Predicting low disease activity and remission using early treatment response to anti-TNF therapy in patients with rheumatoid arthritis: Exploratory analyses from the TEMPO trial

Jeffrey R. Curtis[1], **Shuo Yang**[1], **Lang Chen**[1], **Grace S. Park**[2], **Bojena Bitman**[2], **Brian Wang**[3], **Iris Navarro-Millan**[1], and **Arthur Kavanaugh**[4]

[1]University of Alabama at Birmingham, Birmingham, AL, USA

[2]Amgen Inc., Thousand Oaks, CA, USA

[3]KForce Clinical Research, Tampa, FL, USA

[4]University of California at San Diego, San Diego, CA, USA

Abstract

Objective—To derive and validate decision trees to categorize rheumatoid arthritis (RA) patients 12 weeks after starting etanercept with or without methotrexate into three groups: patients predicted to achieve low disease activity (LDA) at 1 year; patients predicted to not achieve LDA at 1 year; and patients who needed additional time on therapy to be categorized.

Methods—Data from RA patients enrolled in TEMPO were analyzed. Classification and Regression Trees were used to develop and validate decision-tree models with week 12 and earlier assessments that predicted long-term LDA. LDA, defined as DAS28　3.2 or Clinical Disease Activity Index (CDAI)　10.0, was measured at 52 or 48 weeks. Demographics, laboratory data, and clinical data at baseline and through week 12 were analyzed as predictors of response.

Results—Thirty-nine percent (67/172) of patients receiving etanercept and 60% (115/193) of patients receiving etanercept plus methotrexate achieved LDA at week 52. For patients receiving etanercept, 53% were predicted to have LDA, 39% were predicted to not have LDA, and 8% could not be categorized using DAS28 criteria at week 12. For patients receiving etanercept plus methotrexate, 63% were predicted to have LDA, 25% were predicted to not have LDA, and 12% could not be categorized.

Conclusion—Most (80%–90%) patients in TEMPO initiating etanercept with or without methotrexate could be predicted within 12 weeks of starting therapy as likely to have LDA or not at week 52. However, approximately 10%–20% of patients needed additional time on therapy to decide whether to continue treatment.

Key MeSH terms

etanercept; methotrexate; arthritis; rheumatoid; decision tree; prediction

INTRODUCTION

The presentation and disease course are highly variable in patients with rheumatoid arthritis (RA). Unsurprisingly, there is also great variation in the response to both nonbiologic and biologic disease-modifying antirheumatic drugs (DMARDs). Given the chronic nature of RA (with the consequent need for long-term treatment), the expense of newer biologic DMARDs, and the urgency of identifying efficacious treatment to minimize joint damage in individual patients, the ability to predict response to treatment would have substantial clinical and economic impact.

The National Institute for Health and Clinical Excellence (NICE) and the British Society for Rheumatology (BSR) recommend discontinuation of anti-TNF therapies after 6 months in the absence of an adequate response.[1, 2] The American College of Rheumatology (ACR) recommends re-evaluation of patients who have not achieved clinical benefit within 12 weeks of initiating anti-TNF therapy.[3] Given the short period of time in which physicians are expected to make treatment decisions, it has become increasingly important to identify features in individual patients that may assist in decisions to continue or discontinue a treatment regimen.

Etanercept is a human TNF receptor-Fc fusion protein that binds to TNF and inhibits its interaction with cell surface TNF receptors. Etanercept is approved for treatment of moderately to severely active RA. Using data from a pivotal trial of etanercept, the objective of this analysis was to derive and validate a decision tree that was able to categorize patients within 12 weeks after starting etanercept with or without methotrexate into one of three groups: patients predicted to achieve low disease activity (LDA) at 1 year; patients predicted to not achieve LDA at 1 year; and patients who were not able to be categorized at 12 weeks and would need additional time on therapy. Additional analyses substituted the new ACR/EULAR remission definition [4] for the LDA outcome and categorized patients at 12 weeks as being likely or not to achieve remission at 1 year.

METHODS

Patients

Data from patients enrolled in the Trial of Etanercept and Methotrexate with Radiographic Patient Outcomes (TEMPO) [5] were used in this analysis. Patients 18 years or older with active, adult-onset RA were enrolled in TEMPO. Patients in TEMPO received etanercept (25 mg twice weekly [BIW]), methotrexate (7.5 mg escalated to 20 mg oral capsules once weekly [QW] within 8 weeks if patients had any painful or swollen joints), or both.

Definition of LDA

The primary outcome of this retrospective analysis was LDA (DAS28 ≤ 3.2) [6] at week 52 (or week 48 if the DAS28 measurement was missing at week 52). LDA as a goal is consistent with recent treat-to-target recommendations suggesting that while remission is an optimal goal, LDA is acceptable, especially for RA patients with established disease.[7] As a secondary outcome, patients were considered to have LDA if they had Clinical Disease Activity Index (CDAI) ≤ 10 [8] at week 52. An additional secondary outcome required remission using the ACR/EULAR Boolean definition (tender and swollen joint ≤ 1, CRP ≤ 1 mg/L, and patient global assessment ≤ 1 on a scale of 0 to 10).[4]

Predictor variables

Patient demographics (age, sex, race [white vs non-white]), clinical data (tender joint count, swollen joint count, rheumatoid factor status, DAS28 raw score at baseline and change from

baseline score at weeks 4, 8, and 12, Health Assessment Questionnaire Disability Index [HAQ-DI] score, patient pain, Physician Global Assessment, Patient Global Assessment, CDAI, and laboratory data (erythrocyte sedimentation rate, C-reactive protein) at baseline and at each visit through week 12 were included as candidate variables as predictors of LDA at week 52.

