**Supplementary text**

*The EIRA study*

The data collection of the EIRA study started in 1996 and is still on-going. In 2006, a new version of the questionnaire was released, including new variables for analyses (breastfeeding, age at menarche and age at menopause, among others) and increasing the number of controls included for each case. In the first phase of the study (1996-2006), one control was selected for each case, while two controls were selected for each case in the second phase (2006-2014). If a selected control could not be contacted or refused to participate, another control was invited to participate. All these factors explain why the number of controls is about 1.6 times the number of cases. Regarding age of inclusion, initially (1996-2010), cases included were 18-70 years of age but later on (2011-2014) all cases above 18 years were included. Incident cases of RA were included and diagnosed by rheumatologists according to the American College of Rheumatology 1987 criteria for RA. (1) Since 2010, following the introduction of new RA criteria, (2) cases included may fulfill either of the sets of diagnostic criteria.

 *Additive Interaction*

We investigated the presence of interaction as described by Rothman. (3) The biological model of interaction estimates the relative risk (RR) of RA, expressed as odds ratio (OR, with its 95% confidence interval), associated with the exposure to two factors, e.g. (A) and (B). The interaction was calculated in an additive scale. The risk of the disease was calculated in individuals exposed to (A) but not (B), exposed to (B) but not (A) and exposed to both factors (A) and (B). Individuals not exposed to any of the factors were used as the reference category.

If OC use and/or BF presented a protective effect, the risk category should include women with a shorter duration of BF and never OC users for each analysis, as recommended by Knol *et al*. (4) Risk categories for smoking (ever smoking), SE, (any SE allele) and *PTPN22* (any T allele) were coded in the expected direction. Since we found a significant protective association between OC use and the risk of ACPA-positive RA, lack of OC use was considered as the “risk” category for interaction analyses.

The attributable proportion due to interaction (AP) estimates the proportion of the excess risk that is due to the interaction per se (factor A+ Factor B) according to the formula RRAB – RRA – RRB + 1 / RRAB. The AP is the proportion of the incidence among people exposed to two interacting factors that is attributable to the interaction *per se* (i.e. indicating their joint effect apart from the sum of their independent effects). AP is equal to zero in the absence of interaction.

Knowledge of potential interaction between two factors would contribute to disentangle biological mechanisms of importance for the disease. The results obtained from this work indicate a significant interaction between the lack of OC use and another environmental factor: smoking. A potential interaction between OC use and one of the major genetic risk factors, carriage of the SE allele, was observed in the crude model. However, this association turned out not to be significant after adjustments.

**References**

1. 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 Mar; 31(3):315-324.

2. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, 3rd, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum. 2010 Sep; 62(9):2569-2581.

3. Rothman KJ, ed. Epidemiology: An Introduction. 2nd ed: Oxford University Press, USA 2012.

4. Knol MJ, VanderWeele TJ, Groenwold RH, Klungel OH, Rovers MM, Grobbee DE. Estimating measures of interaction on an additive scale for preventive exposures. Eur J Epidemiol. 2011 Jun; 26(6):433-438.