Clinical and Experimental Rheumatology 2010; 28: 258-260.

Evaluation of disease activity assessments in patients with rheumatoid arthritis and an inadequate response to anti-TNF therapy: analyses of abatacept clinical trial data

M. Dougados[1], G. Wells[2], N. Schmidely[3], M. Le Bars[3], P. van Riel[4], D. Aletaha[5], M. Schiff[6], J. Smolen[5]

[1]Rene Descartes University, Medicine Faculty; UPRES–EA 4058; APHP, Hospital Cochin, Rheumatology B Department, Paris, France; [2]University of Ottawa, Ottawa, ON, Canada; [3]Bristol-Myers Squibb, Rueil-Malmaison, France; [4]University Hospital Nijmegen, Nijmegen, The Netherlands; [5]Department of Rheumatology, Medical University of Vienna, Vienna, Austria; [6]University of Colorado, Denver, CO, USA.

Maxime Dougados, MD
George Wells, PhD
Nathalie Schmidely, MSc
Manuela Le Bars, MD, PhD
Piet van Riel, MD, PhD
Daniel Aletaha, MD, MSc
Michael Schiff, MD
Joseph Smolen, MD

Please address correspondence and reprint requests to:
Maxime Dougados,
Hopital Cochin, Rene Descartes University, Service de Rhumatologie B,
27 Rue du Faubourg Saint Jacques, Paris 75014, France.
E-mail:
maxime.dougados@cch.ap-hop-paris.fr

Received on October 9, 2009; accepted in revised form on December 10, 2009.

Key words: Rheumatoid arthritis, disease activity assessment, clinical relevance.

Conflict of interest: Dr Dougados has participated at various symposia organised by Bristol-Myers Squibb and has acted as a consultant in various advisory boards of BMS; Dr Wells has received grants and honoraria from BMS; Dr Schmidely and Dr Bars are both employees of BMS; Dr van Riel has served on the advisory boards and has received research grants from BMS, Wyeth, Merck, Abbott and Schering-Plough; the other co-others have declared no competing interests.

ABSTRACT

Objective. *To assess the ability of efficacy measures that incorporate onset or sustainability to detect treatment effect or reflect patient satisfaction, using exploratory analyses of data from the ATTAIN (Abatacept Trial in Treatment of Anti-TNF INadequate Responders) trial.*

Methods. *218 abatacept- and 99 placebo-treated patients were evaluated. Reporting methods included time to onset (first American College of Rheumatology [ACR] 50 response/Low Disease Activity State [LDAS; DAS28 ≤3.2]) and sustainability of ACR50/LDAS, both assessed according to discriminatory capacity (number of patients needed to study [NNS]) and patient satisfaction with treatment.*

Results. *Efficacy measures incorporating elements of sustainability or onset decreased discriminatory capacity, while sustainability, but not onset of action, was important in reflecting patient satisfaction.*

Conclusions. *Optimal assessment methods depend on whether the outcome of interest is ability to detect treatment effects or to reflect patient satisfaction. Sustainability of response (and possibly, at a lower magnitude, fast onset of action) may be important when evaluating patient satisfaction with RA therapies in patients who have previously failed anti-TNF therapy.*

Introduction

In clinical trials in rheumatoid arthritis (RA), current gold-standard assessments are based on American College of Rheumatology (ACR) responses, and Disease Activity Score 28 (DAS28) and corresponding European League Against Rheumatism (EULAR) response criteria (1). EULAR and ACR recommendations emphasise the importance of examining treatment response and disease status (1), and highlight the value of examining onset and sustainability of a clinical response/disease state.

A *post-hoc* analysis of methotrexate (MTX)-inadequate responders in the AIM (Abatacept in Inadequate responders to MTX) trial (2) showed that the performance of ACR- and DAS28-based criteria incorporating measures of onset and sustainability varied depending on the outcome of interest. The highest discriminatory capacity (*i.e.* lowest number of patients needed to study (NNS) to detect treatment effects) was observed when onset of action and sustainability were not considered, whereas faster onset and increased sustainability of response/status were important in reflecting patient satisfaction (2).