Statistical analysis

Patients were included in this analysis if they had DAS28 assessments at 48 or 52 weeks of therapy and had received etanercept alone or etanercept plus methotrexate. Patients who dropped out early because of unsatisfactory efficacy were considered to be nonresponders. Patients who dropped out of the study for safety reasons were excluded from the analysis.

Classification and Regression Trees (CART) software (Salford Systems, San Diego, CA, USA) was used to develop and validate models for identifying week 12 and earlier assessments that would predict LDA at week 52. The CART model relies on statistically optimum recursive splitting of the patients into subgroups based on critical levels of the prognostic variables. In the general implementation of CART, the dataset is split into the two subgroups that are the most different with respect to the predictor variable outcomes, and subgroups are split further based on the same principle. The percentages of patients with LDA were calculated for each node of the regression tree.

Patients predicted to have LDA at week 52 by the predictor variables were classified as responders and patients predicted to not achieve LDA at week 52 were classified as nonresponders. The remaining patients (who had an approximately 40%–60% predicted likelihood of achieving LDA) had an unclear likelihood of response and were classified as indeterminate responders needing additional time on treatment.

Two analyses were performed: the primary analysis used LDA based on DAS28 at 52 or 48 weeks; secondary analyses used LDA based on CDAI at 52 or 48 weeks and remission at 52 or 48 weeks. A 10-fold cross-validation technique [9, 10] was used to guard against model overfitting, a potential problem in prediction models in which the model fits the dataset used to derive it but would fit other datasets less well. Misclassification penalties of 3:1 were implemented, placing greater emphasis on correctly classifying patients who were predicted to be nonresponders. This procedure optimizes the prediction model for patients predicted to be nonresponders, since it is these patients for whom a decision to change the treatment regimen at 12 weeks would likely be made.

Finally, we examined the tradeoff between the degree of accuracy of a prediction model that might be minimally acceptable to a clinician and the resulting proportion of patients who could be classified with that amount of accuracy. The best-performing decision tree derived from the combination etanercept plus methotrexate data set was evaluated using 1000 simulated data sets with bootstrapping techniques. Patients receiving etanercept plus methotrexate were sampled 1000 times with replacement to generate 1000 bootstrap samples of equal size to the original TEMPO etanercept plus methotrexate arm. For each of the 1000 samples, the performance of the decision tree was evaluated iteratively by varying the level of required accuracy to range from 50% to 100%. Accuracy was defined as the proportion of patients who could be correctly classified by the decision tree. The proportion of the patients in each node of the tree who could be classified with the required amount of accuracy relative to the total sample size was then plotted against the accuracy level for that iteration. The process was then repeated for the remainder of the 1000 data sets. All points were plotted and a LOESS smoothing curve was fitted; this figure described the proportion of the population that could be predicted to have LDA at 1 year as a function of the accuracy required for that prediction.

RESULTS

Patients

Patient demographics and disease characteristics at baseline by treatment group and LDA status at week 52 or 48 as defined by DAS28 are shown in Table 1. Thirty-nine percent of patients receiving etanercept and 60% of patients receiving etanercept plus methotrexate achieved an LDA response at week 52 based on DAS28. Demographic and clinical characteristics were similar across treatment groups and between responders and nonresponders at baseline.

High concordance between LDA assessed by DAS28 and LDA assessed by CDAI at week 52 was demonstrated by the Kappa (K) coefficients: K = 0.64 for patients receiving etanercept; K = 0.78 for patients receiving etanercept plus methotrexate.

Patients receiving etanercept plus methotrexate

LDA in patients receiving etanercept plus methotrexate was predicted by DAS28 at week 12 and change in DAS28 from baseline at week 8 (Figure 1A). Patients receiving etanercept plus methotrexate were categorized into 3 groups by 12 weeks: responders (63% of all patients, 81% accuracy); nonresponders (25% of all patients, 88% accuracy); and patients with an indeterminate likelihood of response (12% of all patients).

Response to therapy in patients receiving etanercept plus methotrexate combination therapy was predicted by CDAI at week 12 and SJC at week 8 (Figure 1B). Patients were categorized at week 12: responders (54% of all patients, 94% accuracy); nonresponders (29% of all patients, 80% accuracy); and patients with an indeterminate likelihood of response (17% of all patients).

In summarizing the two models presented in Figures 1A and 1B, 83% to 88% of patients could be classified as responders or nonresponders by 12 weeks. The accuracy of prediction for these individuals was approximately 85%. For the remainder of the 12% to 17% of patients, additional time on therapy would be necessary to determine their treatment response at 1 year.

The model that predicted remission at 1 year is shown in Figure 2. A total of 26% of patients receiving etanercept plus methotrexate achieved remission. The key predictor variables included tender and swollen joint count (both at week 12), patient pain (at week 12), and CRP (measured at week 4). At 12 weeks, 95% of patients could be classified as responders or nonresponders. The accuracy of prediction for patients classified at week 12 as nonresponders (58% of all patients) was 98%.

Patients receiving etanercept monotherapy

Response to therapy in patients receiving etanercept monotherapy was predicted by DAS28 at week 12 and tender joint count at week 8 (Figure 3A). Patients receiving etanercept monotherapy were categorized into 3 groups by 12 weeks: responders (53% of all patients, 61% accuracy); nonresponders (39% of all patients, 93% accuracy); and patients with an indeterminate likelihood of response (only 8% of all patients).

