Methodology supplementary information

Modelling of the hazard function over time was done using a flexible parametric spline model with 2 degrees of freedom (equivalent to 1 knot). Model selection was chosen by minimising the Akaike Information Criterion. The use of larger numbers of degrees of freedom did not improve the model. This method of spline modelling has been described in detail previously (Royston P, Stata journal 2001: 1; 1, 1-28).

Multiple imputations were performed in Stata using the ICE command. Missing data were present in the following variables: age, disease duration, baseline HAQ, and baseline DAS28 score. The imputation model was constructed separately for the DMARD and anti-TNF cohorts. Age, gender, disease duration, baseline HAQ, baseline DAS28 score, co-morbidity, smoking status, entry year, baseline steroid exposure and prior orthopaedic surgery were all included as predictors within the imputation model. Twenty imputation cycles were performed and the resulting data were analysed using Rubin's rules with the MIM command.

Adjusting for confounders was performed using an inverse probability of treatment weighting score. A probability of treatment (propensity) score was generated using logistic regression. As several covariates were associated with treatment likelihood in a non-linear pattern, or demonstrated an interaction with other covariates, sequential analyses were performed to identify the components of the model. If a non-linear relationship was identified, orthogonal polynomial transformations of the covariates were added to the model until a suitable fit was obtained. The final propensity score included polynomials for age, DAS28 score and HAQ score; interactions were identified between age and DAS28 score, age and prior orthopaedic surgery, entry year and DAS28 score, co-morbidity and DAS28 score, steroid exposure and disease duration, entry year and disease duration, and entry year and steroid exposure. The inverse of the probability (or 1 minus the inverse of the probability in the DMARD cohort) was then used as the treatment weight in the analysis. Weights greater than 20 that would de-stabilise the model were replaced with the value of 20.