Published in final edited form as:

Ann Rheum Dis. 2016 June ; 75(6): 1114–1119. doi:10.1136/annrheumdis-2014-206934.

Validity of early MRI structural damage end points and potential impact on clinical trial design in rheumatoid arthritis

Joshua F Baker^{1,2,3}, Philip G Conaghan⁴, Paul Emery⁴, Daniel G Baker⁵, and Mikkel Østergaard⁶

¹Philadelphia Veterans' Affairs Medical Center, Philadelphia, Pennsylvania, USA

²University of Pennsylvania, School of Medicine, Philadelphia, Pennsylvania, USA

³Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, Pennsylvania, USA

⁴Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, and NIHR Leeds Musculoskeletal Biomedical Research Unit, Leeds, UK

⁵Janssen Research & Development, LLC., Horsham, Pennsylvania, USA

⁶Department of Clinical Medicine, Faculty of Health and Medical Sciences, Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Glostrup Hospital, University of Copenhagen, Copenhagen, Denmark

Abstract

Objective—To evaluate the construct validity of the rheumatoid arthritis MRI score (RAMRIS) erosion evaluation as structural damage end point and to assess the potential impact of incorporation in clinical trials.

Methods—In a randomised trial of early methotrexate-naïve RA (GO-BEFORE), RAMRIS scores were determined from MRIs and van der Heijde-Sharp (vdHS) scores from radiographs, at baseline, week 12, week 24 and week 52. Progression in damage scores was defined as change >0.5. Associations of X-ray and MRI outcomes with clinical features were evaluated for

To request permissions go to: <http://group.bmj.com/group/rights-licensing/permissions>

Correspondence to: Dr Joshua F Baker, Division of Rheumatology, Department of Medicine, 8 Penn Tower Building, 34th and Civic Center Blvd., Hospital of the University of Pennsylvania, Philadelphia, PA 19104, USA; bakerjo@uphs.upenn.edu. JFB and PGC contributed equally.

Contributors All authors made substantial contributions to the conception, design, analysis and interpretation of the data. All authors contributed to the drafting and revision of the work and gave final approval of the version published. The authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Competing interests JFB has nothing to disclose. PGC has done speakers bureaus or consultancies for BMS, Janssen, Merck, Pfizer, and Roche. PE has received consulting fees, speaking fees and/or honoraria from Pfizer, Merck, AbbVie, UCB, Roche, BMS, Lilly and Novartis (less than \$10 000 each). DGB is an employee of Janssen Biotech. MØ has received fees for consultancy or speaker fees and/or research support from Abbott, AbbVie, BMS, Boehringer-Ingelheim, Celgene, Centocor, Eli-Lilly, GSK, Janssen, Merck, Mundipharma, Novo, Pfizer, Schering-Plough, Roche, UCB, Takeda and Wyeth.

Patient consent Obtained.

Ethics approval Exempt Status through University of Pennsylvania.

Provenance and peer review Not commissioned; externally peer reviewed.

convergent validity. Iterative Wilcoxon rank sum tests and tests of proportion estimated the sample size required to detect differences between combination therapy (methotrexate+golimumab) and methotrexate-monotherapy arms in (A) change in damage score and (B) proportion of patients progressing.

Results—Patients with early MRI progression had higher DAS28, C reactive protein (CRP) and vdHS at baseline, and higher 2-year HAQ. Associations were similar to those with 1-year vdHS progression. Differences in change in structural damage between treatment arms achieved significance with fewer subjects when 12-week or 24-week MRI erosion score was the outcome (150 patients; 100 among an enriched sample with baseline-synovitis >5) compared with the 52-week vdHS (275 patients). Differences in the proportion progressing could be detected in 234 total subjects with 12-week MRI in an enriched sample whereas 1-year X-ray required between 468 and 1160 subjects.

Conclusions—Early MRI erosion progression is a valid measure of structural damage that could substantially decrease sample size and study duration if used as structural damage end point in RA clinical trials.