The second model constructed using CDAI is shown in Figure 3B. Response in patients receiving etanercept monotherapy was predicted by HAQ-DI at week 8, change in CDAI at week 4, and CDAI at week 12. Patients were categorized by week 12 as responders (49% of all patients, 81% accuracy); nonresponders (27% of all patients, 91% accuracy); and patients with an indeterminate likelihood of response (24% of all patients). Too few patients in the

etanercept monotherapy arm achieved remission to warrant deriving a prediction model for this treatment group.

Tradeoff of degree of accuracy and the proportion of classifiable patients in a prediction model

As shown in Figure 4, there was a tradeoff between accuracy and the proportion of patients who could be classified with that amount of accuracy. Simulations represented in the cluster of points on the left side of the figure show that only 25% of the patients sampled from the bootstrapped TEMPO study population could be classified with approximately 90% accuracy. Assuming a willingness to tolerate somewhat lesser accuracy of 80% to 85%, the substantial majority of the study population sampled from TEMPO could be classified by week 12; indeed, approximately 60% to 65% of the TEMPO patients could be classified with 85% accuracy, and 85% to 90% of patients could be classified with 80% accuracy.

DISCUSSION

Patients with RA have a heterogeneous pattern of response to currently available therapy, including TNF inhibitors and other DMARDs. This is likely due to the variable course of RA. Researchers have proposed that predicting which patients will develop aggressive versus mild disease is important in order to tailor therapy that will be promising and predictable.[11] Whereas baseline disease activity may assist in selection of the type of treatment used to treat RA, baseline measures are generally inadequate to predict treatment outcome,[12] as we have also shown here (Table 1). For that reason, we built DAS28- and CDAI-based models using early treatment response (through 12 weeks) to show that approximately 80% to 90% of patients in TEMPO could be classified as nonresponders with respect to achieving LDA at 1 year. For the remaining ~15% of patients, additional time on therapy would be needed to determine their longer term treatment response. Substituting the alternate outcome of RA remission for LDA, the proportion of patients able to be classified at week 12 was high (95%); overall model accuracy was similar to the prediction models with LDA as the outcome. The accuracy for predicted nonresponders (98%) using the remission outcome was higher than for the LDA outcome.

Results from a study by Verstappen et al showed that early response to nonbiologic DMARD therapy in the first year, rather than the kind of initial treatment, given predicted disease remission in patients with early RA.[13] Similarly, in patients receiving anti-TNF therapy, the likelihood of continuation of treatment was predicted by the response that they had during the first 3 months of treatment. This finding suggests that decision-making regarding continuation of nonbiologic and biologic DMARD therapy might be considered as early as 3 months,[14] which led to our decision to use the response to treatment through 12 weeks as the key predictor variables.

Aletaha et al [15] analyzed pooled data from several clinical trials of patients with early and established RA.[5, 16–19] Similar to our results, they found that disease activity after 3 months of DMARD or anti-TNF therapy, but not at baseline, determined the treatment response at one year. We have extended those findings to be able to provide a clinically useful, albeit preliminary, decision tree to classify patients as being responders, nonresponders, or those for whom more time (beyond 12 weeks) is needed to predict response at 1 year. In contrast to that prior report, our results incorporate multiple predictors at different time points to improve the validity of prediction and increase the proportion of patients for whom such a prediction can be made by week 12.

Three other studies specifically examined the ability to predict treatment outcomes based on short-term response to treatment with anti-TNF therapies. Gülfe et al found that response to

treatment as early as 6 weeks predicted continuation of current therapy at 3 months.[14] Pocock et al found that a substantial number of patients who had not achieved response at 3 months of treatment were able to continue treatment and achieve a response at 6 months, supporting a need for longer therapeutic trials before discontinuation in some patients.[20] This result is also supported by a previously published study by Kavanaugh et al,[21] which also suggested that some patients who did not respond by 3 months did achieve some response by 6 months. Our results provide guidance on which patients need more time beyond 12 weeks to make a clinical decision, and which patients probably do not. Patients we classified as 'indeterminate,' comprising only approximately 15% of the TEMPO population, likely would benefit from additional time on therapy. In contrast, we were able to accurately predict treatment response for the remainder of the patients (85%) by 12 weeks; for those predicted to be nonresponders at 12 weeks, a decision might be made to switch treatment regimens, and no additional time on therapy would be necessary.

We based our prediction model on clinical and laboratory-based factors. The identification of surrogate markers of response with genetics or proteomics that might assist in predicting treatment response or response to a particular therapy is an attractive possibility but has been elusive. The current lack of genetic markers or biomarkers with the ability to discriminate between patients who will or will not respond to therapy places the emphasis on clinical evaluations and patient-reported measures. We view our LDA response models as providing a useful framework to which more sophisticated biomarker-based predictors might be added.

The CART methodology used in our study has been shown to provide more robust analyses of data containing nonlinear features, colinearity, and interactions than conventional logistic regression analyses.[22] This nonparametric tree-based method of modeling is useful to identify the best predictors of treatment response and has a simple and visual interpretation. The validity and reproducibility of the CART decision trees is enhanced by the cross-validation technique we used, which provides an estimate of how well any classification tree performs on similar but non-identical datasets. However, despite efforts to avoid model overfitting, all prediction models, including these, should be revalidated in an independent data set.

