EXTENDED REPORT

Smoking interacts with genetic risk factors in the development of rheumatoid arthritis among older Caucasian women

L A Criswell, K G Saag, T R Mikuls, J R Cerhan, L A Merlino, R F Lum, K A Pfeiffer, B Woehl, M F Seldin

Ann Rheum Dis 2006;**65**:1163–1167. doi: 10.1136/ard.2005.049676

See end of article for authors' affiliations

Correspondence to: Professor L A Criswell, Division of Rheumatology, University of California, San Francisco, 374 Parnassus Avenue, Box 0500, San Francisco, CA 94143-0500, USA; Lindsey.Criswell@ucsf.edu

Accepted 17 June 2006

Objective: To determine whether the impact of tobacco exposure on rheumatoid arthritis (RA) risk is influenced by polymorphisms at the *HLA-DRB1* and *glutathione S-transferase M1* (*GSTM1*) loci.
Methods: Subjects were participants from a case-control study nested within the Iowa Women's Health Study, a population based, prospective cohort study of postmenopausal women. Incident RA cases (n = 115) were identified and medical records reviewed to confirm RA diagnosis. Controls without RA (n = 466) were matched with RA cases by age and ethnic background. *HLA-DRB1* typing classified subjects according to the presence of alleles encoding the RA ''shared epitope'' (SE) sequence. *GSTM1* was genotyped using a multiplex polymerase chain reaction assay. Conditional logistic regression was used to estimate the odds ratios (ORs) and 95% confidence intervals.
Results: Strong positive associations of smoking (OR = 6.0, p = 0.004), SE positivity (OR = 4.6, p = 0.0006), and *GSTM1* null genotype (OR = 3.4, p = 0.007) with risk of RA, and significant gene-environment interactions (smoking by SE interaction p = 0.034; smoking by *GSTM1* interaction p = 0.047) were observed. Stratified analyses indicated that exposure to tobacco smoke primarily increased the risk of RA among subjects who lacked genetic risk factors for the disease (that is, SE negative or *GSTM1* present).
Conclusions: Although these findings require confirmation in other groups, the results support the importance of considering both genetic and environmental factors, and also their interaction, in studies of complex diseases like RA.

Like most common diseases, the aetiology of rheumatoid arthritis (RA) is multifactorial, with important contributions from multiple genetic and environmental factors.[1] Although recent genome wide screens in RA have substantially increased our knowledge about the multiple genomic regions that contribute to RA susceptibility,[2–5] the *HLA-DRB1* locus in the major histocompatibility complex, and in particular DRB1 alleles encoding a common ''shared'' epitope sequence in the third hypervariable region of the DRB1 chain, is clearly the single most important genetic risk factor for RA.[6 7]

The importance of environmental or other non-genetic risk factors for RA is underlined by the large discordant rate for RA among monozygotic twins, which in most studies is greater than 70%.[8 9] Although the elucidation of specific environmental risk factors for RA has been a major challenge, compelling evidence now supports an aetiological role for exposure to tobacco smoke. Several large and well designed studies document significant associations with disease risk, particularly among men and rheumatoid factor (RF) positive subjects.[10–13] Of interest, RF positivity often precedes the development of RA,[14] and tobacco smoke exposure is associated with RF positivity among subjects without RA.[15 16]

Research related to other diseases associated with tobacco smoke exposure, particularly cancer, indicates that people are not equally susceptible to the deleterious effects of tobacco smoke. For example, genetic polymorphisms of glutathione *S*-transferases (GSTs), which are believed to have an important role in detoxifying carcinogens in tobacco smoke, influence susceptibility to lung and colon cancer.[17 18] Thus, these studies document interactions between genetic and environmental risk factors, consistent with expectations for complex disease. Of note, recent work by Mattey *et al* demonstrated an interaction between exposure to tobacco smoke and the *GSTM1* null polymorphism in determining RA outcome.[19]

Our goal for the current study was to examine the interaction of tobacco smoke exposure with specific genetic risk factors, including the *GSTM1* null polymorphism and the *HLA-DRB1* shared epitope (SE), in determining RA risk. We chose as our study group a unique cohort of older women with well characterised RA, in whom information about exposure to tobacco smoke and other risk factors had been collected before disease onset.

