Published in final edited form as:

Arthritis Care Res (Hoboken). 2012 May ; 64(5): 658–667. doi:10.1002/acr.21600.

Predicting Future Response to Certolizumab Pegol in Rheumatoid Arthritis Patients: Features at 12 Weeks Associated With Low Disease Activity at 1 Year

Jeffrey R. Curtis, MD, MS, MPH¹, Kristel Luijckens, PhD², and Arthur Kavanaugh, MD³

¹Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, Birmingham, AL, USA

²Biostatistics, UCB, Braine l'Alleud, Belgium

³University of California, San Diego, CA, USA

Abstract

Objectives—To determine the prognostic significance of data collected early after starting certolizumab pegol (CZP) to predict low disease activity (LDA) at Week 52.

Methods—Data through Week 12 from 703 CZP-treated patients in the RA Prevention of structural Damage (RAPID 1) trial were used as variables to predict LDA (DAS28 [ESR] 3.2) at Week 52. We identified variables, developed prediction models using classification trees, and tested performance using training and testing datasets. Additional prediction models were constructed using CDAI and an alternate outcome definition (composite of LDA or ACR50).

Results—Using Week 6 and 12 data and across several different prediction models, response (LDA) and nonresponse at 1 year was predicted with relatively high accuracy (70–90%) for most patients. The best performing model predicting nonresponse by 12 weeks was 90% accurate and applied to 46% of the population. Model accuracy for predicted responders (30% of the RAPID1 population) was 74%. The area under the receiver operator curve was 0.76. Depending on the desired certainty of prediction at 12 weeks, ~12–24% of patients required >12 weeks of treatment to be accurately classified. CDAI-based models, and those evaluating the composite outcome (LDA or ACR50), achieved comparable accuracy.

Conclusion—We could accurately predict within 12 weeks of starting CZP whether most established RA patients with high baseline disease activity would likely achieve/not achieve LDA at 1 year. Decision trees may be useful to guide prospective management for RA patients treated with CZP and other biologics.

Many patients with rheumatoid arthritis (RA) exhibit a rapid clinical response to tumor necrosis factor- α (TNF- α) inhibitor therapy, most within 3 months of initiating treatment (1–10). However, limited information is available to determine which factors, measured at baseline or early after starting a new RA medication, will predict a good long-term response, and results are inconsistent. In one study of patients with recent-onset RA treated with methotrexate, a model based on gender, rheumatoid factor, smoking status, disease activity score (DAS), and several genetic polymorphisms was able to predict treatment efficacy (11). However, a later study including patients with early inflammatory

Corresponding author: Jeffrey R. Curtis, MD, MS, MPH, Director, UAB Arthritis Clinical Intervention Program (ACIP), and Co-Director, UAB, Center for Education and Research on Therapeutics (CERTs), Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, FOT 805D, 510 20th Street South, Birmingham, Alabama 35294, USA, Tel: 001 205 975 2176, Fax: 001 205 975 5554, jcurtis@uab.edu.

polyarthritis found routine clinical and laboratory factors to be poor at predicting outcome of treatment with methotrexate (12). Several pharmacogenetic studies have suggested that a single nucleotide polymorphism in the promoter region of TNF- α is significantly associated with response to TNF inhibitor therapy (13–15). However, no single marker has yet emerged that accurately predicts a response in a large cohort of patients. For that reason, clinical and laboratory information is likely to remain paramount in assessing the expected longer-term response to RA medications, ideally in conjunction with yet-to-be-determined genetic or biomarker-based tests.

Even if predictors of long-term response for individual patients can be identified, they may not be easily applicable in clinical practice. Previous reports have shown that U.S. rheumatologists rarely measure the indices for disease evaluation commonly used in clinical trials e.g. American College of Rheumatology response or the Disease Activity Score (e.g. DAS28) (16)(17). Predictive tools thus need to be user friendly and should use data that are easily and consistently measurable in routine clinical practice.

Among several different methods to classify and predict future response, classification and regression trees (CART) have been successfully utilized in a number of therapeutic areas (18–22) to categorize disease and to determine the likelihood of future events. Unlike some statistical methods, CART imposes no prior assumptions on the structure of the data or relationships among variables. This approach has the potential to identify patients likely or unlikely to respond to treatment and to present the predicted likelihood of response in a manner useful to inform RA clinical practice.