There is a trade-off between the accuracy of prediction and the proportion of patients whose treatment response can be predicted with that degree of accuracy. As we have shown in a simulated data example in Figure 4, and using one of the prediction models we derived, there were patients for whom we could predict treatment response or nonresponse with 90% or greater accuracy; only about 25% of a population such as the patients enrolled in TEMPO could be predicted with this amount of accuracy. A higher proportion of patients could have their treatment response predicted if lower amounts of accuracy are acceptable (e.g. 80% to 85%). The degree of accuracy that clinicians require to make a treatment change at 12 weeks for those predicted to be nonresponders is clinician-dependent, but it is reassuring that the prediction accuracy of most of the nonresponder groups in our models was approximately 90%, and was even higher (98%) for the remission outcome. Further illustrating the tradeoff between the accuracy of prediction and the proportion of patients able to be classified, an analysis by van der Heidje [23] showed that patients who did not achieve a change in their DAS28 of greater than 1.2 units at 12 weeks after starting certolizumab pegol had only a 1% likelihood of achieving LDA at 1 year; however, only 13% of patients in the study could be classified as nonresponders using this criterion.

In summary, approximately 80% to 90% of patients in the TEMPO study who initiated etanercept with or without methotrexate could be classified within 12 weeks of starting therapy as likely to have a good response or not at week 52. Additional time on therapy was needed to determine whether to continue or discontinue etanercept for the remaining 10% to

20% of patients. This exploratory decision tree is specific to our study population; we expect that prediction models may need to vary based on the RA patient population (early versus established disease), the level of baseline disease activity (high versus moderate/low), biologic-naïve versus biologic-experienced patients, and perhaps even the specific agent or treatment regimen used. Separate prediction models might be necessary for each of these different RA patient populations; this remains to be tested. However, at least for RA patients whose clinical and disease characteristics are similar to patients enrolled in TEMPO, these models may assist the clinician in assessing the likelihood of response to TNF inhibitor therapy. This could aid clinical decision-making at an earlier time point.. Additional prediction models built with easily measured clinical and laboratory data will likely provide a useful framework to allow for adding predictors of response, including biomarker data.
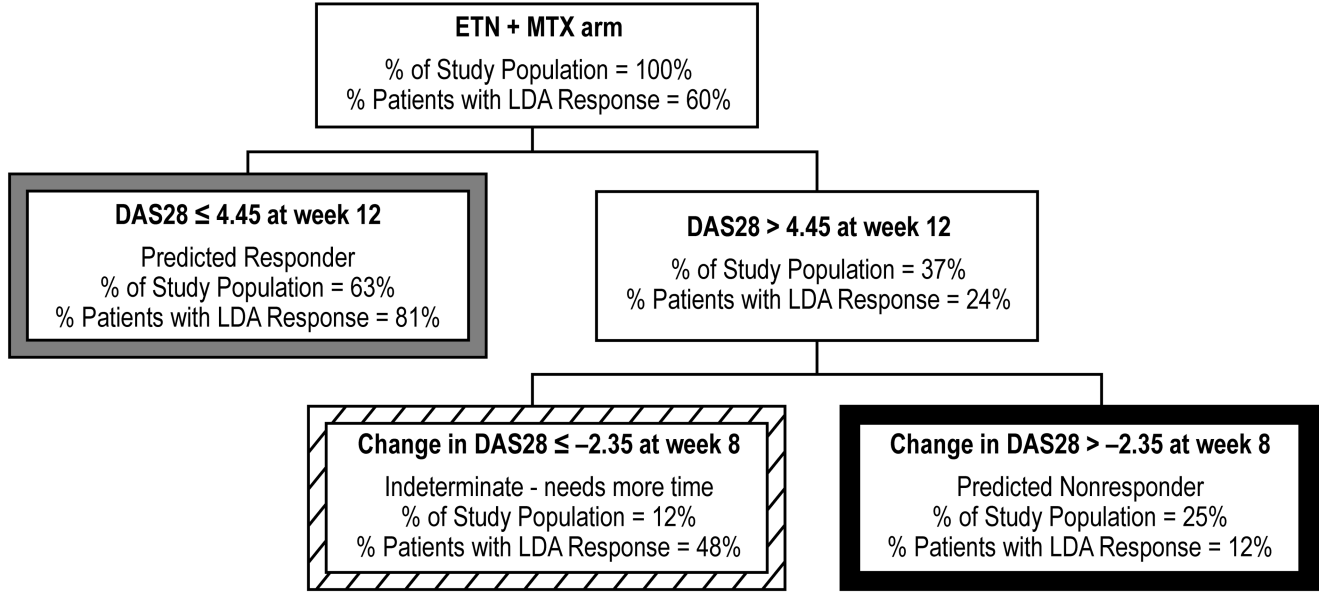
Acknowledgments

REFERENCES

1. National Institute for Health and Clinical Excellence. Adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis. 2007 Oct. Available at: http://guidance.nice.org.uk/TA130/Guidance/pdf/English.

2. Ledingham J, Deighton C. Update on the British Society for Rheumatology guidelines for prescribing TNFalpha blockers in adults with rheumatoid arthritis (update of previous guidelines of April 2001). Rheumatology (Oxford). 2005; 44:157–163. [PubMed: 15637039]

3. Furst DE, Breedveld FC, Kalden JR, et al. Updated consensus statement on biological agents for the treatment of rheumatic diseases, 2007. Ann Rheum Dis. 2007; 66(Suppl 3):iii2–iii22. [PubMed: 17934088]

4. Felson DT, Smolen JS, Wells G, et al. American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. Arthritis Rheum. 2011; 63:573–586. [PubMed: 21294106]

5. Klareskog L, van der Heijde D, de Jager JP, et al. Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: double-blind randomised controlled trial. Lancet. 2004; 363:675–681. [PubMed: 15001324]

6. DAS-score.nl: Disease Activity Score in Rheumatoid Arthritis. Available from: http://www.das-score.nl/www.das-score.nl/.

