 5 Assignment of risk score

	Men (from model 1)		Women (model 2)	
Risk factor	**β coefficient**	**Score**	**β coefficient**	**Score**
Smoking (every 10 pack-years)	+0.20		0.00*,†	
1–10		+1		–
11–20		+2		–
21–30		+3		–
≥31		+4		–
Current smoker	*		+0.45	+2
Alcohol (per unit/day)	−0.16		−0.23	
1–<2		−1		−1
2–<3		−2		−2
≥3		−3		−3
Occupation				
Professional	−1.02	−4	−0.81	−3
Non-manual, non-professional	−0.32	−2	−0.41	−2
Obese (body mass index ≥30)	+0.37	+2	+0.48	+2
Diabetes mellitus	+0.93	+4	+1.45	+6
Parity ≥2			+1.03	+4
Duration of breast feeding (per year)			−0.42	
0.5–<1				−1
1–<1.5				−2
1.5–<2				−3
≥2				−4

*Not included in model.
†β coefficient is from model 1.

men with a score of 5 was approximately double the risk of men who scored 0 (area under receiver operating characteristics curve (AUC)=0.59). The model performed better in women (AUC=0.66). For example, a score of 5 points implied an approximately threefold increased risk of IP, which translates to a 10-year cumulative incidence of >2%. Scores ≥5 were seen in 1159 (8.4%) women in our cohort.

DISCUSSION

In this large population-based prospective cohort, several lifestyle factors were associated with risk of IP. These factors, which are easily ascertained in primary care, can be combined to develop a simple screening tool to identify individuals with an up to sixfold increased risk of IP compared with the population, who could then be targeted for more focused risk assessments. Some of the risk factors we found to be associated with IP are consistent with previous literature and may suggest potential pathogenic pathways. Smoking is thought to interact with the SE to increase the risk of ACPA+ RA.[8 10] Previous literature has noted a weaker association in women[11] and a threshold dose of 10–20 pack-years before the increased risk is apparent.[6] Most women in our cohort had smoked less than this, which explains the lack of association between smoking dose and IP risk in women in our cohort. Alcohol has previously been shown to be protective in case–control studies[9 10 16 17] and one prospective study of women.[18] We have shown that the negative association also holds true for men. Biologically, this may be the result of antioxidant/anti-inflammatory mechanisms including reduction of postprandial oxidative stress,[46] increased urate production, antioxidant properties from polyphenolic flavonoids,[47] or downregulation of the immune response.[48] Others have reported that alcohol is particularly protective in smokers,[17] but we found no significant interaction with smoking in our study (data not shown). The association with BMI is also of interest. Obesity has been found to be a risk factor for IP in case–control studies,[9 22 23] but not in cohort studies.[25 26 28 29] We have previously noted a time-varying risk from obesity in this cohort.[30] In this larger study, we also noted a differential risk by serotype. Obesity markedly increased the risk of seronegative but not seropositive IP, which supports previous literature.[9] The effect of obesity may be mediated through increased availability of oestrogen, or through an altered relationship between leptin and adiponectin contributing to a pro-inflammatory state akin to that noted in CVD.[49–51] Adiponectin levels are also decreased in type 2 DM, and are inversely related to levels of circulating tumour necrosis factor α. The resulting proinflammatory milieu may also explain the strong association observed between IP and DM in our study.[52]

In women, parity of ≥2 was associated with a doubling of risk for IP in our study. Most previous studies have shown either a decreased risk[34 53 54] or no association with parity.[12–15 26 31–33] The apparent contradiction may be due to our relatively older cohort. Episodes of pregnancy may initially have 'protected' these women against RA and this 'protection'

