

LETTERS

DOI 10.1002/art.39981

Reanalysis of the multi-biomarker disease activity score for assessing disease activity in the Abatacept versus Adalimumab Comparison in Biologic-Naive Rheumatoid Arthritis Subjects with Background Methotrexate study: comment on the article by Fleischmann et al

To the Editor:

We read with great interest the report by Fleischmann et al (1), which presents post hoc analyses of multi-biomarker disease activity (MBDA) scores measured using serum samples from AMPLE (Abatacept versus Adalimumab Comparison in Biologic-Naive Rheumatoid Arthritis Subjects with Background Methotrexate), a study sponsored by Bristol-Myers Squibb that compared abatacept versus adalimumab in rheumatoid arthritis (RA) patients with inadequate response to methotrexate (2). The article reported 3 main results: 1) in the first year, mean MBDA scores decreased significantly less with abatacept treatment than with adalimumab, yet clinical responses by Disease Activity Score in 28 joints using the C-reactive protein level (DAS28-CRP) (3) were similar between the 2 treatment groups; 2) RA disease activity category (i.e., low, moderate, or high) as classified by the MBDA score was often discordant with the classification according to the DAS28-CRP, Clinical Disease Activity Index (CDAI) (4), Simplified Disease Activity Index (SDAI) (5), or Routine Assessment of Patient Index Data 3 (6); and 3) radiographic data were interpreted as demonstrating that MBDA scores were not associated with radiographic progression. Based on these 3 results, it was concluded that the MBDA score should not be used to guide decision-making in the management of RA. We wish to demonstrate limitations of these analyses that raise questions regarding the interpretation of these 3 results, as well as the overall conclusion that was reached.

Addressing the 3 results in reverse order, we note that the relationship between radiographic progression and MBDA scores was assessed using a method that seems inadequate to test the desired hypothesis, that the MBDA score is associated with radiographic nonprogression. This analysis, presented in Figure 2D of the article, seemingly shows the proportion of patients whose disease did not progress radiographically within each MBDA category. In other words, one would expect the denominator of this proportion to be the number of patients within each MBDA category, and the numerator to be the number of non-progressors in that category. However, upon careful inspection, this is not what the figure shows. Instead, Figure 2D describes the proportions of patients who were nonprogressors, classified as having low, moderate, or high disease activity by MBDA score. This approach is not informative about the relationship between MBDA scores and radiographic progression because the same denominator was used for each MBDA category. Thus, the proportions reflect only the distribution of MBDA scores among the nonprogressors. A more conventional analysis, such as that described below, or cumulative probability plots showing changes in modified total Sharp scores (7) by MBDA category, would be more informative to determine the likelihood of radiographic progression, conditional on patient MBDA category.

To this point, we reanalyzed the year 1 radiographic outcomes using the data provided in the original publication (1) to determine if the percentage of radiographic progressors increased with increasing MBDA scores, using previously published methods (8). We computed the proportion of radiographic progressors in each MBDA category by dividing the number of progressors in each MBDA category by the number of patients in that category (1). However, the exact numbers of patients with missing data (overall 8 of 189 in the abatacept arm and 4 of 190 in the adalimumab arm), and of patients with radiographic progression (overall 19 of 189 and 21 of 190, respectively), were not provided for the individual MBDA categories. Therefore, to establish the boundaries on all possible results that would be compatible with the data provided, we conducted a sensitivity analysis that varied the distributions of progressors and patients with missing data across MBDA categories. At one extreme (Scenario 1, least conservative), all progressors were assigned to the highest possible MBDA categories. At the other extreme (Scenario 2, most conservative), all progressors were assigned to the lowest possible MBDA categories (Supplementary Tables 1 and 2, on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39981/abstract>).

