CLINICAL SCIENCE

Testing different thresholds for patient global assessment in defining remission for rheumatoid arthritis: are the current ACR/EULAR Boolean criteria optimal?

Paul Studenic ,[1] David Felson ,[2,3] Maarten de Wit ,[4] Farideh Alasti,[1] Tanja A Stamm ,[5] Josef S Smolen,[1] Daniel Aletaha[1]

Handling editor Gerd R Burmester

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-216529).

[1]Department of Internal Medicine 3, Division of Rheumatology, Medical University of Vienna, Wien, Austria
[2]Section of Rheumatology, Boston University School of Medicine, Boston, Massachusetts, USA
[3]Centre for Epidemiology, The University of Manchester, Manchester, UK
[4]Standing Committee of People with Arthritis/Rheumatism in Europe (PARE), EULAR, Zurich, Switzerland
[5]Center for Medical Statistics, Informatics, and Intelligent Systems, Section for Outcomes Research, Medical University of Vienna, Wien, Austria

Correspondence to
Professor Daniel Aletaha, Department of Internal Medicine 3, Division of Rheumatology, Medical University of Vienna, Wien 1090, Austria; daniel.aletaha@meduniwien.ac.at

Received 23 October 2019
Revised 21 December 2019
Accepted 22 January 2020
Published Online First 5 February 2020

To cite: Studenic P, Felson D, de Wit M, *et al*. *Ann Rheum Dis* 2020;**79**:445–452.

ABSTRACT

Objectives This study aimed to evaluate different patient global assessment (PGA) cut-offs required in the American College of Rheumatology/European League Against Rheumatism (ACR/EULAR) Boolean remission definition for their utility in rheumatoid arthritis (RA).

Methods We used data from six randomised controlled trials in early and established RA. We increased the threshold for the 0–10 score for PGA gradually from 1 to 3 in steps of 0.5 (Boolean1.5 to Boolean3.0) and omitted PGA completely (BooleanX) at 6 and 12 months. Agreement with the index-based (Simplified Disease Activity Index (SDAI)) remission definition was analysed using kappa, recursive partitioning (classification and regression tree (CART)) and receiver operating characteristics. The impact of achieving each definition on functional and radiographic outcomes after 1 year was explored.

Results Data from 1680 patients with early RA and 920 patients with established RA were included. The proportion of patients achieving Boolean remission increased with higher thresholds for PGA from 12.4% to 19.7% in early and 5.9% to 12.3% in established RA at 6 months. Best agreement with SDAI remission occurred at PGA cut-offs of 1.5 and 2.0, while agreement decreased with higher PGA (CART: optimal agreement at PGA≤1.6 cm; sensitivity of PGA≤1.5 95%). Changing PGA thresholds at 6 months did not affect radiographic progression at 12 months (mean ΔsmTSS for Boolean, 1.5, 2.0, 2.5, 3.0, BooleanX: 0.35±5.4, 0.38±5.14, 0.41±5.1, 0.37±4.9, 0.34±4.9, 0.27±4.7). However, the proportion attaining HAQ≤0.5 was 90.2%, 87.9%, 85.2%, 81.1%, 80.7% and 73.1% for the respective Boolean definitions.

Conclusion Increasing the PGA cut-off to 1.5 cm would provide high consistency between Boolean with the index-based remission; the integer cut-off of 2.0 cm performed similarly.

Key messages

What is already known about this subject?

- The patient global assessment (PGA) is the most common reason in patients with rheumatoid arthritis (RA) for not reaching American College of Rheumatology/European League Against Rheumatism Boolean remission.
- The PGA has been criticised to not adequately reflect disease activity of RA.

What does this study add?

- A PGA cut-off of 2 (on a 0–10 scale) coincides with a better agreement between Boolean and Simplified Disease Activity Index remission.
- Patients in Boolean remission definition using ≤2 (on a 0–10 scale) on a PGA show good long-term functional and radiologic outcomes.

How might this impact on clinical practice or future developments?

- Using the adapted cut-off in clinical trials and practice will improve evaluation of remission in RA.

Disease activity in rheumatoid arthritis (RA) has been found best reflected in a number of so-called core set variables defined many years ago by the American College of Rheumatology (ACR) and the European League Against Rheumatism (EULAR).[1 2] Irrespective of the use of individual core set variables, composite measures of disease activity comprising several components have better validity than individual components based on the heterogeneity of the disease presentations between and within individual patients[3–5]; in addition, they correlate better with structural and functional outcomes in RA.[3 6]

When the core set variables were defined, remission was more an aspirational than a realistic goal.[7] Today, remission is achievable in a significant proportion of patients and has become a major therapeutic target.[8–10] A clinical definition of remission for RA should reflect no, or at most only minimal, disease activity in terms of inflammation, such as swollen joints or acute phase reactant (APR) levels to prevent structural progression and functional deterioration.[10] The remission definition of the Disease Activity Score using 28 joint counts (DAS28) allows for a significant number of residual swollen joints,[11–13] which cannot be overcome by lowering the cut-off for remission on its scale.[11 14] DAS28 also overweighs the acute phase response,[15] making results from drugs that target interleukin (IL)-6—and thus the APR directly—less comparable with those attained with other compounds.[16 17]

ACR and EULAR provided remission definitions almost one decade ago.[11] Despite the proven validity of the ACR/EULAR remission criteria, the definition of remission is still in discussion and alternative definitions are still frequently used in clinical trials and practice. To attain an ACR/EULAR Boolean remission, a patient must have, among other criteria, a patient global assessment (PGA) score ≤1 (0–10 scale), and this definition has been criticised because patients who have no active joints and a normal C reactive protein (CRP) often have PGA scores exceeding the cut-off of 1.[18] PGA has been incorporated into composite scores and remission definitions to include the patient's perspective in the assessment of disease activity, and it is also recommended for evaluation in clinical trials.[19] Further, the committee developing the remission definition showed that inclusion of PGA improved the discriminant ability of remission criteria to separate effective RA treatments from placebo, suggesting that it represents elements of disease activity missed by other outcome measures. In other words, inclusion of PGA in remission criteria makes it more likely that efficacy of different treatments can be discriminated. However, the PGA sometimes not only reflects symptoms based on inflammatory disease activity but also other factors such as depressive symptoms or functional limitations due to pre-existing joint damage or even comorbidities.[20 21]

The ACR/EULAR index-based Simplified Disease Activity Index (SDAI) remission criteria are slightly less stringent than Boolean remission, given that the sum of several components permits one of them to be slightly elevated (eg, a PGA above 1) if compensated by a lower score of others.[22] Both remission definitions are associated with optimal clinical, functional and structural outcomes[11] and are widely used in clinical trials, where a substantial number of patients today achieve this stringent outcome.[23] To this end, studies have shown that some patients meeting SDAI remission do not meet the more stringent Boolean definition of remission primarily due to the requirement for a PGA of ≤1.[18 24] Since both, the Boolean and the SDAI remission, are recommended by ACR and EULAR, they ideally should be consistent and identify the same patients.