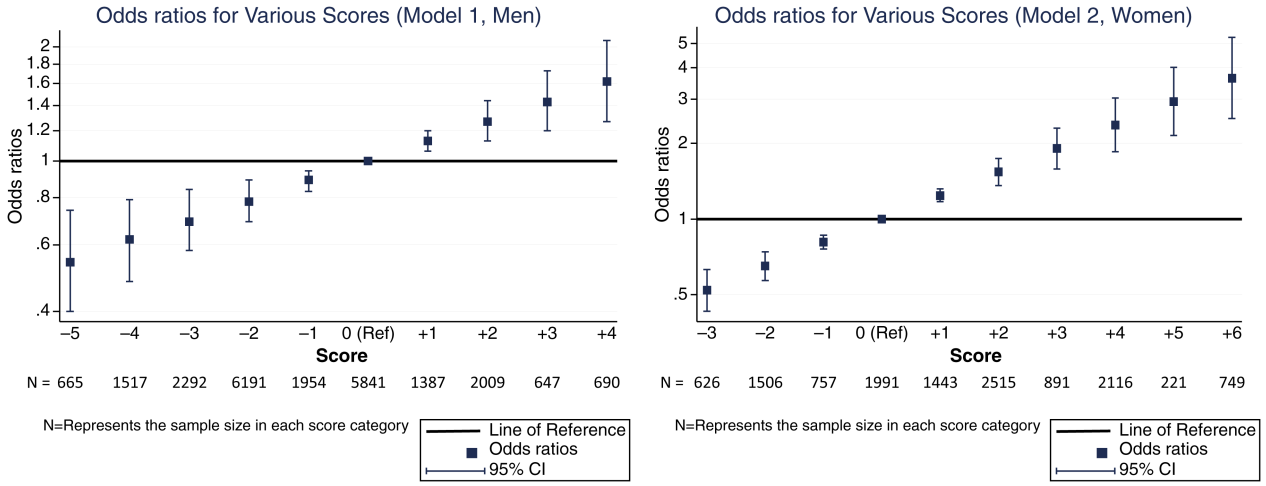

Figure 1 Odds ratios for various scores.

may have waned over time and simply postponed the onset of IP rather than truly reducing risk. This time-varying protective effect has been previously suggested.[34] Conversely, breast feeding was associated with a dose-dependent reduction in risk when adjusted for parity. This is consistent with previously published literature.[12–15] However, the negative association after what would be many years since cessation of breast feeding is difficult to explain. It may again be an example of 'depletion of susceptibles' in that a surge of (proinflammatory) prolactin during the first breastfeeding episode may have unmasked latent RA in susceptible individuals at an early age before their entry into our cohort.[55]

We have therefore confirmed that there are a number of lifestyle factors that can influence the risk of IP. We then attempted to develop a risk score for IP. In developing this, we specifically limited our model to variables easily ascertained at a routine primary care consultation using only pack-years of smoking, alcohol consumption, occupational class, BMI and presence of DM in men, and smoking status, alcohol consumption, occupational class, BMI, presence of DM, parity and duration of breast feeding in women. Some of these factors already form part of the lifestyle advice given for CVD and cancer prevention. With this model, we could identify a number of individuals who were up to six times more likely than the background population to develop IP. Although the absolute risk is small, this model, if validated, would provide a simple population/primary care screen for stratifying populations for more detailed risk assessment approaches such as serological testing or genetic screening. For example, others have noted that persons with two or more first-degree relatives with RA and positive ACPA have a high risk of developing RA over 5 years.[56] Our simple screen would complement such an approach at the population level, first by helping to target modifiable lifestyle factors such as BMI and smoking, but also allowing enhanced screening for ACPA and at-risk genotypes in a higher risk population for potential pharmacological interventions.

Our study has several strengths. Most importantly, these data were derived from a large prospective population-based cohort, and only incident cases were included, so our results are not subject to selection or recall bias. Cases of IP/RA were ascertained by examination by study personnel rather than through self-reported questionnaires or medical record review, and all cases were followed-up and ACR 1987 criteria for RA were applied cumulatively. There are, however, several limitations. Lifestyle data were self-reported and collected cross-sectionally at a single time point, and there may have been increasing misclassification later into the follow-up period. Although every attempt was made to include all incident cases of IP through general practices and speciality rheumatology clinics, there may be incomplete case ascertainment, leading to an underestimation of IP incidence. Also, we were limited to studying risk factors included in the EPIC-Norfolk questionnaire, and hence important predictors such as family history (a clinical surrogate for genetic risk) and occupational exposure (eg, silica) could not be included in our model.[57–59] Our cohort comprised late-middle-aged individuals, and IP cases of earlier onset would have been excluded. Nevertheless, the peak age of onset of IP is in the 6th and later decades,[42] so these data are generalisable to most patients with IP and are also of relevance, as they include the population in which active screening programmes (eg, for CVD) already exist. Lastly, we have based this risk score on just 184 cases, and hence the model will need to be validated in other large cohorts. In spite of including several lifestyle factors in our model, the overall attributable risk from these is likely to be small, as RA has a strong genetic contribution to its development. As expected, the AUCs for our models were marginal.

