Supplementary Online Content

Møller-Bisgaard S, Hørslev-Petersen K, Ejbjerg B, et al. Effect of magnetic resonance imaging vs conventional treat-to-target strategies on disease activity remission and radiographic progression in patients with rheumatoid arthritis. *JAMA*. doi:10.1001/jama.2018.21362

This supplementary material has been provided by the authors to give readers additional information about their work.

eTable 1.

Acquired MRI sequences in the IMAGINE-RA trial

No	Type	Plane	ST (mm)	Gap (mm)	Resolution (mm)*	No slices	FOV (mm)	Averages	Time (min)
I. Applied MRI sequences, when wrist and MCP joints were covered by the same FOV									
1	Localizer	Ax, cor, sag							
2	STIR	Cor	2.5	0.6	0.6x0.6	15	180x145	4	App. 5
3	T1 FFE 3D	Cor	0.8 (0.4)	-	0.8x0.8 (0.4x0.4)	77	180x150	1	App. 5
4	T1 FFE 3D (dyn)	Cor	2 (1)	-	1.5x1.5 (0.75x0.75)	20	180x80	1	6.5 sec. (repeated for 4 minutes)
5	T1 FFE 3D	Cor	0.8 (0.4)	-	0.8x0.8 (0.4x0.4)	77	180x150	1	App. 5
II. Applied MRI sequences, when wrist and MCP joints were scanned separately									
Positioning A (MCP)									
1	Localizer	Ax, cor, sag							
2	STIR	Cor	2.5	0.6	0.6x0.6	15	180x145	4	App. 5
3	T1 FFE 3D	Cor	0.8 (0.4)	-	0.8x0.8 (0.4x0.4)	77	180x150	1	App. 5
Positioning B (wrist)									
4	Localizer	Ax, cor, sag							
5	STIR	Cor	2.5	0.6	0.6x0.6	15	180x145	4	App. 5
6	T1 FFE 3D	Cor	0.8 (0.4)	-	0.8x0.8 (0.4x0.4)	77	180x150	1	App. 5
7	T1 FFE 3D (dyn)	Cor	2 (1)	-	1.5x1.5 (0.75x0.75)	20	180x180	1	6.5 sec. (repeated for 4 minutes)
8	T1 FFE 3D	Cor	0.8 (0.4)	-	0.8x0.8 (0.4x0.4)	77	180x150	1	App. 5
Positioning A (MCP)									
9	Localizer	Ax, cor, sag							
10	T1 FFE 3D	Cor	0.8 (0.4)	-	0.8x0.8 (0.4x0.4)	77	180 x 150	1	App. 5

Abbreviations: Ax, Axial; Cor, Coronal; Dyn, Dynamic; FFE, Fast Field Echo; FOV, Field of view; MCP, Metacarpophalangeal; Sag, Sagittal; ST, Slice thickness; STIR, Short Tau Inversion Recovery; 3D: 3 Dimensional.
MRI of the 2nd to 5th MCP and wrist joints were performed at the same time, if the FOV was large enough to cover both regions. Alternatively, separate scans of MCP and wrist joints using individual FOVs were performed.

* When 2 sets of values are provided, e.g. ”0.8x0.8 (0.4x0.4)”, images were acquired with the first-mentioned slice thickness (e.g. 0.8 mm) but were reconstructed in the last-mentioned slice thickness (e.g. 0.4 mm).

eTable 2.

MRI units used during the study period

Hospital	MRI unit	Coil
Herlev-Gentofte, Herlev	1.0T Philips Panorama	3-channel wrist coil
Bispebjerg-Frederiksberg, Frederiksberg	1.5T Philips Ingenia	8-channel wrist coil
Slagelse	1.5T Siemens Magnetom Avanto	4-channel flex coil
Odense	1.0T Philips Panorama	3-channel wrist coil
King Christian Xth Hospital, Gråsten	1.0T ONI OrthOne	1-channel wrist coil
Aarhus*	3.0T Philips Achieva	8-channel wrist coil
Silkeborg	1.5T Siemens Avanto	16-channel flex coil
Hjørring	1.0T Philips Panorama	3-channel wrist coil

*At Aarhus hospital the coil was changed from an 8-channel knee coil to an 8-channel wrist coil during the study period.

eTable 3.

Composite disease activity scores used in the IMAGINE-RA study

Outcome measure	Definition
ACR-EULAR Boolean remission	At any time point a patient must satisfy all of the following: SJC≤1, TJC≤1, CRP≤1 mg/dL and patient VAS global [on a 0-10 scale] ≤1
CDAI: Clinical Disease Activity Index	SJC (0-28) + TJC (0-28) + PGA (on a 0-10 scale) + evaluator global assessment (on a 0-10 scale), [range 0-76] Remission defined as CDAI≤2.8
DAS28-CRP:	DAS28-CRP = 0.56*√(TJC28) + 0.28*√(SJC28) + 0.36*ln(CRP+1) + 0.014*GH + 0.96 Remission defined as DAS28-CRP<2.6
SDAI: Simple Disease Activity Index	TJC (0-28) + SJC (0-28) + PGA (on a 0-10 scale) + evaluator global assessment (on a 0-10 scale) + CRP mg/dL (0-10), [range 0-86] Remission defined as SDAI≤3.3

Abbreviations: CRP, serum C-Reactive protein; GH, General Health on a 100 mm Visual Analog Scale (0-100); SJC, swollen joint count; TJC, tender joint count; VAS, Visual Analog Scale.

eTable 4.

Secondary outcomes at 0-12, 12 and 12-24 months

		MRI Treat-to-Target[a]		Conventional Treat-to-Target[a]	Difference between groups[b]	P Value[b]
	No		No			
Clinical						
12 months						
Disease activity score in 28 joints C-reactive protein (DAS28-CRP) remission (DAS28-CRP<2.6), No. (%)	89	70 (78.7)	95	72 (75.8)	2.83 (-8.91 to 14.57)	0.64
American College of Rheumatology-European League Against Rheumatism (ACR-EULAR) Boolean remission, No. (%)	89	30 (33.7)	95	30 (31.6)	2.77 (-10.24 to 15.78)	0.68
Disease activity score in 28 joints CRP (DAS28-CRP)	89	2.1 (0.6)	95	2.2 (0.6)	-0.1 (-0.2 to 0.1)	0.41
Function and quality of life						
Baseline to 12 months						
Change in Health Assessment Questionnaire (HAQ)	89	-0.03 (0.20)	95	0.07 (0.20)	-0.09 (-0.15 to -0.03)	0.002
Change in EuroQol-5 dimensions	83	0.02 (0.12)	90	0.02 (0.12)	-0.00 (-0.04 to 0.03)	0.88
Change in Short Form 36 item questionnaire (SF-36) Physical Component Summary	82	0.0 (8.8)	92	0.1 (8.8)	-0.1 (-2.7 to 2.6)	0.97
Radiographic measures (van der Heijde modified Sharp score, vdHSS)						
Baseline to 12 months						
Change in total vdHSS	86	0.6 (2.3)	94	0.4 (2.3)	0.2 (-0.4 to 0.9)	0.48
No radiographic progression[c], No. (%)	85	65 (76.5)	94	78 (83.0)	-6.41 (-18.36 to 5.54)	0.29
12 to 24 months						
Change in total vdHSS	75	0.8 (3.5)	93	1.2 (3.5)	-0.4 (-1.5 to 0.7)	0.46
No radiographic progression[c], No. (%)	75	58 (77.3)	93	62 (66.7)	10.67 (-2.81 to 24.14)	0.12
Magnetic Resonance Imaging (MRI) measures (*RA magnetic resonance imaging scoring system* ***[RAMRIS])***						
12 months						
Osteitis score	88	3.7 (3.2)	92	3.4 (3.2)	0.3 (-0.6 to 1.2)	0.53
Synovitis score	87	5.7 (2.8)	91	5.9 (2.8)	-0.2 (-1.0 to 0.6)	0.59
Baseline to 12 months						
Change in erosion	89	0.1 (1.0)	94	0.3 (1.0)	-0.1 (-0.4 to 0.2)	0.36
No progression in MRI erosion, No. (%)	89	77 (86.5)	93	83 (89.2)	-2.89 (-12.31 to 6.53)	0.55
12 to 24 months						
Change in erosion	74	0.4 (1.7)	94	0.3 (1.7)	0.0 (-0.5 to 0.5)	0.94
No progression in MRI erosion, No. (%)	74	65 (87.8)	94	76 (80.9)	6.99 (-3.91 to 17.88)	0.21

[a]Data are presented as No. (%) for dichotomous data and least squares means (SD) for continuous data.
[b]For endpoints with dichotomous data, adjusted risk differences (RD) and 95% confidence intervals between groups were calculated by Wald tests and data for all 200 patients at 24 months were included, using multiple imputation. For endpoints with continuous data Least

Squares Mean differences between groups were calculated based on repeated-measures mixed linear models (or simple linear models if no repeated measures were collected between 0 and 12 months) adjusted for baseline values.
[c]No increase (change >0) in van der Heijde modified Sharp score (vdHSS) from baseline to month 24.

eTable 5.