Prevention of structural damage in rheumatoid arthritis (RA) has become an important aspect of the management of the disease and a critical outcome in clinical trials of new therapies. For example, the Food and Drug Administration currently requires that clinical studies of new pharmacological therapies demonstrate that a therapy prevents structural damage progression before it may receive an indication for that purpose. Standard X-ray has been the gold standard measure of structural joint damage for nearly two decades.

Studies using X-ray as the primary structural outcome generally require relatively large sample sizes and extended follow-up periods (1 year) to demonstrate significant differences from placebo arms. Improvements in the standard of care have raised ethical questions about the duration of placebo arms in RA trials. In the future, it is likely that sample sizes will necessarily increase in clinical trials in order to detect increasingly small differences between active treatment arms using current structural damage assessments.

New imaging modalities, including musculoskeletal MRI have improved sensitivity and discriminative characteristics over X-ray. We have previously shown that early changes in MRI measures of synovitis, bone oedema and bone erosion are predictive of subsequent X-ray progression.¹ However, it remains to be fully established if early changes in MRI erosion are a valid measure of structural damage resulting from active RA, and how the use of early (12-week and 24-week) MRI measures as a structural end point in a clinical trial setting would impact the efficiency of the study design. In this study we aimed to determine the convergent validity of early MRI erosion progression at 12 weeks and 24 weeks compared with 52-week X-ray progression, and to determine if an early MRI erosion end point would significantly reduce estimated sample sizes in identifying an effective therapy.

Methods

Study setting

This prospective cohort study is ancillary to the GO-BEFORE trial ([Clintrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT00361335) identifier NCT00361335) that included 637 patients with RA with 52 weeks of follow-up. Methods and results of the original trial have been previously reported.²³ The original study compared the efficacy of methotrexate (MTX) or golimumab (GLM) alone to combination therapy with MTX and GLM in MTX and biologic therapy naïve subjects. Three hundred and eighteen total subjects at eligible study sites (based on technical capabilities) participated in the MRI substudy. The trial was conducted according to the principles of the Declaration of Helsinki. As such, all patients provided written informed consent before participating in the study.

Patients 18 years or older who met American College of Rheumatology (ACR) 1987 criteria for RA for at least the past 3 months and had active disease were recruited into the MRI substudy at participating sites. Patient visits occurred at regular 4-week intervals as part of the original trial. Data collection at each visit included independent, blinded assessments of disease activity using the DAS28 with C reactive protein (CRP) (DAS28 (CRP)).

Of the 318 participants, 291 had MRIs of adequate quality to be scored for bone oedema/erosions and 272 were adequate to be scored for synovitis at baseline. At 24 weeks, 280 were scored for bone oedema and erosions, and 268 were scored for synovitis.

MRI

MRIs of the dominant hand at baseline and week 12, week 24, week 52 and week 104 were obtained at participating trial centres. MRIs of the patient's dominant wrist and second to fifth metacarpophalangeal joints were obtained using 1.5 T MRI with contrast enhancement. The MR sequences were as follows: axial T1 fast spin echo precontrast, coronal T1 fast spin echo precontrast, coronal short τ inversion recovery (or T2 fat-suppressed precontrast) and coronal T1 fat-suppressed postcontrast.

Images were scored by two independent readers who were blinded to the image time point or sequence (visit number), patient identity and treatment group. The average score of two readers was determined for synovitis (0–9 for wrist joint, 0–21 for wrist plus metacarpophalangeal joints), bone oedema (osteitis, 0–69) and bone erosions (0–230), using the OMERACT rheumatoid arthritis MRI score (RAMRIS) system.¹⁴ Intrarater and inter-rater reliability coefficients (R) ranged from 0.72 to 0.95 and 0.69 to 0.89, respectively.

