

# Predicting the outcome of ankylosing spondylitis therapy

Nathan Vastesaeger,<sup>1</sup> Désirée van der Heijde,<sup>2</sup> Robert D Inman,<sup>3</sup> Yanxin Wang,<sup>1</sup> Atul Deodhar,<sup>4</sup> Benjamin Hsu,<sup>5</sup> Mahboob U Rahman,<sup>6</sup> Ben Dijkmans,<sup>7</sup> Piet Geusens,<sup>8</sup> Bert Vander Cruyssen,<sup>9</sup> Eduardo Collantes,<sup>10</sup> Joachim Sieper,<sup>11</sup> Jürgen Braun<sup>12</sup>

► Supplementary tables and figures are published online only. To view these files please visit the journal online (<http://ard.bmj.com>).

<sup>1</sup>Schering-Plough Corporation, Kenilworth, New Jersey, USA (Currently Merck, Whitehouse Station, New Jersey, USA)

<sup>2</sup>Leiden University Medical Center, Leiden, The Netherlands

<sup>3</sup>University of Toronto, Toronto, Canada

<sup>4</sup>Oregon Health and Science University, Portland, Oregon, USA

<sup>5</sup>Centocor Research & Development, Malvern, Pennsylvania, USA

<sup>6</sup>Formerly at Centocor Research & Development, Malvern, Pennsylvania, USA (Currently Pfizer, Collegeville, Pennsylvania, USA)

<sup>7</sup>VU Medical Centre, Amsterdam, The Netherlands

<sup>8</sup>University Hasselt, Hasselt, Belgium

<sup>9</sup>Ghent University Hospital, Ghent, Belgium

<sup>10</sup>Reina Sofia Hospital and University Cordoba, Cordoba, Spain

<sup>11</sup>Charite Hospital Berlin, Berlin, Germany

<sup>12</sup>Rheumazentrum Ruhrgebiet, Herne, Germany

## Correspondence to

Dr Nathan Vastesaeger, Allewaertstraat 13, 2000 Antwerpen, Belgium; [nathan.vastesaeger@merck.com](mailto:nathan.vastesaeger@merck.com)

Accepted 22 January 2011

Published Online First

14 March 2011



This paper is freely available online under the BMJ Journals unlocked scheme, see <http://ard.bmj.com/info/unlocked.dtl>

## ABSTRACT

**Objectives** To create a model that provides a potential basis for candidate selection for anti-tumour necrosis factor (TNF) treatment by predicting future outcomes relative to the current disease profile of individual patients with ankylosing spondylitis (AS).

**Methods** ASSERT and GO-RAISE trial data (n=635) were analysed to identify baseline predictors for various disease-state and disease-activity outcome instruments in AS. Univariate, multivariate, receiver operator characteristic and correlation analyses were performed to select final predictors. Their associations with outcomes were explored. Matrix and algorithm-based prediction models were created using logistic and linear regression, and their accuracies were compared. Numbers needed to treat were calculated to compare the effect size of anti-TNF therapy between the AS matrix subpopulations. Data from registry populations were applied to study how a daily practice AS population is distributed over the prediction model.

**Results** Age, Bath ankylosing spondylitis functional index (BASFI) score, enthesitis, therapy, C-reactive protein (CRP) and HLA-B27 genotype were identified as predictors. Their associations with each outcome instrument varied. However, the combination of these factors enabled adequate prediction of each outcome studied. The matrix model predicted outcomes as well as algorithm-based models and enabled direct comparison of the effect size of anti-TNF treatment outcome in various subpopulations. The trial populations reflected the daily practice AS population.

**Conclusion** Age, BASFI, enthesitis, therapy, CRP and HLA-B27 were associated with outcomes in AS. Their combined use enables adequate prediction of outcome resulting from anti-TNF and conventional therapy in various AS subpopulations. This may help guide clinicians in making treatment decisions in daily practice.

Ankylosing spondylitis (AS) is characterised by back pain caused by inflammation of the sacroiliac joints and spine. The management of AS includes non-pharmacological, pharmacological, invasive and surgical interventions that should be tailored to each patient's disease manifestations, current symptoms, clinical findings and prognostic indicators.<sup>1</sup> Non-steroidal anti-inflammatory drugs (NSAID) are recommended as first-line pharmacological treatment, and anti-tumour necrosis factor (TNF) agents are recommended in the case of NSAID failure.<sup>2-6</sup>

Predictors of response to therapy may enable improved patient selection, outcomes and resource

utilisation.<sup>7-8</sup> The recommendations for anti-TNF use in AS are, however, based primarily on inadequate response to conventional therapies and less on the expectation that an anti-TNF agent will be effective in a particular patient.<sup>2</sup> The literature continues to establish predictors of response,<sup>9-14</sup> which are also associated with anti-TNF use in AS.<sup>15</sup> Ideally, these may help clinicians to make evidence-based decisions that maximise the benefits from treatment by targeting subsets of patients most likely to respond;<sup>16</sup> however, single predictors are too weak to be useful for decision-making in the individual patient.

This paper describes the predictor selection and construction of a model that identifies AS subpopulations likely to respond optimally to anti-TNF therapy. In the absence of a 'hard outcome' parameter that can be predicted in AS, such as mortality in cardiovascular disease, the ability and robustness of the predictor model to predict the results of a variety of AS outcome instruments were explored.

In addition, the distribution of AS registry populations encountered in daily rheumatology practice over the prediction model was evaluated.

## PATIENTS AND METHODS

This is a post-hoc analysis of the ASSERT and GO-RAISE trials in adult patients with active AS despite NSAID or disease-modifying antirheumatic drugs (DMARD) and naive to anti-TNF therapy.

In ASSERT, patients were randomly assigned to receive infusions of placebo or 5 mg/kg infliximab at weeks 0, 2, 6, 12 and 18 and were allowed to receive concurrent NSAID but not DMARD or systemic corticosteroids.<sup>5</sup> In GO-RAISE, patients were randomly assigned to receive subcutaneous injections of placebo or 50 or 100 mg golimumab every 4 weeks and could continue concurrent NSAID, DMARD and systemic corticosteroids. For our analysis, week 16 data from GO-RAISE were carried forward to week 24 for placebo patients who received golimumab starting at week 16.<sup>4</sup> Week 24 data were collected between November 2002 and September 2003 in ASSERT and between December 2005 and May 2007 in GO-RAISE.

## Outcome instruments

The Bath ankylosing spondylitis disease activity index (BASDAI) score measures disease activity based on six questions on fatigue, spinal pain,

joint pain/swelling, areas of localised tenderness and morning stiffness.<sup>17</sup> BASDAI50 response is defined as a 50% or greater improvement in the BASDAI score.