Comparison of primary, secondary and exploratory outcomes at 24 months based on the Full analysis set[a]

	MRI T2T	Conventional T2T	Difference between groups	P Value[b]
	(n=98)	(n=99)	(95% CI)	
Primary endpoints				
Clinical:				
DAS28-CRP remission (DAS28-CRP<2.6), No. (%)	64 (85.3%)	83 (88.3%)	OR, 1.03 (0.31 to 3.43)	.958
Radiographic:				
No radiographic progression, No. (%)	49 (66.2%)	58 (62.4%)	OR, 1.19 (0.04 to 39.47)	.922
Secondary and exploratory endpoints				
Clinical				
ACR-EULAR Boolean remission, No. (%)	37 (49.3%)	30 (31.9%)	OR, 4.19 (1.30 to 13.57)	.017
SDAI remission (SDAI≤3.3), No (%)	48 (64.0%)	56 (62.2%)	OR, 1.67 (0.59 to 4.71)	.336
CDAI remission (CDAI≤2.8), No. (%)	53 (69.7%)	59 (64.8%)	OR, 2.75 (0.90 to 8.36)	.075
DAS28-CRP	1.9 (0.1)	2.1 (0.1)	-0.2 (-0.3 to 0.0)	.093
Tender joint count (0-28)	0.2 (0.1)	0.5 (0.1)	-0.2 (-0.6 to 0.1)	.171
Swollen joint count (0-28)	0.0 (0.1)	0.3 (0.1)	-0.3 (-0.5 to -0.0)	.038
Patient's global assessment of disease activity, VAS (0-100)	15.5 (1.8)	21.2 (1.6)	-5.7 (-10.4 to -0.9)	.019
Patient's assessment of pain VAS (0-100)	14.2 (1.7)	18.7 (1.6)	-4.5 (-9.0 to 0.0)	.052
Patient's assessment of fatigue (0-100)	21.8 (1.9)	24.4 (1.7)	-2.6 (-7.7 to 2.4)	.311
Investigator's global assessment of disease activity, VAS (0-100)	4.7 (0.8)	6.9 (0.8)	-2.3 (-4.4 to -0.1)	.041
Radiographic (vdHSS)				
Change in total vdHSS	1.0 (0.3)	1.3 (0.3)	-0.3 (-1.1 to 0.6)	.559
Change in erosion	0.6 (0.2)	0.8 (0.2)	-0.21 (-0.7 to 0.2)	.293
Change in joint space narrowing	0.4 (0.2)	0.5 (0.2)	- 0.0 (-0.6 to 0.5)	.933
MRI (RAMRIS)				
Change in osteitis	-1.8 (0.6)	-0.1 (0.5)	-1.8 (-3.2 to -0.3)	.018
Change in synovitis	-0.5 (0.3)	0.3 (0.3)	-0.8 (-1.8 to 0.1)	.074
Change in tenosynovitis	-0.9 (0.3)	0.3 (0.3)	-1.2 (-2.1 to-0.3)	.007
Change in combined inflammation	-2.9 (1.0)	0.7 (1.0)	-3.6 (-6.4 to -0.8)	.013
Change in erosion	0.5 (0.2)	0.6 (0.2)	-0.1 (-0.6 to 0.4)	.663
Change in joint space narrowing	0.1 (0.2)	0.4 (0.1)	-0.3 (-0.7 to 0.2)	.236
Change in combined damage	0.6 (0.3)	1.0 (0.3)	-0.3 (-1.2 to 0.5)	.395
No progression in MRI erosion, No. (%)	59 (79.7%)	70 (75.3%)	OR, 1.06 (0.02 to 66.59)	.976
Function and quality of life				
Change in Health Assessment Questionnaire	-0.052 (0.024)	0.091 (0.023)	-0.143 (-0.209 to -0.078)	<001

Patient with normal function (HAQ ≤0.5), No. (%)	61 (80.3%)	75 (79.8%)	OR, 0.73 (0.08 to 7.14)	.790
eTable 5 (continued)				
	MRI T2T	**Conventional T2T**	**Difference between groups**	**P Value[b]**
	(n=98)	**(n=99)**	**(95% CI)**	
Change in SF-36 Mental Component Summary	-0.5 (1.0)	-0.9 (0.9)	0.5 (-2.1 to 3.0)	.727
Change in SF-36 Physical Component Summary	1.1 (1.0)	-0.2 (0.9)	1.3 (-1.3 to 4.0)	.330
Change in EuroQol-5 dimensions	0.040 (0.015)	0.019 (0.013)	0.021 (-0.017 to 0.060)	.279
Treatment at year 2				
csDMARD monotherapy, No. (%)	24 (31.6)	68 (71.6)		
csDMARD combination therapy, No. (%)	17 (22.4)	25 (26.3)		
Biologic treatment	35 (46.1)	2 (2.1)		
First biologic, No. (%)	15 (19.7)	2 (2.1)		
Second or later biologic, No. (%)	20(26.3)	NA		

Abbreviations: 95% CI, 95% confidence interval; ACR, American College of Rheumatology; CDAI, Clinical Disease Activity Index; CRP, C-reactive protein; DAS28-CRP, disease activity score in 28 joints CRP-based; EULAR, European League Against Rheumatism; MRI, Magnetic Resonance Imaging; RAMRIS, RA magnetic resonance imaging scoring system; SDAI, Simplified Disease Activity Index; SF-36, Short Form 36 item questionnaire; T2T, treat-to-target; VAS, visual analog scale; vdHSS, van der Heijde modified Sharp score.

[a] Data are presented as least square means (SE) unless stated otherwise. For some variables, fewer patients were included in the analysis due to missing data. At 24 months data were available for minimum 69 patients in the MRI T2T-arm and 85 patients in the conventional T2T-arm.

[b] Analyses are based on repeated measures linear mixed models (continuous outcomes) or repeated measures generalized linear mixed models (binary outcomes), including participants as a random effect, with fixed factors for treatment group (2 levels) and time point (6 levels for disease activity measures and 2 levels for the x-ray imaging outcomes), and the corresponding interaction (time×group), adjusted for the value at baseline.

eTable 6.