We therefore aimed to determine whether an increase of the PGA threshold in the ACR/EULAR Boolean-based criteria might increase its agreement with the ACR/EULAR index-based remission by SDAI without jeopardising good clinical, functional and structural implications, associated with the state of remission.

METHODS

Patients

RA patient data were retrieved from six clinical trials testing the efficacy of tumour necrosis factor inhibitors (TNFi) versus placebo or placebo+methotrexate (MTX) with an observation period between 1 and 2 years (ASPIRE, ATTRACT, PREMIER, DE019, Go Before and Go Forward). The individual trials have been previously reported[25–30] and so has the use of pooled data of these trials obtained from the trial sponsors.[22 31 32] These trials included patients with RA with varying disease durations and treatment histories representing a large spectrum of the disease. ASPIRE (infliximab), Go Before (golimumab) and PREMIER (adalimumab) were trials in MTX-naïve patients with early RA (mean disease duration of the pooled population at baseline 1.5±3.0 years), while ATTRACT (infliximab), DE019 (adalimumab) and Go Forward (golimumab) were performed in MTX-insufficient responders with a mean disease duration of the pooled patients at baseline of 9.7±8.4 years. In all six clinical trials, the patients were asked to provide the assessment of the activity of their RA using a 100 mm visual analogue scale (VAS).[25–30]

Definitions of remission and their modifications

The Boolean definition includes swollen joint counts (SJC), tender joint counts (TJC), PGA (in cm) and CRP levels (in mg/dL) and for a patient to meet remission criteria, all of these must have scores of 1 or less. The SDAI index-based definition of remission sums the scores for the components used in the Boolean definition plus evaluator/physician global assessment, and patients meet this definition if the score is ≤3.3.[11]

We evaluated an expansion of the current Boolean definition of remission by increasing the cut-off of the PGA criterion stepwise (using a 0–10 cm VAS) by 0.5 cm increments from 1 cm to 1.5, 2.0, 2.5 and 3.0 cm. We will refer to them as Boolean1.5, Boolean2.0, Boolean2.5 and Boolean3.0, respectively. Additionally, we omitted the PGA criterion completely from the Boolean definition, labelling this definition as BooleanX; in this definition, only CRP, TJC and SJC need to score ≤1 to attain remission, independent of the PGA value.[33]

Analyses

We assessed agreement of modified Boolean remission rates at 6 and 12 months with the SDAI definition of remission using McNemar's test for agreement. We tested which PGA cut-off in the Boolean remission criteria yielded the best agreement with SDAI remission.[11]

As a next step, we explored the impact of using the modified Boolean remission definitions assessed at 6 months on outcomes at 1 year. Differences in mean radiographic progression (based on the change in modified total Sharp score (mTSS) between baseline and 1 year), number of patients without progression (change in score ≤0), mean functional scores (Health Assessment Questionnaire (HAQ) scores, physical component scores of the Short Form 36 (SF-36)) and patients with normal function (HAQ ≤0.5 at 1 year), were assessed. The distribution of 1-year outcomes was depicted in cumulative frequency plots, separately for patients attaining the various 'modified' remission definitions at 6 months. These analyses were then repeated separately for patients with early and late RA.

To obtain a more sensitive assessment of differences in structural and functional outcomes, we looked at these outcomes for the non-overlapping modified Boolean definition groups (ie, Boolean20 would not include Boolean15 or lower; and analogously for the other definitions). We compared differences in distribution of mTSS changes, HAQ and SF-36 physical component scores at 1 year between discrete modified Boolean definitions (ie, Boolean1.5 only those with PGA of 1.1–1.5, and so on) at the 6-month time point. We used data from patients with early RA only, since numbers of patients with established RA were too few for this analysis.

Furthermore, we conducted a classification and regression tree (CART) analysis to predict SDAI remission in early and established RA based on PGA at weeks 22 and 54 (R rpart package; https://cran.r-project.org/web/packages/rpart/index.html) to determine the PGA cut-off in patients fulfilling the other three Boolean criteria, which shows the highest likelihood of fulfilling the SDAI definition of remission. We then performed receiver operating curve analyses (ROC) to test sensitivity and specificity of all PGA cut-offs between 1 and 2 cm.

Patient and public involvement statement

The place and interpretation of the PGA in defining remission in RA from a patient perspective have repeatedly raised concerns of

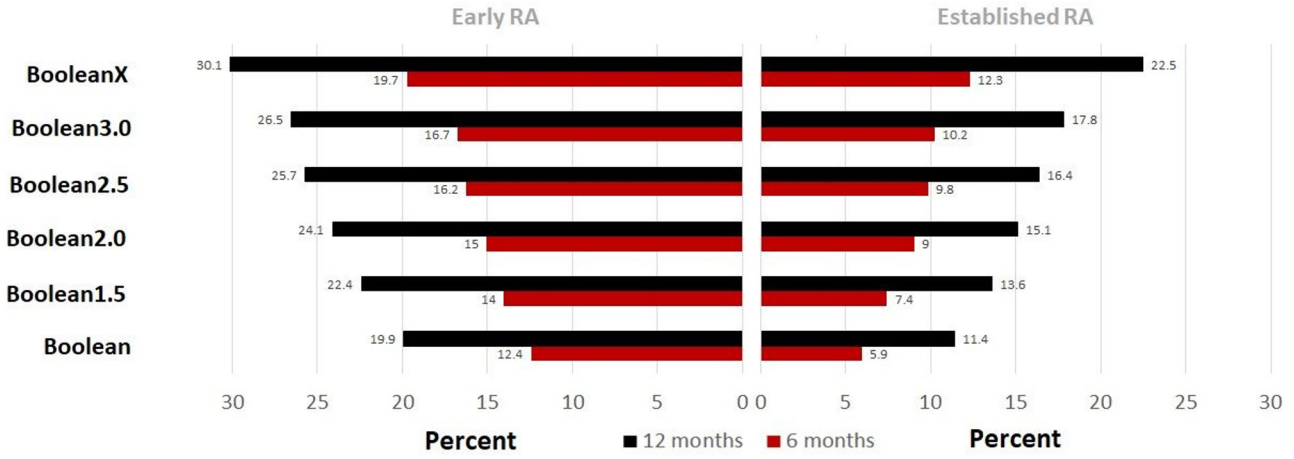

Figure 1 Rates of remission by modified Boolean classifications, using a patient global assessment (PGA) cut-off of 1.0 ('Boolean'), 1.5, 2.0, 2.5, 3.0 cm, or omitting the PGA completely (BooleanX). Rates in % of total, separately depicted at 6 (red bars) and 12 months (black bars) time points; on the left for those in early rheumatoid arthritis (RA) and on the right for established RA.

physicians and patients.[34] For this reason, we included an experienced patient research partner (PRP) in this study (MdW). The PRP was involved throughout the research process and provided critical feedback during all stages of analysis. Face-to-face meetings with the PRP took place in conjunction with EULAR meetings and the PRP will disseminate findings in relevant patient communities.