Assessment of spondyloarthritis (ASAS) 20 response is an improvement of 20% or more in the patient global assessment (PGA), patient assessment of pain, Bath ankylosing spondylitis functional index (BASFI) score and assessment of inflammation. ASAS partial remission is achieved when the value of each of these domains is less than 2 cm on a 10-cm visual analogue scale.<sup>18</sup>

The ankylosing spondylitis disease activity score (ASDAS) measures disease activity state using an algorithm comprising assessment of back pain, morning stiffness duration, joint pain/swelling, PGA and C-reactive protein (CRP).<sup>19, 20</sup> Clinically important and major ASDAS improvements are defined as a decrease of 1.1 units or more and 2.0 units or more, respectively. ASDAS less than 1.3 is the threshold for an inactive disease state.<sup>21</sup>

The association of the following characteristics at baseline with BASDAI50 response and partial remission was studied: age, gender, HLA-B27 status, disease duration, CRP, BASFI, Bath ankylosing spondylitis metrology index (BASMI) score, chest expansion, intermalleolar distance, tragus to wall distance, modified Schobers index, lateral spinal flexion, cervical rotation, PGA, pain assessment, BASDAI, inflammation score, Berlin enthesitis score index and treatment group. MRI, x-rays of the spine and peripheral joint counts were not available for the analysis.

## Statistics

### Study population

The ASSERT and the GO-RAISE datasets were summarised using means±SD and were also combined into a third dataset.

### Outcome predictor selection

Predictors of week 12 BASDAI50 response and week 24 partial remission were identified by comparing the values of the aforementioned baseline characteristics between responders and non-responders and between remitters and non-remitters using Student's *t* test and  $\chi^2$  tests. Variables that differed at  $p=0.1$  were explored further.

Multivariate regression and stepwise selection procedures were used to narrow the number of predictors. The area under the receiver operating characteristics curve (ROC-AUC) and the maximum rescaled  $R^2$  were calculated. The ROC-AUC measures the accuracy of a prediction model as: 90–100% excellent prediction; 80–90% good prediction; 70–80% fair prediction; 60–70% poor prediction and 50–60% failed prediction.<sup>22</sup> The  $R^2$  compares how competing models fit the dataset.<sup>23</sup>

Spearman correlation coefficients were calculated for continuous baseline characteristics, and associations between variables were explored. A variable was selected for the final prediction model if it was retained in stepwise selection in any dataset and for either BASDAI50 response or the partial remission model, provided it did not have a correlation coefficient of 0.4 or greater with another variable. Final predictors were categorised into tertiles or according to a clinically relevant threshold in the matrix model.

### Associations of predictors with outcomes

Associations of predictor variables with BASDAI50, ASAS20, ASDAS clinically important and major improvement, ASAS partial remission and ASDAS inactive disease state were explored using OR and 95% CI of outcomes relative to the categorised predictor variables. OR was interpreted as: 1.5 to 1 weak

association; 2.5 to 1 moderate association; 4 to 1 strong association and 10 to 1 very strong association.<sup>24</sup>

### Matrix model construction

Fitted logistic regression was used to calculate the predicted proportion of patients meeting the outcome criterion according to each subpopulation's value category for the predictors at baseline. These results were organised into a matrix model showing increasing predicted rates of achieving each outcome from left to right, bottom to top.<sup>25</sup> Patient subpopulations with high predicted outcome rates are shown in yellow, those with low rates are shown in red, and those with intermediate rates are shown in orange. The numbers needed to treat (NNT) to realise a target beneficial outcome following anti-TNF treatment was calculated as follows:  $NNT=1/(\text{predicted outcome rate with anti-TNF} - \text{predicted outcome rate with conventional therapy})$  and are presented in matrix models using a white, grey and black colour scheme.

### Algorithm-based models

For each outcome instrument, logistic regression with stepwise selection was used to calculate the model yielding the highest ROC-AUC and  $R^2$  using numeric values for CRP, BASFI, age and enthesitis score; categorical values for treatment and HLA-B27 genotype and their interaction terms. In a similar approach using linear regression, models predicting week 12 ASDAS and BASDAI scores were also calculated. The multiple correlation coefficient (*R*), which represents the correlation between the observed and the predicted values, and the  $R^2$  were calculated, with *R* of 0.1 or less being 'small', *R* of 0.1–0.3 being 'medium', and *R* of 0.3–0.5 being 'large'.<sup>26</sup> The predicted versus the observed change in ASDAS and BASDAI scores were plotted.

### Distribution of two registry AS populations over the prediction model

The ASPECT and the Regisponser studies<sup>15, 27, 28</sup> conducted in 2004–5 in Belgium and Spain, respectively, were used to study the distribution of a daily practice AS population over the model. Cross-sectional data were used from AS patients who had complete data for BASDAI, BASFI, CRP, the presence of enthesitis, age and HLA-B27 status. The percentage of the ASPECT/Regisponser populations falling within each of the predictor value categories in the matrix model is shown for all patients, irrespective of BASDAI score (total registry population), and only for patients with a BASDAI score of 4 or greater (active registry population). The proportion of registry patients corresponding with the NNT categories in the matrix models for various outcome instruments is reported. The OR of BASDAI50 response in the combined dataset were compared with those reported for AS populations treated with anti-TNF therapy in clinical practice.<sup>10–12</sup>

## RESULTS

### Study population

Four hundred and seventy-nine patients treated with anti-TNF agents and 156 treated with placebo in ASSERT or GO-RAISE were included. The characteristics of the datasets are presented in table 1. The mean (SD) ASDAS at baseline was 4.0 (0.8) and median ASDAS (IQR) was 3.9 (3.4–4.5).

### Outcome predictor selection

Age, CRP, HLA-B27, PGA, BASFI, BASDAI, BASMI, cervical rotation, tragus to wall distance, intermalleolar distance,

Berlin enthesitis score and treatment differed significantly ( $p < 0.1$ ) between BASDAI50 responders and non-responders and between partial remitters and non-remitters in ASSERT, GO-RAISE or the combined dataset (see supplementary table 1, available online only).

These variables were further investigated. In stepwise multiple regression analysis (table 2), age, BASFI, enthesitis score, CRP, HLA-B27 and treatment were identified as predictors of BASDAI50 response and ASAS partial remission. BASMI and cervical rotation were identified as predictors of partial remission but not of BASDAI50 response (table 2). High correlation was observed between BASMI, its subcomponents and BASFI

scores but not when other variables were compared (see supplementary table 2, available online only).

Age and BASFI score were significantly higher for HLA-B27-negative than HLA-B27-positive patients, but numeric differences were small and not clinically significant (see supplementary table 3, available online only).

Due to the high correlation between BASMI and BASFI and to limit the total number of predictors to six (which is a reasonable maximum, considering the total number of patients included in the analysis;  $n = 635$ ), BASMI and cervical rotation were not retained in the final model. Age, BASFI, CRP, enthesitis score, treatment and HLA-B27 were retained in at least one of the different stepwise selection models and were therefore retained in the final model.