Baseline characteristics of patients enrolled after the correct final registration at www.clinicaltrials.gov

	Intervention,[a]		
	MRI T2T (n=90)	Conventional T2T (n=88)	All (n=178)
Demographic characteristics			
Women, No. (%)	62 (69)	57 (65)	119 (67)
Age, mean (SD), y	62.0 (9.6)	60.3 (11.1)	61.2 (10.4)
Disease duration, y	9.0 [5.0;16.0]	10.0 [4.0;17.0]	9.0 [4.0;16.3]
Height, mean (SD), cm	169.4 (8.0)	168.4 (9.2)	168.9 (8.6)
Weight, mean (SD), kg	74.6 (14.6)	74.0 (14.9)	74.3 (14.7)
Clinical measures			
IgM-RF positive, No. (%)	79 (88)	78 (89)	157 (88)
CRP, mg/L	7.0 [4.0;8.0]	6.0 [3.0;9.0]	6.5 [3.3;8.8]
DAS28-CRP[b]	2.0 [1.7;2.2]	1.9 [1.6;2.1]	1.9 [1.7;2.2]
DAS28-CRP remission (DAS28-CRP<2.6), No. (%)	79 (88)	84 (95)	163 (92)
ACR-EULAR boolean remission, No. (%)	32 (36)	37 (42)	69 (39)
SDAI[d]	2.7 [1.6;4.0]	2.2 [1.2;3.5]	2.3 [1.4;3.8]
CDAI[e]	1.9 [0.8;3.4]	1.4 [0.6;2.8]	1.6 [0.7;3.2]
Tender joint count (0-28)	0.0 [0.0;0.0]	0.0 [0.0;0.0]	0.0 [0.0;0.0]
Swollen joint count (0-28)	0.0 [0.0;0.0]	0.0 [0.0;0.0]	0.0 [0.0;0.0]
Patient´s global assessment of disease activity, VAS (0-100)	14.5 [6.3;28.8]	9.0 [3.8;22.3]	11.0 [4.0;26.5]
Patient´s assessment of pain, VAS (0-100)	15.0 [5.0;25.5]	9.0 [3.8;20.5]	12.0 [5.0;22.0]
Patient´s assessment of fatigue, VAS (0-100)	19.0 [6.0;31.0]	12.0 [3.8;28.3]	15.0 [4.0;30.0]
Investigator´s global assessment of disease activity, VAS (0-100)	1.0 [0.0;4.0]	2.0 [1.0;5.0]	2.0 [1.0;5.0]
Radiographic measures (vdHSS)			
Total (0-448)	19.0 [8.0;42.0]	14.0 [7.0;36.3]	14.0 [7.0;40.0]
Erosion (0-280)	8.0 [3.0;17.0]	7.0 [3.0;12.8]	7.0 [3.0;16.0]
Joint space narrowing (0-168)	9.0 [2.0;25.0]	8.0 [2.0;22.8]	9.0 [2.0;25.0]
MRI measures (RAMRIS)			
Synovitis (0-21)	6.0 [3.0;8.0]	5.0 [2.5;9.0]	6.0 [3.0;9.0]
Osteitis (0-69)	2.0 [0.0;6.0]	2.0 [0.0;5.3]	2.0 [0.0;6.0]
Tenosynovitis (0-54)	3.0 [0.0;4.8]	3.0 [0.0;6.0]	3.0 [0.0;6.0]
Combined inflammation (0-144)[f]	11.0 [8.0;18.0]	13.0 [6.0;17.8]	12.0 [7.0;18.0]
Erosion (0-230)	4.0 [1.0;10.0]	5.0 [2.0;9.3]	4.0 [2.0;10.0]
Joint space narrowing (0-84)	1.0 [0.0;3.8]	1.0 [0.0;6.0]	1.0 [0.0;4.8]
Combined damage (0-314)[g]	4.5 [2.0;12.0]	6.5 [2.0;14.3]	5.5 [2.0;12.8]
Bone marrow edema present, No. (%)	39 (43)	31 (35)	70 (39)
Function and quality of life			
Health Assessment Questionnaire (0-3)[h]	0.250 [0.000;0.625]	0.125 [0.000;0.375]	0.125 [0.000;0.500]
SF-36 Mental Component Summary score (0-100)[i]	59.0 [54.0;62.0]	58.0 [54.0;61.0]	58.0 [54.0;62.0]
SF-36 Physical Component Summary score (0-100)[i]	48.0 [41.3;51.0]	50.0 [43.0;53.5]	49.0 [42.0;53.0]
EuroQol-5 dimensions (1 to -0.59)[j]	0.852 [0.803;0.855]	0.855 [0.821;1.000]	0.855 [0.805;1.000]
Treatment[k]			
csDMARD monotherapy <maximum dose, No. (%)	43 (48)	53 (60)	96 (54)
csDMARD monotherapy maximum dose, No. (%)	23 (26)	16 (18)	39 (22)
csDMARD combination therapy <maximum dose, No. (%)	24 (27)	19 (22)	43 (24)
NSAID stable dose, No. (%)	8 (9)	6 (7)	14 (8)
Prednisolone stable dose, No. (%)	2 (2)	2 (2)	4 (2)

Abbreviations: ACR, American College of Rheumatology; CDAI, Clinical Disease Activity Index; CRP, C-reactive protein; csDMARD, conventional synthetic disease modifying anti rheumatic drug; DAS28-CRP, disease activity score in 28 joints, CRP-based; EULAR, European League Against Rheumatism; IgM-RF, immunoglobulin M rheumatoid factor; IQR, interquartile range; MRI, Magnetic Resonance

Imaging; RAMRIS, RA magnetic resonance imaging scoring system; SDAI, Simplified Disease Activity Index; SF-36, Short Form 36 item questionnaire; T2T, treat-to-target; VAS, visual analogue scale; vdHSS, van der Heijde modified Sharp score.

[a] Values are median (IQR), unless stated otherwise.
[b] Composite disease activity score (DAS28), based on CRP (See eTable 3).
[c] Boolean based definition based on a 28 joint count of swollen and tender joints, CRP and patient global assessment of disease activity (See eTable 3).
[d] Composite score based on a 28 joint count of swollen and tender joints, patient and investigator global assessment of disease activity and CRP (range 0-86) (See eTable 3).
[e] Composite score based on a 28 joint count of swollen and tender joints, patient and investigator global assessment of disease activity (range 0-76) (See eTable 3).
[f] Sum score of MRI synovitis, osteitis and tenosynovitis.
[g] Sum score of MRI erosion and joint space narrowing.
[h] Patient-related score of physical function (disability), from 0 (best) to 3 (worst).
[i] Patient-related score of physical and mental function based on eight health concepts, with a high score defining a favorable health state.
[j] Patient-related quality of life from 1 (best) to -0.59 (worst).
[k] Treatment algorithm (See Methods section)

eTable 7.

Comparison of primary, secondary and exploratory outcomes at 24 months in patients enrolled after the correct final registration at www.clinicaltrials.gov[a]