RESULTS

Data from 2600 trial patients, 1680 with early RA (mean disease duration: 1.5±3.0 years) and 920 with established RA (mean disease duration: 9.7±8.4 years) were included. As expected, the rates of patients achieving modified Boolean remission increased with an increase in the PGA cut-off from 12.4% (n=208) to 19.7% (n=331) in early RA and 5.9% (n=54) to 12.3% (n=113) in established RA at 6 months and 19.9% (n=335) to 30.1% (n=506) and 11.4% (n=105) to 22.5% (n=207), respectively, at 1 year (figure 1). For both early and late RA, the increase in remission rates was already pronounced when moving the PGA cut-off of from 1.0 to 2.0 cm (+44 patients (+21%) at 6 months) and less when moving the cut-off from 2.0 to 3.0 cm (+29 patients (+14%) at 6 months); however, omitting the PGA criterion completely (BooleanX definition) led to an even larger increase in remission rates compared with the Boolean3.0 category (+50 patients at 6 months; see also online supplementary table 1).

Concordance of modified Boolean remission with SDAI remission

When evaluating the best cut-off for concordance of SDAI and Boolean remission, we found that by increasing the PGA cut-off to 1.5 or 2.0 cm, higher concordance rates between the two definitions were achieved, leading to fewer patients who only fulfilled SDAI remission without fulfilling the respective Boolean remission. The percentage of Boolean remitters (within the SDAI remitters) increased from 74% to 85% when using the Boolean2.0 definition at 6 months, and from 79% to 89% at the 1-year visit. At the same time, however, there was a slight increase in patients fulfilling the Boolean criteria only within the SDAI non-remitter group (from 1.3% to 3.0% at 6 months and from 1.5% to 4.1% at 1 year). Overall, kappa values with SDAI remission were almost identical for the Boolean2.0 definition compared with the traditional Boolean definition (at 6 months: 0.80; 95% CI 0.76 to 0.83, vs 0.78; 0.74 to 0.81; at 1 year: 0.83; 0.80 to 0.86 vs 0.82; 0.80 to 0.85).

When exploring this separately for patients with early and established RA, we found that the concordance between Boolean and SDAI definitions (by means of kappa) was lower in patients with established RA in particular at 6 months, with similar values to early RA at 1 year (figure 2). Regardless of population (early vs late) or time point during the trial (6 months vs 12 months), agreements between the two remission definitions were better when using the Boolean1.5 and 2.0 definition (as seen in the overall data). A further increase in the PGA cut-off beyond 2 cm led to a decrease in concordance; this drop in congruence was very clear when omitting the PGA (lower kappa values than for the traditional Boolean remission). In summary, the increase of the cut point from 1.0 to 2.0 increased the number of patients in remission with a similar overall agreement with the SDAI definition.

Additionally, using CART analyses revealed in patients with SJC, TJC and CRP all <1, that depending on the population (early vs established) or time point of analysis (6 months vs 12 months), the PGA cut-off with the highest likelihood of concurrent SDAI remission ranged between ≤1.1 and ≤1.6 cm (table 1,

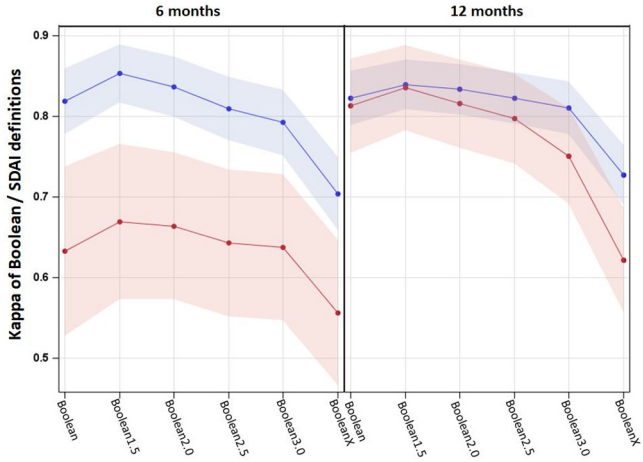

Figure 2 Kappa with CIs between modified Boolean remission categories and Simplified Disease Activity Index (SDAI) remission, separately for early rheumatoid arthritis (RA) (blue line) and established RA (red line) at 6 months and at 12 months.

Table 1 Calculated sensitivity, specificity, positive likelihood ratio (+LR), negative likelihood ratio (–LR) of PGA cut-offs using coinciding SDAI remission as outcome, in patients fulfilling SJC28, TJC28 and CRP all ≤1