7. Smolen JS, Aletaha D, Bijlsma JW, et al. Treating rheumatoid arthritis to target: recommendations of an international task force. Ann Rheum Dis. 2010; 69:631–637. [PubMed: 20215140]

8. Aletaha D, Smolen J. The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. Clin Exp Rheumatol. 2005; 23:S100–S108. [PubMed: 16273793]

9. Breiman L, Spector P. Submodel selection and evaluation in regression: the X-random case. Int Stat Rev. 1992; 60:291–319.
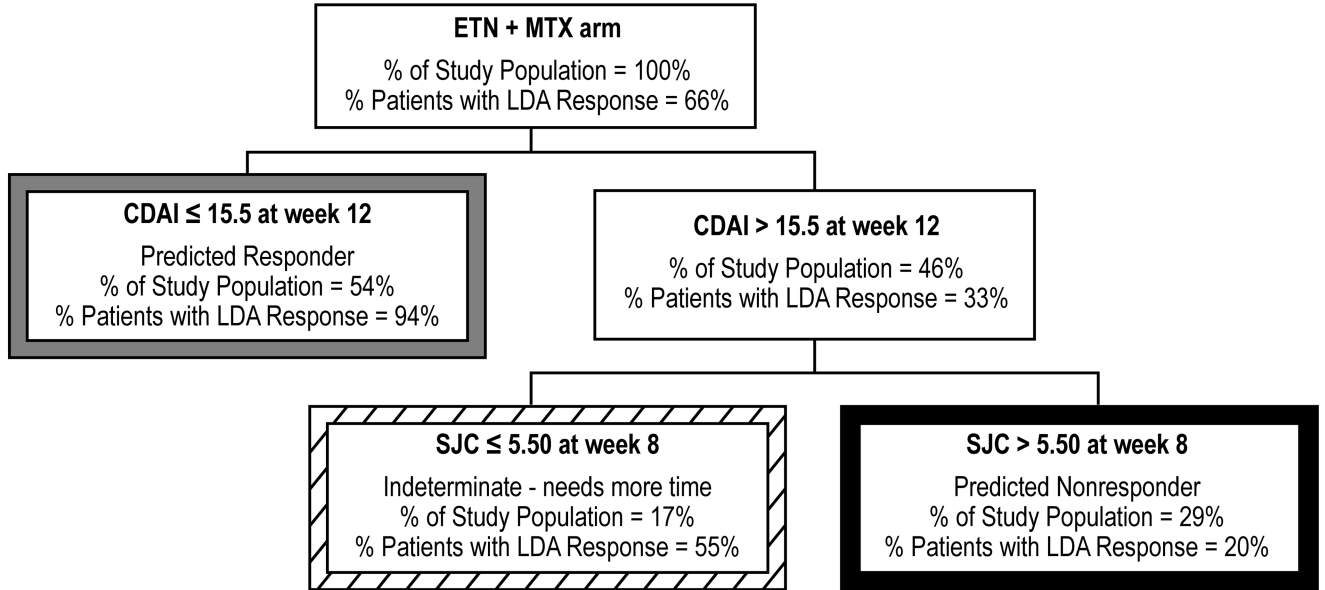
10. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection; Fourteenth International Joint Conference on Artificial Intelligence; 1995; Montreal, Quebec, Canada. 1995.

11. de Vries-Bouwstra J, Le Cessie S, Allaart C, et al. Using predicted disease outcome to provide differentiated treatment of early rheumatoid arthritis. J Rheumatol. 2006; 33:1747–1753. [PubMed: 16845710]

12. Smolen JS, Aletaha D. Challenges of predicting treatment response in patients with rheumatoid arthritis. Nat Clin Pract Rheumatol. 2005; 1:62–63. [PubMed: 16932630]

13. Verstappen SM, van Albada-Kuipers GA, Bijlsma JW, et al. A good response to early DMARD treatment of patients with rheumatoid arthritis in the first year predicts remission during follow up. Ann Rheum Dis. 2005; 64:38–43. [PubMed: 15130899]

14. Gulfe A, Kristensen LE, Geborek P. Six and 12 weeks treatment response predicts continuation of tumor necrosis factor blockade in rheumatoid arthritis: an observational cohort study from southern Sweden. J Rheumatol. 2009; 36:517–521. [PubMed: 19208596]

15. Aletaha D, Funovits J, Keystone EC, et al. Disease activity early in the course of treatment predicts response to therapy after one year in rheumatoid arthritis patients. Arthritis Rheum. 2007; 56:3226–3235. [PubMed: 17907167]

16. 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; 50:3432–3443. [PubMed: 15529377]

17. Lipsky PE, van der Heijde DM, St Clair EW, et al. Infliximab and methotrexate in the treatment of rheumatoid arthritis. Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy Study Group. N Engl J Med. 2000; 343:1594–1602. [PubMed: 11096166]

18. Breedveld FC, Weisman MH, Kavanaugh AF, et al. The PREMIER study: A multicenter, randomized, double-blind clinical trial of combination therapy with adalimumab plus methotrexate versus methotrexate alone or adalimumab alone in patients with early, aggressive rheumatoid arthritis who had not had previous methotrexate treatment. Arthritis Rheum. 2006; 54:26–37. [PubMed: 16385520]