The results of the above-described reanalysis are shown in Figure 1. In Scenario 1, there was a strong and statistically significant association between MBDA category and radiographic progression in both the abatacept and adalimumab arms. At the other extreme, in Scenario 2, the adalimumab results were statistically significant and there was a similar, albeit nonsignificant, trend in the abatacept arm ($P = 0.068$). Given that the actual result must lie at or between the extremes of Scenarios 1 and 2 in this sensitivity analysis, the results as presented in Figure 2D of the article by Fleischmann et al do not support the conclusion that MBDA scores did not reflect radiographic progression status in the AMPLE trial (1). Rather, our reanalysis indicates that the MBDA category is positively associated with radiographic progression in the AMPLE study, as has been reported in other RA cohorts (9–13). We invite replication of this reanalysis using patient-level data.

Second, the MBDA score was initially developed to correlate with the DAS28-CRP, and its RA disease activity categories were established using cutoffs that are specific to the DAS28-CRP (14). Thresholds for DAS28-CRP RA disease activity categories are systematically lower than the corresponding cutoffs in the DAS28 using the erythrocyte sedimentation rate (ESR) (15,16), as noted in the editorial accompanying Fleischmann and colleagues' article (17). The comparison of MBDA scores versus DAS28-CRP in the article used the DAS28-ESR category thresholds, and therefore yielded more discordance than would have been expected between MBDA categories and DAS28-CRP categories using DAS28-CRP thresholds. For both the DAS28-CRP and the other clinical measures examined (e.g., CDAI, SDAI), the claim of no clear association between MBDA scores and commonly used, validated clinical measures was not supported by a statistical test of no association that cross-classified the MBDA category of each patient with his or her clinically defined RA disease activity category.

The copyright line for this article was changed on 16 September 2019 after original online publication.