Future work should involve combining genetic and environmental data in the same cohorts to characterise gene–environment interactions in the development of IP/RA. Studying these within the paradigm of seropositive and seronegative disease, as has been previously suggested, will be crucial to advancing our understanding of the disease.

Acknowledgements NOAR is funded by Arthritis Research UK, Chesterfield, UK (Grant reference 17552); EPIC-Norfolk is funded by The European Commission 'Europe against Cancer' Programme, Cancer Research UK, Medical Research Council with additional support from the Stroke Association, British Heart Foundation, Department of Health, Food Standards Agency and the Wellcome Trust. This report includes independent research supported by the National Institute for Health Research Biomedical Research Unit Funding Scheme. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health.

Contributors All authors contributed to this paper and fulfil the criteria for authorship.

Funding NOAR is funded by Arthritis Research UK, Chesterfield, UK (Grant reference 17552); EPIC-Norfolk is funded by The European Commission 'Europe against Cancer' Programme, Cancer Research UK, Medical Research Council with additional support from the Stroke Association, British Heart Foundation, Department of Health, Food Standards Agency and the Wellcome Trust.

Competing interests None.

Ethics approval Norwich Research Ethics Committee.

Patient consent Obtained.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Alamanos Y, Voulgari PV, Drosos AA. Incidence and prevalence of rheumatoid arthritis, based on the 1987 American College of Rheumatology criteria: a systematic review. *Semin Arthritis Rheum* 2006;36:182–8.
2 Symmons D, Turner G, Webb R, *et al*. The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. *Rheumatology (Oxford)* 2002;41:793–800.
3 Roux CH, Saraux A, Le BE, *et al*. Rheumatoid arthritis and spondyloarthropathies: geographical variations in prevalence in France. *J Rheumatol* 2007;34:117–22.
4 Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447:661–78.
5 Lahiri M, Morgan C, Symmons DP, *et al*. Modifiable risk factors for RA: prevention, better than cure? *Rheumatology (Oxford)* 2012;51:499–512.
6 Costenbader KH, Feskanich D, Mandl LA, *et al*. Smoking intensity, duration, and cessation, and the risk of rheumatoid arthritis in women. *Am J Med* 2006;119:503–9.
7 Kallberg H, Ding B, Padyukov L, *et al*. Smoking is a major preventable risk factor for rheumatoid arthritis: estimations of risks after various exposures to cigarette smoke. *Ann Rheum Dis* 2011;70:508–11.
8 Klareskog L, Stolt P, Lundberg K, *et al*. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. *Arthritis Rheum* 2006;54:38–46.
9 Pedersen M, Jacobsen S, Klarlund M, *et al*. Environmental risk factors differ between rheumatoid arthritis with and without auto-antibodies against cyclic citrullinated peptides. *Arthritis Res Ther* 2006;8:R133.
10 Pedersen M, Jacobsen S, Garred P, *et al*. Strong combined gene-environment effects in anti-cyclic citrullinated peptide-positive rheumatoid arthritis: a nationwide case-control study in Denmark. *Arthritis Rheum* 2007;56:1446–53.
11 Sugiyama D, Nishimura K, Tamaki K, *et al*. Impact of smoking as a risk factor for developing rheumatoid arthritis: a meta-analysis of observational studies. *Ann Rheum Dis* 2010;69:70–81.
12 Brennan P, Bankhead C, Silman A, *et al*. Oral contraceptives and rheumatoid arthritis: results from a primary care-based incident case-control study. *Semin Arthritis Rheum* 1997;26:817–23.

