

RMD
OpenRheumatic &
Musculoskeletal
Diseases

ORIGINAL ARTICLE

Performance of matrices developed to identify patients with early rheumatoid arthritis with rapid radiographic progression despite methotrexate therapy: an external validation study based on the ESPOIR cohort data

Benjamin Granger,^{1,2} Bernard Combe,³ Xavier Le Loet,⁴ Alain Saraux,^{5,6} Francis Guillemain,^{7,8} Bruno Fautrel^{1,9}

To cite: Granger B, Combe B, Le Loet X, *et al.* Performance of matrices developed to identify patients with early rheumatoid arthritis with rapid radiographic progression despite methotrexate therapy: an external validation study based on the ESPOIR cohort data. *RMD Open* 2016;2:e000245. doi:10.1136/rmdopen-2016-000245

► Prepublication history and additional material is available. To view please visit the journal (<http://dx.doi.org/10.1136/rmdopen-2016-000245>).

Received 5 January 2016

Revised 7 March 2016

Accepted 16 April 2016



CrossMark

For numbered affiliations see end of article.

Correspondence to

Dr Benjamin Granger;
benjamin.granger@psl.aphp.fr

ABSTRACT

Introduction: Use of prediction matrices of risk or rapid radiographic progression (RRP) for early rheumatoid arthritis (RA) in clinical practice could help to better rationalise the first line of treatment. Before use, they must be validated in populations that have not participated in their construction. The main objective is to use the ESPOIR cohort to validate the performance of 3 matrices (ASPIRE, BEST and SONORA) to predict patients at high risk of RRP at 1 year of disease despite initial treatment with methotrexate (MTX).

Methods: We selected from the ESPOIR cohort 370 patients receiving MTX or leflunomide (LEF) for ≥ 3 months within the first year of follow-up. Patients were assessed clinically every 6 months, and structural damage progression seen on radiography was measured by the van der Heijde-modified Sharp score (vSHS) at 1 year. RRP was defined as an increase in the vSHS ≥ 5 points during the first year.

Results: At 1 year, the mean vSHS score was 1.7 ± 5.0 and 46 patients had RRP. The ASPIRE matrix had only moderate validity in the ESPOIR population, with area under the receiver operating characteristic curve (AUC) < 0.7 . The AUC for the BEST and SONORA matrices were 0.73 and 0.76. Presence of rheumatoid factor (RF)—or anti-citrullinated protein antibodies (ACPAs) and initial structural damage were always predictive of RRP at 1 year. Disease Activity Score in 28 joints (DAS28) and C reactive protein (ASPIRE threshold) were not associated with RRP.

Conclusions: Matrices to identify patients at risk of RRP tested in the ESPOIR cohort seem to perform moderately. There is no matrix that shows clearly superior performance.

INTRODUCTION

During the past decade, the standard of care for early rheumatoid arthritis (RA) has

Key messages

What is already known about this subject?

- Rapid radiographic progressor matrices can identify early rheumatoid arthritis patients with high risk of structural damage progression despite initiation of methotrexate.

What does this study add?

- The present study validates the good performance of the 3 matrices (BEST, SONORA and ASPIRE) in another population of early RA patients.

How might this impact on clinical practice?

- Rapid radiographic progression matrices are valuable tools to optimise the care of early RA patients by choosing more aggressive therapy than MTX alone.

evolved greatly, combining early referral to the rheumatologist to hasten the RA diagnosis,^{1–3} initiation of rapid disease-modifying anti-rheumatic drugs (DMARDs) within the ‘window of opportunity’^{1 2} and tight control of disease activity based on regular DMARD adaptation according to the ‘treat to target’ strategy.^{2 4–6}

The choice of the first DMARD may be important for RA prognosis and has been addressed by many trials and guidelines. Synthetic DMARDs (sDMARDs) such as methotrexate (MTX) or leflunomide (LEF) are the most recommended as ‘anchor’ therapies because they can be associated with other synthetic or biological DMARDs (bDMARDs) according to the response to

initial monotherapy.^{2 7–9} Additionally, some studies suggested that use of sDMARDs tritherapy or bDMARDs as first-line agents could be efficacious to rapidly achieve remission and block structural damage.^{2 9–13} However, safety concerns and the high costs of biologics appear to limit their use for early RA. Several economic evaluations reported that incremental cost-effectiveness ratios of biologics as first-line treatment for early RA are usually high and largely surpass the generally accepted thresholds.^{14 15}

If biological agents are not recommended for all patients with early RA, they may still be of interest for patients with poor prognostic factors and features of severe disease, for whom MTX is most likely insufficient.^{2 5} Several trials have shown substantial structural progression in a small subset of patients despite rapid initiation of MTX therapy, which led to the development of the concept of rapid radiographic progression (RRP) to identify such patients. RRP is defined as structural damage progression by an increase in the van der Heijde-modified Sharp score (vSHS) ≥ 5 points over 1 year; the cut-off of 5 points corresponds to the destruction of one small joint and to the usually reported smallest detectable difference (SDD).^{16–18} The rationale for RRP patient identification has been validated in two different studies. In the BEST trial, patients with RRP during the first year of follow-up showed increased functional limitations and structural damage progression over 8 years of follow-up, despite a tightly controlled therapeutic strategy.¹⁸ These results were consistent with another study of the ESPOIR cohort, in which patients with RRP during the first year in the cohort—with a definition slightly different from the previous one—showed increased structural damage progression during the second and third years in the cohort.¹⁹

Patient risk stratification with regard to RRP has become highly important and has been addressed by the development of prediction matrices to quantify the risk of RRP at 1 year of follow-up. RRP matrices, based on the same methodology used in cardiology or osteoporosis,^{20 21} were developed within randomised control trials (RCTs)^{16 22 23} or observational cohorts^{24–26} and have involved different sets of baseline characteristics to calculate, at the patient level, the probability of a given patient showing RRP on a 1 year radiograph despite MTX therapy (ie, the probability of being a radiographic MTX inadequate responder). The matrices differ by the nature of the components they use, their thresholds and how they take into account structural damage. These tools have rarely been validated in populations different from those for which they were developed. The only validations were performed with the BRASS cohort, which included patients with established RA,²⁷ and in a cohort of patients with DMARD-naïve early RA recruited at the Department of Rheumatology of the University Hospitals Leuven.^{25 28} The substantial differences between established and early RA resulted in low performance of the three tested matrices (ASPIRE, BEST

and SWEFOT). Therefore, additional validation was needed (bigger cohort, extra matrix—SONORA), which is possible with the ESPOIR cohort. The ESPOIR cohort enrolled patients with early arthritis from the community (with or without unfavourable prognostic factors), most receiving MTX or LEF as first-line treatment.