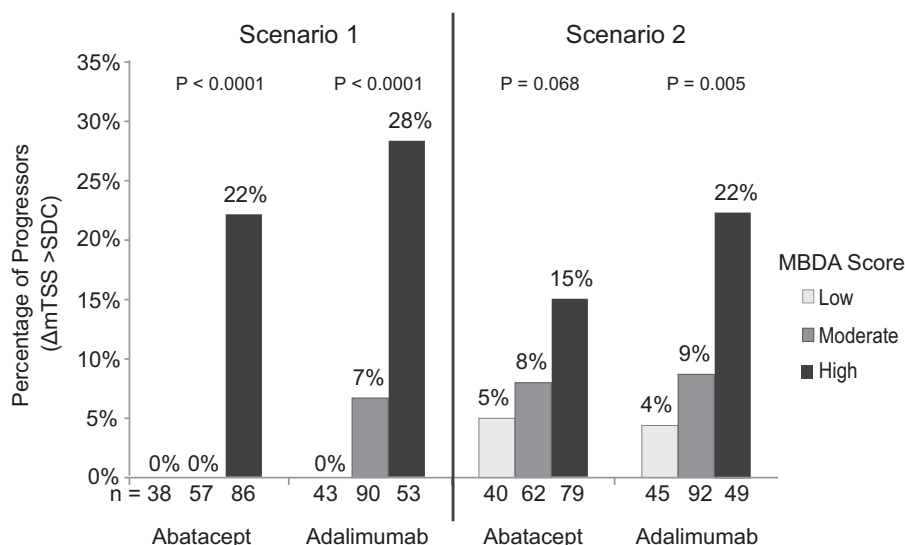


Figure 1. Sensitivity analysis examining the association between radiographic progression and multi-biomarker disease activity (MBDA) score in the AMPLE study (Abatacept versus Adalimumab Comparison in Biologic-Naive Rheumatoid Arthritis Subjects with Background Methotrexate). Percentages of patients with radiographic progression at year 1 were analyzed by year 1 MBDA category for the abatacept arm and adalimumab arm, using published data from Fleischmann et al (1). Progressors were defined as having changes in the modified total Sharp score ($\Delta mTSS$) exceeding the smallest detectable change (SDC) (1). The number of progressors in an individual MBDA category (low, moderate, or high) was determined in 2 steps. First, the total number of progressors in the 3 MBDA categories was calculated by subtracting the total number of nonprogressors, summed from Figure 2D in the report by Fleischmann and colleagues, from the total number of patients, obtained from Supplementary Table 5 in their report (1). Next, the total number of progressors was distributed across the 3 MBDA categories in 2 scenarios, representing the most extreme possibilities compatible with the reported data: Scenario 1 (least conservative) assigned all progressors to the highest possible MBDA categories; Scenario 2 (most conservative) assigned all progressors to the lowest possible MBDA categories (for numerical details, see Supplementary Tables 1 and 2, on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39981/abstract>). N values on the x-axis are the number of patients in the MBDA category, i.e., the sum of progressors (P) and nonprogressors (NP) in that MBDA category. The percentage of patients with radiographic progression in each MBDA category is $100 \times (P/[P + NP])$. Statistical significance was determined by Mantel-Haenszel test for trend, assuming ordinality in the MBDA categories.

Regardless, because the MBDA score was designed to complement, not supplant, clinical assessment, some discordance between clinical and laboratory-based assessments is not only expected (14,18,19), but desirable. Otherwise, the laboratory test would contain no incremental information beyond that offered by clinical evaluation. Several studies have shown that high MBDA scores were associated with radiographic progression and low scores with nonprogression (9–13), even when the patient was classified as having low disease activity or remission based on clinical measures (e.g., the DAS28-CRP) (9,12,20).

The third and last result reported by Fleischmann et al that we wish to address is that the mean decrease in MBDA score was larger for adalimumab-treated patients compared with abatacept-treated patients at month 3 (day 85) and beyond. The approximate difference between the 2 arms was 3–4 units at month 3, 4–5 units at year 1, and 0–1 units at year 2; these differences were statistically significant at month 3 and year 1 but not year 2. While these differences may appear large in Figure 1 of the article, where the y-axis scale ranged from +2 to –16 units, the MBDA score is measured on a 1–100 scale. More importantly, the difference between the 2 treatment arms at month 3 and year 1 is approximately equal to the measurement error of the MBDA score (4.5 units) (21). As was pointed out in the report (1), this small difference in the mean MBDA scores between the 2 treatment arms is likely clinically irrelevant.

For unclear reasons, data on approximately one-fifth of patients were missing from the AMPLE MBDA analysis, such that a systematic bias as to why they were not analyzed cannot be excluded. We also note that 31% of patients had missing MBDA data at year 2 compared with year 1, which exceeds the 8.5% decline in total patient numbers from the end of year 1 to the end of year 2 in the overall AMPLE trial (2,22).

In summary, the analysis by Fleischmann et al has several limitations that raise uncertainties about the interpretations presented in the article. We note that the authors reached the conclusion that the MBDA score is not useful for RA patient management based on interpretations that questioned the validity of the MBDA test in this one study, without considering all the available evidence. We would encourage readers to make the distinction between the scientific validity of a diagnostic test and its clinical utility. There is already a sizable evidence base supporting the development and validation of the MBDA test in diverse RA patient cohorts (9–13,18–20,23–25). A prospective clinical trial is underway to rigorously evaluate its clinical utility and its potential role in RA patient management (26).

Dr. Curtis has received consulting fees, speaking fees, and/or honoraria from Crescendo Bioscience, Inc., Pfizer, and Bristol-Myers Squibb (less than \$10,000 each) and from UCB, Amgen, Janssen, and the CORRONA registry (more than \$10,000 each). Dr. Wright has received consulting fees, speaking fees, and/or honoraria from Medac Pharma (less than

\$10,000) and from AbbVie, Bristol-Myers Squibb, and Crescendo Bioscience, Inc. (more than \$10,000 each). Dr. Strand has received consulting fees, speaking fees, and/or honoraria from AbbVie, Alder, Amgen, Bristol-Myers Squibb, Boehringer Ingelheim, Celgene, Celltrion, the CORRONA registry, Crescendo Bioscience, Inc., GlaxoSmithKline, Janssen, Lilly, Merck, Novartis, Pfizer, Protagen, Regeneron, Samsung, Sandoz, Sanofi, and UCB (less than \$10,000 each). Dr. Davis has received consulting fees from Crescendo Bioscience, Inc. (less than \$10,000). Drs. Hiraya and Sasso are shareholders of Myriad Genetics, Inc.