13 Karlson EW, Mandl LA, Hankinson SE, *et al*. Do breast-feeding and other reproductive factors influence future risk of rheumatoid arthritis? Results from the Nurses' Health Study. *Arthritis Rheum* 2004;50:3458–67.
14 Merlino LA, Cerhan JR, Criswell LA, *et al*. Estrogen and other female reproductive risk factors are not strongly associated with the development of rheumatoid arthritis in elderly women. *Semin Arthritis Rheum* 2003;33:72–82.
15 Pikwer M, Bergstrom U, Nilsson JA, *et al*. Breast feeding, but not use of oral contraceptives, is associated with a reduced risk of rheumatoid arthritis. *Ann Rheum Dis* 2009;68:526–30.
16 Hazes JM, Dijkmans BA, Vandenbroucke JP, *et al*. Lifestyle and the risk of rheumatoid arthritis: cigarette smoking and alcohol consumption. *Ann Rheum Dis* 1990;49:980–2.
17 Kallberg H, Jacobsen S, Bengtsson C, *et al*. Alcohol consumption is associated with decreased risk of rheumatoid arthritis: results from two Scandinavian case-control studies. *Ann Rheum Dis* 2009;68:222–7.
18 Di GD, Alfredsson L, Bottai M, *et al*. Long term alcohol intake and risk of rheumatoid arthritis in women: a population based cohort study. *BMJ* 2012;345:e4230.
19 Bengtsson C, Nordmark B, Klareskog L, *et al*. Socioeconomic status and the risk of developing rheumatoid arthritis: results from the Swedish EIRA study. *Ann Rheum Dis* 2005;64:1588–94.
20 Olsson AR, Skogh T, Wingren G. Aetiological factors of importance for the development of rheumatoid arthritis. *Scand J Rheumatol* 2004;33:300–6.
21 Pedersen M, Jacobsen S, Klarlund M, *et al*. Socioeconomic status and risk of rheumatoid arthritis: a Danish case-control study. *J Rheumatol* 2006;33:1069–74.
22 Symmons DP, Bankhead CR, Harrison BJ, *et al*. Blood transfusion, smoking, and obesity as risk factors for the development of rheumatoid arthritis: results from a primary care-based incident case-control study in Norfolk, England. *Arthritis Rheum* 1997;40:1955–61.
23 Voigt LF, Koepsell TD, Nelson JL, *et al*. Smoking, obesity, alcohol consumption, and the risk of rheumatoid arthritis. *Epidemiology* 1994;5:525–32.
24 Aho K, Heliovaara M. Alcohol, androgens and arthritis. *Ann Rheum Dis* 1993;52:897.
25 Cerhan JR, Saag KG, Criswell LA, *et al*. Blood transfusion, alcohol use, and anthropometric risk factors for rheumatoid arthritis in older women. *J Rheumatol* 2002;29:246–54.
26 Hernandez AM, Liang MH, Willett WC, *et al*. Reproductive factors, smoking, and the risk for rheumatoid arthritis. *Epidemiology* 1990;1:285–91.
27 Pattison DJ, Symmons DP, Lunt M, *et al*. Dietary risk factors for the development of inflammatory polyarthritis: evidence for a role of high level of red meat consumption. *Arthritis Rheum* 2004;50:3804–12.
28 Rodriguez LA, Tolosa LB, Ruigomez A, *et al*. Rheumatoid arthritis in UK primary care: incidence and prior morbidity. *Scand J Rheumatol* 2009;38:173–7.
29 Uhlig T, Hagen KB, Kvien TK. Current tobacco smoking, formal education, and the risk of rheumatoid arthritis. *J Rheumatol* 1999;26:47–54.
30 Goodson NJ, Silman AJ, Pattison DJ, *et al*. Traditional cardiovascular risk factors measured prior to the onset of inflammatory polyarthritis. *Rheumatology (Oxford)* 2004;43:731–6.
31 Hannaford PC, Kay CR, Hirsch S. Oral contraceptives and rheumatoid arthritis: new data from the Royal College of General Practitioners' oral contraception study. *Ann Rheum Dis* 1990;49:744–6.
32 Heliovaara M, Aho K, Reunanen A, *et al*. Parity and risk of rheumatoid arthritis in Finnish women. *Br J Rheumatol* 1995;34:625–8.
33 Pope JE, Bellamy N, Stevens A. The lack of associations between rheumatoid arthritis and both nulliparity and infertility. *Semin Arthritis Rheumatism* 1999;28:342–50.
34 Guthrie KA, Dugowson CE, Voigt LF, *et al*. Does pregnancy provide vaccine-like protection against rheumatoid arthritis? *Arthritis Rheum* 2010;62:1842–8.
35 Khaw KT, Wareham N, Bingham S, *et al*. Combined impact of health behaviours and mortality in men and women: the EPIC-Norfolk prospective population study. *PLoS Med* 2008;5:e12.