6 months						12 months					
Definition	**Cut-off**	**Sensitivity**	**Specificity**	**+LR**	**–LR**	**Definition**	**Cut-off**	**Sensitivity**	**Specificity**	**+LR**	**–LR**
Early	1	88.1 (83.1–92.1)	87.4 (79.7 to 92.9)	7.0 (4.3–11.4)	0.14 (0.09–0.2)	Early	1	87.2 (83.2–90.4)	85.6 (78.9 to 90.9)	6.1 (4.1–9.0)	0.2 (0.1–0.2)
	1.1	90 (85.2–93.6)	84.7 (76.6 to 90.8)	5.9 (3.8–9.1)	0.1 (0.08–0.2)		1.1*	89.1 (85.4–92.1)	84.3 (77.3 to 89.7)	5.7 (3.9–8.2)	0.1 (0.10–0.2)
	1.2	92.7 (88.4–95.8)	82.9 (74.6 to 89.4)	5.4 (3.6–8.2)	0.1 (0.05–0.1)		1.2	89.9 (86.4–92.9)	81.5 (74.2 to 87.4)	4.9 (3.5–6.8)	0.1 (0.09–0.2)
	1.3	94.1 (90.1–96.8)	82.9 (74.6 to 89.4)	5.5 (3.6–8.3)	0.1 (0.04–0.1)		1.3	91.9 (88.6–94.5)	78.8 (71.2 to 85.1)	4.3 (3.2–5.9)	0.1 (0.07–0.1)
	1.4	95.9 (92.3–98.1)	81.1 (72.5 to 87.9)	5.1 (3.4–7.5)	0.1 (0.03–0.10)		1.4	93.6 (90.5–95.9)	76.7 (69.0 to 83.3)	4 (3.0–5.4)	0.1 (0.06–0.1)
	1.5	96.8 (93.5–98.7)	80.2 (71.5 to 87.1)	4.9 (3.4–7.1)	0.04 (0.02–0.08)		1.5	93.9 (90.8–96.1)	73.3 (65.3 to 80.3)	3.5 (2.7–4.6)	0.1 (0.06–0.1)
	1.6*	97.3 (94.1–99.0)	79.3 (70.5 to 86.4)	4.7 (3.3–6.8)	0.04 (0.02–0.08)		1.6	95 (92.2–97.0)	71.2 (63.2 to 78.4)	3.3 (2.6–4.3)	0.1 (0.04–0.1)
	1.7	98.2 (95.4–99.5)	77.5 (68.6 to 84.9)	4.4 (3.1–6.2)	0.02 (0.009–0.06)		1.7	96.1 (93.5–97.8)	70.6 (62.4 to 77.8)	3.3 (2.5–4.2)	0.1 (0.03–0.09)
	1.8	98.2 (95.4–99.5)	73.9 (64.7 to 81.8)	3.8 (2.7–5.1)	0.03 (0.009–0.07)		1.8	96.4 (93.9–98.1)	69.2 (61.0 to 76.5)	3.1 (2.5–4.0)	0.1 (0.03–0.09)
	1.9	98.6 (96.0–99.7)	71.2 (61.8 to 79.4)	3.4 (2.6–4.6)	0.02 (0.006–0.06)		1.9	96.7 (94.2–98.3)	67.1 (58.9 to 74.7)	2.9 (2.3–3.7)	0.1 (0.03–0.09)
	2	98.6 (96.0–99.7)	68.5 (59.0 to 77.0)	3.1 (2.4–4.1)	0.02 (0.01–0.06)		2	96.9 (94.6–98.5)	61.6 (53.2 to 69.6)	2.5 (2.1–3.1)	0.1 (0.03–0.09)
Established	1	73.1 (59.0–84.4)	73.8 (60.9 to 84.2)	2.8 (1.8–4.4)	0.4 (0.2–0.6)	Established	1	83.9 (75.8–90.2)	90.3 (82.4 to 95.5)	8.7 (4.6–16.2)	0.2 (0.1–0.3)
	1.1	75 (61.1–86.0)	73.8 (60.9 to 84.2)	2.9 (1.8–4.5)	0.3 (0.2–0.6)		1.1	83.9 (75.8–90.2)	89.3 (81.1 to 94.7)	7.8 (4.3–14.1)	0.2 (0.1–0.3)
	1.2	76.9 (63.2–87.5)	72.1 (59.2 to 82.9)	2.8 (1.8–4.2)	0.3 (0.2–0.5)		1.2	85.7 (77.8–91.6)	87.1 (78.5 to 93.2)	6.6 (3.9–11.3)	0.2 (0.1–0.3)
	1.3*	84.6 (71.9–93.1)	68.9 (55.7 to 80.1)	2.7 (1.8–4.0)	0.2 (0.1–0.4)		1.3	88.4 (81.0–93.7)	85 (76.0 to 91.5)	5.9 (3.6–9.6)	0.1 (0.08–0.2)
	1.4	84.6 (71.9–93.1)	67.2 (54.0 to 78.7)	2.6 (1.8–3.8)	0.2 (0.1–0.4)		1.4	91.1 (84.2–95.6)	83.9 (74.8 to 90.7)	5.7 (3.5–9.0)	0.1 (0.06–0.2)
	1.5	86.5 (74.2–94.4)	62.3 (49.0 to 74.4)	2.3 (1.6–3.2)	0.2 (0.1–0.4)		1.5*	93.8 (87.5–97.5)	80.7 (71.1 to 88.1)	4.8 (3.2–7.4)	0.1 (0.04–0.2)
	1.6	88.5 (76.6–95.6)	59 (45.7 to 71.4)	2.2 (1.6–3.0)	0.2 (0.09–0.4)		1.6	93.8 (87.5–97.5)	79.6 (69.9 to 87.2)	4.6 (3.1–6.9)	0.1 (0.04–0.2)
	1.7	90.4 (79.0–96.8)	54.1 (40.8 to 66.9)	2 (1.5–2.6)	0.2 (0.07–0.4)		1.7	95.5 (89.9–98.5)	76.3 (66.4 to 84.5)	4 (2.8–5.8)	0.1 (0.02–0.1)
	1.8	92.3 (81.5–97.9)	52.5 (39.3 to 65.4)	1.9 (1.5–2.6)	0.2 (0.06–0.4)		1.8	96.4 (91.1–99.0)	74.2 (64.1 to 82.7)	3.7 (2.6–5.3)	0.05 (0.02–0.1)
	1.9	94.2 (84.1–98.8)	47.5 (34.6 to 60.7)	1.8 (1.4–2.3)	0.1 (0.04–0.4)		1.9	96.4 (91.1–99.0)	73.1 (62.9 to 81.8)	3.6 (2.6–5.0)	0.05 (0.02–0.1)
	2	96.2 (86.8–99.5)	45.9 (33.1 to 59.2)	1.8 (1.4–2.3)	0.1 (0.02–0.3)		2	97.3 (92.4–99.4)	69.9 (59.5 to 79.0)	3.2 (2.4–4.4)	0.04 (0.01–0.1)
All	1	85.2 (80.4–89.2)	82.6 (76.0 to 87.9)	4.9 (3.5–6.8)	0.2 (0.1–0.2)	All	1	86.4 (82.9–89.4)	87.5 (82.6 to 91.4)	6.9 (4.9–9.6)	0.2 (0.1–0.2)
	1.1	87.1 (82.5–90.8)	80.8 (74.1 to 86.4)	4.5 (3.3–6.2)	0.2 (0.1–0.2)		1.1*	88.1 (84.8–90.9)	85.4 (80.2 to 89.6)	6 (4.4–8.2)	0.1 (0.1–0.2)
	1.2	87.5 (82.9–91.2)	80.2 (73.5 to 85.9)	4.4 (3.3–6.0)	0.2 (0.1–0.2)		1.2	88.3 (85.0–91.1)	85.4 (80.2 to 89.6)	6 (4.4–8.2)	0.1 (0.1–0.2)
	1.3	92.3 (88.4–95.1)	77.9 (71.0 to 83.9)	4.2 (3.1–5.5)	0.1 (0.07–0.2)		1.3	89.6 (86.5–92.2)	82.9 (77.5 to 87.4)	5.2 (3.9–6.9)	0.1 (0.1–0.2)
	1.4	87.5 (82.9–91.2)	80.2 (73.5 to 85.9)	4.4 (3.3–6.0)	0.2 (0.1–0.2)		1.4	91.7 (88.8–94.0)	80.8 (75.2 to 85.6)	4.8 (3.7–6.2)	0.1 (0.08–0.1)
	1.5	94.8 (91.5–97.1)	73.8 (66.6 to 80.2)	3.6 (2.8–4.7)	0.1 (0.04–0.1)		1.5	93.8 (91.3–95.8)	76.2 (70.2 to 81.4)	3.9 (3.1–4.9)	0.1 (0.06–0.1)
	1.6*	95.6 (92.4–97.7)	72.1 (64.8 to 78.7)	3.4 (2.7–4.4)	0.1 (0.04–0.1)		1.6	94.7 (92.2–96.5)	74.5 (68.5 to 79.9)	3.7 (3.0–4.6)	0.1 (0.05–0.1)
	1.7	96.7 (93.8–98.5)	69.2 (61.7 to 76.0)	3.1 (2.5–3.9)	0.05 (0.03–0.09)		1.7	94.9 (92.5–96.7)	74.5 (68.5 to 79.9)	3.7 (3.0–4.6)	0.1 (0.05–0.1)
	1.8	97.1 (94.3–98.7)	66.3 (58.7 to 73.3)	2.9 (2.3–3.6)	0.05 (0.02–0.09)		1.8	96.4 (94.3–97.9)	71.1 (64.9 to 76.8)	3.3 (2.7–4.1)	0.1 (0.03–0.08)
	1.9	97.8 (95.2–99.2)	62.8 (55.1 to 70.0)	2.6 (2.2–3.2)	0.04 (0.02–0.08)		1.9	96.4 (94.3–97.9)	70.7 (64.5 to 76.4)	3.3 (2.7–4.0)	0.1 (0.03–0.08)
	2	98.2 (95.7–99.4)	60.5 (52.7 to 67.8)	2.5 (2.1–3.0)	0.03 (0.01–0.07)		2	97 (95.1–98.4)	64.9 (58.4 to 70.9)	2.8 (2.3–3.3)	0.05 (0.03–0.08)