In this study, we aimed to test the performance and validity of the different published RRP prediction matrices in patients with early RA from the ESPOIR cohort.

METHODS

We aimed to assess the performance of different prediction matrices that were developed to identify patients with early RA at substantial risk of rapid structural damage progression, defined as an increase in the vSHS ≥ 5 points between baseline and the 1 year follow-up visit.

Previously published matrices

The patient characteristics in the tested matrices are in online supplementary materials—table 1. Three matrices were developed with RCT populations, all testing the efficiency of a combination of MTX+infliximab versus MTX monotherapy in patients with early RA: ASPIRE,^{16 29} SWEFOT^{23 30} and BEST.^{22 31} In these studies, we only used the matrix developed for patients of the MTX arm. All trials used the same definition of RRP: an increase in the vSHS ≥ 5 points within the first year after treatment initiation, which theoretically corresponds to the destruction of one joint based on vSHS scoring¹⁷ or also five new small erosions in five different joints, or an association of progression in erosion lesions and joint space narrowing. In addition, we used the SONORA matrix based on data for a North American observational cohort of 994 patients with early RA who received a sDMARD.²⁴ In this study, RRP definition was an increase in the vSHS ≥ 3.54 points, which was the SDD. We also used the original ESPOIR matrix as well as a modified version (mESPOIR matrix) developed with the same methodology.²⁶ Although the original ESPOIR matrix assessed structural damage qualitatively—‘typical RA erosion’, yes or no—the mESPOIR matrix used a quantitative vSHS-based assessment for structural damage, categorised in three classes (<5 points, 5–14 points, >14 points). For these two matrices, RRP was defined as an increase in the vSHS ≥ 5 points.

Validation population

The protocol for the ESPOIR cohort study was approved by the Ethics Committee of Montpellier University Hospital, France. All patients gave their signed informed consent to participate in the cohort. Between December 2002 and March 2005, 813 patients with possible RA who were referred by rheumatologists and general practitioners to one of 14 regional centres were included in the ESPOIR cohort.^{32 33} Inclusion criteria were age 18–70 years, more than two swollen joints for >6 weeks and

Table 1 Baseline characteristics and disease evolution of the ESPOIR patients who received MTX or its equivalent, LEF, as a first-line biologic agent for RA, and radiographic structural damage progression during the first year in the cohort

Characteristics	ESPOIR cohort patients who received MTX or LEF		
	All (n=370)	RRP+ (n=41)	RRP- (n=329)
Age, years	49.4±11.4 (51.5)	50.2±10.8 (52.1)	49.3±11.5 (51.4)
Female sex	271 (73.2%)	31 (75.6%)	240 (73%)
Disease duration, weeks*	15.2±8.2 (13.6)	18.6±8.4 (18.6)	14.7±8.1 (13)
2010 ACR/EULAR criteria	316 (85.4%)	38 (92.7%)	278 (84.5%)
Swollen joint count in 28 joints	7.9±5.4 (7)	8.7±6.1 (8)	7.8±5.3 (7)
Tender joint count in 28 joints	8.7±6.9 (7)	8.3±6.3 (7)	8.7±6.9 (7)
ESR, mm/1st h	32.7±25 (26)	32.6±21.0 (30)	32.7±25.4 (25)
CRP, mg/L	24.8±37.7 (11)	26.0±27.9 (14)	24.6±38.7 (11)
DAS28	5.3±1.2 (5.2)	5.3±1.2 (5.4)	5.3±1.2 (5.2)
IgM RF positivity*	204 (55.1%)	29 (70.7%)	175 (53.2%)
ACPA positivity*	185 (50%)	31 (75.6%)	154 (46.8%)
HAQ score	1.0±0.7 (1)	1.0±0.6 (1)	1.0±0.7 (1)
Typical RA erosion	66 (17.8%)	18 (44%)	48 (14.6%)
Duration of DMARD during first year, months	9.06± (3.07)	9.07± (3.20)	8.98± (3.06)
Concomitant glucocorticoids	98 (26%)	7 (17.1%)	91 (27.7%)
vSHS baseline total score*	4.5±6.8 (2)	10.8±9.8 (9)	3.7±5.8 (2)
Δ vSHS total	1.6±5.4 (0)	12.9±8.0 (10)	0.2±2.8 (0)
Progressors (Δ vSHS \geq 1)	126 (34.1%)	41 (100%)	85 (25.8%)

*Significant difference between RRP+ and RRP- (p value <0.05).

Data are mean±SD (median) or no. (%).

ACPA, anti-citrullinated protein antibody (ELISA, DiaSorin, France; positive >50 U/mL); ACR/EULAR, American College of Rheumatology/European League Against Rheumatism; CRP, C reactive protein (normally <10 mg/L); DAS28, Disease Activity Score in 28 joints; DMARD, disease-modifying anti-rheumatic drug; ESR, erythrocyte sedimentation rate; HAQ, Health Assessment Questionnaire; LEF, leflunomide; MTX, methotrexate; RA, rheumatoid arthritis; RF, rheumatoid factor (ELISA, Menarini, France; positive >9 UI/mL); RRP, rapid radiographic progression; vSHS, van der Heijde-modified Sharp score.

<6 months, suspected or confirmed diagnosis of RA, and not taking any DMARDs or steroids except for <2 weeks before enrolment. During the first year, patients were followed every 6 months. At each time point, data were collected on disease activity by the Disease Activity Score in 28 joints (DAS28),³⁴ functional ability by the Health Assessment Questionnaire (HAQ),³⁵ radiography of the hand and foot (anteroposterior views) and therapeutic regimen. Treatment strategies were not protocol-based in the ESPOIR cohort, and patients received usual care by their rheumatologist.

This study involved data for ESPOIR patients with an RA diagnosis according to their rheumatologist and initiation of a first sDMARD such as MTX or LEF with demonstrated structural efficacy for at least 3 months during the first year of follow-up in the cohort. Patients who initially received MTX started the drug 27.2±15.1 weeks after disease onset. Those who initially received LEF received it 30.1±18.1 weeks after disease onset.²⁶ Treatment duration was 37.2±12.3 (39.4) weeks during the first year for patients under MTX and 39.6±13.5 (43.1) for patients under LEF. No statistical difference between RRP+ and RRP- patients was found for treatment duration. Some patients received insufficient dosage of another treatment (as other sDMARD or bDMARD);²⁶ this point has not been considered in the analysis. As detailed elsewhere,²⁶ radiographs were read pairwise by a well-trained investigator blinded to clinical evaluation (intraclass correlation coefficient 0.99, SDD

0.966).⁶ Structural damage was assessed qualitatively by the presence of typical RA erosions, based on their location and aspect, and quantified by the vSHS.^{36 37} RRP was defined as change in vSHS (Δ vSHS) \geq 5 at 1 year.^{16 18 22 26}

Statistical analysis

Baseline characteristics and disease evolution were described for all patients and by RRP status by mean±SD (median) or frequency (percentage) as appropriate (table 1). Baseline characteristics and disease evolution were compared between RRP+ and RRP- by Mann-Whitney U test (for numerical data) and Fisher's exact test (for categorical data) (table 1). To test the relevance of RRP predictors used with the previous cohorts, we used Fisher's exact test for univariate analysis, then logistic regression analysis to determine predictors of RRP as the outcome variable (table 2).