BASFI was categorised into 4.5 or less (35%), 4.5–6.5 (31%) and over 6.5 (34% of patients). CRP was categorised into 0.6 mg/dl or less (corresponding with the upper limit of normal (ULN) 32%), ULN to 2 mg/dl (34%) and over 2 mg/dl (33%). An age cut-off of 40 years yielded the highest ROC-AUC; 46% of patients were 40 years old or less and 54% were over 40 years old. Enthesitis was present (enthesitis score  $> 0$ ) in 64% and absent (enthesitis score 0) in 36%. Additional information leading to the selection of age and enthesitis categories is provided in supplementary table 4, available online only.

The ROC-AUC and  $R^2$  of the different models presented in table 2 indicate that the accuracy of the predicted BASDAI50 response and predicted partial remission was similar when models with many predictor variables were compared with models with few variables. In addition, they show that the final set of predictors predicts BASDAI50 response and partial remission in the three datasets reasonably well. The relationship between the week 12 BASDAI50 response and week 24 partial remission is shown in supplementary table 5, available online only.

**Associations of predictor variables with outcomes**

The OR (95% CI) of achieving an outcome relative to the value category of a predictor variable is presented in table 3.

HLA-B27 was more strongly associated with large improvements and disease states (BASDAI50, ASDAS major improvement, ASAS partial remission, ASDAS inactive disease) than

**Table 1** Characteristics at baseline of the ASSERT, the GO-RAISE and the combined datasets

Variables at baseline	ASSERT	GO-RAISE	Combined
N	279	356	635
Male (%)	80.6	71.6	75.6
Age (years)	39.8±10.2	39.3±12.1	39.5±11.3
BASDAI score	6.4±1.6	6.7±1.5	6.6±1.5
BASFI score	5.8±2.0	5.1±2.4	5.4±2.2
BASMI score	4.0±2.1	3.5±2.2	3.7±2.1
Cervical rotation (°)	45.8±21.9	48.8±20.3	47.5±21.0
Intermalleolar distance (cm)	95.3±30.6	100.8±24.8	98.3±27.7
Lateral flexion (cm)	11.2±11.0	11.0±5.8	11.1±8.5
Tragus to wall distance (cm)	17.0±6.3	14.2±5.7	15.4±6.1
CRP (mg/dl)	2.4±2.7	1.8±2.0	2.1±2.4
Enthesitis, presence of (%)	63.3	63.7	63.5
HLA-B27 positive (%)	87.1	83.4	85.0
PGA (cm)	6.8±1.8	7.0±1.8	6.9±1.8
Concomitant medications for placebo patients			
NSAID (%)	90.7	92.3	–
Methotrexate (%) (mean dose (mg/week))	–	19.2 (14.2)	–
Sulphasalazine (%) (mean dose (g/day))	–	30.8 (1.8)	–
Oral corticosteroids (%) (mean dose (mg/day))	–	16.7 (7.2)	–
Anti-inflammatory/antirheumatic drugs (%)	38.5	–	–

Values are mean±SD for continuous variables or percentages for categorical variables unless otherwise specified. BASDAI, Bath ankylosing spondylitis disease activity index; BASFI, Bath ankylosing spondylitis functional index; BASMI, Bath ankylosing spondylitis metrology index; CRP, C-reactive protein; NSAID, non-steroidal anti-inflammatory drug; PGA, patient global assessment.

**Table 2** Multivariate regression models for BASDAI50 response and ASAS partial remission using variables with  $p < 0.1$  on univariate screening, stepwise selection and final selection of variables for both the ASSERT and the GO-RAISE datasets

	ASSERT			GO-RAISE			Combined			
	All	Select	Final	All	Select	Final	All	Select	Final	Final category
Age	√, x	√, x	√, x	√, x		√, x	√, x		√, x	$> / \leq 40$ years
BASDAI score	√, x			√, x			√, x			
BASFI score	√, x	√	√, x	√, x	√, x	√, x	√, x	√, x	√, x	Tertiles
BASMI score	√, x	x		√, x			√, x			
Cervical rotation	√, x			√, x			√, x	x		
Intermalleolar distance	√, x			√, x			√, x			
Tragus to wall distance	√, x			√, x			√, x			
Enthesitis score	√, x		√, x	√, x	x	√, x	√, x	√, x	√, x	$0 / \geq 1$
CRP	√, x	√, x	√, x	√, x	√	√, x	√, x	√, x	√, x	Tertiles
PGA	√, x			√, x			√, x			
HLA-B27	√, x		√, x	√, x	√	√, x	√, x	√, x	√, x	+ / –
Treatment	√, x	√, x	√, x	√, x	√, x	√, x	√, x	√, x	√, x	Anti-TNF/placebo
$R^2$	0.40, 0.34	0.36, 0.27	0.37, 0.27	0.27, 0.24	0.23, 0.18	0.24, 0.22	0.29, 0.26	0.27, 0.25	0.28, 0.24	0.32, 0.28
ROC-AUC (95% CI)	0.83, 0.84	0.81, 0.80	0.82, 0.80	0.77, 0.78	0.75, 0.74	0.75, 0.77	0.78, 0.79	0.77, 0.79	0.77, 0.78	0.80 (0.76–0.83), 0.77 (0.73–0.82)

$R^2$  and the area under the receiver operating characteristics curve (ROC-AUC) are presented for the different models. √ represents Bath ankylosing spondylitis disease activity index (BASDAI) 50, x represents assessment of spondyloarthritis (ASAS) partial remission.  $R^2$  and ROC-AUC values are presented for BASDAI50 response, ASAS partial remission. CRP, C-reactive protein; PGA, patient global assessment.