The objective of this analysis was to determine the prognostic significance of clinical and laboratory data measured at baseline and during the first 12 weeks of therapy with the PEGylated anti-TNF certolizumab pegol (CZP) to predict low disease activity (LDA) at 1 year. The goal of the prediction model was to categorize patients 12 weeks after starting CZP to facilitate clinical management at that time point. Patients were categorized at 12 weeks into 3 groups:

1. Those predicted to have a low likelihood of achieving LDA at Week 52 with CZP therapy, and who might consider alternate treatment options at Week 12
2. Those predicted to have a high likelihood of achieving LDA at Week 52 with CZP therapy and who are likely to be reassured at Week 12 about their future response
3. Those with an uncertain likelihood of achieving LDA at Week 52, who need longer than 12 weeks of CZP therapy (e.g. up to 6 months) before being able to accurately predict a 1-year outcome.

MATERIALS AND METHODS

Patients

Participants included in this analysis were patients with active RA who were randomized to treatment with CZP in the RA Prevention of structural Damage (RAPID) 1 clinical trial (NCT00152386) (23). In brief, eligible patients were aged ≥ 18 years and had active disease at screening and baseline as defined by ≥ 9 tender joints and ≥ 9 swollen joints and erythrocyte sedimentation rate (ESR) (Westergren) ≥ 30 mm/h or C-reactive protein (CRP) >15 mg/L. Patients were required to have received methotrexate for ≥ 6 months, at stable dosage of ≥ 10 mg/week for ≥ 2 months prior to baseline.

Protocol

The design and primary results of the RAPID 1 trial have been published (23). The intent-to-treat (ITT) population in RAPID 1 comprised 982 patients, randomized 2:2:1 to receive treatment with subcutaneous CZP (at an initial dose of 400 mg at Weeks 0, 2, and 4 followed by 200 mg or 400 mg every 2 weeks) in combination with methotrexate, or placebo in combination with methotrexate, for 1 year (i.e., 52 weeks). Patients who failed to achieve an ACR20 response at Weeks 12 and 14 were withdrawn from RAPID 1 at Week 16 (per study protocol). Patients who withdrew at Week 16 or who successfully completed RAPID 1 were offered enrollment in an open-label extension (OLE) study of certolizumab pegol 400 mg every 2 weeks (NCT00175877). Institutional review boards or ethics committees approved the protocol at each center. All patients gave written consent, and the study was conducted in accordance with the Declaration of Helsinki.

Patients initially randomized to placebo were not included in this analysis. A total of 783 patients were randomized to treatment with CZP, and the eligible patient population for this analysis (n=703) excluded those who withdrew during the first year of the trial for reasons other than efficacy. For patients who withdrew due to efficacy and reconsented into the OLE, the OLE data gathered at Week 52 were used.

For patients who withdrew due to efficacy but did not consent into the OLE, or for any patients using rescue medication, nonresponder imputation was used for ACR20/50/70, and last observation carried forward (LOCF) imputation was used for all continuous data (e.g., DAS28, further categorized as LDA or not) in order to capture CZP response in a conservative way.

Data analysis

The outcome of the prediction model was LDA at Year 1 (i.e., Week 52) after starting CZP therapy. LDA was defined as DAS28 (ESR)4 of ≤ 3.2 . Although other clinical outcomes are desirable (e.g., remission or DAS28 < 2.6) (22), LDA was chosen to be a proxy for doing “well enough,” such that a patient and physician would likely choose to continue therapy. Achieving LDA is also in line with the recent European League Against Rheumatism (EULAR) recommendations, which state that LDA is an appropriate goal for patients with long-standing disease and high disease activity as they are unlikely to achieve remission (24).

A probability plot based on logistic regression was constructed showing the absolute DAS28 at Week 12, and the change in DAS28 at Week 12, as a predictor of LDA at 1 year. With the intent to improve upon use of this single measure assessed at only 1 time point, a decision tree-based approach was used for the prediction model. The advantages of a tree-based approach are that it can be applied to the treatment of individual patients because each individual is categorized uniquely within the tree, it is easy to understand, it can incorporate multiple predictors measured at different times, and it can be described in a visual fashion without the need for calculations that would require a computer software tool.

An a-priori model was constructed based only on clinical input that used DAS change from baseline of ≥ 1.2 units at Weeks 4 and 12 as predictors. Several different prediction models were subsequently derived and tested using CART software (Salford Systems, San Diego, CA, USA). Since any data-based, empiric prediction model is likely to suffer from overfitting (i.e., the model adequately fits to the data used to develop the model, but the model fails to provide accurate predictions for new subjects or datasets) (25), a split sample approach (i.e., separate training/testing datasets) was used. The prediction model was built on a two-thirds random sample of the data (training dataset) and tested on the remaining one third of the data (testing dataset). In this article, we show only the more conservative

performance from the testing datasets, and not the typically better performance from the training datasets. It has previously been shown that CART modeling can perform better than conventional logistic regression analysis, especially when the data contain nonlinear features, colinearity, and interactions(25,26).