	MRI T2T	Conventional T2T	Difference between groups	P Value[b]
	(n=88)	(n=87)	(95% CI)	
Primary endpoints				
Clinical:				
DAS28-CRP remission (DAS28-CRP<2.6), No. (%)	61 (87.1%)	72 (87.8%)	OR, 1.29 (0.37 to 4.51)	.690
Radiographic:				
No radiographic progression, No. (%)	45 (65.2%)	50 (61.7%)	OR, 1.24 (0.03 to 45.72) [c]	.907
Secondary and exploratory endpoints				
Clinical				
ACR-EULAR boolean remission, No. (%)	35 (50.0%)	24 (29.3%)	OR, 5.70 (1.59 to 20.45)	.008
SDAI remission (SDAI≤3.3), No. (%)	46 (65.7%)	46 (59.0%)	OR, 2.05 (0.71 to 5.89)	.183
CDAI remission (CDAI≤2.8), No. (%)	51 (71.8%)	49 (62.0%)	OR, 3.33 (1.05 to 10.58)	.042
DAS28-CRP	1.9 (0.1)	2.1 (0.1)	-0.2 (-0.4 to -0.0)	.033
Tender joint count (0-28)	0.2 (0.1)	0.5 (0.1)	-0.3 (-0.7 to 0.1)	.121
Swollen joint count (0-28)	0.0 (0.1)	0.3 (0.1)	-0.3 (-0.6 to -0.0)	.028
Patient ´s global assessment of disease activity, VAS (0-100)	15.0 (1.8)	21.6 (1.7)	-6.6 (-11.6 to -1.7)	.009
Patient´s assessment of pain VAS (0-100)	14.0 (1.8)	18.8 (1.7)	-4.9 (-9.7 to -0.0)	.049
Patient´s assessment of fatigue (0-100)	21.6 (2.0)	25.9 (1.9)	-4.3 (-9.7 to 1.1)	.116
Investigator´s global assessment of disease activity, VAS (0-100)	4.4 (0.9)	7.6 (0.8)	-3.2 (-5.6 to -0.9)	.007
Radiographic (vdHSS)				
Change in total vdHSS	1.1 (0.4)	1.3 (0.3)	-0.2 (-1.2 to 0.7)	.661
Change in erosion	0.6 (0.2)	0.8 (0.2)	-0.2 (-0.7 to 0.3)	.428
Change in joint space narrowing	0.5 (0.2)	0.5 (0.2)	-0.0 (-0.6 to 0.6)	.936
MRI (RAMRIS)				
Change in osteitis	-1.7 (0.6)	0.0 (0.5)	-1.7 (-3.3 to -0.2)	.031
Change in synovitis	-0.6 (0.4)	0.4 (0.3)	-0.9 (-1.9 to 0.0)	.058
Change in tenosynovitis	-1.0 (0.4)	0.4 (0.3)	-1.4 (-2.4 to -0.4)	.006
Change in combined inflammation	-3.0 (1.1)	0.9 (1.1)	-3.9 (-6.9 to -0.9)	.012
Change in erosion	0.5 (0.2)	0.5 (0.1)	-0.0 (-0.4 to 0.4)	.948
Change in joint space narrowing	0.1 (0.1)	0.2 (0.1)	-0.1 (-0.2 to 0.1)	.342
Change in combined damage	0.6 (0.2)	0.7 (0.2)	-0.1 (-0.6 to 0.4)	.740
No progression in MRI erosion, No. (%)	55 (79.7%)	62 (75.6%)	OR, 1.04 (0.01 to 86.20)	.987
Function and quality of life				
Change in Health Assessment Questionnaire	-0.064 (0.026)	0.098 (0.025)	-0.162 (-0.232 to -0.092)	<.001
Patient with normal function (HAQ≤0.5), No. (%)	59 (83.1%)	65 (79.3%)	OR, 3.04 (0.22 to 42.13)	.406

Change in SF-36 Mental Component Summary	0.1 (1.0)	-0.8 (1.0)	0.9 (-1.9 to 3.7)	.535
eTable 7 (continued)				
	MRI T2T	**Conventional T2T**	**Difference between groups**	**P Value[b]**
	(n=88)	**(n=87)**	**(95% CI)**	
Change in SF-36 Physical Component Summary	1.6 (1.1)	0.1 (1.0)	1.5 (-1.3 to 4.4)	.295
Change in EuroQol-5 dimensions	0.049 (0.015)	0.013 (0.014)	0.036 (-0.006 to 0.077)	.089

Abbreviations: 95% CI, 95% confidence interval; ACR, American College of Rheumatology; CDAI, Clinical Disease Activity Index; CRP, C-reactive protein; DAS28-CRP, disease activity score in 28 joints, CRP-based; EULAR, European League Against Rheumatism; MRI, Magnetic Resonance Imaging; RAMRIS, RA magnetic resonance imaging scoring system; SDAI, Simplified Disease Activity Index; SF-36, Short Form 36 item questionnaire; T2T, treat-to-target; VAS, visual analogue scale; vdHSS, van der Heijde modified Sharp score.

[a] Data are presented as least square means (SE) unless stated otherwise. For some variables, fewer patients were included in the analysis due to missing data. At 24 months data were available for minimum 64 patients in the MRI T2T arm and 75 patients in the conventional T2T arm.

[b] Analyses are based on repeated measures linear mixed models (continuous outcomes) or repeated measures generalized linear mixed models (binary outcomes), including participants as a random effect, with fixed factors for treatment group (2 levels) and time point (6 levels for disease activity measures and 2 levels for the x-ray imaging outcomes), and the corresponding interaction (time×group), adjusted for the value at baseline.

[c] Baseline was not included in the model, since the model failed to converge when the level at baseline was included.

eTable 8.

Comparison of primary, secondary and exploratory outcomes outcomes at 24 months in the per protocol population[a]

	MRI T2T	Conventional T2T	Difference between groups	P Value[b]
	(n=73)	(n=92)	(95% CI)	
Primary endpoints				
Clinical:				
DAS28-CRP remission (DAS28-CRP<2.6), No. (%)	62 (84.9%)	81 (88.0%)	OR, 1.00 (0.30 to 3.35)	.996
Radiographic:				
No radiographic progression, No. (%)	48 (65.8%)	58 (63.0%)	OR, 1.26 (0.03 to 46.45)	.901
Secondary and exploratory endpoints				
Clinical				
ACR-EULAR boolean remission, No. (%)	37 (50.7%)	28 (30.4%)	OR, 5.03 (1.46 to 17.38	.011
SDAI remission (SDAI≤3.3), No. (%)	48 (65.8%)	56 (62.2%)	OR, 2.10 (0.72 to 6.11)	.174
CDAI remission (CDAI≤2.8), No. (%)	52 (71.2%)	59 (64.8%)	OR, 3.31 (1.04 to 10.53)	.043
DAS28-CRP	1.9 (0.1)	2.1 (0.1)	-0.2 (-0.3 to 0.0)	.087
Tender joint count (0-28)	0.2 (0.1)	0.5 (0.1)	-0.2 (-0.6 to 0.1)	.177
Swollen joint count (0-28)	0.0 (0.1)	0.3 (0.1)	-0.3 (-0.6 to -0.0)	.031
Patient ´s global assessment of disease activity, VAS (0-100)	15.5 (1.9)	21.3 (1.7)	-5.8 (-10.8 to -0.8)	.023
Patient´s assessment of pain VAS (0-100)	14.0 (1.8)	18.6 (1.6)	-4.6 (-9.2 to 0.1)	.055
Patient´s assessment of fatigue (0-100)	21.7 (2.0)	24.1 (1.8)	-2.4 (-7.7 to 2.9)	.375
Investigator´s global assessment of disease activity, VAS (0-100)	4.3 (0.8)	6.9 (0.8)	-2.6 (-4.8 to -0.3)	.025
Radiographic (vdHSS)				
Change in total vdHSS	0.7 (0.2)	1.3 (0.2)	-0.6 (-1.3 to 0.1)	.070
Change in erosion	0.5 (0.2)	0.8 (0.2)	-0.3 (-0.8 to 0.1)	.166
Change in joint space narrowing	0.2 (0.1)	0.5 (0.1)	-0.3 (-0.6 to 0.0)	.055
MRI (RAMRIS)				
Change in osteitis	-1.8 (0.6)	-0.1 (0.5)	-1.8 (-3.3 to -0.3)	.021
Change in synovitis	-0.7 (0.4)	0.1 (0.3)	-0.8 (-1.8 to 0.2)	.114
Change in tenosynovitis	-1.0 (0.4)	0.1 (0.3)	-1.2 (-2.1 to -0.2)	.021
Change in combined inflammation	-3.8 (1.1)	0.0 (1.0)	-3.8 (-6.8 to -0.9)	.011
Change in erosion	0.5 (0.1)	0.5 (0.1)	-0.0 (-0.4 to 0.4)	.956
Change in joint space narrowing	0.1 (0.1)	0.2 (0.1)	-0.1 (-0.3 to 0.1)	.282
Change in combined damage	0.6 (0.2)	0.7 (0.2)	-0.1 (-0.6 to 0.4)	.703
No progression in MRI erosion, No. (%)	59 (80.8%)	68 (74.7%)	OR, 2.07 (0.03 to 144.50)	.736
Function and quality of life				
Change in Health Assessment Questionnaire	-0.038 (0.026)	0.092 (0.024)	-0.130 (-0.199 to -0.060)	<.001
Patient with normal function (HAQ≤0.5), No. (%)	60 (82.2%)	72 (79.1%)	OR, 3.40 (0.15 to 76.59)	.442

Change in SF-36 Mental Component Summary	-1.9 (1.0)	-3.0 (0.9)	1.1 (-1.4 to 3.7)	.390
eTable 8 (continued)				
	MRI T2T	**Conventional T2T**	**Difference between groups**	**P Value[b]**
	(n=73)	**(n=92)**	**(95% CI)**	
Change in SF-36 Physical Component Summary	0.7 (0.9)	-2.7 (0.8)	3.4 (1.1 to 5.7)	,004
Change in EuroQol-5 dimensions	0.012 (0.014)	-0.024 (0.012)	0.037 (-0.000 to 0.074)	.051

Abbreviations: 95% CI, 95% confidence interval; ACR, American College of Rheumatology; CDAI, Clinical Disease Activity Index; CRP, C-reactive protein; DAS28-CRP, disease activity score in 28 joints, CRP-based; EULAR, European League Against Rheumatism; MRI, Magnetic Resonance Imaging; RAMRIS, RA magnetic resonance imaging scoring system; SDAI, Simplified Disease Activity Index; SF-36, Short Form 36 item questionnaire; T2T, treat-to-target; VAS, visual analogue scale; vdHSS, van der Heijde modified Sharp score.