Cut-offs (to be interpreted as ≤) marked with * were proposed in the classification and regression tree (CART) analyses.
CRP, C reactive protein; PGA, patient global assessment; SDAI, Simplified Disease Activity Index; SJC, swollen joint counts; TJC, tender joint counts .

'*' marks). In ROC analyses, sensitivity and specificity characteristics of PGA cut-offs in 0.1 cm increments from 1.0 to 2.0 are outlined in table 1, supporting cut-offs of CART analyses. The retrieved lower sensitivity and specificity of PGA at 6 months in patients with established RA compared with early RA is in line with the general worse agreement (lower kappa) of Boolean definitions and SDAI definitions in this population. When aiming for high sensitivity of the PGA criterion in modified Boolean definition to coincide with the SDAI definition for all patients with RA, 1.5 seems to be an appropriate cut-off, resulting in similar sensitivity at both time points (95% at 6 months, and 94% at 12 months).

Structural and functional implications of remission definitions

We studied the distribution of HAQ scores at 1 year and of X-ray progression (ΔmTSS) separately for patients in the different Boolean definitions. The radiographic outcomes were independent of the PGA cut-off, and score changes were similar between different definitions (mean ΔmTSS for Boolean1.5, 2.0, 2.5, 3.0 and BooleanX were: 0.38±5.14, 0.41±5.1, 0.37±4.9, 0.34±4.9 and 0.27±4.7). These Boolean definitions also led to similar fraction of patients progressing during the first year (defined as ΔmTSS>0; for Boolean1.0, 1.5, 2.0, 2.5, 3.0 and BooleanX: 39.3%, 39.4%, 38.6%, 38.1%, 37.5% and 37.3%).

In contrast to the expected radiographic data, higher PGA thresholds were accompanied by higher HAQ scores, with BooleanX showing the highest level of functional impairment. The proportion achieving a good functional outcome defined as HAQ≤0.5 was 90.2%, 87.9%, 85.2%, 81.1%, 80.7% and 73.1% for Boolean, 1.5, 2.0, 2.5, 3.0 and BooleanX, respectively; mean HAQ scores were 0.15±0.31, 0.19±0.37, 0.22±0.39, 0.26±0.42, 0.27±0.43 and 0.37±0.52, respectively. Scores in established RA were generally worse than for early RA but the distribution over different Boolean classifications remained similar (figure 3 depicts results for early RA; and online supplementary figure 1 oneestablished RA). The SF-36 physical component scores were distributed like the HAQ scores and were worse when the PGA was completely omitted (green line). The distribution of scores was likewise similar in established RA but appears generally worse than in early.

We have also explored these distribution plots in non-overlapping groups of modified Boolean remitters, so that every patient is attributed to only one definition (eg, Boolean2.0 remitters would not include Boolean1.0 or Boolean1.5 remitters in this analysis). We found distinct distributions of scores on HAQ and SF-36 physical components (online supplementary figure 2). The rate of progressors in mTSS was not different between Boolean and BooleanX patients (40.6% vs 32.0%; p=0.264). However, as the remission threshold for PGA increased, the proportion with good functional outcomes (defined as HAQ≤0.5) decreased and this proportion dropped further when PGA was completely removed (HAQ≤0.5 (n) in Boolean, 1.5, 2.0, 2.5, 3.0 and BooleanX: 92.3% (193), 75.9% (22), 47.1% (8), 33.3% (7), 77.8% (7), 37.3% (19)).