The change in MRI erosion at 12 weeks and 24 weeks was calculated by comparing with baseline scores. Where MRI erosion change scores were dichotomised due to non-normal distributions, a change in MRI erosion score of >0.5 was considered structural damage progression as previously described.¹⁵ A change of >0.5 demonstrated the best test characteristics in predicting later radiographic progression.¹

Radiographs of hands and feet

Radiographs of hands and feet were performed at baseline, week 24 and week 52. Radiographs were scored by two blinded readers using the van der Heijde-Sharp (vdHS) method. Change from baseline in vdHS scores at 24 weeks and 52 weeks was determined using centralised readers and standardised methods, as previously described.⁶⁷ Where analysed as a dichotomous variable, X-ray progression was defined as a change in vdHS score of >0.5 . This cut-off has previously been chosen to reduce misclassification error.^{18–10} Also studied were two other frequently used definitions of progression: a change in vdHS of >0 ¹¹ and >3 units.¹²

Statistical analysis

Data were analysed with STATA V.11 software (StataCorp, LP, College Station, Texas, USA). To assess the convergent validity of MRI erosion as a study end point, descriptive statistics were used to evaluate the differences in disease activity and disease severity measures between subjects who demonstrated progression in the RAMRIS erosion score or the vdHS score. Analyses were performed with an intention-to-treat approach to mimic the original trial design.

For sample size analyses, we based analyses on the assumption that combination therapy with MTX and GLM was superior to MTX monotherapy in preventing structural damage progression. Therefore, we aimed to determine the study sample size (assuming equal size treatment arms) that would be required to demonstrate the true difference between the combination therapy and MTX monotherapy arms in terms of structural damage progression. The 50 mg and 100 mg doses of GLM were similar in radiographic outcomes (not shown) and estimates of sample size and therefore combined.

Since the change in MRI erosion scores and change in vdHS scores over the study period were highly skewed and exhibiting kurtosis, sample size calculations using a mean-comparison tool were not considered appropriate. The sample size calculation was therefore approached in two ways. First, by varying the number of randomly selected subjects included in the analysis by 25 subjects at a time, a bootstrap method using iterative Wilcoxon rank sum tests (25 iterations) was used to assess differences in the change in MRI erosion and vdHS scores between the MTX monotherapy and combination therapy arms. The total sample size and the average z-statistic and corresponding p value were recorded. It was expected that increasing sample size would increase the probability that the radiographic outcome would be significantly different between the two treatment arms.

Estimated study sample sizes for X-ray and MRI structural outcomes were also determined by dichotomising each outcome as previously described. The proportion of subjects that demonstrated progression of structural damage on MRI/X-ray within the MTX monotherapy arms and combination arms was determined. Sample size calculations were performed to estimate required sample sizes using early MRI or X-ray progression as the dichotomous trial end point at 80% power. We defined structural damage progression as a change in RAMRIS erosion score or vdHS score >0.5 . We also explored differences using another commonly used cut-off for progression; namely a change of >0 unit and >3 units for vdHS.

In order to determine the impact of enriching the study population with subjects with active joint inflammation as assessed by baseline MRI synovitis or bone oedema, the sample sizes for MRI outcomes were also estimated after excluding subjects with a baseline synovitis score <5 (below the 25th centile for the study). In a previous study, a cut-off for baseline RAMRIS synovitis score of 5 was established as the best definition of an inflammatory activity acceptable state.¹³ The effect of excluding subjects with a bone oedema score 5 was also explored.

Results

Characteristics of the study population are shown in table 1, and have been previously described.¹⁶¹⁰¹⁴¹⁵ On average, subjects from the trial had relatively short disease duration and high disease activity at baseline. Subjects in the MRI substudy had baseline characteristics similar to the total study population and had high mean synovitis and bone oedema scores (table 1).