**Table 3** Associations of predictor variables with selected outcome instruments

	ASDAS clinically important improvement 3 months			ASAS20 response 3 months		
	OR	Lower 95% CI	Upper 95% CI	OR	Lower 95% CI	Upper 95% CI
Anti-TNF vs conventional	17.7	10.0	31.3	5.8	3.7	9.0
HLA-27+ vs HLA-27-	1.3	0.7	2.2	1.7	1.1	2.9
Age ≤40 vs >40 years	2.2	1.5	3.4	1.7	1.2	2.4
CRP high vs low	9.3	5.5	16.0	2.2	1.4	3.4
CRP moderate vs low	3.5	2.2	5.7	1.2	0.8	1.9
Berlin enthesitis score 0 vs >0	0.9	0.6	1.3	1.2	0.8	1.6
BASFI low vs high	2.2	1.3	3.6	1.5	1.0	2.4
BASFI moderate vs high	2.2	1.3	3.6	1.9	1.2	3.0
	ASDAS major improvement 3 months			BASDAI50 response 3 months		
	OR	Lower 95% CI	Upper 95% CI	OR	Lower 95% CI	Upper 95% CI
Anti-TNF vs conventional	14.2	6.5	31.1	8.7	4.9	15.6
HLA-27+ vs HLA-27-	2.4	1.2	4.7	2.5	1.4	4.5
Age ≤40 or >40 years	1.8	1.2	2.9	1.9	1.3	2.8
CRP high vs low	15.0	7.9	28.6	3.6	2.2	5.8
CRP moderate vs low	4.1	2.2	7.8	2.3	1.4	3.7
Berlin enthesitis score 0 vs >0	1.6	1.0	2.5	1.4	1.0	2.1
BASFI low vs high	1.635	0.9	2.8	3.4	2.1	5.5
BASFI moderate vs high	1.4	0.8	2.4	3.0	1.8	4.8
	ASDAS inactive disease 6 months			ASAS partial remission 6 months		
	OR	Lower 95% CI	Upper 95% CI	OR	Lower 95% CI	Upper 95% CI
Anti-TNF vs conventional	46.5	6.4	339.6	16.8	5.2	54.4
HLA-27+ vs HLA-27-	2.4	1.0	5.5	2.2	1.0	5.0
Age ≤40 or >40 years	1.6	1.0	2.6	1.6	1.0	2.6
CRP high vs low	2.3	1.3	4.2	2.1	1.2	3.7
CRP moderate vs low	1.6	0.9	3.0	1.3	0.8	2.4
Berlin enthesitis score 0 vs >0	1.5	0.9	2.4	1.3	0.8	2.1
BASFI low vs high	3.2	1.7	5.9	4.1	2.2	7.6
BASFI moderate vs high	1.8	1.0	3.5	2.6	1.4	4.8

ASDAS, ankylosing spondylitis disease activity score; ASAS, assessment of spondyloarthritis; BASDAI, Bath ankylosing spondylitis disease activity index; BASFI, Bath ankylosing spondylitis functional index; CRP, C-reactive protein; TNF, tumour necrosis factor.

with small improvements (ASAS20, ASDAS clinically important improvement). Age was more strongly associated with improvement than with disease states. Enthesitis showed weak associations with all outcome instruments. The BASFI score was strongly associated with disease state and BASDAI50 improvement but less so with ASDAS and ASAS20 improvements. The very strong association between CRP and ASDAS improvement is striking, albeit reasonable given that CRP is an intrinsic component of ASDAS. A strong association was also seen between CRP and BASDAI50. Finally, very strong associations between anti-TNF therapy and all outcomes were seen with OR ranging from 5.8 to 46.5.

**Matrix model construction**

Matrix models using the six predictor variables were created for all outcome instruments (figure 1A–F) and show a good spread of outcome rates over the different subpopulations defined by the predictor value categories. The strength of associations between predictor and outcome instrument is reflected in the differences between outcome rates in these subpopulations. Differences of 22% or less were seen when rates of large improvement and disease states were compared between similar HLA-B27-positive versus negative patients. Differences of 14% or less were seen when small improvements were compared between genotypes. Differences in improvement rates were larger than differences in rates of disease state when older and younger patients were compared, whereas the association of BASFI led to larger differences in disease state. Differences in outcome rates related to the presence of enthesitis were small. The association of CRP with ASDAS improvement led to major differences in outcomes;

for example, ASDAS major improvement in HLA-B27-positive patients aged 40 years or less who had BASFI of 4.5 or less and no enthesitis was 81% if their CRP was over 2 mg/dl but only 22% if their CRP was normal.

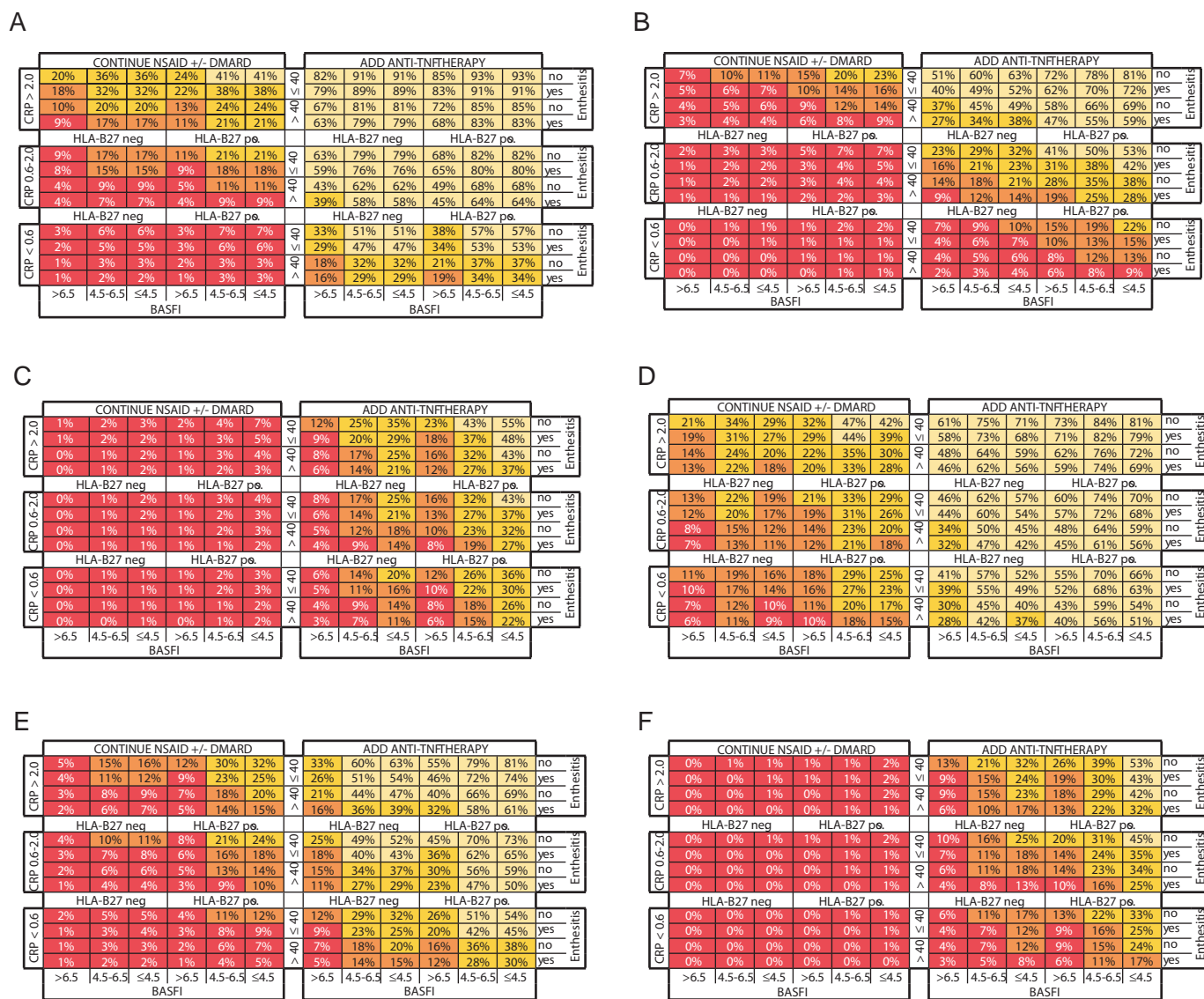
Differences in response rates exceeding 50% were observed when anti-TNF was compared with conventional therapy. The robustness of response to anti-TNF therapy is further highlighted by figure 2A–F, which indicates that almost all subpopulations have a NNT of less than five to achieve small improvements. High NNT indicate that large improvements and inactive disease states are difficult to achieve in some subpopulations. The ROC–AUC (R<sup>2</sup>) for the matrix model of ASAS20 response, BASDAI50 response, ASAS partial remission, ASDAS clinically important and major improvement, and ASDAS inactive disease was 0.74 (0.28), 0.80 (0.32), 0.77 (0.28), 0.84 (0.44), 0.84 (0.39) and 0.79 (0.25), respectively.