The performance of previously published RRP matrices (BEST, SONORA, ASPIRE) with patients with early RA from the ESPOIR cohort was tested by several statistics (table 3). The SWEFOT matrix was not tested due to lack of information about the model parameters estimation available in the reference article. Two models were used for each previously published matrix. The first was based on only the published values of the different estimated parameters (intercept and regression coefficients). The second was recalibrated (with a new estimate of the intercept, the regression coefficients

Table 2 Association of the previously published matrix determinants with RRP status on univariate and multivariate analysis

	Univariate analysis		p Value	Multivariate analysis		
	RRP+ (n=41)	RRP- (n=329)		Estimate	SE	p Value
ESPOIR matrix						
Swollen joint count 14–20	6 (14.6)	32 (9.7)	0.08	0.28	0.53	0.60
Swollen joint count >20	4 (9.8)	12 (3.7)		1.25	0.68	0.07
CRP<35 mg/L	25 (61.0)	168 (51.1)	0.04	0.83	0.57	0.15
CRP ≥35 mg/L	12 (29.3)	72 (21.9)		0.86	0.63	0.17
ACPA positivity	31 (75.6)	154 (46.8)	0.0007	1.11	0.40	0.006
Typical RA erosion	18 (43.9)	48 (14.6)	<0.0001	1.31	0.37	0.0004
mESPOIR matrix						
Swollen joint count >13	10 (24.4)	44 (13.4)	0.09	0.89	0.46	0.05
ACPA positivity	31 (75.6)	154 (46.8)	0.0007	1.22	0.41	0.003
Baseline SHS 5–14	19 (46.3)	48 (14.6)	<0.0001	2.15	0.41	<0.0001
Baseline SHS>14	10 (24.4)	18 (5.5)		2.34	0.51	<0.0001
SONORA matrix						
DAS28≤3.2	2 (4.9)	14 (4.3)	0.69	0.41	0.87	0.64
ACPA≤20 UI/L	10 (24.4)	170 (51.7)	0.001	-1.09	0.40	0.007
Baseline vSHS=0	3 (7.3)	112 (34.0)	<0.0001	-2.80	0.67	<0.0001
Baseline vSHS=1–5	9 (21.9)	151 (45.9)		-1.97	0.47	<0.0001
Baseline vSHS=6–10	13 (31.7)	32 (9.7)		-0.06	0.46	0.90
ASPIRE matrix						
Swollen joint count 10–17	11 (26.8)	79 (24.0)	0.72	0.25	0.40	0.54
Swollen joint count ≥17	4 (9.8)	26 (7.9)		0.45	0.60	0.45
CRP 6–30 mg/L	21 (51.2)	134 (40.7)	0.09	0.92	0.48	0.06
CRP≥30 mg/L	13 (31.7)	84 (25.5)		1.41	0.61	0.02
ESR 21–50 mm/1st h	18 (43.9)	107 (32.5)	0.37	-0.07	0.42	0.86
ESR ≥50 mm/1st h	8 (19.5)	81 (24.6)		-1.03	0.60	0.09
RF 80–200 UI	6 (14.6)	23 (7.0)	0.02	1.08	0.52	0.04
RF≥200 UI	10 (24.4)	42 (12.8)		0.97	0.43	0.02
BEST matrix						
CRP 10–35 mg/L	13 (31.7)	108 (32.8)	0.55	0.03	0.43	0.95
CRP≥35 mg/L	12 (29.3)	72 (21.9)		0.11	0.44	0.81
RF+ or ACPA+	4 (9.8)	51 (15.5)	0.007	0.05	0.65	0.94
RF+ and ACPA+	28 (68.3)	139 (42.2)		0.94	0.43	0.03
Radiographic erosions 1–3	8 (19.5)	123 (37.4)	<0.0001	0.75	0.63	0.23
Radiographic erosions ≥4	29 (70.7)	73 (22.2)		2.48	0.56	<0.0001
SWEFOT matrix						
Current smoker	14 (34.1)	158 (48.0)	0.09	-0.51	0.36	0.15
CRP 10–35 mg/L	13 (31.7)	108 (32.8)	0.55	0.03	0.41	0.94
CRP≥35 mg/L	12 (29.3)	72 (21.9)		0.32	0.42	0.46
Radiographic erosions	18 (43.9)	48 (14.6)	<0.0001	1.47	0.35	<0.0001

ACPA, anti-citrullinated protein antibody; CRP, C reactive protein; DAS28, Disease Activity Score in 28 joints; ESR, erythrocyte sedimentation rate; RA, rheumatoid arthritis; RF, rheumatoid factor; RRP, rapid radiographic progression; vSHS, van der Heijde-modified Sharp score.

remaining the same) on a subsample of the ESPOIR cohort population, representing one-third of the total population, to take into account a possible systematic error of prediction. Then their performance was estimated on the remaining two-thirds of the population.

The likelihood of the fit of models was assessed by the Bayesian information criteria. The overall significance of the model was assessed by computation of the Nagelkerke pseudo- R^2 . The calibration of the model was assessed by the Hosmer-Lemeshow goodness-of-fit test by comparing the expected and observed event rates in subgroups of the population (deciles); a significant p value indicated that the model did not fit the observed data. The calibration assessed the degree of agreement

between the predicted and observed probability. The discriminatory property was tested by the area under the receiver operating characteristic curve (AUC) of sensitivity versus 1-specificity as a predictive model to identify cases (RRP+) and non-cases (RRP-). The 95% CI of the AUC was estimated with a bootstrap procedure based on 1000 replications. The mean predicted probability was calculated for cases (RRP+) and non-cases (RRP-). The discriminatory ability of the models was tested pairwise for each model relative to the ESPOIR model by the integrated discrimination improvement statistic calculated as: $(\text{average predicted probability cases} - \text{average predicted probability controls})_{\text{new model}} - (\text{average predicted probability cases} - \text{average predicted probability controls})_{\text{original ESPOIR model}}$