36 Day N, Oakes S, Luben R, *et al*. EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. *Br J Cancer* 1999;80 (Suppl 1):95–103.
37 Bingham SA, Welch AA, McTaggart A, *et al*. Nutritional methods in the European prospective investigation of cancer in Norfolk. *Public Health Nutr* 2001;4: 847–58.
38 Shohaimi S, Luben R, Wareham N, *et al*. Residential area deprivation predicts smoking habit independently of individual educational level and occupational social class. A cross sectional study in the Norfolk cohort of the European Investigation into Cancer (EPIC-Norfolk). *J Epidemiol Community Health* 2003;57:270–6.
39 Wareham NJ, Jakes RW, Rennie KL, *et al*. Validity and repeatability of a simple index derived from the short physical activity questionnaire used in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. *Public Health Nutr* 2003;6:407–13.
40 Symmons DP, Barrett EM, Bankhead CR, *et al*. The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. *Br J Rheumatol* 1994;33:735–9.
41 Arnett FC, Edworthy SM, Bloch DA, *et al*. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315–24.
42 Symmons DP, Silman AJ. The Norfolk Arthritis Register (NOAR). *Clin Exp Rheumatol* 2003;21:S94–9.
43 Humphreys JH, Verstappen SM, Hyrich KL, *et al*. The incidence of rheumatoid arthritis in the UK: comparisons using the 2010 ACR/EULAR classification criteria and the 1987 ACR classification criteria. Results from the Norfolk Arthritis Register. *Ann Rheum Dis* 2013;72:1315–20.
44 Rosenbaum S, Skinner RK, Knight IB, *et al*. A survey of heights and weights of adults in Great Britain, 1980. *Ann Hum Biol* 1985;12:115–27.
45 Davey SG, Hart C, Hole D, *et al*. Education and occupational social class: which is the more important indicator of mortality risk? *J Epidemiol Community Health* 1998;52:153–60.
46 Puddey IB, Croft KD. Alcohol, stroke and coronary heart disease. Are there anti-oxidants and pro-oxidants in alcoholic beverages that might influence the development of atherosclerotic cardiovascular disease? *Neuroepidemiology* 1999;18:292–302.
47 Covas MI, Gambert P, Fit M, *et al*. Wine and oxidative stress: Up-to-date evidence of the effects of moderate wine consumption on oxidative damage in humans. *Atherosclerosis* 2010;208:297–304.
48 Waldschmidt TJ, Cook RT, Kovacs EJ. Alcohol and inflammation and immune responses: summary of the 2005 Alcohol and Immunology Research Interest Group (AIRIG) meeting. *Alcohol* 2006;38:121–5.
49 Longcope C, Jaffee W, Griffing G. Production rates of androgens and oestrogens in post-menopausal women. *Maturitas* 1981;3:215–23.
50 Otero M, Lago R, Gomez R, *et al*. Towards a pro-inflammatory and immunomodulatory emerging role of leptin. *Rheumatology (Oxford)* 2006;45:944–50.
51 Fantuzzi G. Adiponectin and inflammation: consensus and controversy. *J Allergy Clin Immunol* 2008;121:326–30.
52 She JX. Susceptibility to type I diabetes: HLA-DQ and DR revisited. *Immunol Today* 1996;17:323–9.
53 Hazes JM, Dijkmans BA, Vandenbroucke JP, *et al*. Pregnancy and the risk of developing rheumatoid arthritis. *Arthritis Rheum* 1990;33:1770–5.
54 Spector TD, Roman E, Silman AJ. The pill, parity, and rheumatoid arthritis. *Arthritis Rheum* 1990;33:782–9.
55 Brennan P, Silman A. Breast-feeding and the onset of rheumatoid arthritis. *Arthritis & Rheumatism* 1994;37:808–13.
56 Nielen MM, van SD, Reesink HW, *et al*. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. *Arthritis Rheum* 2004;50:380–6.
57 Khuder SA, Peshimam AZ, Agraharam S. Environmental risk factors for rheumatoid arthritis. *Rev Environ Health* 2002;17:307–15.
58 Stolt P, Yahya A, Bengtsson C, *et al*. Silica exposure among male current smokers is associated with a high risk of developing ACPA-positive rheumatoid arthritis. *Ann Rheum Dis* 2010;69:1072–6.
59 Stolt P, Kallberg H, Lundberg I, *et al*. Silica exposure is associated with increased risk of developing rheumatoid arthritis: results from the Swedish EIRA study. *Ann Rheum Dis* 2005;64:582–6.