Convergent validity of MRI as structural end point

Subjects with MRI progression at 12 weeks and 24 weeks had higher 2-year HAQ scores, greater CRP levels, greater DAS28 (CRP) and greater vdHS scores at baseline compared with subjects who did not progress on MRI. MRI progression at 24 weeks was also associated with a lower likelihood of achieving an ACR50 response by 24 weeks in the study. These associations were similar in magnitude to those seen between clinical measures and X-ray progression at 1 year in the same subjects (table 2). Early progression in the MRI erosion score was associated with greater HAQ scores at 104 weeks, while progression on X-ray at 52 weeks was not associated.

Effect of use of RAMRIS erosion scores on calculated sample size

The change in MRI erosion score at 12 weeks was significantly different among those receiving combination therapy and MTX monotherapy (median 0 (−0.5, 0.5) vs 0 (0, 0.52) $p=0.02$, $N=190$). Similarly, among 194 subjects, the change at 24 weeks was significantly different between combination and monotherapy groups (median 0 (−0.5, 0.5) vs 0 (0, 0.76) $p=0.01$). Among 420 subjects with available data at 52 weeks, there was significantly less change in vdHS among those subjects in the combination group (median 0 (−0.12, 0.5) vs 0 (0, 1.5) $p=0.02$).

Figure 1 demonstrates the increased likelihood of identifying a statistical difference ($p<0.05$) between treatment groups with increasing total sample size among a random sample of subjects from the clinical trial. This figure demonstrates that the Wilcoxon rank sum tests assessing differences in treatment groups will achieve statistical significance with fewer subjects when MRI erosion score is the outcome assessed. Using 24-week or 52-week changes in vdHS as the outcome, the test did not, on average, achieve statistical significance until there were 300 or 275 total study subjects included in the analysis, respectively. In contrast, 12-week and 24-week changes in MRI erosion score required a total study sample size of 150 (ie, approximately 75 per study arm). Furthermore, when the study sample is

enriched with subjects with a synovitis score >5 at baseline, the total study sample was further reduced to 100 (50 per arm) for 12-week and 24-week MRI outcomes.

Sample size calculations using dichotomous progression outcomes

The dichotomisation of study end points results in a loss of information and reduced study power. Therefore, higher sample size estimates were seen overall using this method. Over all possible cut-offs, the discrimination of the effective therapy was greater for change in MRI erosion at 12 weeks (area under the curve (AUC) 0.60) and 24 weeks (AUC 0.61), compared with the change in vdHS at 52 weeks (AUC 0.56).

The calculated sample sizes per arm were 324 and 581 using an outcome of vdHS progression (>0.5) at 24 weeks and 52 weeks, respectively. At 52 weeks, the estimated sample size was 391 using an outcome of vdHS progression >3 (17/141 (12%) progressed in the MTX group and 17/279 (6%) progressed in the combination group) and 234 using an outcome of vdHS progression >0 (66/141 (47% progressed in the MTX group and 94/279 (34%) progressed in the combination group). Table 3 and figure 2 illustrate the overall reduction in sample size with the use of MRI progression as the structural damage outcome. For example, the calculated sample size (per group) using a dichotomous early MRI progression outcome at 12 weeks was 229. The sample size is further reduced with exclusion of subjects with a low synovitis or bone oedema score at baseline. For example, 117 subjects per treatment group enriched with baseline synovitis scores >5 would have 80% power to detect a difference in the proportion of subjects with RAMRIS erosion progression at 12 weeks.

Discussion

This study overall demonstrates the construct validity and increased efficiency of the use of MRI erosion in clinical trials studying the efficacy of new therapies in RA in preventing structural damage. We believe that these observations support the use of this modality in clinical trials as the primary study end point.

Associations between clinical and MRI measures have been demonstrated,¹⁶ and MRI measures of inflammation and greater CRP have been previously shown to predict change in the MRI erosion score.¹⁷ The current study goes further to evaluate and establish the construct validity of *progression* in the RAMRIS erosion score at early time points in a clinical trial. We previously showed that early RAMRIS erosion score progression predicted radiographic progression at 1 year. In this study, RAMRIS erosion score progression at 12 weeks and 24 weeks was also associated with other measures that would be expected to correlate with structural joint damage progression such as CRP, baseline structural damage and ACR response. These associations with clinical measures were similar to associations seen with X-ray progression at 52 weeks. Thus, progressive changes in MRI erosion score are very likely to be measuring the destructive consequences of the ongoing inflammatory disease.