Table 3 Likelihood and discriminative properties of different models

Model	BIC	R ² Nagelkerke	Hosmer- Lemeshow test	AUC (95% CI)	Mean predicted probability cases/non-cases	Integrated discrimination improvement (p value)
ESPOIR	265.89	0.17	0.95910	0.75 (0.70 to 0.84)	0.20/0.10	
mESPOIR	239.59	0.30	0.54740	0.82 (0.76 to 0.89)	0.29/0.09	-0.101 (p=0.3)
SONORA	264.55	0.00	0.00709	0.76 (0.68 to 0.83)	0.33/0.22	-0.008 (p=0.5)
SONORA _{recalculated}	262.76	0.00	0.00019	0.76 (0.69 to 0.83)	0.06/0.04	0.079 (p=0.7)
BEST	1234.82	0.00	0.00023	0.73 (0.65 to 0.81)	0.90/0.78	-0.010 (p=0.5)
BEST _{recalibrated}	449.76	0.00	0.00008	0.73 (0.64 to 0.80)	0.61/0.40	-0.102 (p=0.4)
ASPIRE ESR	269.79	0.00	0.00635	0.54 (0.45 to 0.64)	0.17/0.15	0.091 (p=0.7)
ASPIRE ESR _{recalibrated}	464.10	0.00	0.00085	0.54 (0.46 to 0.64)	0.45/0.43	0.082 (p=0.6)
ASPIRE CRP	268.91	0.00	0.24880	0.62 (0.54 to 0.69)	0.21/0.18	0.067 (p=0.6)
ASPIRE CRP _{recalibrated}	574.03	0.00	0.00052	0.62 (0.41 to 0.60)	0.59/0.52	0.035 (p=0.6)

Integrated discrimination improvement is calculated as (average predicted probability cases—average predicted probability controls)_{new model}—(average predicted probability cases—average predicted probability controls)_{original ESPOIR model}. A positive value indicates that the new model provides an improvement over the original ESPOIR model.

AUC, area under the receiver operating characteristic curve; BIC, Bayesian Information Criterion; CRP, C reactive protein; ESR, erythrocyte sedimentation rate.

A positive value indicated that the new model provided an improvement over the original ESPOIR model. The p values were calculated as described by Pencina *et al.*³⁶ For each model, the observed probability of RRP was plotted for groups of participants by predicted probability of RRP. The groups in the model were defined by cut-offs for quartiles of predicted probability of RRP; therefore, owing to a large number of equivalent values, the construction of four groups led to large differences in the number of participants in each group. The plots represent the observed proportion of RRP and the 95% CI by group.

All analyses involved the use of R 2.15.2 for Windows (R Foundation, Vienna, Austria).

RESULTS

Patient characteristics

The selected patients have been described elsewhere.²⁶ Briefly, 370 patients (45.5%) from the initial ESPOIR cohort started MTX (n=335, mean dose 17.5 mg/week) or LEF (n=35, mean dose 20 mg/week) as first-line DMARDs, referred to as 'sDMARD-treated patients'. No statistical difference was found for the treatment dose between RRP+ and RRP- patients. Their main characteristics are given in table 1. Among the 126 patients (34.1%) with disease progression (ie, Δ vSHS \geq 1 (SDD)), 41 had RRP (ie, Δ vSHS \geq 5), representing 11.1% of all sDMARD-treated patients and 32.5% of patients with disease progression. Baseline characteristics between RRP+ and RRP- were significantly different for disease duration, which was longer for RRP+ than RRP- patients. Significant differences were found concerning the IgM rheumatoid factor (RF) positivity and the anti-citrullinated protein antibodies (ACPAs) positivity with a more frequent positivity among RRP+. Another significant difference is found concerning the baseline score

vSHS with a higher score among RRP+. Moreover, we do not find any difference concerning the part of the population which fulfils the 2010 classification criteria for RA.

Relevance of the determinants

The RRP predictors identified in the previously published matrices were tested in the ESPOIR sDMARD population (table 2). Substantial differences were noted in terms of the statistical significance of these predictors. The presence of RF—or ACPAs, when tested—and initial structural damage (qualitatively or quantitatively assessed) were always predictive of RRP at 1 year. However, DAS28 was not associated with RRP-positive status, as compared with what was observed in the SONORA cohort. C reactive protein (CRP) level and RRP-positive status were associated significantly with the ASPIRE matrix thresholds, but the association did not persist with the other matrices.

Likelihood of models with the ESPOIR cohort

According to the Bayesian information criterion, not surprisingly, the more likely models were the mESPOIR model (table 3). If we consider only models not developed on the ESPOIR population, we found that SONORA seems to be the more likely model. We found the same results when we analysed the pseudo-R², with the highest pseudo-R² (0.2) for the mESPOIR model. The Hosmer-Lemeshow test indicates the fit of the model to the data: all models had a poor fit except for the two ESPOIR models, which was predictable, and the ASPIRE CRP model.

Performance of models with the ESPOIR cohort

The discriminative power of the matrices, determined by receiver operating characteristic (ROC) analysis, is given in figure 1 and table 3. The BEST and SONORA

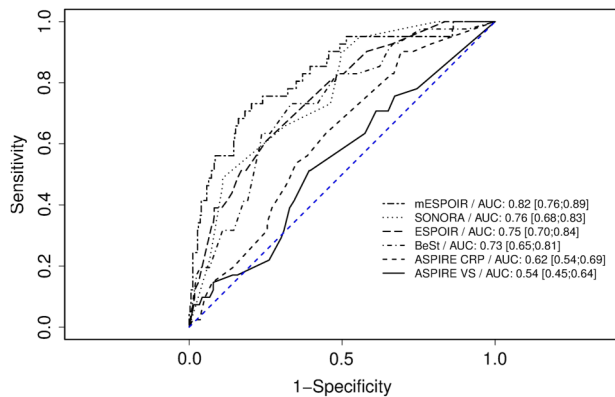


Figure 1 Performance on the different matrices with the ESPOIR patients who initially received methotrexate or leflunomide by ROC curve analysis. Data are area under the ROC curve (95% CIs). AUC, area under the ROC curve; CRP, C reactive protein; ROC, receiver operating characteristic.

matrices showed interesting discriminative capacities with the ESPOIR cohort, with AUC values of 0.73 and 0.76, respectively—close to that of the original ESPOIR matrix. The ASPIRE CRP matrix showed only fair discriminative capacity (0.62) and the ASPIRE erythrocyte sedimentation rate (ESR) matrix only low capacity, with an AUC close to 0.5.

In pairwise comparisons by the integrated discrimination improvement test (table 3), any model compared with the ESPOIR model showed a significantly larger difference in predicted probabilities between cases and non-cases, which suggests no significant better discrimination between cases and non-cases.