Patients with insufficient response to MTX alone do not adequately represent the current RA patient population, as over the past decade many patients have received anti-tumour necrosis factor (TNF) therapies. In clinical trials, patients who have experienced prior anti-TNF therapy exhibit consistently lower ACR responses than MTX-inadequate responders or MTX-naïve patients (3). To investigate whether these observations also apply to patients who have failed anti-TNF agents, we explored the performance of efficacy measures that incorporate onset or sustainability using data from the Phase III ATTAIN (Abatacept Trial in Treatment of Anti-TNF INadequate Responders) trial (4). These difficult-to-treat patients have long-standing RA and high levels of disease activity. The efficacy and safety of abatacept in this patient population has been reported previously (reviewed in (5-7)).

Methods

Database

These analyses were exploratory assessments of the 6-month, randomised, double-blind, placebo-controlled ATTAIN trial of abatacept (plus at least one disease-modifying antirheumatic drug) in patients with active RA and inadequate response to ≥3 months of anti-TNF. The trial design has been reported previously (4). For these analyses, abatacept was considered the 'active' drug.

Disease activity assessments

Treatment response was assessed by ACR50, and disease status by DAS28 Low Disease Activity State (LDAS; DAS28 ≤3.2). To determine the significance of onset of action, proportions of

Table I. Performance of ACR and DAS28 criteria based on onset of action according to discriminatory capacity and ability to reflect patient satisfaction.

Onset occurred in the first:	ACR50			LDAS (DAS28 ≤3.2)		
	Patients, n (%) (abatacept + placebo, n=317)	NNS	LR+ (95% CI)	Patients, n (%) (abatacept + placebo, n=317)	NNS	LR+ (95% CI)
1 month	24 (7.6)	405	2.79 (0.85, 9.11)	11 (3.5)	387	1.79 (0.39, 8.12)
2 months	53 (16.7)	487	2.14 (1.05, 4.37)	30 (9.5)	539	1.31 (0.58, 2.94)
3 months	79 (24.9)	234	2.39 (1.33, 4.30)	42 (13.3)	315	1.69 (0.82, 3.51)
4 months	91 (28.7)	276	2.13 (1.28, 3.57)	51 (16.1)	203	1.86 (0.94, 3.65)
5 months	110 (34.7)	133	2.44 (15.0, 3.97)	68 (21.5)	82	2.61 (1.35, 5.03)
6 months	119 (37.5)	90	2.49 (1.56, 3.97)	74 (23.3)	76	2.88 (1.50, 5.52)

Categories overlap, so patients experiencing a response in the first month will also be counted in all other categories; Abatacept and placebo treatment groups were pooled for patient satisfaction, and not pooled for NNS; A lower NNS value indicates greater discriminatory capacity; Higher LR+ values indicate greater probability of satisfaction with treatment (9); ACR: American College of Rheumatology; DAS28: Disease Activity Score 28; LDAS: Low Disease Activity State (DAS28 [C-reactive protein] ≤3.2); NNS: number of patients needed to study; LR+: positive likelihood ratio; CI: confidence interval.

Table II. Performance of ACR and DAS28 criteria based on sustainability of response according to discriminatory capacity and ability to reflect patient satisfaction.

Response or status achieved for:	ACR50			LDAS (DAS28 ≤3.2)		
	Patients, n (%) (abatacept + placebo, n=317)	NNS	LR+ (95% CI)	Patients, n (%) (abatacept + placebo, n=317)	NNS	LR+ (95% CI)
≥1 visit over 6 months	119 (37.5)	90	2.49 (1.56, 3.97)	74 (23.3)	76	2.88 (1.50, 5.52)
≥2 consecutive visits	56 (17.7)	116	21.5 (3.02, 153.04)	38 (12.0)	94	14.73 (2.05, 105.74)
≥3 consecutive visits	34 (10.7)	71	Infinite* (–)	18 (5.7)	283	Infinite* (–)
≥4 consecutive visits	25 (7.9)	107	Infinite* (–)	12 (3.8)	835	Infinite* (–)
≥5 consecutive visits	16 (5.1)	205	Infinite* (–)	8 (2.5)	757	Infinite* (–)
6 consecutive visits	8 (2.5)	NA	Infinite* (–)	2 (0.6)	NA	Infinite* (–)