To our knowledge, this is the first study to specifically evaluate the estimated sample size requirements for a clinical trial of anti-TNF therapy using MRI erosion as a structural

damage end point. Sample size estimates using early MRI end points were consistently smaller. If this study were repeated, it is highly likely that significant differences in MRI erosion at 12 weeks and 24 weeks would be identified between the treatment and placebo groups at 12 weeks in a study of only 75 subjects per group. These data suggest that the total study size to compare two treatments might be reduced substantially by incorporating MRI into clinical trial design. A previous study suggested that it may be possible to discriminate an effective therapy at 3 months with conventional radiography in certain settings.¹⁸ While our study does not rule out the potential use of conventional radiography to follow short-term changes in some circumstances, it indicates that in this particular clinical trial setting, it was inferior to MRI.

Furthermore, enrichment of the study population with subjects who had more active synovitis or more bone oedema at baseline on MRI, acted to further reduce the estimated sample sizes (to approximately 40–50 per group). The observation that enrichment of the study population using active disease on MRI can reduce the number of participants has even greater implications when one considers that this clinical trial population was composed of selected subjects with clinically active disease.

Overall, MRI was shown to have improved sensitivity to change compared with radiographs and these changes are likely to more efficiently discriminate between the effective and ineffective treatment arms. It is important to note that improvements in the methodology and MRI analysis technology have been made and other studies have demonstrated discrimination of the effective therapy with even smaller numbers than the conservative estimates presented here.^{19,20}

These observations have several major implications for clinical trial design. First, the convergent validity this study establishes suggests that the early change in MRI erosion score is meaningful. We have already reported the importance and predictive validity of corresponding early changes in MRI inflammatory measures. Second, these observations demonstrate that the use of MRI erosion change score at 12 weeks or 24 weeks, particularly in an enriched sample with active synovitis on MRI, would shorten the length of clinical trials, and reduce the sample sizes required to demonstrate significant differences between the treatment arms.

There are several limitations worth noting. There was some missing data as a result of dropouts during the trial as well as missing data for synovitis scores due to lack of contrast enhancement in some subjects (approximately 10%). Clinical characteristics were similar, however, among those who did not have synovitis scores performed. Synovitis was scored from coronal images, and more information would have been available from having additional axial images. For this study, an a priori cut-off for MRI progression was used. Future study might identify a more ideal cut-off for change in MRI erosion in discriminating treatment groups. Finally, this analysis was performed as an ancillary study to a previously completed clinical trial. Therefore, aspects of the original study design, including allowing for early escape, could potentially influence the magnitude of difference observed between the modalities over different follow-up periods. Overall, the limitations of the current study are felt to most likely result in underestimation of the improved discrimination using more

optimal MRI acquisitions, such as axial synovitis views. Further prospective studies using data from recent MRI RCTs should help to corroborate the evidence presented here.

In conclusion, these data support the incorporation of MRI erosion scores into clinical trial designs to study the effects of new therapies on structural joint damage progression in RA. We encourage regulatory agencies to consider MRI as a valid and efficient structural damage end point that, if used in trials, would decrease the sample size and durations of clinical trials, enhancing the development of novel therapies.

Acknowledgments

Funding JFB is supported by a Veterans Affairs Clinical Science Research and Development Career Development Award (IK2 CX000955).