The discriminative power was analysed graphically by evolution of the observed proportion of RRP according to the predicted proportion of RRP by the matrix model, stratified by quartiles (figures 2 and 3). As a general rule, a clear gradient should be observed between these two measurements, and as observed in figure 3, this is the case for all matrices with the ESPOIR population, except for the two ASPIRE matrices. Whatever the quartile of predicted proportion of RRP estimated by the BEST matrix, the predicted proportion of RRP in the ESPOIR population was always relatively higher.

DISCUSSION

We assessed the performance of three matrix risk models to predict RRP at 1 year in patients with early RA despite csDMARD therapy. No matrix clearly showed superior performance compared with the others. All matrices could identify cases (RRP-positive) and non-cases, but the ability was lower for matrices developed in a clinical trial population, such as ASPIRE and BEST, than those developed in a clinic-based RA population such as SONORA.

One of the major strengths of the study is the use of the ESPOIR population data,^{32 33} an early arthritis cohort in which most of the participants had early RA

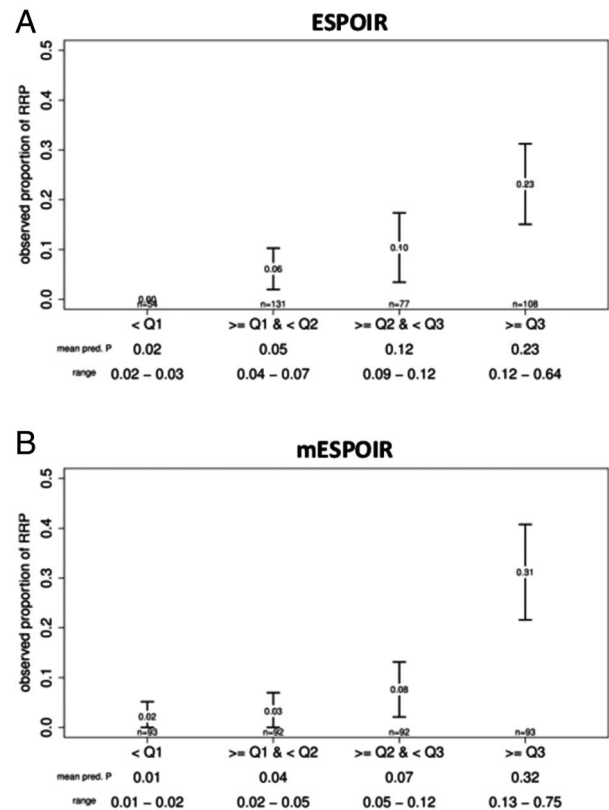


Figure 2 Discriminative power of the two ESPOIR matrices. The figure presents the observed proportions of rapid radiographic progression (RRP) of rheumatoid arthritis by groups of predicted probabilities of RRP for each model. Groups were based on quartiles; error bars represent the 95% CI. Value at the centre of the error bar represents the observed proportion of RPP for a group. The mean predicted proportion of RPP by group and the range are noted at the bottom of the figure.

according to the 1987 American College of Rheumatology (ACR) or 2010 ACR/European League Against Rheumatism (ACR/EULAR) criteria.³⁷ For this study, we focused on patients receiving MTX within the first year of the disease, to be consistent with the trials that were used for matrix development. We also included patients receiving LEF, a drug of similar efficacy, which is recommended along with MTX as a first-line agent by all national or international clinical practice guidelines.^{2 9} For these reasons, the ESPOIR sample perfectly corresponds to the target population of an RRP matrix, that is, patients with early RA receiving sDMARDs in real-life settings for whom assessing whether such a therapy could be suboptimal in terms of structural damage progression is highly relevant clinically. Compared with a previously published matrix validation study performed in an established RA population,²⁷ our study appears more relevant. One other recent study showed only fair performance of the matrices in a single hospital-based early RA cohort; the small sample (n=74) and the low number of RRP+ patients could explain the difference of their results and ours.²⁸

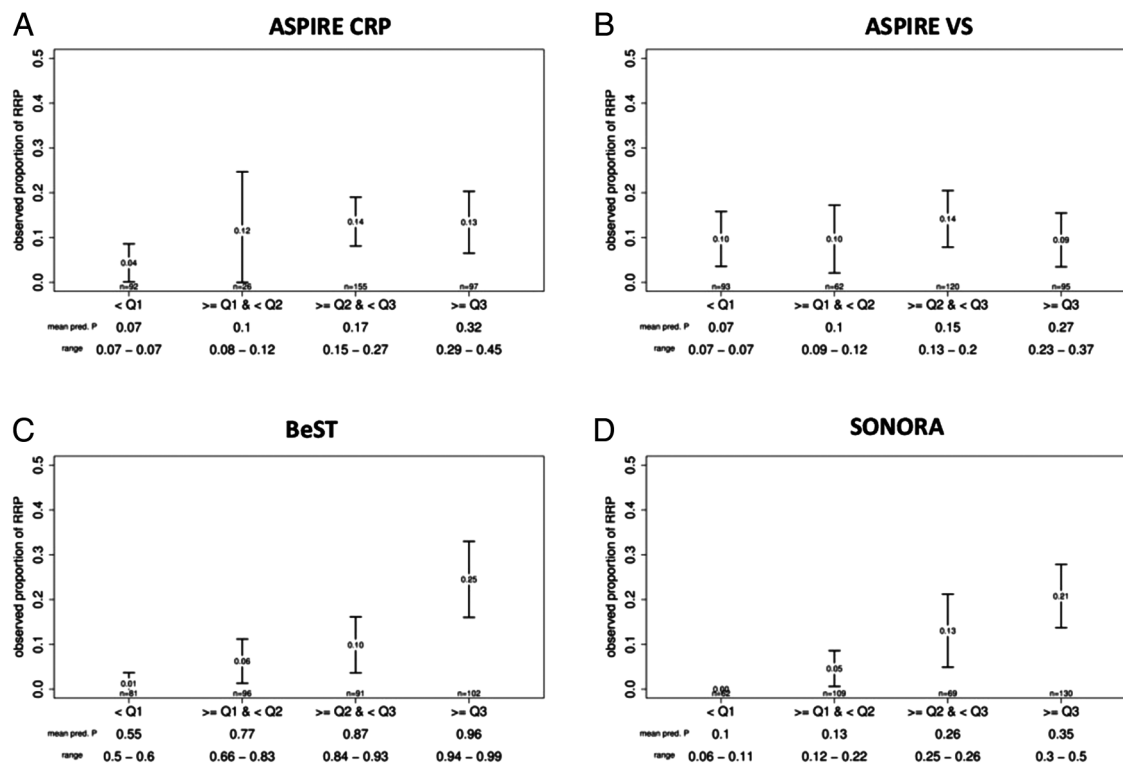


Figure 3 Discriminative power of the ASPIRE, BEST and SONORA matrices. The figure presents the observed proportions of rapid radiographic progression (RRP) of rheumatoid arthritis by groups of predicted probabilities of RRP for each model. Groups were based on quartiles; error bars represent the 95% CI. Value at the centre of the error bar represents the observed proportion of RRP for a group. The mean predicted proportion of RRP by group and the range are noted at the bottom of the figure.