[a] Data are presented as least square means (SE) unless stated otherwise. For some variables, fewer patients were included in the analysis due to missing data. At 24 months data were available for minimum 64 patients in the MRI T2T arm and 77 patients in the conventional T2T arm.

[b] Analyses are based on repeated measures linear mixed models (continuous outcomes) or repeated measures generalized linear mixed models (binary outcomes), including participants as a random effect, with fixed factors for treatment group (2 levels) and time point (6 levels for disease activity measures and 2 levels for the x-ray imaging outcomes), and the corresponding interaction (time×group), adjusted for the value at baseline.

Table 9.

Sensitivity analysis: Comparison of primary outcomes at 24 months based on the full analysis set applying descriptive statistics for each group and contrast between groups based on Odds Ratio applying "worst case", "best case imputation" and multiple-imputation (MI) analyses, using the "Markov chain Monte Carlo" (MCMC) simulation statement.

	MRI T2T	Conventional T2T	Difference between groups	P Value[a]
	(n=100)	(n=100)	(95% CI)	
Best case imputation				
Clinical:				
DAS28-CRP remission (DAS28-CRP<2.6), No. (%)	89 (89.0%)	89 (89.0%)	OR, 1.41 (0.45 to 4.46)	0.554
Radiographic:				
No radiographic progression, No. (%)	75 (75.0%)	65 (65.0%)	OR, 2.06 (0.13 to 32.53)	0.608
Worst case imputation				
Clinical:				
DAS28-CRP remission (DAS28-CRP<2.6), No. (%)	64 (64.0%)	83 (83.0%)	OR, 0.25 (0.09 to 0.73)	0.011
Radiographic:				
No radiographic progression, No. (%)	49 (49.0%)	58 (58.0%)	OR, 0.74 (0.05 to 11.47)	0.828
Multiple imputation				
Clinical:				
DAS28-CRP remission (DAS28-CRP<2.6), No. (%)	89 (89.0%)	89 (89.0%)	OR, 1.04 (0.31 to 3.51)	0.670
Radiographic:				
No radiographic progression, No. (%)	75 (75.0%)	65 (65.0%)	OR, 1.32 (0.05 to 37.21)	0.670

[a]Analyses are based on repeated-measures logistic regression models.

When using Pearson´s chi-square approximation with a 1-sided significance level of .05 "Best case" imputation at 24 months was .000, P=1.000 for achieving the co-primary end point: DAS28-CRP remission (DAS28-CRP<2.6) and 2.381, P=.123 for achieving the co-primary end point: no radiographic progression (change in total van der Heijde modified Sharp SHS (vdHSS) score≤0. "Worst case" imputation was 9.267, P=.002 for achieving the co-primary end point: DAS28-CRP remission (DAS28-CRP<2.6), favoring the conventional T2T arm, and 1.628, P=.202 for achieving the co-primary end point: no radiographic progression (change in total van der Heijde modified Sharp SHS (vdHSS) score≤0

eTable 10.

Primary outcomes at 24 months stratified by center

	N	MRI Treat-to-Target (n=100)	N	Conventional Treat-to-Target (n=100)	Difference between groups, RD (95% CI) [a]	P value [b]
Primary endpoints						
DAS28-CRP remission (DAS28-CRP<2.6)						0.92
Gentofte Hospital, No. (%)	11	10 (91)	12	11 (92)	1.49 (-9.35 to 12.32)	
Glostrup Hospital, No. (%)	12	10 (83)	20	16 (80)	1.96 (-9.01 to 12.93)	
Slagelse Hospital, No. (%)	20	17 (85)	15	13 (87)	4.44 (-5.92 to 14.79)	
Odense, Hjørring, Silkeborg and Frederiksberg Hospitals [c], No. (%)	12	10 (83)	15	12 (80)	1.77 (-9.87 to 13.41)	
Graasten Hospital, No. (%)	12	10 (83)	21	20 (95)	-18.57 (-29.83 to -7.32)	
Aarhus Hospital, No. (%)	8	7 (88)	11	11 (100)	-1.71 (-16.34 to 12.92)	
No radiographic progression						0.88
Gentofte Hospital, No. (%)	12	7 (58)	11	4 (36)	20.42 (-6.21 to 47.04)	
Glostrup Hospital, No. (%)	12	9 (75)	20	12 (60)	9.30 (-12.47 to 31.08)	
Slagelse Hospital, No. (%)	20	13 (65)	15	9 (60)	-0.48 (-21.92 to 20.96)	
Odense, Hjørring, Silkeborg and Frederiksberg Hospitals [c], No. (%)	11	6 (55)	15	10 (67)	-12.06 (-36.46 to 12.34)	
Graasten Hospital, No. (%)	12	8 (67)	21	15 (71)	-10.39 (-31.85 to 11.06)	
Aarhus Hospital, No. (%)	7	6 (86)	11	8 (73)	-2.74 (-26.89 to 21.41)	

Abbreviations: 95% CI, 95% confidence interval; CRP, C-reactive protein; DAS28-CRP, disease activity score in 28 joints CRP-based; MRI, Magnetic Resonance Imaging; RD, risk difference.

[a] To assess risk differences between groups, data for all 200 patients at 24 months were included by applying multiple imputation.

[b] P-value for interaction between group and center.

[c] Odense, Hjørring, Silkeborg and Frederiksberg Hospitals were merged in order to ensure at least five patients in each group.

eTable 11.

Safety data (collected adverse events; full analysis set)[a]

	MRI treat-to-target (n=100)	**Conventional treat-to-target** (n=100)	Total
Serious Adverse Events (SAE)	19	7	26
Patients with serious infection	3[b]	3[c]	6
Cancer	3[d]	1[e]	4
Death	1[f]	1[g]	2
Other	12[h]	2[i]	14
Discontinuation from the study due to SAE	6[j]	1[k]	7
Discontinuation from the study due to AE	2[l]	0	2

Values are numbers

[a] Seventeen patients (17%) in the MRI-treat-to-target-group and 6 patients (6%) in the conv-treat-to-target-group experienced SAEs

[b] Two patients had pneumonia and one patient had pneumonia complicated by empyema

[c] Erysipelas, pneumonia, infection without focus

[d] Esophageal cancer, squamous cell carcinoma, transitional cell carcinoma

[e] Squamous cell carcinoma with a component of mucoepidermoid carcinoma

[f] Fall accident and subsequent immobilisation leading to pneumonia sepsis and dead.