Jeffrey R. Curtis, MD, MS, MPH
 University of Alabama at Birmingham
 Grace C. Wright, MD, PhD
 New York University Langone Medical Center
 New York, NY
 Vibeke Strand, MD
 Stanford University School of Medicine
 Stanford, CA
 Charles S. Davis, PhD
 CSD Biostatistics, Inc.
 Oro Valley, AZ
 Elena Hiraya, MD, PhD
 Eric H. Sasso, MD
 Crescendo Bioscience
 South San Francisco, CA

- Fleischmann R, Connolly SE, Maldonado MA, Schiff M. Estimating disease activity using multi-biomarker disease activity scores in rheumatoid arthritis patients treated with abatacept or adalimumab. *Arthritis Rheumatol* 2016;68:2083–9.
- Weinblatt ME, Schiff M, Valente R, van der Heijde D, Citera G, Zhao C, et al. Head-to-head comparison of subcutaneous abatacept versus adalimumab for rheumatoid arthritis: findings of a phase IIIb, multinational, prospective, randomized study. *Arthritis Rheum* 2013;65:28–38.
- Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts: development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;38:44–8.
- Aletaha D, Nell VP, Stamm T, Uffmann M, Pflugbeil S, Machold K, et al. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. *Arthritis Res Ther* 2005;7:R796–806.
- Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G, et al. A Simplified Disease Activity Index for rheumatoid arthritis for use in clinical practice. *Rheumatology (Oxford)* 2003;42:244–57.
- Pincus T, Bergman MJ, Yazici Y. RAPID3—an index of physical function, pain, and global status as “vital signs” to improve care for people with chronic rheumatic diseases. *Bull NYU Hosp Jt Dis* 2009;67:211–25.
- Van der Heijde DM. How to read radiographs according to the Sharp/van der Heijde method [corrected and republished in *J Rheumatol* 2000;27:261–3]. *J Rheumatol* 1999;26:743–5.
- Emery P, Genovese MC, van Vollenhoven R, Sharp JT, Patra K, Sasso EH. Less radiographic progression with adalimumab plus methotrexate versus methotrexate monotherapy across the spectrum of clinical response in early rheumatoid arthritis. *J Rheumatol* 2009;36:1429–41.
- Van der Helm-van Mil AH, Knevel R, Cavet G, Huizinga TW, Haney DJ. An evaluation of molecular and clinical remission in rheumatoid arthritis by assessing radiographic progression. *Rheumatology (Oxford)* 2013;52:839–46.
- Markus IM, Dirven L, van den Broek M, Bijkerk C, Han KH, Ronday HK, et al. A multibiomarker disease activity score for rheumatoid arthritis predicts radiographic joint damage in the BeSt study. *J Rheumatol* 2014;41:2114–9.
- Hambardzumyan K, Bolce R, Saevarsdottir S, Cruickshank SE, Sasso EH, Chernoff D, et al. Pretreatment multi-biomarker disease activity score and radiographic progression in early RA: results from the SWEFOT trial. *Ann Rheum Dis* 2015;74:1102–9.
- Li W, Sasso EH, van der Helm-van Mil AH, Huizinga TW. Relationship of multi-biomarker disease activity score and other risk factors with radiographic progression in an observational study of patients with rheumatoid arthritis. *Rheumatology (Oxford)* 2016;55:357–66.
- Hirata S, Li W, Kubo S, Fukuyo S, Mizuno Y, Hanami K, et al. Association of the multi-biomarker disease activity score with joint destruction in patients with rheumatoid arthritis receiving tumor necrosis factor- α inhibitor treatment in clinical practice. *Mod Rheumatol* 2016. E-pub ahead of print.