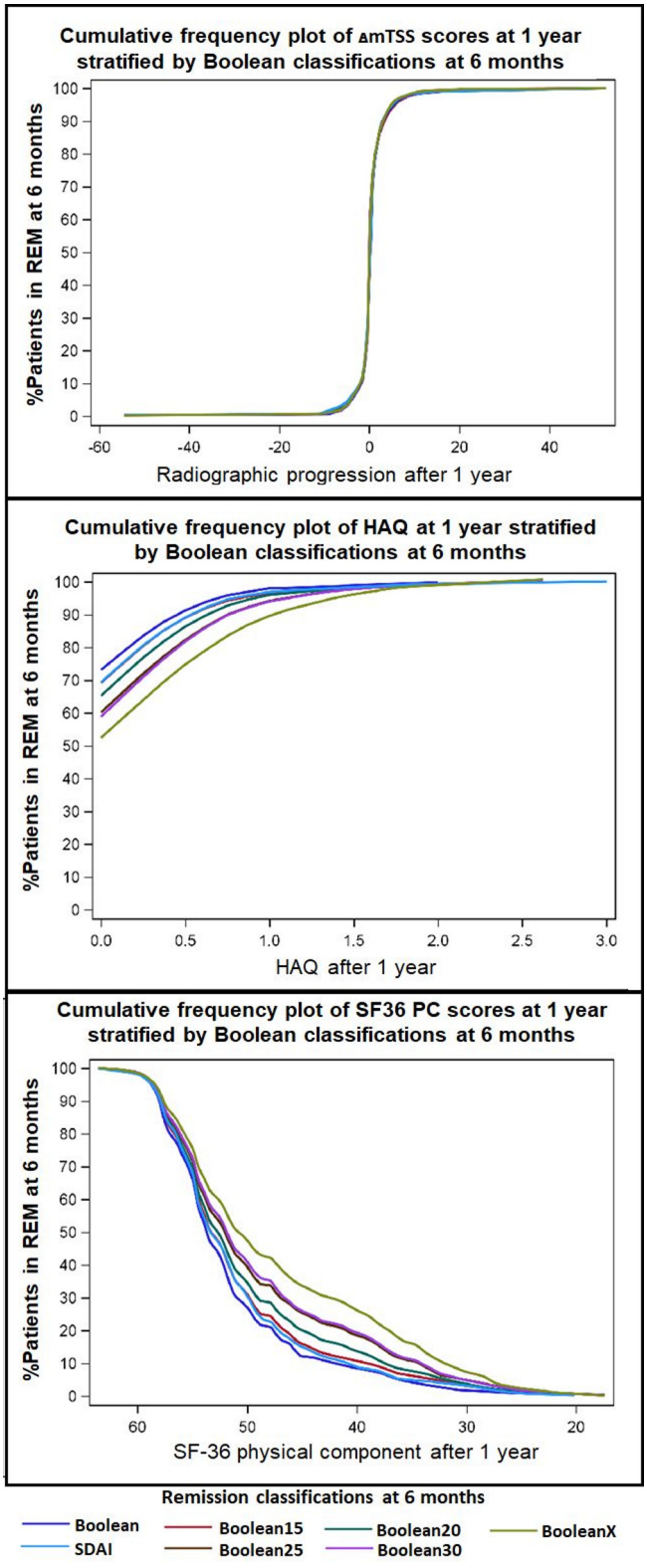

Figure 3 Cumulative frequency distribution of X-ray change (mTSS), HAQ and SF-36 physical component scores in patients with early rheumatoid arthritis (RA), separately by categorisation in modified Boolean remission definitions and Simplified Disease Activity Index (SDAI) remission (overlapping groups).

DISCUSSION

Pooling six different large clinical trials, we evaluated the role of PGA, or its cut-off, in the Boolean remission definition, as well as its impact on outcomes. We used the SDAI remission definition, which is the ACR/EULAR index-based remission criterion, as the comparator in our analyses. Maintaining SJC, TJC and CRP at their maximum cut-point of 1, we tested different levels of PGA as the fourth component of the Boolean criteria to see if higher PGA scores would change overall outcomes.

Generally, in our population, around 40% of the patients showed radiographic progression, in accordance with other

studies. Nevertheless, the mean ᴀmTSS was low, in line with observations of a secular trend of lower progression rates.[35] The observed somewhat high rates of progression in remission can be explained by the latency (or carry-over) effect of disease activity on radiographic progression.[36] Furthermore, since SJC and CRP are associated with joint damage,[13 37] we did not expect to see differences in damage progression rates when higher PGA scores were a component of the Boolean remission criteria, and this was observed in our analyses. In contrast, physical function as assessed by the HAQ, but also by SF-36, deteriorated with increasing the threshold for PGA. However, the difference in good functional outcomes was small when comparing 1, 1.5 and 2 cm ratings of the PGA (about 5% difference in proportions of normative HAQ), while this difference was much larger when PGA was completely excluded. Since remission ought to encompass clinical, structural and functional remission,[11] the omission of the PGA from Boolean criteria is not in line with an optimal understanding of remission. On the other hand, many more patients (+20% in early RA at 6 months) can be classified as in remission by Boolean criteria when the threshold for the PGA is increased from 1 to 2 cm, without a major loss of good outcomes. Still, one may ask if the PGA should be included at all in a definition of remission of inflammation, since functional outcomes, for whatever reason are worse, independent of differences in radiographic progression. Other studies have, however, shown that the HAQ has only a minor influence on PGA score, suggesting there is little reverse causation, whereas pain is the greatest driver of PGA.[20 38] This integration of patient-derived factors and more objective markers provides a robust overall assessment of disease activity. An exclusion would constitute a step back in disease activity assessment. In addition, studies informing the work developing the definition of improvement[1] showed that PGA was usually the outcome measure that best discriminated disease modifying antirheumatic drugs (DMARDs) from placebo, suggesting that PGA provides information on inflammation and its response to treatment. Omitting PGA would compromise the ability to detect treatment efficacy.

One goal of this study was to increase the concordance between two equally applicable definitions of remission. While this may seem to be circular, it can also be seen as a strength, since both definitions have been confirmed to coincide with high predictive validity for the inhibition of bad outcome.[11 39] This constitutes a main reason for targeting remission in the treatment of patients with RA. All trials included in this study have been conducted in the last decade and investigated MTX and TNFi, although nowadays many other DMARD classes are available. In particular, Jak inhibitors have shown fast response; however, Jak-inhibitor trials of the last years outlined 6-month and 12-month Boolean remission rates between 7% and 23%,[40–43] similar to our patients with early RA (12%–20%).

Based on the comprehensive interpretation of the results from the kappa, CART and ROC analyses, increasing the PGA cut-off to 1.5 cm would provide the highest consistency between Boolean and index-based remission, while the integer cut-off of 2 cm (or 2/10) would also allow the use of an integer-based numerical rating scales. We acknowledge that a 2 cm cut-off, instead of 1.5 cm, harbours the risk of lower specificity for remission. However, when considering that in patients, who score a PGA≤1 cm, a smallest detectable difference for the PGA ranging between 1.3 and 1.8 cm has been reported.[44] Another study outlined even a smallest detectable difference of 2.3 cm in the PGA.[45] This suggested new cut-off would discount the stringency of the PGA in the remission context, while keeping the patient perspective as a core element of RA disease activity evaluation, without compromising long-term structural outcomes.

A cut-off beyond 2 cm would not only jeopardise agreement with the index definition and be associated with poorer long-term function but also require other factors to be considered. While mostly pain and partly fatigue influence PGA irrespective of disease activity,[38] pain and fatigue may also reflect active inflammation and thus disease activity in many patients.[46]

Although to a much smaller extent than PGA, it needs to be noted that also joint swelling and CRP levels may not always be accurate: joint swelling may often be doubtful, observer-dependent or related to concomitant diseases, such as osteoarthritis, and increased CRP may be caused by other concomitant diseases, such as undetected infection.[47–49] Analogously, SJC and CRP levels may be elevated even though a patient is in RA remission.[18] Furthermore, certain drugs, such as IL-6- and Jak inhibitors, may normalise CRP irrespective of clinical improvement[50] (and, thus, lead to potential undertreatment with the consequence of joint damage progression and irreversible disability). This may be even more misleading than a high patient global which still necessitates a physician's attention. Its relation to inflammation can be well differentiated from a relation to non-inflammatory abnormalities by most rheumatologists using a patient-centred approach.

This patient-centred approach needs to accompany any clinical consultation and should address the background to situations, where the PGA may indeed be unduly high.[51] The fact that fatigue, pain, anxiety and function influence the variance of the PGA in a state of near remission[33 52 53] also shows that the score represents factors, that would not be covered otherwise and may be influenced by inflammation. Some lack of specificity may be caused by the question phrasing (eg, DAS used to include a PGA on global health, not specifying arthritis-related symptoms). When the PGA does not specify arthritis-related symptoms, it may lead to a misimpression that RA is active while in reality other factors may explain a patient's score.