19. Bathon JM, Martin RW, Fleischmann RM, et al. A comparison of etanercept and methotrexate in patients with early rheumatoid arthritis. N Engl J Med. 2000; 343:1586–1593. [PubMed: 11096165]

20. Pocock JM, Vasconcelos JC, Ostor AJ. Assessment of anti-TNF-alpha efficacy in rheumatoid arthritis: is 3 months sufficient? Rheumatology (Oxford). 2008; 47:1073–1076. [PubMed: 18503089]

21. Kavanaugh A, Klareskog L, van der Heijde D, et al. Improvements in clinical response between 12 and 24 weeks in patients with rheumatoid arthritis on etanercept therapy with or without methotrexate. Ann Rheum Dis. 2008; 67:1444–1447. [PubMed: 18535115]

22. Wolfe F, Flowers N, Burke TA, et al. Increase in lifetime adverse drug reactions, service utilization, and disease severity among patients who will start COX-2 specific inhibitors: quantitative assessment of channeling bias and confounding by indication in 6689 patients with rheumatoid arthritis and osteoarthritis. J Rheumatol. 2002; 29:1015–1022. [PubMed: 12022317]

23. van der Heijde DMFM, Schiff M, Keystone EC, et al. Probability to achieve low disease activity at 52 weeks in rheumatoid arthritis (RA) patients treated with certolizumab pegol (CZP) depends on time to and level of initial response [abstract]. Arthritis Rheum. 2009; 60(Suppl 10):992.

A. Response based on DAS28

ETN + MTX arm
% of Study Population = 100%
% Patients with LDA Response = 60%

DAS28 ≤ 4.45 at week 12
Predicted Responder
% of Study Population = 63%
% Patients with LDA Response = 81%

DAS28 > 4.45 at week 12
% of Study Population = 37%
% Patients with LDA Response = 24%

Change in DAS28 ≤ –2.35 at week 8
Indeterminate - needs more time
% of Study Population = 12%
% Patients with LDA Response = 48%

Change in DAS28 > –2.35 at week 8
Predicted Nonresponder
% of Study Population = 25%
% Patients with LDA Response = 12%

B. Response based on CDAI

ETN + MTX arm
% of Study Population = 100%
% Patients with LDA Response = 66%

CDAI ≤ 15.5 at week 12
Predicted Responder
% of Study Population = 54%
% Patients with LDA Response = 94%

CDAI > 15.5 at week 12
% of Study Population = 46%
% Patients with LDA Response = 33%

SJC ≤ 5.50 at week 8
Indeterminate - needs more time
% of Study Population = 17%
% Patients with LDA Response = 55%

SJC > 5.50 at week 8
Predicted Nonresponder
% of Study Population = 29%
% Patients with LDA Response = 20%

Figure 1.
CART decision tree of DAS28 and CDAI response in patients receiving etanercept. CART decision trees with **A)** DAS28 3.2 (responders) or > 3.2 (nonresponders) at 52 or 48 weeks as the outcome variable and with **B)** CDAI 10 (responders) or > 10 (nonresponders) at 52 or 48 weeks as the outcome variable are shown for patients receiving etanercept plus methotrexate combination therapy. Tree nodes of predicted responders, nonresponders, and patients with indeterminate outcomes are identified by borders of gray, black, and diagonal patterns, respectively. ETN, Etanercept; LDA, Low disease activity; DAS28, Disease