Within each tree, patients were categorized into “terminal nodes,” which were the lowest branches. Nodes with a likelihood of achieving LDA much greater than 50% were considered “predicted responder” nodes, and those with a predicted likelihood of achieving LDA much less than 50% were considered “predicted nonresponder” nodes. The misclassification rate for each node was therefore the proportion of patients who achieved LDA in a node classified as a predicted nonresponder, and the proportion of patients in a predicted responder node who did not achieve LDA. The accuracy of classification for patients in each node was determined as 100 minus the misclassification rate. For example, if the proportion of patients who achieve LDA in a predicted nonresponder node is 14%, then the accuracy of classification in that node is 86%.

In the first model, described as CART Model 1, demographic and disease characteristics at baseline, drug level, clinical, and related laboratory data from CZP-treated patients from baseline, Weeks 4, 6, 8, and 12 were used as candidate variables. Each variable was represented as the absolute value, absolute change from baseline, and percent change from baseline. The ability of Model 1 to discriminate between patients who did or did not achieve LDA at Year 1 was assessed using area under the receiver-operator curve (AUC). With the intent to reduce the misclassification of Model 1, especially for patients predicted to be nonresponders, a related model (Model 2) was constructed that implemented a 3:1 misclassification penalty. The aim of this model was to reduce the misclassification for predicted nonresponders to 10%. This penalty, available as a standard option within the CART software, allowed for misclassifying a nonresponder at Week 12 who was actually a responder at 1 year to be set to be 3 times worse (in terms of model building) than misclassifying in the reverse direction. The goal of applying this penalty was to increase the certainty of prediction at 12 weeks for the patients predicted to be nonresponders.

Sensitivity analyses

With recognition that many physicians do not calculate DAS28 measures in real time, a “simplified” model (Model 3) was derived. This prediction model was derived de novo, and more complex variables such as DAS were disallowed. Only variables that were more easily measured in clinical practice, such as the Clinical Disease Activity Index (CDAI), were included as potential predictors. The outcome of this model was the same i.e., LDA by DAS28ESR at 1 year.

As patients who start with high disease activity might be satisfied enough if they reach an ACR50, even if they did not achieve LDA, we derived a final model (Model 4). The construction of this prediction model was similar to Model 1, but used a composite outcome (LDA and/or ACR50) at 1 year. Finally, to test the hypothesis that tree-based prediction models derived from RA patients with similar key characteristics (e.g. established disease) might generalize across anti-TNF therapies, we used RAPID1 data to assess the performance of a recently published CART-based model that was derived using from patients with established RA initiating etanercept (33).

RESULTS

Patient characteristics

Of 783 patients randomized to treatment with CZP in RAPID 1 (ITT population), 703 were included in this analysis; 80 patients who withdrew for reasons other than efficacy were

excluded. The population of patients represented in this analysis was similar to the overall RAPID 1 population (Table 1). Disease activity at the start of the trial was high; patients had a median DAS28 of 7.0, and median RA disease duration was 6.1 years. A total of 486 randomly selected patients were used in the training dataset to build the models. The remaining 217 patients were included in the testing dataset to evaluate model performance, and these individuals form the basis of all results reported below. At Year 1, 35% (n = 76) of patients included in the testing dataset achieved LDA.

Model performance

The probability of achieving LDA at Year 1 based on the absolute DAS28, and on the change from baseline in DAS28 scores at Week 12, is shown in Figures 1A and 1B, respectively. Only patients in the right- and left-most tails of both of these figures could be predicted with high probability to be in LDA at 1 year.

As an enhancement to this approach, results from the a priori tree-based model are shown in Figure 2. At Week 12, 77% of patients had at least some clinical response to CZP therapy (i.e., they had a DAS28 change from baseline of >1.2 units at Week 12). The patients who did not achieve a DAS28 change from baseline of at least 1.2 units at Week 12 had a very low (6%) likelihood of LDA at 1 year. However, only 23% of patients could be classified as being nonresponders using this criterion. Moreover, this approach could not identify responder patients who were likely to achieve LDA, demonstrated by the observation that there were no nodes where patients had a >50% predicted likelihood of response. Overall, the a priori model correctly classified 65% of patients and had an AUC of 0.68.

Using CART Model 1 (Figure 3), which included data from baseline and Weeks 4, 6, 8, and 12 as potential predictors, LDA at Year 1 was predicted with 65–86% accuracy for 88% (n = 192) of the population (Figure 3). This 88% included 54% (n = 117) of the study population predicted to be nonresponders and 35% (n = 75) predicted to be responders. Of the 54% of patients who still had a DAS28 >4.5 at Week 12, 86% did not achieve LDA at Year 1, and the remaining 14% were misclassified. Only 12% of patients required further time on treatment to accurately predict the outcome. Overall, 76% of patients were correctly classified, and the model had good discrimination with an AUC of 0.73.