[g] Died from squamous cell carcinoma with a component of mucoepidermoid carcinoma

[h] Pneumonitis, appendicitis, allergic reaction (rash), allergic reaction (rash), spleen rupture, lower back pain, hospitalized under the suspicion of deep venous thrombosis, pleural effusion, liver enzyme elevation, new diagnosed diabetes type 2, atrial fibrillation, gastroenteritis

[i] Hospitalization due to abdominal pain, humerus fracture

[j] Pneumonitis due to methotrexate, rash after sulfasalazine start, esophageal cancer, squamous cell carcinoma, pneumonia complicated by empyema, pneumonia

[k] Squamous cell carcinoma with a component of mucoepidermoid carcinoma

[l] Persistent elevated alanine aminotransferase, recurrent upper respiratory tract infections

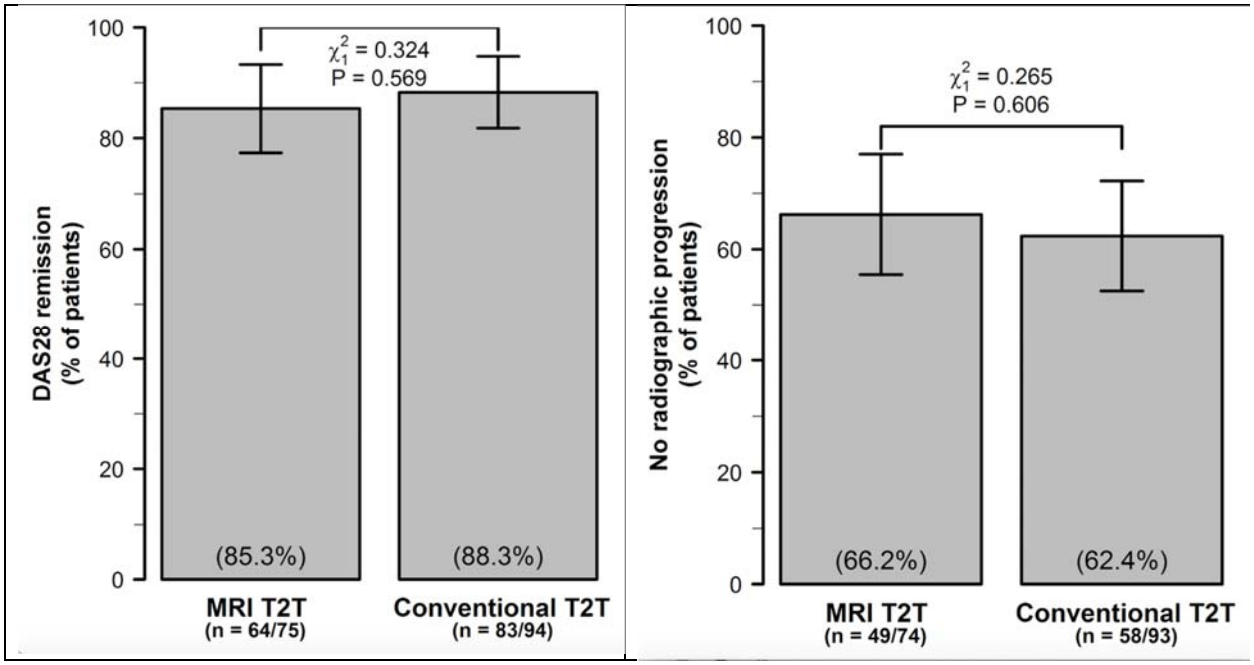
eTable 12.

Summary narratives for malignancies and deaths

Summary narratives for malignancies (n=4)
1) Adenocarcinoma
A 64-year-old male in the MRI T2T arm was diagnosed with adenocarcinoma of the esophagus approximately14 months after study start. Medical history included no co-morbidities. The patient was receiving methotrexate 25mg weekly at the time of the event and concomitant medication included folic acid, cholecalciferol and calcium. The patient was withdrawn from the study shortly after the occurrence of the event. He was treated with curative surgery and chemotherapy. Approximately 5 months after the occurrence of the event, the patient was reported recovered with no sequelae. The investigator considered the event not to be related to the study drug.
2) Squamous cell carcinoma
A 77-year-old male in the MRI T2T arm was diagnosed with squamous cell carcinoma of the lip approximately 9 months after study start. Medical history included duodenal ulcer, syncope and unspecified chest pain. The patient was receiving methotrexate 25mg weekly, sulfasalazine 2000mg daily and hydroxychloroquine 400mg daily at the time of the incident and concomitant medication included folic acid, cholecalciferol and calcium. The patient was withdrawn from the study shortly after the occurrence of the event. The affected area had resolved without intervention at time of planned surgery and the patient was followed up after 5 months and was reported recovered with no sequelae. The investigator considered the event not to be related to the study drug.
3) Transitional cell carcinoma
A 69-year-old male in the MRI T2T arm was diagnosed with transitional cell carcinoma (urothelial carcinoma) in the bladder approximately 1 year and 7 months after study start. Medical history included bilateral surgery of the meniscus, knee arthrosis and a conservatively treated prolapsed intervertebral disc. The patient was receiving methotrexate 25mg weekly, sulfasalazine 1000mg daily and hydroxychloroquine 200mg daily at the time of the incident and concomitant medication included folic acid, acetylsalicylic acid, and a statin. The patient was withdrawn shortly after the occurrence of the event. The patient received chemotherapy and the cancer was found to be inoperable. The patient did not recover during the follow up period. The investigator considered the event not to be related to the study drug.
4) Squamous cell carcinoma with a component of mucoepidermoid carcinoma
A 65–year-old female in the conventional T2T arm was diagnosed with squamous cell carcinoma with a component of mucoepidermoid carcinoma of the alveolar process approximately 10 months after study start. Medical history included no co-morbidities. The patient was receiving methotrexate 12,5mg weekly at the time of the event and concomitant medication included folic acid, cholecalciferol and calcium. The patient underwent surgery and intended curative chemotherapy and radiotherapy. In the follow up period approximately 14 months after the event the patient had not recovered and died from the incident.
Summary narratives for deaths (n=2)
5) Sepsis
A 77-year-old female in the MRI T2T arm was excluded from the study approximately 8 months after study start because she did not want to follow treatment decision rules. Approximately 6 months after exclusion the patient had a fall accident was hospitalized and had pneumonia, subsequent sepsis and died. Medical history included hypertension, nephrectomy (benign tumor), tumor of the breast (benign), bilateral elbow fracture, neurolysis median and ulnar nerves, excision and synovectomy of the right ulna and instability in the cervical spine. The patient was receiving methotrexate 25mg weekly, sulfasalazine 1000mg daily and hydroxychloroquine 200mg daily at the time of the incident and concomitant medication included folic acid, verapamil, enalapril hydrochlorothiazide. The investigator considered the event not to be related to the study drug.
6) Squamous cell carcinoma with a component of mucoepidermoid carcinoma
See ”4)”

eFigure.

Proportion of patients achieving co-primary outcomes at 24 months, based on the intention-to-treat population with no data imputation to replace missing data and assessed using Pearson's chi-square approximation with a 1-sided significance level of 0.05



The bar graph illustrates the proportion of completers in the MRI-guided treat-to-target (MRI-T2T) arm and in the conventional treat-to-target (conventional-T2T) arm who achieved co-primary endpoints: DAS28-CRP remission (DAS28-CRP<2.6) and no radiographic progression (change in total van der Heijde modified Sharp SHS (vdHSS) score≤0) at 24 months, based on the intention-to-treat population with no data imputation to replace missing data and assessed using Pearson's chi-square approximation with a 1-sided significant level of 0.05. Error bars indicate 95% confidence limits. Of the 76 and 95 completers in the MRI-T2T-arm and conventional-T2T-arm, 75 patients in the MRI-T2T-arm and 94 patients in the conventional-T2T-arm were included in the DAS28 remission analysis, and 74 and 93 patients in the radiographic progression analysis. This was due to missing data.

eAppendix.

IMAGINE-RA:

Sensitivity Analyses for the Intention To Treat (ITT) Population With Missing Outcome Data (Data Missing Not At Random & Tipping Point Analysis)

Missing outcome data threaten the validity of many clinical trials. However, loss to follow-up is often hard to avoid, thus a framework for "intention to treat analysis" that depends on making plausible assumptions about the missing data and including all participants in sensitivity analyses need to be applied.