activity score based on 28 joints; TJC, Tender joint count; CDAI, Clinical Disease Activity Index; HAQ-DI, Health Assessment Questionnaire Disability Index.
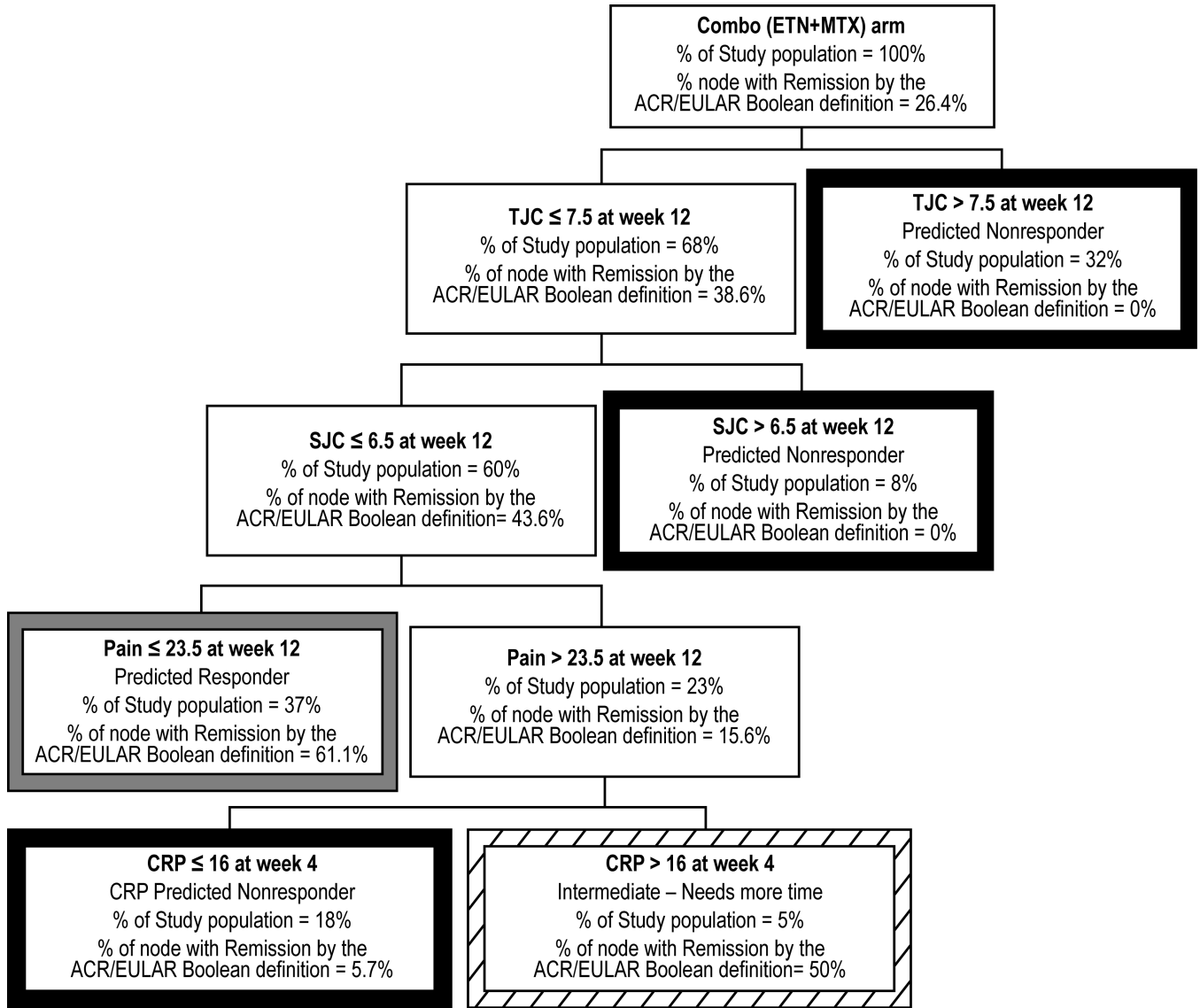
Figure 2.
CART decision tree in patients receiving etanercept plus methotrexate with ACR/EULAR remission at 52 or 48 weeks as the outcome variable. Tree nodes of predicted responders, nonresponders, and patients with indeterminate outcomes are identified by borders of gray, black, and diagonal patterns, respectively. ETN, Etanercept; MTX, Methotrexate; SJC, Swollen joint count; TJC, Tender joint count; CRP, C-reactive protein (in mg/L).

A. Response based on DAS28

ETN only arm
% of Study Population = 100%
% Patients with LDA Response = 39%

DAS28 ≤ 4.45 at week 12
Predicted Responder
% of Study Population = 53%
% Patients with LDA Response = 61%

DAS28 > 4.45 at week 12
% of Study Population = 47%
% Patients with LDA Response = 14%

TJC ≤ 4.50 at week 8
Indeterminate - needs more time
% of Study Population = 8%
% Patients with LDA Response = 46%

TJC > 4.50 at week 8
Predicted Nonresponder
% of Study Population = 39%
% Patients with LDA Response = 7%

B. Response based on CDAI

ETN arm only
% of Study Population = 100%
% Patients with LDA Response = 50%

CDAI ≤ 15.5 at week 12
Predicted Responder
% of Study Population = 43%
% Patients with LDA Response = 78%

CDAI > 15.5 at week 12
% of Study Population = 57%
% Patients with LDA Response = 29%

Change in CDAI ≤ -33.0 at week 4
Predicted Responder
% of Study Population = 6%
% Patients with LDA Response = 90%

Change in CDAI > -33.0 at week 4
% of Study Population = 51%
% Patients with LDA Response = 22%

HAQ-DI ≤ 1.25 at week 8
Indeterminate - needs more time
% of Study Population = 24%
% Patients with LDA Response = 37%

HAQ-DI > 1.25 at week 8
Predicted Nonresponder
% of Study Population = 27%
% Patients with LDA Response = 9%

Figure 3.
CART decision tree of DAS28 and CDAI response in patients receiving etanercept plus methotrexate. CART decision trees with **A)** DAS28 3.2 (responders) or > 3.2 (nonresponders) at 52 or 48 weeks as the outcome variable and with **B)** CDAI 10 (responders) or > 10 (nonresponders) at 52 or 48 weeks as the outcome variable are shown for patients receiving etanercept. Tree nodes of predicted responders, nonresponders, and patients with indeterminate outcomes are identified by borders of gray, black, and diagonal patterns, respectively. ETN, Etanercept; MTX, Methotrexate; LDA, Low disease activity;