References

1. Baker JF, Ostergaard M, Emery P, et al. Early MRI measures independently predict 1-year and 2-year radiographic progression in rheumatoid arthritis: secondary analysis from a large clinical trial. *Ann Rheum Dis*. 2013; 73:1968–74. [PubMed: 23904470]
2. Emery P, Fleischmann RM, Moreland LW, et al. Golimumab, a human anti-tumor necrosis factor alpha monoclonal antibody, injected subcutaneously every four weeks in methotrexate-naïve patients with active rheumatoid arthritis: twenty-four-week results of a phase III, multicenter, randomized, double-blind, placebo-controlled study of golimumab before methotrexate as first-line therapy for early-onset rheumatoid arthritis. *Arthritis Rheum*. 2009; 60:2272–83. [PubMed: 19644849]
3. Ostergaard M, Emery P, Conaghan PG, et al. Significant improvement in synovitis, osteitis, and bone erosion following golimumab and methotrexate combination therapy as compared with methotrexate alone: a magnetic resonance imaging study of 318 methotrexate-naïve rheumatoid arthritis patients. *Arthritis Rheum*. 2011; 63:3712–22. [PubMed: 22127693]
4. Ostergaard M, Peterfy C, Conaghan P, et al. OMERACT Rheumatoid Arthritis Magnetic Resonance Imaging Studies. Core set of MRI acquisitions, joint pathology definitions, and the OMERACT RA-MRI scoring system. *J Rheumatol*. 2003; 30:1385–6. [PubMed: 12784422]
5. Baker JF, Ostergaard M, George M, et al. Greater body mass independently predicts less radiographic progression on X-ray and MRI over 1–2 years. *Ann Rheum Dis*. 2014; 73:1923–8. [PubMed: 25091439]
6. Emery P, Fleischmann R, van der Heijde D, et al. The effects of golimumab on radiographic progression in rheumatoid arthritis: results of randomized controlled studies of golimumab before methotrexate therapy and golimumab after methotrexate therapy. *Arthritis Rheum*. 2011; 63:1200–10. [PubMed: 21305524]
7. Kremer J, Ritchlin C, Mendelsohn A, et al. Golimumab, a new human anti-tumor necrosis factor alpha antibody, administered intravenously in patients with active rheumatoid arthritis: forty-eight-week efficacy and safety results of a phase III randomized, double-blind, placebo-controlled study. *Arthritis Rheum*. 2010; 62:917–28. [PubMed: 20131276]
8. Emery P, Breedveld F, van der Heijde D, et al. Two-year clinical and radiographic results with combination etanercept-methotrexate therapy versus monotherapy in early rheumatoid arthritis: a two-year, double-blind, randomized study. *Arthritis Rheum*. 2010; 62:674–82. [PubMed: 20187135]
9. Aletaha D, Smolen JS. Joint damage in rheumatoid arthritis progresses in remission according to the Disease Activity Score in 28 joints and is driven by residual swollen joints. *Arthritis Rheum*. 2011; 63:3702–11. [PubMed: 21953215]
10. Baker JF, Conaghan PG, Smolen JS, et al. Development and validation of modified disease activity scores in rheumatoid arthritis: Superior correlation with MRI synovitis and X-ray progression. *Arthritis Rheum*. 2014; 66:794–802.