Several statistics were used to compare models. There are two major types of parameters: parameters assessing the fit and parameters assessing the discriminative power. If the ESPOIR and mESPOIR matrices are set aside since they have been developed on the same data as the validation cohort, the most adequate model was the SONORA model with the parameter of fit criteria, and the ASPIRE model according the discriminative properties.

The performance of the matrices varied in the ESPOIR population. Such variations might be explained by several factors. First, the study population was an unselected early RA population with a minimal level of disease activity and no specific severity characteristics or predictors required;^{32 33} this situation resulted in a favourable outcome overall for most of the patients in the ESPOIR cohort.³⁸ As a consequence, the radiographic structural progression in the ESPOIR population was lower—mean 1.6 vSHS points—than in the MTX arms for the SWEFOT, ASPIRE or BEST populations—mean 2.7, 3.7 and 5, respectively.^{16 31 39}

Another source of variation was the nature and framing of the tested RRP determinants. Some just reflect disease parameters at a certain point in time such as inflammation of clinical or biologic parameters, while others reflect more disease intrinsic characteristics such as erosion or autoantibody status. Expectedly, the predictive performance of the latter is better than that of the former at baseline. Although ESR or CRP level, immunologic

status (presence of RF or ACPA) and baseline radiographic findings were always present in the matrix, other variables were not always present, especially swollen joint count. Smoking status was present only in the SWEFOT matrix;²³ this determinant was not predictive of RRP in the ESPOIR cohort and has even been associated with a more favourable outcome and remission.⁴⁰ The variable framing is also a matter of variation. Baseline radiographic characteristics were expressed qualitatively—erosive disease or not—in some matrix tools. Although interesting from a clinical perspective, the framing raises the question of definition and reproducibility of the notions of erosive disease or ‘typical RA erosion’, as mentioned in the 1987 ACR criteria. Recently, van der Heijde *et al.*⁴¹ proposed a definition of erosive disease in early RA based on the number with erosive changes, namely >3 joints. The quantitative expression of structural damage—number of erosions or vSHS value—has been used in other matrix tools and appeared to be highly predictive of further structural damage, with better AUC value and fit than matrices involving qualitative information. This outcome was clearly expected according to the paradigm ‘who has eroded joints will have eroded joints’. However, such quantitative assessment is not performed in daily care, which limits the applicability of these RRP matrices in daily practice.

In addition, the threshold used for each continuous variable is questionable. Several methods have been

used, mainly based on variable distributions, so proposed matrices are too much tailored on the patient population in which they were conceived. This is particularly true for the matrices developed in an RCT population. Another question which limits the applicability of these RRP matrices in daily practice is that X-ray damage is not even seen in 70% of MTX-treated patients, and because X-ray damage is perhaps not the most relevant outcome for patients, the RRP is not a relevant outcome to try to predict. Perhaps we should focus on the prediction of clinical response. However, we should think that variables found to predict clinical response to MTX treatment overlap with those associated with rapid progression. Furthermore, about 10% of patients got RRP when treated with csDMARD monotherapy and hence could benefit from more intense therapy, such as combo DMARDs or bDMARD. Thus, RRP is an important hallmark of aggressive or severe RA.

Potentially interesting determinants have not been tested in the matrices developed to date. Proteomics of genomic biomarkers, such as serum interleukin 6, metalloproteinases, shared epitope or eventually their combination, were not included, although several studies demonstrated their ability to predict disease severity.¹⁹ Recently, a for-profit company proposed a commercial kit of serum biomarkers associated with a patented specific algorithm, showing high correlation with disease activity and severity.^{42–43} Although interesting at the group level and for clinical research, the applicability and cost-effectiveness of such a marketed biomarker kit remains to be confirmed.⁴⁴

Finally, no matrix included early treatment response. The ‘treat to target’ principle⁵ has been widely acknowledged and is included in many clinical practice guidelines.² This situation is reinforced by the results of several trials involving dynamic therapeutic schemes—step-up trials—which reported close efficacy of MTX, combination therapies or biologics for early RA.^{30–45–46} Several studies have shown that early therapeutic response (ie, at 3 or 6 months) is highly predictive of further treatment success or failure. The use of matrices could help avoid MTX failure and be included in future guidelines.

Matrices have limitations. The logistic regression models from which they are derived provide estimates of the probability of having a RRP at one year conditionally of a combination of predictors for every single cell of the matrix; however, since RRP is less frequent with modern therapies, the right-top cells, corresponding to the most at-risk categories, are usually weakly populated, which results in wide CIs of the estimates. Such lack of power deserves further research.

CONCLUSION

Matrices seem important to consider in optimising the care of early RA by identifying patients who will respond poorly to MTX. An alternative option could be biologic

agents; however, their place as first-line RA therapy has been ruled out by several economic evaluations reporting high incremental cost-effectiveness ratios largely above the generally accepted thresholds.^{15–16} Yet another option could be a combination of sDMARDs.^{9–11} Matrix tools could be an adequate option to identify patients with early RA at risk of inadequate response to MTX, who could benefit from more aggressive therapies such as biologic agents, although for now the predicting performance was modest. Although differences between matrices are observed, this external validation study does not allow one to state that one matrix is superior to the others.