DAS28, Disease activity score based on 28 joints; CDAI, Clinical Disease Activity Index; SJC, Swollen joint count.
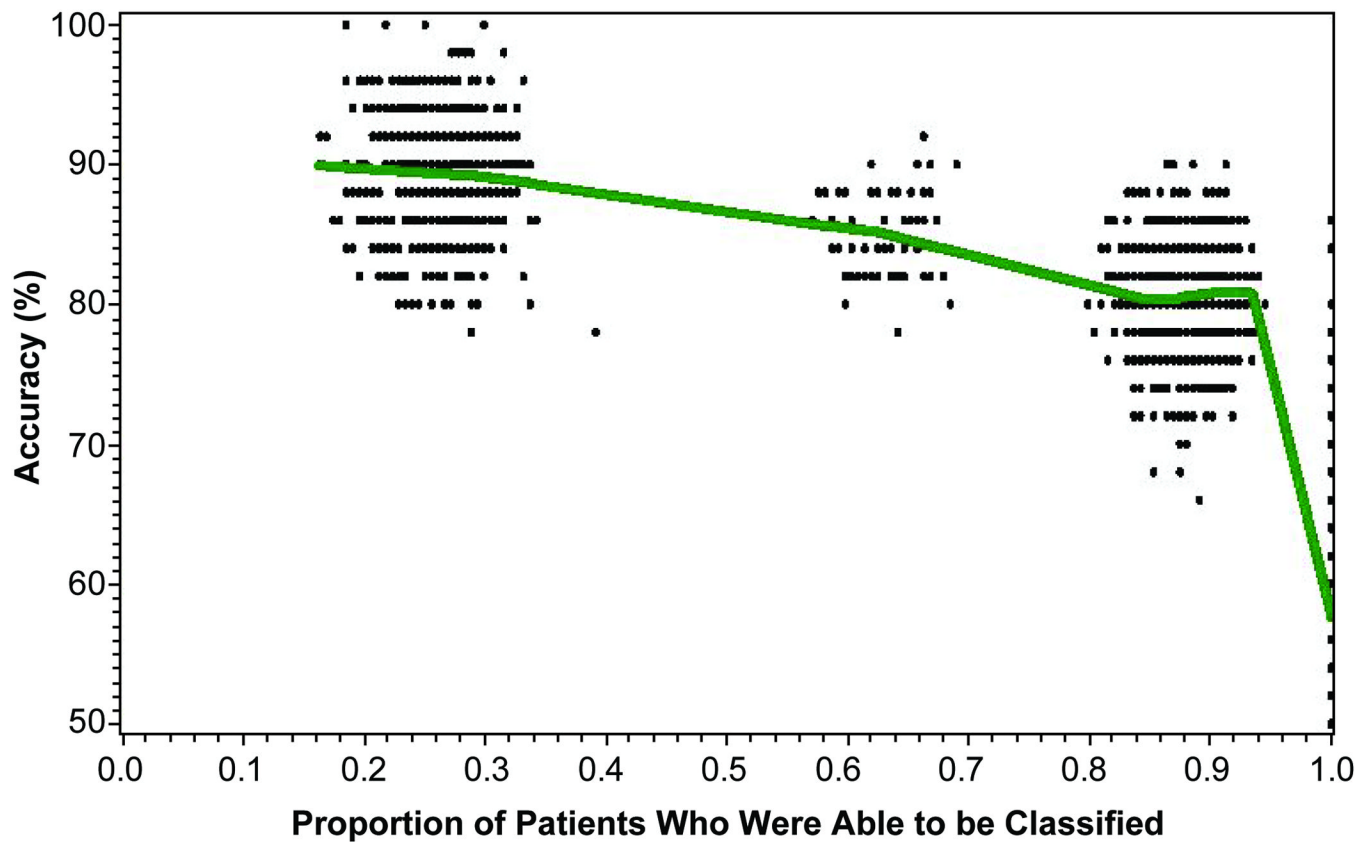
Figure 4.
Application of prediction model shown in Figure 1A to 1000 bootstrap samples of TEMPO patients receiving etanercept plus methotrexate (n = 193 patients in each sample, sampled with replacement). The accuracy required of the prediction model was iteratively varied at set threshold levels between 50% and 100%. Each data point represents the proportion of each bootstrap sample that could be classified with that amount of accuracy, and a Loess curve line was fitted through the data points. This simulated model illustrates the tradeoff of misclassification rate versus the proportion of patients who can be classified with the prediction model. With the requirement for greater accuracy (left-hand side of figure), fewer patients could be classified at week 12.

Table 1

Patient demographics and disease characteristics at baseline stratified by LDA (DAS28 ≤ 3.2) at week 52 or 48

Characteristic	Etanercept		Etanercept + Methotrexate	
	DAS28 ≤ 3.2 (n = 67)	DAS28 > 3.2 (n = 105)	DAS28 ≤ 3.2 (n = 115)	DAS28 > 3.2 (n = 78)
Sex, n (%)				
Female	46 (69)	84 (80)	81 (70)	62 (80)
Male	21 (31)	21 (20)	34 (30)	16 (20)
Age, mean years (SD)	50.9 (15.0)	53.5 (13.3)	51.4 (12.9)	52.6 (11.9)
Race, n (%)				
White	66 (99)	103 (98)	114 (99)	76 (97)
Non-White	1 (1)	2 (2)	1 (1)	2 (3)
RF positive, n (%)	46 (69)	79 (75)	90 (78)	61 (78)
DAS28ESR, mean score (SD)	6.4 (0.9)	7.1 (0.8)	6.5 (1.1)	7.1 (0.8)
CDAI, mean score (SD)	43.6 (12.0)	49.4 (12.4)	43.3 (12.3)	52.1 (13.3)
ESR, mean mm/hr (SD)	30.6 (20.9)	45.1 (22.3)	37.2 (23.4)	44.3 (25.9)
TJC, mean no. (SD)	17.1 (6.4)	19.0 (6.5)	16.5 (6.9)	20.4 (6.4)
SJC, mean no. (SD)	14.4 (5.7)	15.9 (5.7)	13.3 (5.2)	17.3 (6.3)
HAQ-DI, mean score (SD)	1.5 (0.6)	1.9 (0.6)	1.6 (0.6)	1.9 (0.6)

CDAI, clinical disease activity index; DAS28ESR, disease activity score based in 28 joints with ESR as marker of inflammation; ESR, erythrocyte sedimentation rate; HAQ-DI, Health Assessment Questionnaire Disability Index; LDA, low disease activity; RF, rheumatoid factor; SD, standard deviation; SJC, swollen joint count; TJC, tender joint count