**Algorithm-based models**

The formulae of the models using selected predictor variables and/or their interaction terms are presented in supplementary table 6, available online only.

The values for ROC–AUC and R<sup>2</sup> were very similar to those of the matrix models. Comparisons of R<sup>2</sup> show that the model to predict week 12 ASDAS fitted the combined dataset best. Values for R indicate that the association of the algorithm-based model with week 12 ASDAS was higher than that with BASDAI. Supplementary figures 1a and b (available online only) show the predicted versus the observed changes in ASDAS and BASDAI scores and also illustrate that the prediction of ASDAS is more accurate than the prediction of BASDAI.





**Figure 1** Matrix presentation of outcome rates of different patient subpopulations (%) defined by the categorised predictor variables: (A) ankylosing spondylitis disease activity score (ASDAS) clinically important improvement, (B) ASDAS major improvement, (C) assessment of spondyloarthritis (ASAS) partial remission, (D) ASAS 20 response, (E) Bath ankylosing spondylitis disease activity index (BASDAI) 50 response and (F) ASDAS inactive disease. BASFI, Bath ankylosing spondylitis functional index; CRP, C-reactive protein; DMARD, disease-modifying antirheumatic drug; NSAID, non-steroidal anti-inflammatory drug; TNF, tumour necrosis factor.

**Distribution of a cross-sectional AS registry population over the model**

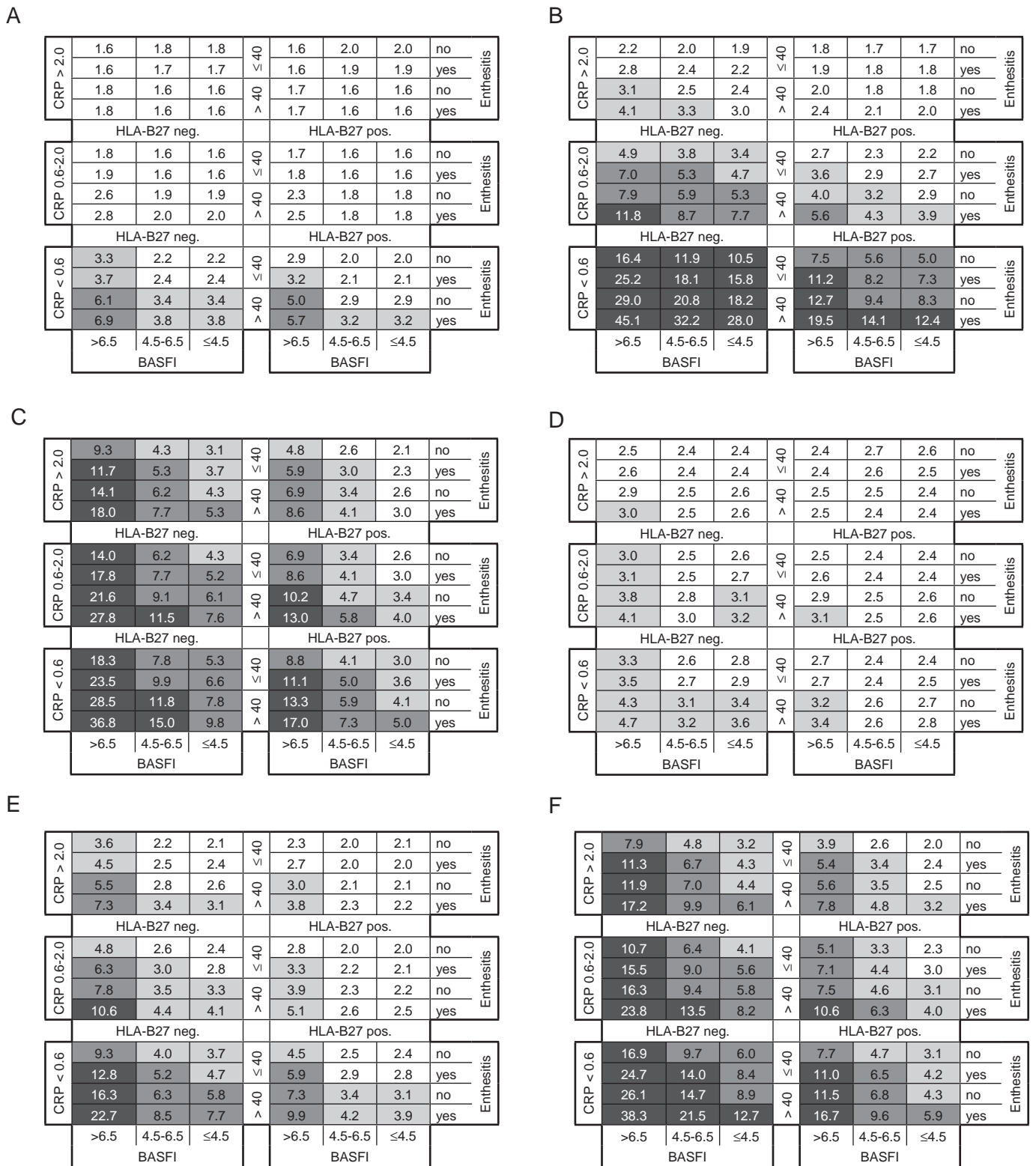
Of the 1760 AS patients in the total registry population, 1051 (59.7%) had an elevated BASDAI score of 4 or greater (ie, the active registry population). The distribution of CRP in the total/active populations, respectively, was: 56.6%/51.0% for patients with CRP less than ULN; 29.8%/33.6% for those with a CRP level of ULN to 2 mg/dl; and 13.6%/15.3% for those with CRP greater than 2 mg/dl. The distribution of BASFI was: 53.9%/32.9% for patients with a score less than 4.5; 22.5%/30.4% for those with a score of 4.5–6.5; and 23.6%/36.7% for those with a score greater than 6.5. Approximately 83% of patients were HLA-B27 positive, and approximately 33% of patients were 40 years old or younger in both the total and the active populations. Enthesitis was present in 16% and 21% of the total and the active registry populations, respectively. The percentage of the total and the active registry patients falling into each cell of the matrix

is shown in figure 3A,B. The percentage of registry patients falling into the different NNT categories for each outcome instrument (figure 2A–F) is reported in supplementary table 7, available online only. For example, for ASDAS clinically important improvement, 82.2%, 9.3% and 8.5% of the active registry population fell into the NNT less than three, three to five and five to 10 categories, respectively. NNT greater than 10 was not observed for this outcome (figure 2A), therefore 0% of registry patients fell into this category.