All the analyses presented for the IMAGINE-RA trial was based on the ITT principle:

"*The ITT principle asserts that the effect of a treatment policy can be best assessed by evaluating on the basis of the intention to treat a subject (that is, the planned treatment regimen) rather than the actual treatment given. It has the consequence that subjects allocated to a treatment group should be followed up, assessed and analysed as members of that group irrespective of their compliance to the planned course of treatment.*" (ICH E9)

In the prespecified Statistical Analysis Plan (SAP) for the IMAGINE-RA we applied the ITT framework suggested by Ian R. White and colleagues (BMJ 2011;342:d40) on how to perform an ITT analysis with incomplete observations:

1) We attempted to follow up all randomized participants, even if they withdrew from allocated treatment (incl. poor adherence to protocol)
2) We performed the main analysis of all observed data that are valid under a plausible assumption about the missing data
3) We performed sensitivity analyses to explore the effect of departures from the assumption made in the main analysis
4) We accounted for all randomized participants, at least in the sensitivity analyses.

The four points above clearly illustrate that even when using the "ITT population", the potential bias due to missing data depends on the mechanism causing the data to be missing, and thus the analytical methods applied to amend the missingness require careful planning and attention.

The aim of this supplementary material is to explicitly report how the sensitivity analysis were handled in the IMAGINE-RA trial analyses, incl. the changes that were introduced following the constructive and thorough peer review.

The IMAGINE-RA investigators define missing data as "values that are not available for analysis despite various efforts to get in contact with the participants." Following up participants with chronic diseases who withdraw from randomized treatment (in accordance with ethics; i.e. patients can withdraw consent at any time) is difficult but is important to pursue. A trial that does not attempt to follow participants after treatment withdrawal cannot claim to follow the intention to treat principle (White, et al, BMJ 2011;342:d40).

There is no analytic approach that can assuredly produce unbiased estimates of treatment effects when relevant data are missing and therefore the IMAGINE-RA investigators placed increased emphasis on strategies to explore how robust various statistical inference statements are.

The observed data collected during the 2-year IMAGINE-RA Trial is presented in **Supplementary Table 1** (below).

Supplementary Table 1. Data collected for the two co-primary outcomes (clinical remission and no radiographic progression, respectively)

	Response: "Yes"	**Response: "No"**	**Response: "Don't know"**	**Total**
Clinical Remission:				
MRI Treat-to-Target group	64	11	25	100
Conventional Treat-to-Target group	83	11	6	100
No Radiographic Progression:				
MRI Treat-to-Target group	49	25	26	100
Conventional Treat-to-Target group	58	35	7	100

The data analysis for comparison of groups that is presented in Table 2 in the manuscript (adjusted Risk Difference) is based on Multiple Imputation. Multiple Imputation is considered a valid approach to missing data in an ITT population if the missingness only depends on the observed data; then missing data are 'Missing At Random' (MAR) given the observed data. Another option that is considered equally useful (i.e. valid) when investigators have collected repeated measures, would be to use repeated measures maximum likelihood models (i.e. Generalized Linear Mixed Models); analyses based on this approach are available in eTable 5.

As a consequence, both of the analyses for co-primary outcomes presented in the manuscript Table 2 and eTable 5, respectively, are "valid" in the case that missing data can be considered under the "MAR assumption".

The MAR assumption, however, may not always be clinically plausible (Sterne et al, BMJ 2009;338:b2393). If the mechanism depends on the missing data, and this dependency remains even given the observed data, then data are classified as 'Missing Not At Random' (MNAR). Consequently, the real issue among experienced statisticians is that the MAR and MNAR conditions cannot be distinguished based on the observed data because - by definition - the missing data are unknown, and it can therefore not be assessed if the observed data can predict the unknown data.

Sensitivity and tipping point analyses – assuming data was MNAR

Sensitivity analyses, was performed to assess the robustness of the primary analyses (Table 2 in the Manuscript, as well as eTable 5) included repeated-measures and multiple-imputation analyses, which used a model-based approach for missing data as explained above.

The following sensitivity analyses were performed to assess the potential impact that MNAR may have on the estimated results. In the following we use single imputation, which often result in an underestimation of the variability because each unobserved value carries the same weight in the analysis as the known, observed values. If large proportions of data are missing it ought to be considered just to report the results of the complete case analysis and then discuss the resulting interpretative limitations of the trial results. If the proportions of missing data are large (for example, more than 20%, as was the case for the MRI Treat-to-Target group in IMAGINE-RA [**Supplementary Table 1,** above]) on a co-primary outcome, then trial results may only be considered as hypothesis generating results. A rare exception would be if the underlying mechanism behind the missing data can be described as 'Missing Completely At Random' MCAR.

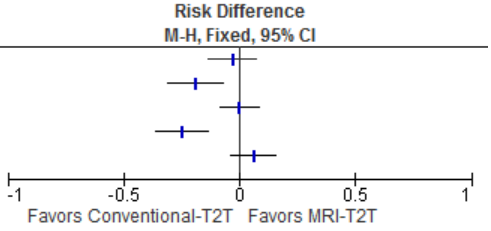
Best- and Worst-Case Imputations

Best- and Worst-Case Imputation, as well as Best-Worst, and Worst-Best Case Imputation sensitivity analyses may show the full theoretical range of uncertainty and conclusions ought to be related to this range of uncertainty. As presented in the **Supplementary Figure 1** (below), sensitivity analyses may show how assumptions, different from those made in the primary analysis, influence the results obtained (Little et al, NEnglJMed. 2012;367[14]: 1355–60). Inferential statistics ought to be predefined as described in the SAP, but additional sensitivity analyses might be warranted and valid.

Supplementary Figure 1. Best- and Worst-Case Imputation, as well as Best-Worst, and Worst-Best Case Imputation sensitivity analyses (created using Review Manager)

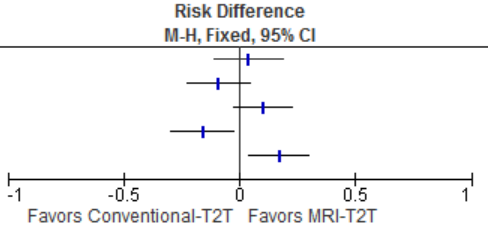
A: Clinical Remission

Study or Subgroup	MRI-T2T Events	MRI-T2T Total	Conventional-T2T Events	Conventional-T2T Total	Risk Difference M-H, Fixed, 95% CI
A Data 'As Observed'	64	75	83	94	-0.03 [-0.13, 0.07]
B 'Worst Case Imputation'	64	100	83	100	-0.19 [-0.31, -0.07]
C 'Best Case Imputation'	89	100	89	100	0.00 [-0.09, 0.09]
D Worst-Best Case	64	100	89	100	-0.25 [-0.36, -0.14]
E Best-Worst Case	89	100	83	100	0.06 [-0.04, 0.16]



B: No Radiographic Progression

Study or Subgroup	MRI-T2T Events	MRI-T2T Total	Conventional-T2T Events	Conventional-T2T Total	Risk Difference M-H, Fixed, 95% CI
A Data 'As Observed'	49	74	58	93	0.04 [-0.11, 0.18]
B 'Worst Case Imputation'	49	100	58	100	-0.09 [-0.23, 0.05]
C 'Best Case Imputation'	75	100	65	100	0.10 [-0.03, 0.23]
D Worst-Best Case	49	100	65	100	-0.16 [-0.30, -0.02]
E Best-Worst Case	75	100	58	100	0.17 [0.04, 0.30]