Results from CART Model 2 (Figure 4), which was a variation of Model 1 that applied a penalty to reduce the misclassification rate of the predicted nonresponders, showed that the initial DAS28 threshold selected by CART at 12 weeks was higher (4.9). Model 2 was able to classify 46% of patients as predicted nonresponders with only 10% misclassification (i.e., 90% accuracy); however, the proportion of patients who could not be classified accurately at 12 weeks doubled to 24%. Of the 30% of patients predicted to be responders, 74% achieved LDA at 1 year. Model 2 was shown to have an improved discrimination ability compared with Model 1 (AUC 0.76) and was able to correctly classify more of the overall patient population within 12 weeks of starting treatment (79% versus 76%).

CART Model 3 (Figure 5) shows the results from the sensitivity analysis that included more simple predictors (e.g., CDAI, and not DAS28). A total of 55% (n = 119) of patients were predicted to be nonresponders and 20% (n = 43) were predicted to be responders. The accuracy of the nonresponder prediction was 80%. For 25% of patients, further time on treatment was necessary to accurately predict the outcome. Model 3 had a similar discrimination (AUC 0.63) and accuracy to Model 1 and was able to correctly classify 71% of the overall patient population within 12 weeks of starting treatment.

In CART model 4 predicting the composite outcome measure of LDA and/or ACR50 at 1 year (Appendix 1), 45% of patients were classified as responders with 85% accuracy, 39%

of patients were predicted to be nonresponders with 78% accuracy, and the remaining 16% of patients were not able to be well classified at 12 weeks. Model 4 achieved similar performance to the best model (Model 2), and had an AUC of 0.77 and an ability to correctly classify 77% of the overall patient population within 12 weeks of starting therapy.

Finally, the replication of a previously published CART-derived decision tree is shown in Appendix 2. The prediction model represented by this tree had similar performance when applied to the RAPID1 data as its original performance using TEMPO data.

DISCUSSION

A major challenge in the management of RA is the prediction of long-term response to therapy. At present, only limited information is available to determine which factors, if any, will predict a good long-term response. The ability to predict achievement of LDA at 1 year would enable physicians to tailor treatment early during the course of therapy, thereby improving outcomes and potentially minimizing patient exposure to ineffective therapies. Furthermore, achieving LDA is a treatment goal supported by the recent EULAR guidelines (24). Using clinically applicable models, we show that, 12 weeks after start of CZP therapy, we could accurately classify the vast majority (~88%) of patients from RAPID 1, a study population with mainly high baseline disease activity, as likely to achieve or not achieve LDA at 1 year. Across several prediction models, we also found that patients with an early response to treatment at Weeks 4, 6 and 8 had an even greater likelihood of achieving LDA at 1 year. Approximately 12–25% of patients could not be classified accurately at Week 12 and needed treatment longer to determine with a high degree of certainty the likelihood of achieving LDA at Week 52. At least for the types of patients enrolled in RAPID 1, our results identify which patients can likely be switched at 12 weeks if they are predicted to be nonresponders (i.e., patients having a very low predicted likelihood of achieving LDA) with a relatively high degree of accuracy. Indeed, these data highlight the possibility to use such models as a negative predictability tool – identifying those unlikely to achieve LDA – and this is perhaps the patient population a treating physician would most like to identify early so that treatment can be altered. In our best performing model, Model 2, this prediction was made with 90% accuracy. Whether a 90% level of certainty, or similar amounts of certainty, is sufficient to give physicians enough confidence to make treatment changes at 12 weeks is a matter of individual judgment. Other factors, including patient and physician preferences and access to alternative therapies, are also likely to play important roles in the decision to switch RA treatments (27–30).

Results using an alternate model in which DAS was replaced by CDAI (CART Model 3) had somewhat lower discrimination and accuracy than Model 1. This lower performance was likely to be a consequence of only moderate correlation between the CDAI (used for the predictor variables) and the DAS28 (used for the outcome). Performance of this model would likely have been better if we had used CDAI to define LDA, rather than the DAS28. Finally, CART Model 4 (similar to Model 1, but with a composite outcome [LDA and/or ACR50] at 1 year) had the numerically best discrimination of all prediction models.