A detailed comparison of associations between predictors and outcomes reported from comparable analyses performed in AS populations treated with anti-TNF therapy in clinical practice<sup>10–12</sup> is provided in supplementary table 8, available online only.

**DISCUSSION**

Our analyses show that CRP, HLA-B27 genotype, BASFI, age, enthesitis and choice of therapy are independent predictors



**Figure 2** Matrix presentation of numbers needed to treat for one patient to respond to anti-tumour necrosis factor treatment according to different outcome instruments at 12 or 24 weeks: (A) ankylosing spondylitis disease activity score (ASDAS) clinically important improvement, (B) ASDAS major improvement, (C) assessment of spondyloarthritis (ASAS) partial remission, (D) ASAS 20 response, (E) Bath ankylosing spondylitis disease activity index (BASDAI) 50 response, and (F) ASDAS inactive disease. BASFI, Bath ankylosing spondylitis functional index; CRP, C-reactive protein.

of a variety of outcome instruments, and that the combination of these six variables adequately predicted clinical improvement following therapy and subsequent disease states in the ASSERT and the GO-RAISE datasets separately and combined.

**Model validity**

The goal of our analysis was to create a practical, evidence-based model that can help guide clinicians in making informed treatment choices for AS patients. The predictive variables identified in these randomised studies have been shown to be associated

**A**

CRP > 2.0	0.1%	0.1%	0.3%	≤ 40	0.7%	1.0%	1.8%	no	Enthesitis
	0.1%	0.1%	0.0%		0.2%	0.3%	0.2%	yes	
	0.3%	0.1%	0.8%	> 40	2.0%	1.8%	2.4%	no	
	0.2%	0.1%	0.1%		0.6%	0.2%	0.2%	yes	
HLA-B27 neg.				HLA-B27 pos.					
CRP 0.6-2.0	0.1%	0.1%	0.4%	≤ 40	0.8%	1.4%	4.2%	no	Enthesitis
	0.1%	0.2%	0.1%		0.3%	0.2%	1.1%	yes	
	0.9%	0.3%	1.2%	> 40	4.3%	4.1%	6.5%	no	
	0.2%	0.2%	0.2%		1.2%	1.0%	0.7%	yes	
HLA-B27 neg.				HLA-B27 pos.					
CRP < 0.6	0.5%	0.6%	1.3%	≤ 40	2.0%	2.0%	10.2%	no	Enthesitis
	0.2%	0.1%	0.4%		0.3%	0.7%	1.5%	yes	
	1.3%	0.9%	3.8%	> 40	5.6%	6.0%	14.0%	no	
	0.6%	0.1%	0.3%		1.1%	1.0%	2.2%	yes	
HLA-B27 neg.				HLA-B27 pos.					
BASFI				BASFI					
>6.5				>6.5					
4.5-6.5				4.5-6.5					
≤4.5				≤4.5					

**B**

CRP > 2.0	0.2%	0.1%	0.2%	≤ 40	1.1%	1.6%	1.0%	no	Enthesitis
	0.1%	0.1%	0.0%		0.4%	0.6%	0.1%	yes	
	0.5%	0.1%	0.6%	> 40	3.0%	2.0%	1.5%	no	
	0.3%	0.2%	0.2%		1.0%	0.4%	0.0%	yes	
HLA-B27 neg.				HLA-B27 pos.					
CRP 0.6-2.0	0.2%	0.0%	0.2%	≤ 40	1.1%	2.0%	3.0%	no	Enthesitis
	0.1%	0.3%	0.0%		0.6%	0.4%	1.0%	yes	
	1.5%	0.5%	0.9%	> 40	6.6%	6.0%	4.3%	no	
	0.3%	0.3%	0.1%		2.0%	1.5%	0.9%	yes	
HLA-B27 neg.				HLA-B27 pos.					
CRP < 0.6	0.8%	0.9%	1.0%	≤ 40	3.4%	2.8%	5.3%	no	Enthesitis
	0.3%	0.1%	0.5%		0.6%	1.1%	1.3%	yes	
	2.0%	0.9%	2.2%	> 40	7.9%	7.3%	6.7%	no	
	1.0%	0.1%	0.3%		1.8%	1.2%	1.6%	yes	
HLA-B27 neg.				HLA-B27 pos.					
BASFI				BASFI					
>6.5				>6.5					
4.5-6.5				4.5-6.5					
≤4.5				≤4.5					

**Figure 3** Cross-sectional application of the ASPECT/Regisponder populations over the matrix grid: percentage of (A) the total registry population (including all patients irrespective of Bath ankylosing spondylitis disease activity index (BASDAI) score), and of (B) the active registry population (including only patients with a BASDAI score of ≥4) defined by the categorised predictor variables. BASFI, Bath ankylosing spondylitis functional index; CRP, C-reactive protein.

with response and remission in other datasets and outside of a randomised controlled setting, which lends support to the external validity of the model.<sup>8-14</sup>

The associations of age, CRP, HLA-B27 and BASFI with BASDAI50 response in ASSERT/GO-RAISE are very similar to those in previous reports.<sup>10-12</sup> Figure 3 further indicates that the 72 subpopulations characterised by the baseline values for predictors reasonably represent the AS population in clinical practice. This may support the value of our model in daily practice.

There are, however, several weaknesses of our data indicating that validation of the model is necessary. The association of the enthesitis score with outcomes was not investigated in previous reports, and comparisons between anti-TNF and conventional treatment were not performed. Our algorithms and models originate from studies designed and powered to show the superiority of anti-TNF therapy over placebo, and identifying predictors of response was not a formal endpoint of those studies. The blinded, controlled design of the trials may have led to outcomes different from those observed in clinical practice,

and other data sources may have led to the development of different models. Finally, the cross-sectional registry data do not provide any insight into the model's ability to predict outcomes adequately in daily practice.