PATIENTS AND METHODS

Study group

RA cases and controls were participants in the Iowa Women's Health Study (IWHS), which is a prospective cohort study established in 1986.[20–22] The study was approved by the Institutional Review Board at the University of Iowa, Iowa City, and subjects provided informed consent. Details of the identification and validation of incident RA cases in this cohort have been published previously.[13 23–25] In brief, potential incident RA cases were initially identified by self report in follow up surveys in this cohort, and confirmation of RA diagnosis was established through review of medical records and documentation of American College of

Abbreviations: GST, glutathione *S*-transferase; IWHS, Iowa Women's Health Study; OR, odds ratio; RA, rheumatoid arthritis; RF, rheumatoid factor; SE, shared epitope

Rheumatology criteria or a diagnosis from a board certified/eligible rheumatologist.[26 27] A total of 158 incident RA cases were identified in this cohort. Buccal DNA samples were obtained from 123 incident RA cases. The other 35 incident RA cases were not sampled owing to death (n = 21), refusal (n = 11) or loss to follow up (n = 3). For one of the sampled cases smoking information was missing and seven of the sampled cases were not successfully genotyped; thus the total number of incident RA cases studied was 115.

Controls were randomly selected from IWHS participants who reported no history of RA, and were matched with RA cases by age and ethnic background—namely, Dutch, English/Scottish/Welsh, German, Irish, Swedish, Norwegian, Southern European (Italian, Greek, etc), Eastern European (Polish, Russian, Czech, etc), other Central or Western European, Asian, African American, American Indian, and other. We initially identified up to eight potential controls who matched each of the 158 RA cases, and successfully enrolled a total of 524 controls, who completed a questionnaire and provided a buccal DNA sample. Fifty eight of these controls were excluded owing to missing smoking data (n = 6) or missing genotypes (n = 52), leaving a total of 466 controls available for the analysis (a range of 1–6 controls for each case).

Exposure to tobacco smoke

Detailed information about exposure to tobacco smoke was obtained for all cases and controls as part of the baseline IWHS survey in 1986. Smoking data included age started, age stopped, and number of cigarettes smoked per day. Respondents who reported smoking on the baseline questionnaire were classified as current smokers. Data collected about smoking status in follow up surveys (1992 and 1997) indicated that smoking status was remarkably stable over time (data not shown).

HLA-DRB1 genotyping

The assignment of *HLA-DRB1* alleles was determined using DNA sequencing of exon 2. Briefly, all *DRB1* exon 2 sequences were amplified and sequenced using the AlleleSEQR HLA-DRB1 reagent kit and protocol (Forensic Analytical, Hayward, CA, USA). Sequencing was performed using an ABI 3700 (PE Applied Biosystems, Foster City, CA, USA). The sequences were analysed using the Sequencing Analysis, MT Navigator, and MatchTools software programs (PE Applied Biosystems), which enable assignment of *DRB1* genotypes based on a recent library file of *DRB1* alleles. The allele library file (DRB1.L257) was obtained from the International Immunogenetics database (ftp://ftp.ebi.ac.uk/pub/databases/imgt/mhc/hla/ (accessed 21 June 2006)). To resolve ambiguous results for some exon 2 sequences, an additional sequence reaction was similarly performed and analysed for codon 86 sequences using the Forensic Analytical reagent kit and protocol. This method detects all SE positive alleles, including *0101, *0102, *0401, *0404, *0405, *0408, *0413, *1001, and *1402. Subjects were classified as SE positive (that is, one or two SE copies) or SE negative (that is, no SE copies) based on their *HLA-DRB1* genotype.

GSTM1 genotyping

The *GSTM1*0* (null allele) was identified using a multiplex polymerase chain reaction technique, modified from a previous method.[28] Three primers were used: (p1) FAM-5′-CGCCATCTTGTGCTACATTGCCCG-3′ (final reaction concentration, 25 pmol/l); (p2) 5′-ATCTTCTCCTCTTCTGTCTC3′ (50 pmol/l); and (p3) 5′-TTCTGGATTGTAGCAGATCA-3′ (25 pmol/l). The primer pair p1 and p3 amplified a 230 bp fragment in the presence of the *GSTM1* allele (no polymerase chain reaction product for null allele); the primers p1 and p2 amplified a 158 bp fragment as an internal control. Size fragments were analysed on an ABI 3700 automated sequencer and alleles determined using the Applied Biosystems GeneScan, version 3.5 and Genotyper, version 2.6 software packages.