*An infinite LR+ generally indicates that all patients are satisfied in the presence of a positive clinical outcome, however, infinite values should be interpreted with caution when the n-number of patients with a positive clinical outcome is low; A lower NNS value indicates greater discriminatory capacity; Higher LR+ values indicate greater probability of observing satisfaction with treatment (9); Abatacept and placebo treatment groups were pooled for patient satisfaction, and not pooled for NNS; ACR: American College of Rheumatology; DAS: Disease Activity Score 28; LDAS: Low Disease Activity State (DAS28 [C-reactive protein] ≤3.2); NNS: Number of patients needed to study; LR+: positive likelihood ratio; CI: confidence interval.

patients achieving their first ACR50 and LDAS within 1, 2, 3, 4, 5 or 6 months were assessed. To determine the significance of sustainability of treatment response and disease status, proportions of patients achieving ACR50 and LDAS for at least 1, 2, 3, 4, 5 or 6 consecutive visits over 6 months were calculated. Equal weighting was applied for each visit.

Evaluation of disease activity assessments

Discriminatory capacity was calculated based on number of patients required per treatment arm to perform a two-arm 1:1 randomised study comparing active treatment with placebo, based on a difference similar to that observed in ATTAIN. Numbers of patients required were calculated with a standard testing procedure (α=0.05 (two-tailed), β=0.20, *chi*-squared test for binary variables and student's *t*-test for continuous variables). Lower NNS indicates greater discriminatory capacity.

Satisfaction with treatment was assessed at Month 6. Patients were asked 'how would you rate your satisfaction with the treatment you received?', with answers rated on a 5-point scale from 1 (excellent) to 5 (poor) (2). Responses were dichotomised as 1–3 (favourable) versus 4–5 (not favourable). In previous analyses exploring alternative cut-offs for dichotomisation, results were not affected by cut-off choice (2). Data from the abatacept and placebo groups were pooled at each timepoint.

To assess the ability of clinical measures that incorporate onset and sustainability (based on ACR and DAS28 criteria) to reflect patient satisfaction, positive likelihood ratios (LR+) (8) were calculated, with higher values indicating that reporting measures reflect patient satisfaction (2). Based on the literature, LR+ >2 is considered of relevant prognostic value (9). An infinite LR+ generally indicates good prognostic value, although results should be interpreted with caution when a low proportion of patients achieve success.

Statistical analysis

The sample size was based on the original primary efficacy analyses (4). Statistical testing is, therefore, inappropriate for our analyses, which are considered exploratory. Analyses are based on patients with data available at the visit of interest (as-observed).

Results

Patient disposition

Data are presented from 317 patients (abatacept 218, placebo 99) who completed the double-blind period and entered the long-term ATTAIN exten-

sion. Demographic and baseline clinical characteristics were comparable across treatment groups (4).

Treatment response and disease activity assessments

Of the 317 patients in the pooled abatacept/placebo population, 18% achieved an ACR50 response and 14% achieved LDAS at Month 6. Seventy percent reported 'favourable' treatment satisfaction at Month 6.

For discriminatory capacity, there was a trend toward increasing NNS when considering earlier onset, for both ACR50 and LDAS (Table I). NNS was lower if the response or status was achieved within the first month compared with occurrence in the second month; however, this may reflect the smaller patient numbers achieving ACR50 or LDAS during Month 1. For patient satisfaction with treatment, LR+ was <3 regardless of time to onset for both ACR50 and LDAS. There was no apparent consistent pattern between time of onset and ability to reflect patient satisfaction, although LR+ values for LDAS were <2 for onset at Months 1–4, and >2 after Month 4.

For both ACR50 and LDAS, NNS was generally higher with increasing sustainability, although data were unavailable for six consecutive visits (Table II). NNS for ACR50 was higher for responses sustained over at least two visits compared with at least three or four visits. By contrast, sustainability of response was an important factor for reflecting patient treatment satisfaction for both treatment response and disease status: LR+ scores for ACR50 and LDAS increased progressively from 2.49 and 2.88 (sustained for at least 1 visit), respectively, to infinite (sustained for at least 3 visits), although these results should be interpreted with caution due to the low patient numbers with these responses.

