Appendix. Detailed statistical methods.

Threshold value of CRP was determined to obtain as much upward discordance (CRP > threshold while SAA < 10 mg/L) than downward discordance (CRP < threshold while SAA ≥ 10 mg/L). This threshold was found by dichotomy in converging the McNemar odds ratio towards 1. The CRP threshold was set at 15 mg/L for the first iteration, an obviously too high value indicated by more downward than upward discordance. The CRP threshold was therefore decreased to (15+0)/2=7.5 mg/L for the next iteration. If there were still more downward than upward discordances at this threshold (McNemar OR > 1), this threshold was still too high and again lowered to (7.5+0)/2=3.75 for the next iteration. If there were more upward than downward discordances at 7.5 mg/L (McNemar OR < 1), the CRP threshold was now too low and increased to (7.5+15)/2=11.25 for the next iteration. If there were as many upward than downward discordances at 7.5 mg/L (McNemar OR = 1), the CRP threshold was deemed appropriate. This process was iterated until finding the CRP threshold corresponding to as many upward than downward discordances (McNemar OR = 1).

Several patients had more than one blood test over the study period. To avoid bias due to the over-influence of these patients, only one sample per patients could be analyzed simultaneously. We used within patient resampling to produce unbiased statistics [1]: 500 lots were made up by randomly drawing one blood test per patient, then statistics of all lots were combined to produce the final statistics. We looked for an interaction between age and the CRP threshold in order to derive different thresholds if appropriate.

Once the CRP threshold was defined, we first assessed the reproducibility of the concordance between CRP < threshold and SAA < 10 mg/L within clusters of samples from the same patient at the same colchicine dose (0 mg/d, ≤ 1 mg/d, ≤ 2 mg/d, > 2 mg/d). For this analysis, a pair of samples was randomly selected from each cluster with multiple samples. Each pair was classified as concordant if CRP and SAA were both over or both under the threshold, and

discordant otherwise. The reproducibility of the classification of the pairs of samples was then assessed. This process was repeated 500 times and the statistics of all lots were pooled to produce the final statistics.

We then assessed the overall concordance between CRP < threshold and SAA < 10 across all patients. For this analysis, one sample was randomly selected from each patient and the concordance of CRP and SAA was then assessed with Cohen's kappa index. Again this was performed 500 times to pool the results and produce the final statistics. Finally, we used logistic regression to look for predictors of concordance among predetermined variables: with or without colchicine, M694V homozygosity or not, Sephardic origins or not.

All statistical analyses were performed with Stata 9.2 (SataCorp, Mountainview, TX).

Reference:

D. Follmann, M. Fay, Exact inference for complex clustered data using within-cluster resampling. J. Biopharm. Stat. 20 (2010) 850-69.