Statistical methods

Univariate associations between smoking, the SE, *GSTM1*, and RA risk were examined using contingency table analysis. Owing to the matched case-control design, we used conditional logistic regression to estimate the impact of smoking, presence of the SE, *GSTM1* null homozygosity, and interactions between these factors on RA risk in multivariable analyses. We also calculated odds ratios (ORs) to describe associations of smoking with RA risk among strata defined by specific genotypes.

RESULTS

Owing to features of the source population (that is, IWHS), all study subjects were Caucasian women. Mean age at the time of enrolment into the IWHS was about 60 years for RA cases and controls (61.0 *v* 60.3, p = 0.06). Mean age at RA onset for cases was 67.5 years. Sixty three per cent of RA cases were RF positive and 31% had clear documentation of erosive disease.

Table 1 shows univariate associations of smoking, SE positivity, and *GSTM1* null homozygosity with RA risk. All three exposures were positively associated with RA risk, with ORs ranging from 1.4 to 1.8. The frequencies of SE positivity and *GSTM1* null homozygosity among controls were in the range for Caucasian populations.[29–31] The prevalence of smoking among controls was representative of the source population (Iowa), and somewhat lower than the American national average.[32]

Multivariable analysis using conditional logistic regression accounted for the matched design, and table 2 shows the results obtained. In addition to the three exposures of interest, the model also included interaction terms between smoking and both of the genetic risk factors. All three exposures were significantly associated with RA risk in this multivariable analysis, as were both of the gene-environment interactions. Adjustment for other exposures shown to be associated with RA in this cohort[24 25 33] did not alter these

Table 1 Results of univariate analysis of smoking, shared epitope positivity, and *glutathione S-transferase M1 (GSTM1)* null homozygosity and RA risk

Exposure	Patients with RA No (%)	Controls No (%)	OR (95% CI)*
Tobacco smoke, ever	48 (41.7)	159 (34.1)	1.38 (0.91 to 2.10)
Tobacco smoke, current	22 (24.7)	59 (16.1)	1.71 (0.98 to 2.98)
Shared epitope positive	63 (60.6)	109 (45.8)	1.82 (1.14 to 2.91)
GSTM1 null homozygosity	58 (54.2)	200 (43.2)	1.56 (1.02 to 2.37)

*Odds ratios (OR) and 95% confidence intervals (CIs) describe univariate associations between each of the exposures and RA risk.

Table 2 Results of conditional logistic regression analysis of associations of smoking, shared epitope positivity, and *glutathione S-transferase M1 (GSTM1)* null homozygosity with RA risk

Independent variable	OR (95% CI)	p Value
Smoking	6.0 (1.8 to 20.0)	0.0036
Shared epitope (SE) positivity*	4.6 (1.9 to 11.2)	0.0006
GSTM1 null homozygosity	3.4 (1.4 to 8.2)	0.0074
Smoking-SE interaction	0.2 (0.06 to 0.90)	0.034
Smoking-*GSTM1* null interaction	0.3 (0.07 to 0.98)	0.047

*SE positive alleles include HLA-DRB1*0101, *0102, *0401, *0404, *0405, *0408, *0413, *1001, and *1402.

results. In particular, adjustment for cruciferous vegetable intake, which has been reported to interact with *GSTM1* in the presence of smoking related cancers, did not alter these results (data not shown).

To illustrate further the impact of tobacco exposure on RA risk and to better understand the nature of the gene-environment interactions we calculated ORs describing associations with smoking according to strata defined by specific genetic risk factors (table 3). As suggested by the full multivariable analysis described above, these stratified analyses indicate variation in risk associated with smoking for subjects with different genetic risk factors. For example, exposure to tobacco smoke was associated with an increased risk of RA among SE negative subjects but not among SE positive subjects. Similarly, exposure to tobacco smoke was associated with a significantly increased risk of RA among women with the *GSTM1* present genotype, but not among women with the *GSTM1* null genotype. Thus, in both situations the impact of smoking on RA risk operates primarily among subjects who lack these genetic risk factors.