When comparing results across the models, we see a clear compromise between the accuracy of prediction and the proportion of patients who could be classified at Week 12. In the a-priori model (Figure 2), for example, the misclassification rate for patients predicted to be nonresponders was very low, 6%. However, only 23% of eligible RAPID 1 patients could be classified with that high level of accuracy. The a-priori model also had suboptimal discrimination and calibration for the entire study population. Using the CART-based, data-driven approach, the misclassification rate for patients predicted to be nonresponders shown in Model 1 (Figure 3) was slightly higher at 14%, but twice as many patients (54%) could be

classified. Results from Model 2 (Figure 4) were even better and had a misclassification rate of only 10% for a group of patients who comprised 46% of the eligible RAPID 1 population. These data illustrate the concession between the certainty of classification and the proportion of all patients able to be classified.

Prior RA data evaluating single predictors at a fixed time point indicate that the level of disease activity at baseline and after the first 3 months of treatment is significantly related to the level of disease activity at 1 year (Figure 1)(31,32). While probability plots offer important insights, they have limitations—they provide probabilities for only 1 specific time point and 1 variable. To improve upon this, using data collected within the first 12 weeks of CZP therapy, we constructed several models using CART to predict which patients would achieve LDA (DAS28 $\leq$ 3.2) at 1 year. The benefit of this approach is that it allows for inclusion of multiple predictors and allows patients to be further classified based upon whether they had a very early response (4–6 weeks). Furthermore, these early time points reflect visit intervals at which a patient could be reasonably assessed in clinical practice; measuring predictor variables of response at shorter time points (e.g., 2 weeks after initiating anti-TNF therapy) may not be realistically feasible outside of controlled trials.

This study has several limitations. Despite the use of a random split sample methodology possible with the large RAPID 1 dataset, with separate testing and training datasets, the potential for overfitting (i.e., model fails to provide accurate predictions when applied to new subjects/datasets) remains. Furthermore, model accuracy reported herein may be specific to the biologic or biologic class examined (CZP, or perhaps only anti-TNF therapies), or applicable only to the types of patients recruited to RAPID 1 (i.e., those with high disease activity and established RA). To address this generalizability issue, we replicated a recently published decision tree prediction model that used clinical assessments at Week 12 and earlier to predict LDA at 1 year derived from RA patients treated with etanercept in the TEMPO trial (33). The prediction model built from the TEMPO data appeared to perform similarly well using RAPID1 data. Based upon this empiric example, we would suggest that the prediction models represented in our results (which had much more data than in the previous analysis (33)) will perform well for established RA patients with high disease activity treated with other anti-TNF agents. Additional replication studies, in RA patient populations receiving differing drug treatments, will be useful to confirm whether these models can be applied more broadly. However, we suspect that different prediction models will be required for patients with early RA and for those who start with lower disease activity than RAPID1 patients.

Despite these caveats, our results suggest that it may be possible to develop predictive tools that are user friendly and could be easily and consistently applied in clinical practice. Using an analytic framework like CART, biomarker and pharmacogenetic information could be added to prediction tools and likely would complement clinical assessments to prospectively guide the management of individual RA patients. Designing a trial around the concept of predicting response at 12 weeks or earlier, and altering therapy for patients predicted to be nonresponders, would be an optimal approach and should be tested. Prediction models similar to the type we have developed also have the potential to improve treatment to target approaches (24) by identifying groups of patients who should be switched to an alternate strategy more quickly, which ultimately could improve patient outcomes.

SIGNIFICANCE AND INNOVATIONS

- Classification and regression trees (CART) have been successfully utilized in a number of therapeutic areas to categorize disease state and to determine the likelihood of future events.

- The ability to quickly predict achievement of outcomes in patients with rheumatoid arthritis (RA) at 1 year would enable physicians to tailor treatment early during the course of therapy.
- Using CART, it was possible to predict one-year response within the first 12 weeks of starting therapy with CZP.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This analysis was funded by UCB. We acknowledge the editorial services of Shelley Lindley from PAREXEL, which were funded by UCB.

Disclosures/Financial Interests

Dr. Curtis has research grants and/or consulting/honoraria from Amgen, Abbott, UCB, Pfizer, BMS, Centocor, Genentech/Roche, and CORRONA. He also receives support from the NIH (AR053351) and AHRQ (R01HS018517). Dr. Luijters is an employee of UCB. Prof. Kavanaugh has research grants and/or consulting/honoraria from UCB, Amgen, Abbott, Centocor, Roche, BMS, and Celgene.