Discussion

Results from this exploratory analysis in abatacept-treated patients with RA and inadequate response to anti-TNF therapy are generally consistent with previous observations in MTX-inadequate responders (2). These results further demonstrate that the performance of methods of reporting treatment response/disease status varies depending on the outcome of interest (*e.g.* ability to detect treatment effects or reflect patient satisfaction).

As observed in MTX-inadequate responders (2), discriminatory capacity was reduced (*i.e.* NNS increased) when faster onset of action or greater sustainability were considered. Given the higher numbers of patients required to detect treatment effects using ACR- and DAS28-based onset and sustainability responses in a TNF-refractory population, it may not be beneficial to consider these aspects when designing clinical trials. Also consistent with previous findings in MTX-inadequate responders, the sustainability of response or disease status was important in reflecting patient satisfaction. Interestingly, onset of action did not appear to be important for determining treatment satisfaction in patients with previous anti-TNF therapy failure. These data differ from previous reports in MTX-inadequate responders, where onset of action did affect satisfaction (2). This could be explained by the fact that those patients who are refractory to conventional DMARDs and some biological therapies may be more willing to accept a therapy with a slower onset of action, if it provides a sustained treatment effect.

Several limitations of this analysis should be noted. This was an exploratory analysis of one trial evaluating a single compound (abatacept). Consequently, any recommendations arising from our findings require corroboration in analyses from databases evaluating other compounds (*e.g.* anti-TNF agents). In addition, for most analyses the observed LR+ values were inconclusive as the values were <10 or even <5, the thresholds often used to reflect 'relevant' values for diagnostic purposes (9). However, there is no clear, accepted definition of relevant LR+ thresholds. Finally, the numbers of patients achieving very early onset (within 1 month) or maximal sustainability (outcome achieved for at least six consecutive visits) are too low for valid comparisons.

In conclusion, our analyses suggest that the optimal assessment method depends on whether the outcome of interest is ability to detect treatment effects or ability to reflect patient satisfaction. Sustainability of response (particularly with respect to patient satisfaction) may be an important factor to consider when assessing the efficacy of RA therapies for patients with previous anti-TNF therapy failure. These findings are consistent with recent EULAR/ACR recommendations highlighting the importance of sustainability of clinical response or disease state (1).

Acknowledgements

This study was supported in part by an unrestricted research grant-in-aid from Bristol-Myers Squibb. Editorial assistance was provided by Medicus International and funded by Bristol-Myers Squibb. Under direction of the authors, editorial assistance was provided at the first draft stage, and during subsequent revision of the manuscript.

References

1. ALETAHA D, LANDEWE R, KARONITSCH T *et al*.: Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations. *Ann Rheum Dis* 2008; 67: 1360-4.
2. DOUGADOS M, SCHMIDELY N, LE BARS M *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: 484-9.
3. SMOLEN JS, ALETAHA D, KOELLER M, WEISMAN MH, EMERY P: New therapies for treatment of rheumatoid arthritis. *Lancet* 2007; 370: 1861-74.
4. GENOVESE MC, BECKER JC, SCHIFF M *et al*.: Abatacept for rheumatoid arthritis refractory to tumor necrosis factor alpha inhibition. *N Engl J Med* 2005; 353: 1114-23.
5. CHOY EH: Selective modulation of T-cell co-stimulation: a novel mode of action for the treatment of rheumatoid arthritis. *Clin Exp Rheumatol* 2009; 27: 510-8
6. ROZELLE AL, GENOVESE MC: Efficacy results from pivotal clinical trials with abatacept. *Clin Exp Rheumatol* 2007; 25 (Suppl. 46): S30-4
7. MICHAUD K, BOMBARDIER C, EMERY P: Quality of life in patients with rheumatoid arthritis: does abatacept make a difference? *Clin Exp Rheumatol* 2007; 25 (Suppl. 46): S35-45.
8. PERERA R, HENEGHAN C: Making sense of diagnostic test likelihood ratios. *ACP J Club* 2007; 146: A8-9.
9. DEEKS JJ: Systematic reviews in health care: systematic reviews of evaluations of diagnostic and screening tests. *BMJ* 2001; 323: 157-62.