DISCUSSION

Two of the factors examined in this study, the *HLA-DRB1* SE and cigarette smoking, are well established RA risk factors. The association of *GSTM1* with RA risk or outcome has only been examined in a few previous studies. Those studies suggest an association of the *GSTM1* null allele with RA risk and outcome, particularly radiographic disease progression.[19 34] Thus, our results confirm important preliminary work related to this functional polymorphism. Interest in *GSTM1* and other GSTs relates to their role in the metabolism of potentially toxic substances, such as the constituents of tobacco smoke. Subjects with reduced or absent expression of critical components in this pathway may be exposed to a wide range of reactive oxygen species or other potentially pathogenic substances. Thus, antioxidants have been shown in this cohort and other populations to be associated with a decreased risk of RA.[24 35 36]

A unique feature of our study was the examination of two genetic and one environmental factor and their interaction in a well designed population based, nested case-control study. A particular strength is that controls were sampled from the same population that generated the cases, minimising the potential for selection bias, and cases and controls were matched by ethnicity, minimising the potential for population stratification, both of which have been cited as potential problems with many genetic association studies. Although interactions between genes and environmental factors are considered to be fundamentally important in complex disease aetiology, few studies have examined them. In part, this reflects the challenges of defining the specific genetic and non-genetic determinants of complex diseases such as RA.

The results of our study suggest that exposure to tobacco smoke increases the risk of RA among older Caucasian women who have not inherited the most well established genetic risk factor for the disease—namely, the *HLA-DRB1* SE (table 3). In contrast, this exposure was not associated with an increased risk of disease among SE positive women. Of interest, previous work by Hutchinson *et al* demonstrated that exposure to heavy smoking was particularly important for subjects without a family history of RA.[11] Although results of SE (or other) genotyping were not available for that study, it is tempting to speculate that the patients without a family history of RA were less likely to have inherited the SE or other genetic risk factors.

More recently, Padyukov *et al* studied associations of the SE and smoking with RA in a population based case-control study in Sweden.[37] Our results were quite similar to theirs for associations with the SE and smoking as main effects, as well as significant interaction between these factors. In contrast, the results of stratified analyses were notably different. In the Swedish population, smoking was associated with RA risk primarily among SE positive subjects. However, there were important differences between the two study populations that may account for these differences. In particular, because the Swedish study did not focus on elderly onset disease the average age of RA onset in that population was 20 years lower than in our study. There are also likely to be important differences in other potentially relevant exposures between Sweden and America that may influence RA risk.

Our results also suggest that exposure to cigarette smoke is a more important RA risk factor among subjects who have the *GSTM* present genotype. This finding was unexpected, as we hypothesised that smoking would be a stronger risk factor for RA among those who do not express *GSTM1* (null

Table 3 Odds ratios describing the risk of RA associated with exposure to tobacco smoke among strata defined by specific genetic risk factors

Genetic factor*	Case	Control	OR (95% CI)
SE positive			
Non-smoker	39	65	1 (reference)
Smoker	24	44	0.91 (0.48 to 1.72)
SE negative			
Non-smoker	23	90	1 (reference)
Smoker	18	39	1.81 (0.88 to 3.71)
GSTM1 null genotype			
Non-smoker	37	140	1 (reference)
Smoker	21	60	1.32 (0.72 to 2.45)
GSTM1 present genotype			
Non-smoker	22	166	1 (reference)
Smoker	27	97	2.10 (1.13 to 3.89)

*SE, shared epitope (includes HLA-DRB1*0101, *0102, *0401, *0404, *0405, *0408, *0413, *1001 and *1402 alleles); *GSTM1*, *glutathione S-transferase M1*.

genotype), as these subjects would be expected to have a greater exposure to reactive oxygen species and other compounds in cigarette smoke. On the other hand, these results are consistent with our findings related to smoking and SE status—namely, that smoking influenced RA risk primarily among those who lacked this genetic risk factor. Moreover, it may be that lack of *GSTM1* expression itself is such a strong risk factor for RA (main effect) that cigarette smoking adds no additional risk or only a small amount (there are multiple sources of oxidative load), but in the lower risk setting of *GSTM1* expression, cigarette smoking can overwhelm protective mechanisms and thereby increase RA risk. These provocative results clearly require confirmation in other studies.