References

1. Keystone EC, Genovese MC, Klareskog L, Hsia EC, Hall ST, Miranda PC, et al. Golimumab, a human antibody to tumour necrosis factor {alpha} given by monthly subcutaneous injections, in active rheumatoid arthritis despite methotrexate therapy: the GO-FORWARD Study. *Ann Rheum Dis.* 2009; 68(6):789–796. [PubMed: 19066176]
2. Weinblatt ME, Kremer JM, Bankhurst AD, Bulpitt KJ, Fleischmann RM, Fox RI, et al. A trial of etanercept, a recombinant tumor necrosis factor receptor:Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate. *N Engl J Med.* 1999; 340(4):253–259. [PubMed: 9920948]
3. Bathon JM, Martin RW, Fleischmann RM, Tesser JR, Schiff MH, Keystone EC, et al. A comparison of etanercept and methotrexate in patients with early rheumatoid arthritis. *N Engl J Med.* 2000; 343(22):1586–1593. [PubMed: 11096165]
4. Klareskog L, van der Heijde DM, de Jager JP, Gough A, Kalden J, Malaise M, 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(9410):675–681. [PubMed: 15001324]
5. Smolen JS, Kay J, Doyle MK, Landewe R, Matteson EL, Wollenhaupt J, et al. Golimumab in patients with active rheumatoid arthritis after treatment with tumour necrosis factor alpha inhibitors (GO-AFTER study): a multicentre, randomised, double-blind, placebo-controlled, phase III trial. *Lancet.* 2009; 374(9685):210–221. [PubMed: 19560810]
6. Bombardieri S, Ruiz AA, Fardellone P, Geusens P, McKenna F, Unnebrink K, et al. Effectiveness of adalimumab for rheumatoid arthritis in patients with a history of TNF-antagonist therapy in clinical practice. *Rheumatology (Oxford).* 2007; 46(7):1191–1199. [PubMed: 17504821]
7. Abe T, Takeuchi T, Miyasaka N, Hashimoto H, Kondo H, Ichikawa Y, et al. A multicenter, double-blind, randomized, placebo controlled trial of infliximab combined with low dose methotrexate in Japanese patients with rheumatoid arthritis. *J Rheumatol.* 2006; 33(1):37–44. [PubMed: 16395748]
8. Ramiro S, van Tubergen AM, Landewe RB. RAPID and FAST4WARD trials: certolizumab pegol for rheumatoid arthritis. *Expert Rev Clin Immunol.* 2010; 6(5):713–720. [PubMed: 20828279]
9. Fleischmann R, Vencovsky J, van Vollenhoven R, Borenstein D, Box J, Coteur G, et al. Efficacy and safety of certolizumab pegol monotherapy every 4 weeks in patients with rheumatoid arthritis failing previous disease-modifying antirheumatic therapy: the FAST4WARD study. *Ann Rheum Dis.* 2009; 68(6):805–811. [PubMed: 19015206]