The predictors retained in the step-wise selection procedures differed between ASSERT and GO-RAISE (table 2), and enthesitis was not associated with ASAS20 response (table 3, supplementary table 6, available online only). As such, some final predictor variables are redundant for certain datasets or for certain outcomes. However, independent of the dataset used and the outcome instrument predicted, the ROC-AUC of the six selected predictors combined remains close to 0.80, indicating good accuracy of prediction.

### Comparison of different outcome instruments

Interestingly, although final predictors were selected for their ability to predict BASDAI50 response and ASAS partial remission, these predictors were more accurate in predicting week 12 ASDAS improvement and inactive disease. Our single component analysis shows that this is due to a stronger association of CRP and therapy with the ASDAS scoring system (table 2). The difference in strength of association between predictors and outcome instruments is relevant for trial design in AS. The stronger association of anti-TNF therapy with ASDAS than with traditional outcomes indicates that the ASDAS scoring system may be a more powerful tool than current outcome instruments in showing the efficacy of biological agents. The associations identified may also improve patient selection in studies.

The inclusion of CRP as a component in the ASDAS formula may explain partly why outcomes assessed with ASDAS were very strongly associated with baseline CRP. However, although BASFI is a component of ASAS20 response and ASAS partial remission criteria, the association between BASFI and these outcomes was not as strong as that between CRP and ASDAS outcomes.

In subpopulations with normal CRP, BASDAI50 response and ASAS partial remission rates following anti-TNF treatment were higher than ASDAS major improvement and ASDAS inactive disease rates, and absolute differences with response to conventional therapy led to higher NNT. Differences between ASAS20 response and ASDAS clinically important improvement were also present but smaller. This may indicate that outcomes in patients with normal CRP may be better assessed with an outcome instrument based only on patient-reported outcomes. These findings are in concordance with validation sets of the ASDAS in which discrimination of ASDAS was better than that of BASDAI in patients with elevated CRP and equal to BASDAI in patients with normal CRP.<sup>20</sup>

### Selected predictors

Although BASDAI was not a predictor of response in our datasets, it was in previous reports.<sup>8–11</sup> This may be due to a homogeneous selection of study patients based on elevated BASDAI scores as part of the inclusion criteria. BASFI was retained as a predictor in this and previous studies.<sup>10–12</sup> The correlation between BASDAI and BASFI is relevant for selecting candidates for anti-TNF therapy in AS, as shown in the AS registries. The proportion of patients in the lowest BASFI category is much higher in the total than the active registry population. The high correlation between BASFI and BASDAI is due to the exclusion of patients with low BASDAI in the active registry population. Patients with a BASDAI less than 4, however, may still have

other clinical characteristics that are associated with response and remission in addition to a low BASFI score. For example, 658 (37.4%) of all registry patients were HLA-B27 positive and had CRP elevation greater than ULN; of these, 214 (32.5%) had a BASDAI less than 4. These patients have not been studied in clinical trials and are currently not recommended for anti-TNF therapy.

Our data show that somewhat worse outcomes can be expected in patients with an elevated enthesitis score. Because of the lack of agreement on how enthesitis should be measured,<sup>29</sup> enthesitis was assessed only as present or absent and was not scored in the registries. This explains why enthesitis is present in the majority of patients in randomised studies but only in a minority of patients in registries. The differences in response and remission to anti-TNF therapy were not large when similar patients with and without enthesitis were compared. As anti-TNF agents are very effective in patients with well-defined enthesitis,<sup>30</sup> patients with peripheral manifestations having worse enthesitis may be a reflection of more severe disease in general.<sup>31–32</sup>

HLA-B27-positive patients responded better to anti-TNF treatment in our study and in previous reports.<sup>8–12</sup> It is unclear whether this is a function of HLA-B27 facilitating earlier and correct diagnosis or the disease biology differing in HLA-B27-positive versus negative AS patients.

Age was an independent predictor of outcome in the ASSERT study, and significant differences were seen when age was compared between responders and remitters in the GO-RAISE study and the combined dataset. The importance of age in response prediction has been shown previously.<sup>8–10–11</sup> Although disease duration has been shown to be relevant for outcome prediction,<sup>8</sup> disease duration was not retained in our dataset because age can be more precisely determined than disease duration and may be more useful for prediction.

Our data confirm the association of elevated CRP levels with good response to anti-TNF therapy.<sup>9–14</sup> As the registry data show that AS patients with normal CRP constitute approximately half the AS population, recognising suitable candidates for anti-TNF treatment among such patients may be challenging. Other inflammatory biomarkers and MRI may help in predicting response to therapy,<sup>14–29</sup> and may be especially useful in distinguishing responders from non-responders in patients with low CRP.<sup>14–33</sup>

### Subpopulations with robust response to anti-TNF treatment

Anti-TNF therapy is recommended for patients who have sustained elevated disease activity despite conventional therapy and should be prescribed based on expert opinion.<sup>2</sup> Our prediction model may help guide that expert opinion. The data show that the continuation of conventional therapy in the face of sustained elevated disease activity will be unlikely to result in improvement. The differential responses in ASAS20 and ASDAS clinically important improvement rates from using anti-TNF versus continued conventional treatment and the resulting low NNT indicate that anti-TNF treatment is a clinically sound choice in all subpopulations with elevated disease activity. Given the lack of good alternatives, the treating physician should therefore consider a defined trial period with an anti-TNF agent if disease activity is not controlled with NSAID.<sup>2</sup> Large improvements and remission may, however, not be achievable therapeutic goals for all patients.

In conclusion, our analysis shows that a model combining age, HLA-B27 genotype, CRP level, functional status and the



presence of enthesitis at baseline enables a good prediction of the response to anti-TNF or conventional therapy in AS, as measured by various outcome instruments. This may help clinicians choose more appropriate therapies for patients in daily practice and also help improve patient selection and protocol design for clinical studies.

**Acknowledgements** The authors would like to thank Jennifer Han and Robert Achenbach of Centocor Ortho Biotech Services, LLC, for their assistance with revising and preparing the manuscript.

