**Appendix 1**

**Missing data management**

To account for missing data in the confounders and predictors in the analysis, we used multiple imputation by chained equations (MICE). MICE performs sequential regression multivariate imputation, and the procedure is as follows. We initially replaced all missing values by random sampling with replacement on the observed values. The sequential regression then proceeded as follows. Beginning with the first confounder or predictor, say x1, we regressed it on all other variables restricted to the subjects where we observed x1. From this regression, we obtained the posterior predictive distribution, from which we used to draw samples from to replace the missing values in x1. We then moved to the next confounder or predictor, say x2, and regressed it on all other variables restricted to the subjects where we observed x2 and using the imputed values of x1. We then repeated this procedure for all confounders and predictors in the model, and then continued looping through the variables many times to stabilise the results. Altogether, this gave a single imputed data set. We repeated this procedure multiple times to produce multiple imputed data sets. We performed our analysis on each of the imputed data sets to investigate the sensitivity of our conclusions relative to the complete case analysis. The conclusions drawn from our analyses did not differ using the imputed data sets, and so we present here only results from the observed data.