10. Keystone E, Curtis JR, Fleischmann RM, Furst D, Khanna D, Smolen JS, et al. Rapid improvement in the signs and symptoms of rheumatoid arthritis following certolizumab pegol treatment predicts better long-term outcomes: post-hoc analysis of a randomized controlled trial. *Arthritis Res Ther*. 2011; 38(6):990–996.
11. Wessels JA, van der Kooij SM, Le Cessie S, Kievit W, Barerra P, Allaart CF, et al. A clinical pharmacogenetic model to predict the efficacy of methotrexate monotherapy in recent-onset rheumatoid arthritis. *Arthritis Rheum*. 2007; 56(6):1765–1775. [PubMed: 17530705]
12. Hider SL, Silman AJ, Thomson W, Lunt M, Bunn D, Symmons DP. Can clinical factors at presentation be used to predict outcome of treatment with methotrexate in patients with early inflammatory polyarthritis? *Ann Rheum Dis*. 2009; 68(1):57–62. [PubMed: 18292102]
13. Mugnier B, Balandraud N, Darque A, Roudier C, Roudier J, Reviron D. Polymorphism at position -308 of the tumor necrosis factor alpha gene influences outcome of infliximab therapy in rheumatoid arthritis. *Arthritis Rheum*. 2003; 48(7):1849–1852. [PubMed: 12847678]
14. Padyukov L, Lampa J, Heimburger M, Ernestam S, Cederholm T, Lundkvist I, et al. Genetic markers for the efficacy of tumour necrosis factor blocking therapy in rheumatoid arthritis. *Ann Rheum Dis*. 2003; 62(6):526–529. [PubMed: 12759288]
15. Seitz M, Wirthmuller U, Moller B, Villiger PM. The -308 tumour necrosis factor-alpha gene polymorphism predicts therapeutic response to TNFalpha-blockers in rheumatoid arthritis and spondyloarthritis patients. *Rheumatology (Oxford)*. 2007; 46(1):93–96. [PubMed: 16720636]
16. Flowers N, Wolfe F. What do rheumatologists do in their practices? [abstract]. *Arthritis Rheum*. 1998; 41(suppl):S337.
17. Schoels M, Aletaha D, Smolen JS, Bijlsma JW, Burmester G, Breedveld FC, et al. Follow-up standards and treatment targets in rheumatoid arthritis: results of a questionnaire at the EULAR 2008. *Ann Rheum Dis*. 2010; 69(3):575–578. [PubMed: 19389716]
18. Bukkapatnam S, Komanduri R, Yang H, Rao P, Lih WC, Malshe M, et al. Classification of atrial fibrillation episodes from sparse electrocardiogram data. *J Electrocardiol*. 2008; 41(4):292–299. [PubMed: 18367198]
19. Mockel M, Danne O, Muller R, Vollert JO, Muller C, Lueders C, et al. Development of an optimized multimarker strategy for early risk assessment of patients with acute coronary syndromes. *Clin Chim Acta*. 2008; 393(2):103–109. [PubMed: 18423399]
20. Kohrt HE, Olshen RA, Bermas HR, Goodson WH, Wood DJ, Henry S, et al. New models and online calculator for predicting non-sentinel lymph node status in sentinel lymph node positive breast cancer patients. *BMC Cancer*. 2008; 8:66. [PubMed: 18315887]
21. Duvvuri SK, Finkelstein J. Computerized decision support for asthma management. *Value Health*. 2006; 9(3):A101–A102.
22. Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LH, Funovits J, et al. American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Arthritis Rheum*. 2011; 63(3):573–586. [PubMed: 21294106]
23. Keystone E, van der Heijde D, Mason D, Landewe R, van Vollenhoven R, Combe B, et al. Certolizumab pegol plus methotrexate is significantly more effective than placebo plus methotrexate in active rheumatoid arthritis: findings of a fifty-two-week, phase III, multicenter, randomized, double-blind, placebo-controlled, parallel-group study. *Arthritis Rheum*. 2008; 58(11):3319–3329. [PubMed: 18975346]
24. Smolen JS, Landewe R, Breedveld FC, Dougados M, Emery P, Gaujoux-Viala C, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs. *Ann Rheum Dis*. 2010; 69(6):964–975. [PubMed: 20444750]
25. Breiman, L.; Friedman, J.; Stone, CJ.; Olshen, RA. *Classification and Regression Trees*. Boca Raton, FL: CRC Press LLC; 1984.
26. Steinberg, D.; Colla, P. *CART: tree-structured non-parametric data analysis*. San Diego CA: Salford Systems; 1995.

27. Wolfe F, Pincus T, O'Dell J. Evaluation and documentation of rheumatoid arthritis disease status in the clinic: which variables best predict change in therapy. *J Rheumatol*. 2001; 28(7):1712–1717. [PubMed: 11469486]
28. Emery P, Reginster JY, Appelboom T, Breedveld FC, Edelman E, Kekow J, et al. WHO Collaborating Centre consensus meeting on anti-cytokine therapy in rheumatoid arthritis. *Rheumatology (Oxford)*. 2001; 40(6):699–702. [PubMed: 11426031]
29. 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(2):157–163. [PubMed: 15637039]
30. Furst DE, Keystone EC, Fleischmann R, Mease P, Breedveld FC, Smolen JS, et al. Updated consensus statement on biological agents for the treatment of rheumatic diseases, 2009. *Ann Rheum Dis*. 2010; 69(Suppl 1):i2–i29. [PubMed: 19995740]
31. Aletaha D, Funovits J, Keystone EC, Smolen JS. Disease activity early in the course of treatment predicts response to therapy after one year in rheumatoid arthritis patients. *Arthritis Rheum*. 2007; 56(10):3226–3235. [PubMed: 17907167]
32. Dougados M, Schmidely N, Le Bars M, Lafosse C, Schiff M, Smolen JS, et al. Evaluation of different methods used to assess disease activity in rheumatoid arthritis: analyses of abatacept clinical trial data. *Ann Rheum Dis*. 2009; 68(4):484–489. [PubMed: 19074177]
33. Curtis JR, Yang S, Chen L, Park GS, Bitman B, Wang B, et al. Predicting low disease activity and remission using early treatment response to antitumour necrosis factor therapy in patients with rheumatoid arthritis: exploratory analyses from the TEMPO trial. *Ann Rheum Dis*. 2011

Figure 1.

Graph showing probability of LDA at 1 year (A) versus absolute DAS28 and/or (B) vs. DAS28 improvement from baseline at Week 12.