Author affiliations

¹Université Pierre et Marie Curie (UPMC)—Paris 6, GRC 08, Pierre Louis Institute of Epidemiology and Public Health, Paris, France

²Department of Biostatistics, Public Health and Medical Information, AP-HP Pitié Salpêtrière Hospital, Paris, France

³Department of Rheumatology, Montpellier I University; Lapeyronie Hospital, Montpellier, France

⁴Rheumatology Department, Rouen University Hospital & INSERM U905, Institute for Research and Innovation in Biomedicine, Rouen University, Rouen, France

⁵Department of Rheumatology, Brest University, La Cavale Blanche University Hospital, Brest, France

⁶INSERM ESPRI, ERI29 Université Bretagne Occidentale, Brest, France

⁷EA 4360 APEMAC, Lorraine University, Paris-Descartes University, Nancy, France

⁸Faculty of Medicine, CS 50184, 54505 Vandoeuvre-lès-Nancy & Inserm, CIC-EC, CHU de Brabois, 54505 Vandoeuvre-lès-Nancy, Nancy, France

⁹Department of Rheumatology, Pitié Salpêtrière Hospital, Paris, France

Twitter Follow Alain Saraux at @alain.saraux

Acknowledgements The authors thank the French rheumatologists who referred their patients to the ESPOIR cohort in the following rheumatology departments: Amiens (P. Fardellone, P. Boumier), Bordeaux (T. Schaevebeke), Brest (A. Saraux), Lille (R.M. Flipo), Montpellier (B. Combe), Paris-Bicêtre (X. Mariette), Paris-Bichat (O. Meyer), Paris-Cochin (M. Dougados), Paris-La Pitié (B. Fautrel), Paris-St Antoine (F. Berenbaum), Rouen (X. Le Loët, O. Vittecoq), Strasbourg (J. Sibilia), Toulouse (A. Cantagrel) and Tours (P. Goupille). They are grateful to N. Rincheval for data management and expert monitoring; to S. Martin for performing all the centralised assays of CRP, IgM rheumatoid factors and anti-CCP antibodies; and to S. Harvard and L. Smales for translation and copyediting. An unrestricted grant from Merck Sharp and Dohme (MSD) was allocated for the first 5 years of the ESPOIR cohort study. Two additional grants from the INSERM were obtained to support part of the biological database. The French Society for Rheumatology, Abbott, Amgen and Wyeth also supported the ESPOIR cohort study. They also thank Drs or Prs N Vastesaeger, G Trape, CR Allaart, K Visser, J Smolen, D Aletaha, P Verschuere, C Bombardier, R Van Vollenhoven, S Saevarsdottir and S Lillegraven for thoughtful discussions about the RRP prediction matrices.

Contributors BG has carried out the analysis and interpretation of data and is one of the two main contributors involved in drafting the manuscript. BF has made substantial contributions to the conception, design and acquisition of data and is one of the two main contributors involved in drafting of the manuscript. BC, XLL, AS and FG have made substantial contributions to the conception and design, acquisition of data, and carried out a critical revision of the paper. All the authors read and approved the final manuscript.

Funding Institutional support: unrestricted grant from: - GERCER (Groupe d'Etudes et de Recherches Cliniques En Rhumatologie), Paris, France. - Merck/MSD, USA.

Competing interests None declared.

Patient consent Obtained.

Ethics approval Ethics Committee of Montpellier University Hospital, France.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