Although the population based nature of the IWHS cohort is a study strength, this feature also limited the number of incident RA cases available and restricted the sample to postmenopausal Caucasian women. This also limited our ability to study specific SE alleles or genotypes. Further, the older age of the cohort meant that a number of the incident RA cases had died or were too frail to participate in this genetic study. However, comparison of demographic and clinical characteristics of the incident cases who were not included in the current genetic study with those who were included suggests that their non-participation was not likely to have introduced bias. Further, comparison of the controls examined in the current study with the entire IWHS cohort, including rates of smoking, suggests that the particular control group studied did not introduce bias.[13]

The older age of the study population, in particular the mean age at RA onset among the incident cases of nearly 67, is also of interest because it is widely assumed that genetic susceptibility factors are more important for subjects with early onset disease. For example, the genetic contribution to breast cancer is substantially greater among women with early onset disease.[38 39] It is also possible that the impact of smoking, a non-genetic risk factor, is relatively large in this RA cohort, given their advanced age. However, research in this area remains limited for RA and other complex diseases.

The limited sample size and restriction of the current study to Caucasian women means that additional work will be required to confirm these findings and extend this work to men and non-Caucasian populations. Given the higher rate of smoking among men, this is an important question. In addition, the SE does not appear to be as important a risk factor for RA among some non-Caucasian populations, such as African and certain Hispanic American groups.[40 41] Thus, further examination of the SE, in conjunction with smoking or other epidemiological risk factors, is required to clarify the impact of this genetic risk factor for non-Caucasian populations.

In summary, we observed an association of smoking, two genetic risk factors, and their interaction, with the risk of RA in a population based case-control study of older Caucasian women. Our results suggest that the impact of smoking on RA risk is substantially greater among women who have not inherited specific genetic risk factors for the disease. The interaction of genetic susceptibility with environmental risk factors has been insufficiently studied, yet failure to account for this complexity is likely to impede progress in the identification of genetic and environmental determinants of RA and other complex diseases.

ACKNOWLEDGEMENTS

This work was supported by a Clinical Science grant from the Arthritis Foundation, the National Institutes of Health (K-24-AR-02175), the National Cancer Institute (R01 CA39741), and the Rosalind Russell Medical Research Center for Arthritis.

Authors' affiliations

L A Criswell, R F Lum, K A Pfeiffer, Rosalind Russell Medical Research Center for Arthritis, Department of Medicine, University of California, San Francisco, USA
K G Saag, Division of Clinical Immunology and Rheumatology, University of Alabama, Birmingham, USA
T R Mikuls, Section of Rheumatology and Omaha VA Medical Center, University of Nebraska, Omaha, USA
J R Cerhan, Department of Epidemiology, Mayo Clinic College of Medicine, Rochester, USA
L A Merlino, College of Public Health, University of Iowa, Iowa City, USA
B Woehl, M F Seldin, Rowe Program in Human Genetics, University of California, Davis, USA

Competing interests: None.