Overall accuracy of model: 65%

$$(PRN \times [AC/100]) + (MT \times [AC/100]) + (PNR \times [AC/100])$$

$$(23.5 \times [70.6/100]) + (53.5 \times [50.0/100]) + (23.0 \times [94.0/100])$$

*Actual responders = achieved LDA at 1 year. AC = node accuracy; PNR = % of patients predicted to be a nonresponder (data from predicted nonresponder nodes); MT = % of patients needing more time on treatment for classification

Figure 2.

A-priori model: constructed based only on clinical input using DAS28 change from baseline of >1.2 units at Weeks 4 and 12 as predictors.

Overall accuracy of model: 76%

$(PR \times [AC/100]) + (MT \times [AC/100]) + (PNR \times [AC/100])$
 $(34.6 \times [65.3/100]) + (11.5 \times [56.0/100]) + (53.9 \times [86.3/100])$

*Actual responders = achieved LDA at 1 year. AC = node accuracy; PR = % of patients predicted to be a responder (data from predicted responder node); PNR = % of patients predicted to be a nonresponder (data from predicted nonresponder node); MT = % of patients needing more time on treatment for classification

Figure 3.

CART Model 1: all potential predictors (measured at baseline and at Weeks 4, 6, 8, and 12) included to predict low disease activity (DAS28 ≤ 3.2) at 1 year.

Variables and cut points were derived empirically using CART. Results shown are from the testing dataset only since performance of the models using the training dataset was generally superior.

Overall accuracy of model: 79%

$$(PR \times [AC/100]) + (MT \times [AC/100]) + (PNR \times [AC/100])$$

$$(30.0 \times [73.8/100]) + (24.4 \times [66.0/100]) + (45.6 \times [89.9/100])$$

*Actual responders = achieved LDA at 1 year. AC = node accuracy; PR = % of patients predicted to be a responder (data from predicted responder node); PNR = % of patients predicted to be a nonresponder (data from predicted nonresponder node); MT = % of patients needing more time on treatment for classification

Figure 4.

CART Model 2: variation on Model 1 (Figure 3), 3:1 misclassification cost applied to maximize accuracy for predicted nonresponders.

Outcome is low disease activity (DAS28 ≤3.2) at 1 year. Variables and cut points were derived empirically using CART. Results shown are from the testing dataset only since performance of the models using the training dataset was generally superior.

^aCategorized CDAI: high being >22

Overall accuracy of model: 71%

$$(PR \times [AC/100]) + (MT \times [AC/100]) + (PNR \times [AC/100])$$

$$(19.8 \times [65.1/100]) + (25.3 \times [56.4/100]) + (54.8 \times [79.8/100])$$

*Actual responders = achieved LDA at 1 year. AC = node accuracy; PR = % of patients predicted to be a responder (data from predicted responder node); PNR = % of patients predicted to be a nonresponder (data from predicted nonresponder node); MT = % of patients needing more time on treatment for classification

Figure 5.

CART Model 3: selected predictors easy to measure in daily practice.

Variables and cut points were derived empirically using CART but included only those that were more easily measured in clinical practice (e.g. CDAI, SJC; not DAS28). Results shown are from the testing dataset only since performance of the models using the training dataset was generally superior.

Table 1

Baseline characteristics

	RAPID 1 ITT[†] N = 982	Analysis population[‡] N = 703
Patient demographics and characteristics [*]		
Age, years	52.0 (11.6)	51.7 (11.5)
Gender (% female)	83.2	82.6
Disease duration, years	6.1 (4.3)	6.1 (4.3)
Concomitant methotrexate dose, mg/week	13.6 (4.2)	13.6 (4.1)
Rheumatoid factor (% positive [≥ 14 IU/mL])	81.8	81.5
Disease activity status [*]		
Tender joint count (0–68)	30.7 (12.9)	31.0 (12.9)
Swollen joint count (0–66)	21.5 (9.8)	21.8 (10.0)
Patient's assessment of arthritis pain, 0–100 mm VAS	63.1 (18.9)	62.7 (18.7)
HAQ-DI	1.7 (0.61)	1.7 (0.60)
DAS28 (median)	7.0	7.0
CRP, mg/L (geometric mean [CV])	14.7 (144.2)	14.2 (141.4)

^{*} Given as mean (SD) unless otherwise specified.

[†] (Keystone et al 2008).

[‡] The population represented in this analysis comprise 393 patients treated with CZP 200 mg plus methotrexate plus 390 patients treated with CZP 400 mg plus methotrexate in the RAPID 1 study; 80 patients withdrew due to reasons other than lack of efficacy and were excluded. RAPID 1 patients initially randomized to placebo also were excluded.

CRP = C-reactive protein; CV = coefficient of variation; HAQ-DI = Health Assessment Questionnaire, Disability Index; VAS = visual analog scale.