Other factors that influence the outlined remission criteria need to be considered on application to the respective patient. This has been clearly stated in the treat-to-target recommendations, where recommendation 5 states: "The choice of the (composite) measure of disease activity and the target value should be influenced by comorbidities, patient factors and drug-related risks" and certain comorbidities, such as fibromyalgia, are explicitly mentioned.

Our findings suggest that modifying the cut-off for PGA in the Boolean criteria for remission to 2 (on a scale of 0–10) results in better agreement with the SDAI-based ACR/EULAR definition of remission than when using the current PGA definition of 1. This change should be strongly considered.

Twitter Paul Studenic @Stiddyo

Acknowledgements We thank Abbvie and Centocor (Janssen) for kindly providing patient level data.

Contributors Study design: PS, DF, JSS and DA. Analyses of data: PS, FA and TAS. Interpretation of data: PS, DF, MdW, FA, TA, JSS and DA. Writing and editing of the manuscript: PS, DF, MdW, TAS, JSS and DA.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval No additional review board approval was obtained due to secondary data analysis of pooled clinical trial data.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. Data have been provided by the respective sponsors of the trials. Any requests for individual patient level data will have to be addressed to these sponsors directly.

ORCID iDs
Paul Studenic http://orcid.org/0000-0002-8895-6941
David Felson http://orcid.org/0000-0002-2668-2447
Maarten de Wit http://orcid.org/0000-0002-8428-6354
Tanja A Stamm http://orcid.org/0000-0003-3073-7284

REFERENCES

1 Felson DT, Anderson JJ, Boers M, *et al*. The American College of rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. The Committee on outcome measures in rheumatoid arthritis clinical trials. *Arthritis Rheum* 1993;36:729–40.
2 Scott D, van Riel P, van der Heijde D. *Assessing disease activity in rheumatoid arthritis - The EULAR handbook of standard methods. On behalf of the EULAR Standing Committee for International Clinical Studies Including Therapeutic Trials - ESCISIT (Chairman: Smolen JS)*. Zürich: EULAR, 1993.
3 Goldsmith CH, Smythe HA, Helewa A. Interpretation and power of a pooled index. *J Rheumatol* 1993;20:575–8.
4 Ward MM, Guthrie LC, Alba MI. Clinically important changes in individual and composite measures of rheumatoid arthritis activity: thresholds applicable in clinical trials. *Ann Rheum Dis* 2015;74:1691–6.
5 Aletaha D, Smolen J. The simplified disease activity index (SDAI) and the clinical disease activity index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheumatol* 2005;23:S100–8.
6 van der Heijde DM, van 't Hof MA, van Riel PL, *et al*. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. *Ann Rheum Dis* 1990;49:916–20.
7 Emery P, Salmon M. Early rheumatoid arthritis: time to aim for remission? *Ann Rheum Dis* 1995;54:944–7.
8 Singh JA, Saag KG, Bridges SL, *et al*. 2015 American College of rheumatology guideline for the treatment of rheumatoid arthritis. *Arthritis Rheumatol* 2016;68:1–26.
9 Smolen JS, Landewé R, Bijlsma J, *et al*. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis* 2017;76:960–77.
10 Smolen JS, Breedveld FC, Burmester GR, *et al*. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international Task force. *Ann Rheum Dis* 2016;75:3–15.
11 Felson DT, Smolen JS, Wells G, *et al*. American College of Rheumatology/European League against rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Ann Rheum Dis* 2011;70:404–13.
12 van der Heijde D, Klareskog L, Boers M, *et al*. Comparison of different definitions to classify remission and sustained remission: 1 year tempo results. *Ann Rheum Dis* 2005;64:1582–7.
13 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.
14 Schoels M, Alasti F, Smolen JS, *et al*. Evaluation of newly proposed remission cut-points for disease activity score in 28 joints (DAS28) in rheumatoid arthritis patients upon IL-6 pathway inhibition. *Arthritis Res Ther* 2017;19:155.
15 Bakker MF, Jacobs JWG, Verstappen SMM, *et al*. Tight control in the treatment of rheumatoid arthritis: efficacy and feasibility. *Ann Rheum Dis* 2007;66(Suppl 3):iii56–60.
16 Emery P, Keystone E, Tony HP, *et al*. Il-6 receptor inhibition with tocilizumab improves treatment outcomes in patients with rheumatoid arthritis refractory to anti-tumour necrosis factor biologicals: results from a 24-week multicentre randomised placebo-controlled trial. *Ann Rheum Dis* 2008;67:1516–23.
17 Cohen SB, Emery P, Greenwald MW, *et al*. Rituximab for rheumatoid arthritis refractory to anti-tumor necrosis factor therapy: results of a multicenter, randomized, double-blind, placebo-controlled, phase III trial evaluating primary efficacy and safety at twenty-four weeks. *Arthritis Rheum* 2006;54:2793–806.
18 Studenic P, Smolen JS, Aletaha D. Near misses of ACR/EULAR criteria for remission: effects of patient global assessment in Boolean and index-based definitions. *Ann Rheum Dis* 2012;71:1702–5.
19 Aletaha D, Landewe R, Karonitsch T, *et al*. Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations. *Ann Rheum Dis* 2008;67:1360–4.
20 Studenic P, Radner H, Smolen JS, *et al*. Discrepancies between patients and physicians in their perceptions of rheumatoid arthritis disease activity. *Arthritis Rheum* 2012;64:2814–23.
21 Radner H, Yoshida K, Tedeschi S, *et al*. Different rating of global rheumatoid arthritis disease activity in rheumatoid arthritis patients with multiple morbidities. *Arthritis Rheumatol* 2017;69:720–7.
22 Mack ME, Hsia E, Aletaha D. Comparative assessment of the different American College of Rheumatology/European League against rheumatism remission definitions for rheumatoid arthritis for their use as clinical trial end points. *Arthritis Rheumatol* 2017;69:518–28.
23 Genovese MC, Kremer J, Zamani O, *et al*. Baricitinib in patients with refractory rheumatoid arthritis. *N Engl J Med* 2016;374:1243–52.
24 Thiele K, Huscher D, Bischoff S, *et al*. Performance of the 2011 ACR/EULAR preliminary remission criteria compared with DAS28 remission in unselected patients with rheumatoid arthritis. *Ann Rheum Dis* 2013;72:1194–9.
25 St Clair EW, van der Heijde DMFM, Smolen JS, *et al*. Combination of infliximab and methotrexate therapy for early rheumatoid arthritis: a randomized, controlled trial. *Arthritis Rheum* 2004;50:3432–43.
26 Breedveld FC, Weisman MH, Kavanaugh AF, *et al*. The premier study: a multicenter, randomized, double-blind clinical trial of combination therapy with adalimumab plus methotrexate versus methotrexate alone or adalimumab alone in patients with early, aggressive rheumatoid arthritis who had not had previous methotrexate treatment. *Arthritis Rheum* 2006;54:26–37.
27 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-naive 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.
28 Maini R, St Clair EW, Breedveld F, *et al*. Infliximab (chimeric anti-tumour necrosis factor alpha monoclonal antibody) versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: a randomised phase III trial. attract Study Group. *Lancet* 1999;354:1932–9.
29 Keystone EC, Kavanaugh AF, Sharp JT, *et al*. Radiographic, clinical, and functional outcomes of treatment with adalimumab (a human anti-tumor necrosis factor monoclonal antibody) in patients with active rheumatoid arthritis receiving concomitant methotrexate therapy: a randomized, placebo-controlled, 52-week trial. *Arthritis Rheum* 2004;50:1400–11.
30 Keystone EC, Genovese MC, Klareskog L, *et al*. Golimumab, a human antibody to tumour necrosis factor {alpha} given by monthly subcutaneous injections, in active rheumatoid arthritis despite methotrexate therapy: the GO-FORWARD Study. *Ann Rheum Dis* 2009;68:789–96.
31 Aletaha D, Alasti F, Smolen JS. Rheumatoid arthritis near remission: clinical rather than laboratory inflammation is associated with radiographic progression. *Ann Rheum Dis* 2011;70:1975–80.
32 Aletaha D, Funovits J, Smolen JS. Physical disability in rheumatoid arthritis is associated with cartilage damage rather than bone destruction. *Ann Rheum Dis* 2011;70:733–9.
33 Ferreira RJO, Duarte C, Ndosi M, *et al*. Suppressing inflammation in rheumatoid arthritis: does patient global assessment blur the target? A practice-based call for a paradigm change. *Arthritis Care Res* 2018;70:369–78.
34 van Tuyl LHD, Hewlett S, Sadlonova M, *et al*. The patient perspective on remission in rheumatoid arthritis: 'You've got limits, but you're back to being you again'. *Ann Rheum Dis* 2015;74:1004–10.
35 Rahman MU, Buchanan J, Doyle MK, *et al*. Changes in patient characteristics in anti-tumour necrosis factor clinical trials for rheumatoid arthritis: results of an analysis of the literature over the past 16 years. *Ann Rheum Dis* 2011;70:1631–40.
36 Aletaha D, Funovits J, Breedveld FC, *et al*. Rheumatoid arthritis joint progression in sustained remission is determined by disease activity levels preceding the period of radiographic assessment. *Arthritis Rheum* 2009;60:1242–9.
37 Vastesaeger N, Xu S, Aletaha D, *et al*. A pilot risk model for the prediction of rapid radiographic progression in rheumatoid arthritis. *Rheumatology* 2009;48:1114–21.
38 Karpouzas GA, Strand V, Ormseth SR. Latent profile analysis approach to the relationship between patient and physician global assessments of rheumatoid arthritis activity. *RMD Open* 2018;4:e000695.
39 Paulshus Sundlisæter N, Aga A-B, Olsen IC, *et al*. Clinical and ultrasound remission after 6 months of treat-to-target therapy in early rheumatoid arthritis: associations to future good radiographic and physical outcomes. *Ann Rheum Dis* 2018;77:1421–5.
40 Fleischmann R, Mysler E, Hall S, *et al*. Efficacy and safety of tofacitinib monotherapy, tofacitinib with methotrexate, and adalimumab with methotrexate in patients with rheumatoid arthritis (oral strategy): a phase 3b/4, double-blind, head-to-head, randomised controlled trial. *Lancet* 2017;390:457–68.
41 Burmester GR, Kremer JM, Van den Bosch F, *et al*. Safety and efficacy of upadacitinib in patients with rheumatoid arthritis and inadequate response to conventional synthetic disease-modifying anti-rheumatic drugs (SELECT-NEXT): a randomised, double-blind, placebo-controlled phase 3 trial. *Lancet* 2018;391:2503–12.
42 Fleischmann RM, Genovese MC, Enejosa JV, *et al*. Safety and effectiveness of upadacitinib or adalimumab plus methotrexate in patients with rheumatoid arthritis over 48 weeks with switch to alternate therapy in patients with insufficient response. *Ann Rheum Dis* 2019;78:1454–62.
43 Taylor PC, Keystone EC, van der Heijde D, *et al*. Baricitinib versus placebo or adalimumab in rheumatoid arthritis. *N Engl J Med* 2017;376:652–62.