11. Genovese MC, Fleischmann R, Kivitz AJ, et al. Sarilumab plus methotrexate in patients with active rheumatoid arthritis and inadequate response to methotrexate: results of a phase III study. *Arthritis Rheumatol.* 2015; 67:1424–37. [PubMed: 25733246]
12. van der Helm-van Mil AH, Knevel R, Cavet G, et al. An evaluation of molecular and clinical remission in rheumatoid arthritis by assessing radiographic progression. *Rheumatology (Oxford).* 2013; 52:839–46. [PubMed: 23287359]
13. Gandjbakhch F, Haavardsholm EA, Conaghan PG, et al. Determining a magnetic resonance imaging inflammatory activity acceptable state without subsequent radiographic progression in rheumatoid arthritis: results from a followup MRI study of 254 patients in clinical remission or low disease activity. *J Rheumatol.* 2014; 41:398–406. [PubMed: 24334647]
14. Baker JF, Baker DG, Toedter G, et al. Associations between vitamin D, disease activity, and clinical response to therapy in rheumatoid arthritis. *Clin Exp Rheumatol.* 2012; 30:658–64. [PubMed: 22776409]
15. Baker JF, Mehta NN, Baker DG, et al. Vitamin D, metabolic dyslipidemia, and metabolic syndrome in rheumatoid arthritis. *Am J Med.* 2012; 125:1036.e9–15. [PubMed: 22800875]
16. Emery P, van der Heijde D, Ostergaard M, et al. Exploratory analyses of the association of MRI with clinical, laboratory and radiographic findings in patients with rheumatoid arthritis. *Ann Rheum Dis.* 2011; 70:2126–30. [PubMed: 21926186]
17. Palosaari K, Vuotila J, Takalo R, et al. Bone oedema predicts erosive progression on wrist MRI in early RA--a 2-yr observational MRI and NC scintigraphy study. *Rheumatology (Oxford).* 2006; 45:1542–8. [PubMed: 16670155]
18. Bruynesteyn K, Landewe R, van der Linden S, et al. Radiography as primary outcome in rheumatoid arthritis: acceptable sample sizes for trials with 3 months' follow up. *Ann Rheum Dis.* 2004; 63:1413–8. [PubMed: 15037445]
19. Peterfy C, Emery P, Tak PP, et al. MRI assessment of suppression of structural damage in patients with rheumatoid arthritis receiving rituximab: results from the randomised, placebo-controlled, double-blind RA-SCORE study. *Ann Rheum Dis.* 2014 Published Online First: 29 oct 2014. 10.1136/annrheumdis-2014-206015
20. Bowes MA, Guillard G, Gill E. Novel quantification of MRI provides a more sensitive outcome measure than RAMRIS. *Arthritis Rheumatol.* 2014; 66:S517.

Figure 1.

Iterative rank sum tests in random populations of varying size to estimate sample sizes required to demonstrate significant ($p < 0.05$) differences between the golimumab+MTX combination and MTX monotherapy arms using different study outcomes: (1) the change in X-ray at 52 weeks, (2) the change in MRI erosion score at 12/24-weeks, and (3) the change in MRI erosion score at 12/24 weeks among those with synovitis scores >5 at baseline. Abbreviations: MTX, Methotrexate; vdHS, van der Heijde-Sharp.

Figure 2.

Sample size estimations (per group) based on dichotomous X-ray and MRI end points for vdHS or rheumatoid arthritis MRI score (RAMRIS) at 80% power. All change scores dichotomised at a change of >0.5 unless otherwise noted. Abbreviations: vdHS, van der Heijde-Sharp.

Table 1
Basic characteristics of the study population

	GO-BEFORE Study	
	All Subjects	MRI Sub-Study
N	637	291
Age (years)	49.5 (12.3)	49.0 (12.6)
Female, N (%)	528 (82.9)	237 (81.4)
Race		
Asian, N (%)	117 (18.4)	75 (25.8)
White, N (%)	461 (72.4)	190 (65.3)
Disease duration (years)	1.2 (0.5, 4)	1.2 (0.6, 3.3)
DAS28(CRP)	5.72 (1.06)	5.60 (1.06)
CRP (mg/dL)	1.3 (0.5, 3.4)	1.2 (0.5, 2.7)
HAQ, baseline	1.53 (0.67)	1.56 (0.69)
HAQ, 104 weeks	0.76 (0.72)	0.76 (0.75)
CCP positive, N (%)	468 (73.5)	222 (76.3)
vdHS, baseline	6 (2, 21)	5.5 (2, 21.5)
RAMRIS scores		
Synovitis, baseline	–	9.5 (5.0)
Bone oedema, baseline	–	10.0 (10.0)
Bone erosion, baseline	–	14.5 (10, 22.5)

CCP, cyclic citrullinated peptide antibody; CRP, C reactive protein; DAS28, disease activity score of 28 joints; HAQ, health assessment questionnaire; RAMRIS, rheumatoid arthritis MRI score; RCT, randomized clinical trial; TNF, tumour necrosis factor; vdHS, van der Heijde-Sharp.