**Funding** BVC is a postdoctoral researcher supported by the FWO Flanders.

**Competing interests** None.

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

- Zochling J, van der Heijde D, Burgos-Vargas R, *et al.* ASAS/EULAR recommendations for the management of ankylosing spondylitis. *Ann Rheum Dis* 2006;**65**:442–52.
- Braun J, Davis J, Dougados M, *et al.* First update of the international ASAS consensus statement for the use of anti-TNF agents in patients with ankylosing spondylitis. *Ann Rheum Dis* 2006;**65**:316–20.
- Davis JC Jr, Van Der Heijde D, Braun J, *et al.* Recombinant human tumor necrosis factor receptor (etanercept) for treating ankylosing spondylitis: a randomized, controlled trial. *Arthritis Rheum* 2003;**48**:3230–6.
- Inman RD, Davis JC Jr, van der Heijde D, *et al.* Efficacy and safety of golimumab in patients with ankylosing spondylitis: results of a randomized, double-blind, placebo-controlled, phase III trial. *Arthritis Rheum* 2008;**58**:3402–12.
- van der Heijde D, Dijkmans B, Geusens P, *et al.* Efficacy and safety of infliximab in patients with ankylosing spondylitis: results of a randomized, placebo-controlled trial (ASSERT). *Arthritis Rheum* 2005;**52**:582–91.
- van der Heijde D, Kivitz A, Schiff MH, *et al.* Efficacy and safety of adalimumab in patients with ankylosing spondylitis: results of a multicenter, randomized, double-blind, placebo-controlled trial. *Arthritis Rheum* 2006;**54**:2136–46.
- Gossec L, Dougados M, Phillips C, *et al.* Dissemination and evaluation of the ASAS/EULAR recommendations for the management of ankylosing spondylitis: results of a study among 1507 rheumatologists. *Ann Rheum Dis* 2008;**67**:782–8.
- Rudwaleit M, Listing J, Brandt J, *et al.* Prediction of a major clinical response (BASDAI 50) to tumour necrosis factor alpha blockers in ankylosing spondylitis. *Ann Rheum Dis* 2004;**63**:665–70.
- de Vries MK, van Eijk IC, van der Horst-Bruinsma IE, *et al.* Erythrocyte sedimentation rate, C-reactive protein level, and serum amyloid A protein for patient selection and monitoring of anti-tumor necrosis factor treatment in ankylosing spondylitis. *Arthritis Rheum* 2009;**61**:1484–90.
- Glinthorg B, Ostergaard M, Krogh NS, *et al.* Predictors of treatment response and drug continuation in 842 patients with ankylosing spondylitis treated with anti-tumour necrosis factor: results from 8 years' surveillance in the Danish nationwide DANBIO registry. *Ann Rheum Dis* 2010;**69**:2002–8.
- Lord PA, Farragher TM, Lunt M, *et al.* Predictors of response to anti-TNF therapy in ankylosing spondylitis: results from the British Society for Rheumatology Biologics Register. *Rheumatology (Oxford)* 2010;**49**:563–70.
- Rudwaleit M, Claudepierre P, Wordsworth P, *et al.* Effectiveness, safety, and predictors of good clinical response in 1250 patients treated with adalimumab for active ankylosing spondylitis. *J Rheumatol* 2009;**36**:801–8.
- Stone MA, Payne U, Pacheco-Tena C, *et al.* Cytokine correlates of clinical response patterns to infliximab treatment of ankylosing spondylitis. *Ann Rheum Dis* 2004;**63**:84–7.
- Visvanathan S, Wagner C, Marini JC, *et al.* Inflammatory biomarkers, disease activity and spinal disease measures in patients with ankylosing spondylitis after treatment with infliximab. *Ann Rheum Dis* 2008;**67**:511–17.
- Vander Cruyssen B, Ribbens C, Boonen A, *et al.* The epidemiology of ankylosing spondylitis and the commencement of anti-TNF therapy in daily rheumatology practice. *Ann Rheum Dis* 2007;**66**:1072–7.
- Smith MD, Ahern MJ. Pharmaceutical Benefits Scheme criteria for the use of tumour necrosis factor-alpha inhibitors in the treatment of ankylosing spondylitis in Australia: are they evidence based? *Intern Med J* 2006;**36**:72–6.
- Garrett S, Jenkinson T, Kennedy LG, *et al.* A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. *J Rheumatol* 1994;**21**:2286–91.
- Anderson JJ, Baron G, van der Heijde D, *et al.* Ankylosing spondylitis assessment group preliminary definition of short-term improvement in ankylosing spondylitis. *Arthritis Rheum* 2001;**44**:1876–86.
- Lukas C, Landewé R, Sieper J, *et al.* Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. *Ann Rheum Dis* 2009;**68**:18–24.
- van der Heijde D, Lie E, Kvien TK, *et al.* ASDAS, a highly discriminatory ASAS-endorsed disease activity score in patients with ankylosing spondylitis. *Ann Rheum Dis* 2009;**68**:1811–18.
- Machado P, Landewé R, Lie E, *et al.* Ankylosing Spondylitis Disease Activity Score (ASDAS): defining cut-off values for disease activity states and improvement scores. *Ann Rheum Dis* 2011;**70**:47–53.
- Tape TG. Interpreting diagnostic tests: ROC curves – the area under a ROC curve. <http://jim.unmc.edu/dxtests/ROC3.htm>. (accessed 25 Jan 2010).
- Hosmer DW, Lemeshow S. *Applied logistic regression*, 2nd edn. Hoboken, NJ: John Wiley & Sons, 2000.
- Rosenthal JA. Qualitative descriptors of strength of association and effect size. *J Soc Serv Res* 1996;**21**:37–59.
- 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.
- Cohen J. *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
- Collantes E, Zarco P, Muñoz E, *et al.* Disease pattern of spondyloarthropathies in Spain: description of the first national registry (REGISPONSER) extended report. *Rheumatology (Oxford)* 2007;**46**:1309–15.
- Vander Cruyssen B, Muñoz-Gomariz E, Font P, *et al.* Hip involvement in ankylosing spondylitis: epidemiology and risk factors associated with hip replacement surgery. *Rheumatology (Oxford)* 2010;**49**:73–81.
- Gladman DD, Inman RD, Cook RJ, *et al.* International spondyloarthritis interobserver reliability exercise – the INSPIRE study: II. Assessment of peripheral joints, enthesitis, and dactylitis. *J Rheumatol* 2007;**34**:1740–5.
- Maksymowych WP, Mallon C, Morrow S, *et al.* Development and validation of the Spondyloarthritis Research Consortium of Canada (SPARCC) Enthesitis Index. *Ann Rheum Dis* 2009;**68**:948–53.
- Heuft-Dorenbosch L, van Tubergen A, Spoorenberg A, *et al.* The influence of peripheral arthritis on disease activity in ankylosing spondylitis patients as measured with the Bath Ankylosing Spondylitis Disease Activity Index. *Arthritis Rheum* 2004;**51**:154–9.
- Song IH, Rudwaleit M, Listing J, *et al.* Comparison of the Bath Ankylosing Spondylitis Disease Activity Index and a modified version of the index in assessing disease activity in patients with ankylosing spondylitis without peripheral manifestations. *Ann Rheum Dis* 2009;**68**:1701–7.
- Rudwaleit M, Schwarzklose S, Hilgert ES, *et al.* MRI in predicting a major clinical response to anti-tumour necrosis factor treatment in ankylosing spondylitis. *Ann Rheum Dis* 2008;**67**:1276–81.