- Curtis JR, van der Helm-van Mil AH, Knevel R, Huizinga TW, Haney DJ, Shen Y, et al. Validation of a novel multibiomarker test to assess rheumatoid arthritis disease activity. *Arthritis Care Res (Hoboken)* 2012;64:1794–803.
- Inoue E, Yamanaka H, Hara M, Tomatsu T, Kamatani N. Comparison of Disease Activity Score (DAS)28- erythrocyte sedimentation rate and DAS28- C-reactive protein threshold values. *Ann Rheum Dis* 2007;66:407–9.
- Fleischmann R, van der Heijde D, Koenig AS, Pedersen R, Szumski A, Marshall L, et al. How much does Disease Activity Score in 28 joints ESR and CRP calculations underestimate disease activity compared with the Simplified Disease Activity Index? *Ann Rheum Dis* 2015;74:1132–7.
- Davis JM III. The Multi-Biomarker Disease Activity (MBDA) test for rheumatoid arthritis: is it a valid measure of disease activity? [editorial]. *Arthritis Rheumatol* 2016;68:2061–6.
- Bakker MF, Cavet G, Jacobs JW, Bijlsma JW, Haney DJ, Shen Y, et al. Performance of a multi-biomarker score measuring rheumatoid arthritis disease activity in the CAMERA tight control study. *Ann Rheum Dis* 2012;71:1692–7.
- Hirata S, Dirven L, Shen Y, Centola M, Cavet G, Lems WF, et al. A multi-biomarker score measures rheumatoid arthritis disease activity in the BeSt study. *Rheumatology (Oxford)* 2013;52:1202–7.
- Hambardzumyan K, Bolce RJ, Saevarsdottir S, Forslund K, Wallman JK, Cruickshank SE, et al. Association of a multibiomarker disease activity score at multiple time-points with radiographic progression in rheumatoid arthritis: results from the SWEFOT trial. *RMD Open* 2016;2:e000197.
- Chernoff D, Bolce RJ, Hwang CC, Wang X, Kivitz A, Curtis JR. Examination of diurnal and daily variation of the Multi-Biomarker Disease Activity (MBDA) score in RA to establish the minimally important difference [abstract]. *Arthritis Rheumatol* 2016;68 Suppl 10. URL: <http://acrabstracts.org/abstract/examination-of-diurnal-and-daily-variation-of-the-multi-biomarker-disease-activity-mbda-score-in-ra-to-establish-the-minimally-important-difference/>.
- Schiff M, Weinblatt ME, Valente R, van der Heijde D, Citera G, Elegbe A, et al. Head-to-head comparison of subcutaneous abatacept versus adalimumab for rheumatoid arthritis: two year efficacy and safety findings from AMPLE trial. *Ann Rheum Dis* 2014;73:86–94.
- Eastman PS, Manning WC, Qureshi F, Haney D, Cavet G, Alexander C, et al. Characterization of a multiplex, 12-biomarker test for rheumatoid arthritis. *J Pharm Biomed Anal* 2012;70:415–24.
- Hirata S, Defranoux N, Kentaro H, Yamaoka K, Tanaka Y. A multi-biomarker disease activity score for monitoring rheumatoid arthritis. *Curr Biomark Find* 2015;5:69–78.
- Van Vollenhoven RF, Bolce R, Hambardzumyan K, Saevarsdottir S, Forslund K, Petersson IF, et al. Enhancement of patient recruitment in rheumatoid arthritis clinical trials using a multi-biomarker disease activity score as an inclusion criterion. *Arthritis Rheumatol* 2015;67:2855–60.
- Crescendo Bioscience, sponsor. Prospective outcomes study: Vectra DA guided care compared to usual care. *ClinicalTrials.gov* identifier: NCT02832297; 2016.