Abbreviations: T2T, Treat-to-Target

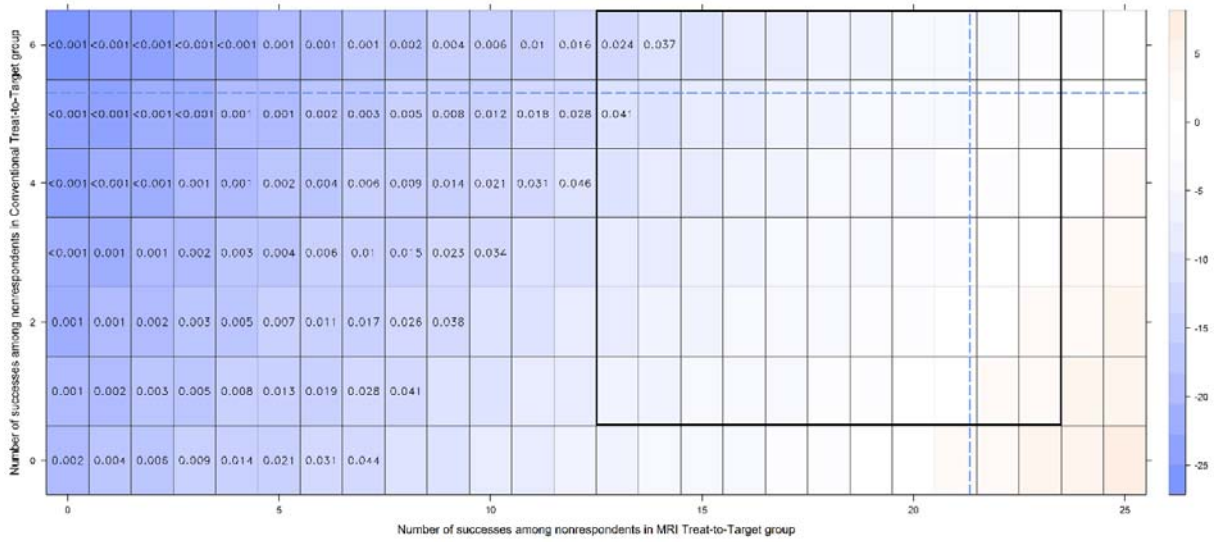
In **Supplementary Figure 1** we present five sensitivity analyses for each of the co-primary outcomes: The **A** plot being clinical remission, whereas the **B** plot is the no radiographic progression. The sensitivity analyses presented explore the possible impact of MNAR, which reveals that the primary data may be influenced by attrition bias.

Tipping point analyses

We also apply a more advanced, but graphical, display that reveal the effects of all possible combinations of the values of missing data in the first arm (MRI Treat-to-Target group) and the second arm (Conventional Treat-to-Target group) of a two-arm study on various quantities of interest, typically, on p-values and point estimates. As explained by Liublinska and Rubin (Statistics in Medicine, 2014), the displays are based on the idea of 'tipping-point' (TP) analysis, first introduced in 2009, as a method of assessing the impact of missing data on a study's conclusions about some quantity of interest. A tipping point is defined as a particular combination of missing data values that would change the study's conclusions, as summarized by its p-value. We illustrate this by a detailed display in conjunction with multiple imputation (MI) of missing data.

We anticipate that the displays (**Supplementary Figure 2** and **Supplementary Figure 3**) enable practitioners to identify whether alternative assumptions about the missingness mechanism would change the study's conclusions and thereby allow them to assess the strength of the study's evidence. The ideal scenario is when the display shows no tipping points, that is, when all combinations of missing values lead to the same conclusion of the study.

Supplementary Figure 2. Tipping-point display for clinical remission at 24 months.

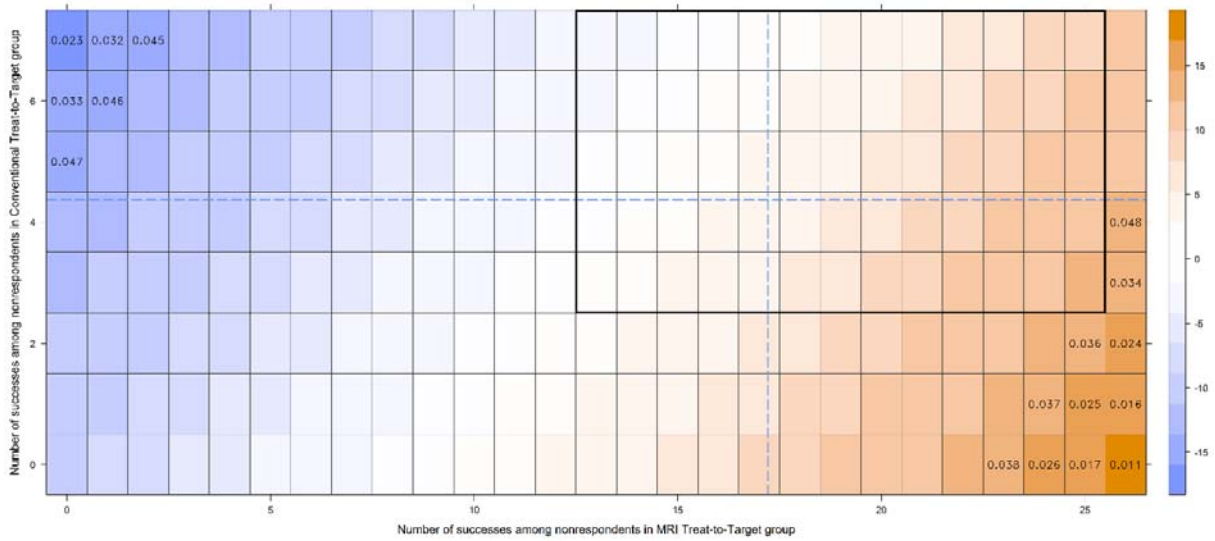


The tipping-point analysis display for clinical remission at 24 months, show the estimated risk differences (%) using a heat map. Axes represents the number of successes (clinical remissions) that could be observed among non-respondents in the MRI Treat-to-Target group and in the Conventional Treat-to-Target group. Each combination corresponds to a value of the estimated risk differences. Its magnitude and direction are represented using a color palette that changes from dark blue (large negative values) to dark orange (large positive values), with white representing zero (or close to zero) risk difference. The individual p-values are shown when these are <0.050, highlighting combinations that result in rejecting the null

hypothesis that the risk difference is 0, with the border between empty cells and non-empty cells indicating the tipping points of the study. Vertical and horizontal dashed lines in blue corresponds to observed success (clinical remission) rates in the MRI Treat-to-Target group and Conventional Treat-to-Target group, 85.3% and 88.3%. The square connects minimum and maximum number of successes among 100 imputations for non-respondents in the MRI Treat-to-Target group and the Conventional Treat-to-Target group.

Interpretation of the **Supplementary Figure 2**, is that the missing data pattern observed possibly could be covering the fact that the Conventional Treat-to-Target approach is more likely to achieve clinical remission than the MRI Treat-to-Target (as indicated by three combinations resulting in a statistical p-values was obtained in the multiple imputation datasets, and that no tipping point exists for MRI Treat-to-Target being statistically significant better than Conventional Treat-to-Target). Thus, our conclusion from **Supplementary Figure 2** is that it is highly unlikely that MRI Treat-to-Target is superior to Conventional Treat-to-Target in terms of clinical remission.

Supplementary Figure 3. Tipping-point display for no radiographic progression at 24 months.



Tipping-point display for no radiographic progression at 24 months, show the estimated risk differences (%). Its magnitude and direction are represented using a color palette that changes from dark blue (large negative values) to dark orange (large positive values), with white representing zero (or close to zero) risk difference. The individual p-values are shown when these are <0.050, highlighting combinations that result in rejecting the null hypothesis that the risk difference is 0, with the border between empty cells and non-empty cells indicating the tipping points of the study. Vertical and horizontal dashed lines in blue corresponds to observed success (no radiographic progression) rates in the MRI Treat-to-Target group and Conventional Treat-to-Target group, 66.2% and 62.4%. The square connects minimum and maximum number of successes among 100 imputations for non-respondents in the MRI Treat-to-Target group and the Conventional Treat-to-Target group.

Interpretation of the **Supplementary Figure 3**, is that the missing data pattern is unlikely to have an impact of the outcome of the trial in terms of radiographic remission (no statistically significant findings was obtained in any of the multiple imputation datasets). Thus, our conclusion from **Supplementary Figure 3** is that it is highly likely that MRI Treat-to-Target is similar to Conventional Treat-to-Target group, in terms of Radiographic remission.