REFERENCES

- Emery P, Breedveld FC, Dougados M, *et al*. Early referral recommendation for newly diagnosed rheumatoid arthritis: evidence based development of a clinical guide. *Ann Rheum Dis* 2002;61:290–7.
- Smolen JS, Landewé R, Breedveld FC, *et al*. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. *Ann Rheum Dis* 2014;73:492–509.
- Gaujoux-Viala C, Gossec L, Cantagrel A, *et al*. Recommendations of the French Society for Rheumatology for managing rheumatoid arthritis. *Jt Bone Spine Rev Rhum* 2014;81:287–97.
- Grigor C, Capell H, Stirling A, *et al*. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. *Lancet* 2004;364:263–9.
- Smolen JS, Aletaha D, Bijlsma JW, *et al*. Treating rheumatoid arthritis to target: recommendations of an international task force. *Ann Rheum Dis* 2010;69:631–7.
- Lukas C, Combe B, Ravaud P, *et al*. Favorable effect of very early disease-modifying antirheumatic drug treatment on radiographic progression in early inflammatory arthritis: Data from the Etude et Suivi des polyarthrites indifférenciées récentes (study and followup of early undifferentiated polyarthritis). *Arthritis Rheum* 2011;63:1804–11.
- Schoels M, Knevel R, Aletaha D, *et al*. Evidence for treating rheumatoid arthritis to target: results of a systematic literature search. *Ann Rheum Dis* 2010;69:638–43.
- Le Loët X, Berthelot JM, Cantagrel A, *et al*. Clinical practice decision tree for the choice of the first disease modifying antirheumatic drug for very early rheumatoid arthritis: a 2004 proposal of the French Society of Rheumatology. *Ann Rheum Dis* 2006;65:45–50.
- Gaujoux-Viala C, Nam J, Ramiro S, *et al*. Efficacy of conventional synthetic disease-modifying antirheumatic drugs, glucocorticoids and tofacitinib: a systematic literature review informing the 2013 update of the EULAR recommendations for management of rheumatoid arthritis. *Ann Rheum Dis* 2014;73:510–15.
- Boers M, Verhoeven AC, Markusse HM, *et al*. Randomised comparison of combined step-down prednisolone, methotrexate and sulphasalazine with sulphasalazine alone in early rheumatoid arthritis. *Lancet* 1997;350:309–18.
- Moreland LW, O'Dell JR, Paulus HE, *et al*. A randomized comparative effectiveness study of oral triple therapy versus etanercept plus methotrexate in early aggressive rheumatoid arthritis: the treatment of Early Aggressive Rheumatoid Arthritis Trial. *Arthritis Rheum* 2012;64:2824–35.
- Kuriya B, Arkema EV, Bykerk VP, *et al*. Efficacy of initial methotrexate monotherapy versus combination therapy with a biological agent in early rheumatoid arthritis: a meta-analysis of clinical and radiographic remission. *Ann Rheum Dis* 2010;69:1298–304.
- Nam JL, Ramiro S, Gaujoux-Viala C, *et al*. Efficacy of biological disease-modifying antirheumatic drugs: a systematic literature review informing the 2013 update of the EULAR recommendations for the management of rheumatoid arthritis. *Ann Rheum Dis* 2014;73:516–28.
- Finckh A, Bansback N, Marra CA, *et al*. Treatment of very early rheumatoid arthritis with symptomatic therapy, disease-modifying antirheumatic drugs, or biologic agents: a cost-effectiveness analysis. *Ann Intern Med* 2009;151:612–21.
- van der Velde G, Pham B, Machado M, *et al*. Cost-effectiveness of biologic response modifiers compared to disease-modifying antirheumatic drugs for rheumatoid arthritis: a systematic review. *Arthritis Care Res (Hoboken)* 2011;63:65–78.
- Vastesaeger N, Xu S, Aletaha D, *et al*. A pilot risk model for the prediction of rapid radiographic progression in rheumatoid arthritis. *Rheumatology (Oxford)* 2009;48:1114–21.
- Sharp JT, Wolfe F, Mitchell DM, *et al*. The progression of erosion and joint space narrowing scores in rheumatoid arthritis during the first twenty-five years of disease. *Arthritis Rheum* 1991;34:660–8.
- van den Broek M, Dirven L, de Vries-Bouwstra JK, *et al*. Rapid radiological progression in the first year of early rheumatoid arthritis is predictive of disability and joint damage progression during 8 years of follow-up. *Ann Rheum Dis* 2012;71:1530–3.
- Tobón G, Sarau A, Lukas C, *et al*. First-year radiographic progression as a predictor of further progression in early arthritis: results of a large national French cohort. *Arthritis Care Res (Hoboken)* 2013;65:1907–15.
- Conroy RM, Pyorala K, Fitzgerald AP, *et al*. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J* 2003;24:987–1003.
- Kanis JA, Johnell O, Oden A, *et al*. FRAX and the assessment of fracture probability in men and women from the UK. *Osteoporos Int* 2008;19:385–97.
- Visser K, Goekoop-Ruiterman YP, de Vries-Bouwstra JK, *et al*. A matrix risk model for the prediction of rapid radiographic progression in patients with rheumatoid arthritis receiving different dynamic treatment strategies: post hoc analyses from the BeSt study. *Ann Rheum Dis* 2010;69:1333–7.
- Saevarsdottir S, Rezaei H, Geborek P, *et al*, SWEFOT study group. Current smoking status is a strong predictor of radiographic progression in early rheumatoid arthritis: results from the SWEFOT trial. *Ann Rheum Dis* 2015;74:1509–14.
- Bombardier C, Li X, Gregersen P, *et al*. A risk model for the prediction of radiographic progression: results from SONORA study. *Arthritis Rheum* 2009;61(Suppl):S276.
- Dumez A, Vanderschueren G, Lateur L, *et al*. Effectiveness of initial treatment allocation based on expert opinion for prevention of rapid radiographic progression in daily practice of an early RA cohort. *Ann Rheum Dis* 2011;70:634–7.
- Fautrel B, Granger B, Combe B, *et al*. Matrix to predict rapid radiographic progression of early rheumatoid arthritis patients from the community treated with methotrexate or leflunomide: results from the ESPOIR cohort. *Arthritis Res Ther* 2012;14:R249.
- Lillegraven S, Prince F, Shadick N, *et al*. The Performance of Matrix-Based Risk Models for Rapid Radiographic Progression in An Observational Cohort of Established Rheumatoid Arthritis Patients. *Arthritis Rheum* 2011;63(Suppl):S1010.
- De Cock D, Vanderschueren G, Meyfroidt S, *et al*. The performance of matrices in daily clinical practice to predict rapid radiologic progression in patients with early RA. *Semin Arthritis Rheum* 2014;43:627–31.
- St Clair EW, van der Heijde DM, Smolen JS, *et al*. Combination of infliximab and methotrexate therapy for early rheumatoid arthritis: a randomized, controlled trial. *Arthritis Rheum* 2004;50:3432–43.
- van Vollenhoven RF, Ernestam S, Geborek P, *et al*. Addition of infliximab compared with addition of sulfasalazine and hydroxychloroquine to methotrexate in patients with early rheumatoid arthritis (SWEFOT trial): 1-year results of a randomised trial. *Lancet* 2009;374:459–66.
- Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Allaart CF, *et al*. Comparison of treatment strategies in early rheumatoid arthritis: a randomized trial. *Ann Intern Med* 2007;146:406–15.
- Combe B, Benessiano J, Berenbaum F, *et al*. The ESPOIR cohort: a ten-year follow-up of early arthritis in France: methodology and baseline characteristics of the 813 included patients. *Joint Bone Spine* 2007;74:440–5.
- Combe B, Rinciveau N. Early lessons from the recent-onset rheumatoid arthritis cohort ESPOIR. *Joint Bone Spine* 2015;82:13–17.
- Prevo ML, van 't Hof MA, Kuper HH, *et al*. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;38:44–8.
- Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the Health Assessment Questionnaire, Disability and Pain Scales. *J Rheumatol* 1982;9:789–93.
- Pencina MJ, D'Agostino RBSr, D'Agostino RB Jr, *et al*. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;27:157–72; discussion 207–12.
- Fautrel B, Combe B, Rinciveau N, *et al*. Level of agreement of the 1987 ACR and 2010 ACR/EULAR rheumatoid arthritis classification criteria: an analysis based on ESPOIR cohort data. *Ann Rheum Dis* 2012;71:386–9.

38. Combe B, Rincheval N, Benessiano J, *et al.* Five-year favorable outcome of patients with early rheumatoid arthritis in the 2000s: data from the ESPOIR cohort. *J Rheumatol* 2013;40:1650–7.
39. Rezaei H, Saevarsdottir S, Forslind K, *et al.* In early rheumatoid arthritis, patients with a good initial response to methotrexate have excellent 2-year clinical outcomes, but radiological progression is not fully prevented: data from the methotrexate responders population in the SWEFOT trial. *Ann Rheum Dis* 2011;71:186–91.
40. Vesperini V, Lukas C, Fautrel B, *et al.* Association of tobacco exposure and reduction of radiographic progression in early rheumatoid arthritis: results from a French multicenter cohort. *Arthritis Care Res (Hoboken)* 2013;65:1899–906.
41. van der Heijde D, van der Helm-van Mil AH, Aletaha D, *et al.* EULAR definition of erosive disease in light of the 2010 ACR/EULAR rheumatoid arthritis classification criteria. *Ann Rheum Dis* 2013;72:479–81.
42. Eastman PS, Manning WC, Qureshi F, *et al.* Characterization of a multiplex, 12-biomarker test for rheumatoid arthritis. *J Pharm Biomed Anal* 2012;70:415–24.
43. Li W, Sasso EH, Emerling D, *et al.* Impact of a multi-biomarker disease activity test on rheumatoid arthritis treatment decisions and therapy use. *Curr Med Res Opin* 2013;29:85–92.
44. Smolen JS, Aletaha D. Forget personalised medicine and focus on abating disease activity. *Ann Rheum Dis* 2013;72:3–6.
45. Allaart CF, Goekoop-Ruiterman YP, de Vries-Bouwstra JK, *et al.* FARR study group. Aiming at low disease activity in rheumatoid arthritis with initial combination therapy or initial monotherapy strategies: the BeSt study. *Clin Exp Rheumatol* 2006;24(Suppl 43):S–77–82.
46. Soubrier M, Puechal X, Sibilia J, *et al.* Evaluation of two strategies (initial methotrexate monotherapy vs its combination with adalimumab) in management of early active rheumatoid arthritis: data from the GUEPARD trial. *Rheumatology (Oxford)* 2009;48:1429–34.