44 Studenic P, Stamm T, Smolen JS, *et al*. Reliability of patient-reported outcomes in rheumatoid arthritis patients: an observational prospective study. *Rheumatology* 2016;55:41–8.
45 Masri KR, Shaver TS, Shahouri SH, *et al*. Validity and reliability problems with patient global as a component of the ACR/EULAR remission criteria as used in clinical practice. *J Rheumatol* 2012;39:1139–45.
46 Gossec L, Dougados M, Dixon W. Patient-reported outcomes as end points in clinical trials in rheumatoid arthritis. *RMD Open* 2015;1:e000019.
47 Meersseman P, Van de Vyver C, Verbruggen G, *et al*. Clinical and radiological factors associated with erosive radiographic progression in hand osteoarthritis. *Osteoarthritis Cartilage* 2015;23:2129–33.
48 Khan NA, Spencer HJ, Nikiphorou E, *et al*. Intercentre variance in patient reported outcomes is lower than objective rheumatoid arthritis activity measures: a cross-sectional study. *Rheumatology* 2017;56:1395–400.
49 Turk M, Pope JE. Physician global assessments for disease activity in rheumatoid arthritis are all over the MAP! *RMD Open* 2018;4:e000578.
50 Song S-NJ, Iwahashi M, Tomosugi N, *et al*. Comparative evaluation of the effects of treatment with tocilizumab and TNF-α inhibitors on serum hepcidin, anemia response and disease activity in rheumatoid arthritis patients. *Arthritis Res Ther* 2013;15:R141.
51 Ferreira RJO, Carvalho PD, Ndosi M, *et al*. Impact of patient's global assessment on achieving remission in patients with rheumatoid arthritis: a multinational study using the Meteor database. *Arthritis Care Res* 2019;71:1317–25.
52 Egsmose EL, Madsen OR. Interplay between patient global assessment, pain, and fatigue and influence of other clinical disease activity measures in patients with active rheumatoid arthritis. *Clin Rheumatol* 2015;34:1187–94.
53 Inanc N, Yilmaz-Oner S, Can M, *et al*. The role of depression, anxiety, fatigue, and fibromyalgia on the evaluation of the remission status in patients with rheumatoid arthritis. *J Rheumatol* 2014;41:1755–60.