Table 2
Clinical characteristics of MRI/X-ray progressors and non-progressors within the MRI substudy (convergent validity)

12 week MRI						
MRI 12 weeks (N=231)			X-ray 52 weeks (N=231)			
	Progression (N=46)	No progression (N=185)	p Value	Progression (N=61)	No Progression (N=170)	p Value
Baseline DAS28(CRP)	5.89 (1.03)	5.56 (1.03)	0.06	5.86 (1.09)	5.54 (1.05)	0.05
Baseline CRP	1.8 (0.8, 3.6)	1 (0.5, 2.6)	0.01	2.4 (0.8, 4.4)	1 (0.4, 2.4)	0.0001
Baseline vdHS	11 (3.5, 27.5)	4.5 (2, 13)	0.004	8.5 (3, 24.5)	4.5 (2, 14.5)	0.01
HAQ score (week 104)*	0.88 (0.66)	0.64 (0.66)	0.03	0.77 (0.67)	0.65 (0.66)	0.2
24 week ACR50, N (%)	17/49 (37%)	85/183 (47%)	0.2	22/61 (36%)	80/168 (48%)	0.1
24 week MRI						
MRI 24 weeks (N=238)			X-ray 52 weeks (N=238)			
	Progression (N=52)	No Progression (N=186)	P Value	Progression (N=62)	No progression (N=176)	P Value
Baseline DAS28(CRP)	5.81 (1.13)	5.53 (1.06)	0.1	5.79 (1.10)	5.52 (1.06)	0.09
Baseline CRP	1.8 (0.95, 3.4)	1 (0.4, 2.5)	0.002	2 (0.8, 4.2)	1 (0.4, 2.4)	0.001
Baseline vdHS	15 (3.25, 29.8)	4.5 (2, 13)	0.002	9 (2.9, 26)	4.5 (2, 15)	0.01
HAQ score (week 104)*	0.82 (0.65)	0.63 (0.64)	0.07	0.72 (0.65)	0.66 (0.65)	0.5
24 Week ACR50, N (%)	14 (28%)	86 (47%)	0.01	21 (34%)	79 (45%)	0.1

Among those with available data for both outcomes.

* Adjusted for baseline HAQ.

ACR, American College of Rheumatology; CRP, C reactive protein; vdHS, van der Heijde-Sharp.

Table 3

Proportion of subjects treated with combination therapy or methotrexate monotherapy who progressed >0.5 on X-ray (vdHS) and MRI (RAMRIS erosion) and the estimated sample size (per group) to achieve statistical significance at 80% power

	MRI progression (12 weeks)		MRI progression (24 weeks)		X-ray progression (52 weeks)	
	No	Yes	No	Yes	No	Yes
All participants	(N=190)		(N=194)		(N=420)	
MTX monotherapy	46	16 (26%)	47	17 (27%)	97	44 (31%)
Combination MTX+GLM	109	19 (15%)	110	20 (15%)	213	66 (24%)
p value	0.07		0.06		0.1	
Estimated group N	229		196		581	
Synovitis >5	(N=147)		(N=152)			
MTX monotherapy	30	15 (33%)	32	16 (33%)		
Combination MTX+GLM	85	17 (17%)	86	18 (17%)		
p Value	0.02		0.03			
Estimated group N	117		127			
Bone oedema >5	(N=109)		(N=110)			
MTX monotherapy	20	11 (35%)	20	12 (38%)		
Combination MTX+GLM	64	14 (18%)	65	13 (17%)		
p Value	0.05		0.02			
Estimated group N	110		80			

GLM, golimumab; MTX, methotrexate; RAMRIS, rheumatoid arthritis MRI score; vdHS, van der Heijde-Sharp.