REFERENCES

1 **Seldin MF**, Amos CI, Ward R, Gregersen PK. The genetics revolution and the assault on rheumatoid arthritis. *Arthritis Rheum* 1999;**42**:1071–9.
2 **Cornelis F**, Faure S, Martinez M, Prud'homme J-F, Fritz P, Dib C, *et al*. New susceptibility locus for rheumatoid arthritis suggested by a genome-wide linkage study. *Proc Natl Acad Sc USAi* 1998;**95**:10746–50.
3 **Jawaheer D**, Seldin MF, Amos CI, Chen WV, Shigeta R, Etzel C, *et al*. Screening the genome for rheumatoid arthritis susceptibility genes. A replication study and combined analysis of 512 multicase families. *Arthritis Rheum* 2003;**48**:906–16.
4 **John S**, Shephard N, Liu G, Zeggini E, Cao M, Chen W, *et al*. Whole-genome scan, in a complex disease, using 11,245 single-nucleotide polymorphisms: comparison with microsatellites. *Am J Hum Genet* 2004;**75**:54–64.
5 **Amos CI**, Chen WV, Lee A, Li W, Kern M, Lundsten R, *et al*. High-density SNP analysis of 642 Caucasian families with rheumatoid arthritis identifies two new linkage regions on 11p12 and 2q33. *Genes Immun* 2006;**7**:277–86.
6 **Silman AJ**, Pearson JE. Epidemiology and genetics of rheumatoid arthritis. *Arthritis Res* 2002;**4**(suppl 3):S265–72.
7 **Newton JL**, Harney SM, Wordsworth BP, Brown MA. A review of the MHC genetics of rheumatoid arthritis. *Genes Immun* 2004;**5**:151–7.
8 **Aho K**, Koskenvuo M, Tuominen J, Kaprio J. Occurrence of rheumatoid arthritis in a nationwide series of twins. *J Rheumatol* 1986;**13**:899–902.
9 **Silman AJ**, MacGregor AJ, Thomson W, Holligan S, Carthy D, Farhan A, *et al*. Twin concordance rates for rheumatoid arthritis: results from a nationwide study. *Br J Rheumatol* 1993;**32**:903–7.
10 **Uhlig T**, Hagen KB, Kvien TK. Current tobacco smoking, formal education, and the risk of rheumatoid arthritis. *J Rheumatol* 1999;**26**:47–54.
11 **Hutchinson D**, Shepstone L, Moots R, Lear JT, Lynch MP. Heavy cigarette smoking is strongly associated with rheumatoid arthritis (RA), particularly in patients without a family history of RA. *Ann Rheum Dis* 2001;**60**:223–7.
12 **Karlson EW**, Lee IM, Manson JE, Buring JE, Hennekens CH. A retrospective cohort study of cigarette smoking and risk of rheumatoid arthritis in female health professionals. *Arthritis Rheum* 1999;**42**:910–17.
13 **Criswell LA**, Merlino L, Cerhan JR, Mikuls T, Mudano A, Burma M, *et al*. Cigarette smoking and risk of elderly onset rheumatoid arthritis: results from the Iowa Women's Health Study. *Am J Med* 2002;**112**:465–71.
14 **Aho K**, Heliovaara M, Maatela J, Tuomi T, Palosuo T. Rheumatoid factor antedating clinical rheumatoid arthritis. *J Rheumatol* 1991;**18**:1282–4.
15 **Mathews JD**, Whittingham S, Hooper BM, Mackay IR, Stenhouse NS. Association of autoantibodies with smoking, cardiovascular morbidity, and death in the Brusselton population. *Lancet* 1973;**2**:754–8.
16 **Tuomi T**, Heliovaara M, Palosuo T, Aho K. Smoking, lung function, and rheumatoid arthritis. *Ann Rheum Dis* 1990;**49**:753–6.
17 **Seidegard J**, Pero RW, Markowitz MM, Roush G, Miller DG, Beattie EJ. Isoenzyme(s) of glutathione transferase (class Mu) as a marker for the susceptibility to lung cancer: a follow up study. *Carcinogenesis* 1990;**11**:33–6.
18 **Chenevix-Trench G**, Young J, Coggan M, Board P. Glutathione S-transferase M1 and T1 polymorphisms: susceptibility to colon cancer and age of onset. *Carcinogenesis* 1995;**16**:1655–7.
19 **Mattey DL**, Hutchinson D, Dawes PT, Nixon NB, Clarke S, Fisher J, *et al*. Smoking and disease severity in rheumatoid arthritis: association with polymorphism at the glutathione S-transferase M1 locus. *Arthritis Rheum* 2002;**46**:640–6.
20 **Folsom AR**, Kaye SA, Prineas RJ, Potter JD, Gapstur SM, Wallace RB. Increased incidence of carcinoma of the breast associated with abdominal adiposity in postmenopausal women. *Am J Epidemiol* 1990;**131**:794–803.
21 **Folsom AR**, Kaye SA, Sellers TA, Hong C-P, Cerhan JR, Potter JD, *et al*. Body fat distribution and 5-year risk of death in older women. *JAMA* 1993;**269**:483–7.
22 **Folsom AR**, Mink PJ, Sellers TA, Hong CP, Zheng W, Potter JD. Hormonal replacement therapy and morbidity and mortality in a prospective study of postmenopausal women. *Am J Public Health* 1995;**85**:1128–32.
23 **Cerhan JR**, Saag KG, Criswell LA, Merlino LA, Mikuls T. Blood transfusion, alcohol use and anthropometric risk factors for elderly onset rheumatoid arthritis in women. *J Rheumatol* 2002;**29**:246–54.
24 **Cerhan JR**, Saag KG, Merlino LA, Mikuls TR, Criswell LA. Antioxidant micronutrients and risk of rheumatoid arthritis in a cohort of older women. *Am J Epidemiol* 2003;**157**:345–54.

25 **Merlino LA**, Cerhan JR, Criswell LA, Mikuls TR, Saag KG. Vitamin D intake and rheumatoid arthritis: results from the Iowa Women's Health Study. *Arthritis Rheum* 2004;**50**:72–7.
26 **Arnett FC**, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, *et al*. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;**31**:315–24.
27 **MacGregor AJ**, Bamber S, Silman AJ. A comparison of the performance of different methods of disease classification for rheumatoid arthritis. Results of an analysis from a nationwide twin study. *J Rheumatol* 1994;**21**:1420–6.
28 **Zhong S**, Wyllie AH, Barnes D, Wolf CR, Spurr NK. Relationship between the GSTM1 genetic polymorphism and susceptibility to bladder, breast and colon cancer. *Carcinogenesis* 1993;**14**:1821–4.
29 **Fries JF**, Wolfe F, Apple R, Erlich H, Bugawan T, Holmes T, *et al*. HLA-DRB1 genotype associations in 793 white patients from a rheumatoid arthritis inception cohort: frequency, severity, and treatment bias. *Arthritis Rheum* 2002;**46**:2320–9.
30 **Hengstler JG**, Arand M, Herrero ME, Oesch F. Polymorphisms of N-acetyltransferases, glutathione S-transferases, microsomal epoxide hydrolase and sulfotransferases: influence on cancer susceptibility. *Recent Results Cancer Res* 1998;**154**:47–85.
31 **Garte S**, Gaspari L, Alexandrie AK, Ambrosone C, Autrup H, Autrup JL, *et al*. Metabolic gene polymorphism frequencies in control populations. *Cancer Epidemiol Biomarkers Prev* 2001;**10**:1239–48.
32 Health United States, 1990. Hyattsville, Maryland: Public Health Service, 1991.
33 **Mikuls T**, Cerhan JR, Criswell LA, Merlino L, Mudano AS, Burma M, *et al*. Coffee, tea and caffeine consumption and risk of rheumatoid arthritis. *Arthritis Rheum* 2002;**46**:83–91.
34 **Mattey DL**, Hassell AB, Plant M, Dawes PT, Ollier WR, Jones PW, *et al*. Association of polymorphism in glutathione S-transferase loci with susceptibility and outcome in rheumatoid arthritis: comparison with the shared epitope. *Ann Rheum Dis* 1999;**58**:164–8.
35 **Heliovaara M**, Knekt P, Aho K, Aaran RK, Alfthan G, Aromaa A. Serum antioxidants and risk of rheumatoid arthritis. *Ann Rheum Dis* 1994;**53**:51–3.
36 **Comstock GW**, Burke AE, Hoffman SC, Helzlsouer KJ, Bendich A, Masi AT, *et al*. Serum concentrations of alpha tocopherol, beta carotene, and retinol preceding the diagnosis of rheumatoid arthritis and systemic lupus erythematosus. *Ann Rheum Dis* 1997;**56**:323–5.
37 **Padyukov L**, Silva C, Stolt P, Alfredsson L, Klareskog L. A gene-environment interaction between smoking and shared epitope genes in HLA-DR provides a high risk of seropositive rheumatoid arthritis. *Arthritis Rheum* 2004;**50**:3085–92.
38 **Pharoah PD**, Day NE, Duffy S, Easton DF, Ponder BA. Family history and the risk of breast cancer: a systematic review and meta-analysis. *Int J Cancer* 1997;**71**:800–9.
39 **Anderson H**, Bladstrom A, Olsson H, Moller TR. Familial breast and ovarian cancer: a Swedish population-based register study. *Am J Epidemiol* 2000;**152**:1154–63.
40 **McDaniel DO**, Alarcon GS, Pratt PW, Reveille JD. Most African-American patients with rheumatoid arthritis do not have the rheumatoid antigenic determinant (epitope). *Ann Intern Med* 1995;**123**:181–7.
41 **Teller K**, Budhai L, Zhang M, Haramati N, Keiser HD, Davidson A. HLA-DRB1 and DQB typing of Hispanic American patients with rheumatoid arthritis: the ''shared epitope'' hypothesis may not apply. *J Rheumatol* 1996;**23**:1363–8.