



HAL
open science

Efficacy of a tight-control and treat-to-target strategy in axial spondyloarthritis: results of the open-label, pragmatic, cluster-randomised TICOSPA trial

Anna Molto, Clementina López-Medina, Filip van den Bosch, Annelies Boonen, Casper Webers, Emanuelle Dernis, Floris van Gaalen, Martin Soubrier, Pascal Claudepierre, Athan Baillet, et al.

► To cite this version:

Anna Molto, Clementina López-Medina, Filip van den Bosch, Annelies Boonen, Casper Webers, et al.. Efficacy of a tight-control and treat-to-target strategy in axial spondyloarthritis: results of the open-label, pragmatic, cluster-randomised TICOSPA trial. *Annals of the Rheumatic Diseases*, 2021, pp.annrheumdis-2020-219585. 10.1136/annrheumdis-2020-219585 . hal-03222492

HAL Id: hal-03222492

<https://hal.sorbonne-universite.fr/hal-03222492>

Submitted on 30 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.









Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License



OPEN ACCESS

CLINICAL SCIENCE

Efficacy of a tight-control and treat-to-target strategy in axial spondyloarthritis: results of the open-label, pragmatic, cluster-randomised TICOSPA trial

Anna Molto ^{1,2} Clementina López-Medina ^{2,3} Filip E Van den Bosch ⁴,
 Annelies Boonen ^{5,6} Casper Webers,^{5,6} Emanuelle Dernis,⁷ Floris A van Gaalen,⁸
 Martin Soubrier,⁹ Pascal Claudepierre,^{10,11} Athan Baillet ^{12,13}
 Mirian Starmans-Kool,¹⁴ Anneke Spoorenberg,¹⁵ Peggy Jacques ^{4,16}
 Philippe Carron ^{16,17} Rik Joos,¹⁸ Jan Lenaerts,¹⁹ Laure Gossec ^{20,21}
 Sophie Pouplin,²² Adeline Ruysen-Witrand,^{23,24} Laetitia Sparsa,²⁵
 Astrid van Tubergen ²⁶ Désirée van der Heijde ²⁷ Maxime Dougados^{1,2}

Handling editor Josef S Smolen

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/annrheumdis-2020-219585>).

For numbered affiliations see end of article.

Correspondence to

Dr Anna Molto, Rheumatology, Hospital Cochin, Paris 75014, France; anna.molto@aphp.fr

Received 24 November 2020
 Revised 16 April 2021
 Accepted 18 April 2021
 Published Online First 6 May 2021



► <http://dx.doi.org/10.1136/annrheumdis-2021-220603>



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Molto A, López-Medina C, Van den Bosch FE, et al. *Ann Rheum Dis* 2021;**80**:1436–1444.

ABSTRACT

Objectives To compare the benefits of a tight-control/treat-to-target strategy (TC/T2T) in axial spondyloarthritis (axSpA) with those of usual care (UC).

Methods Pragmatic, prospective, cluster-randomised, controlled, open, 1-year trial (NCT03043846). 18 centres were randomised (1:1). Patients met Axial Spondylo Arthritis International Society (ASAS) criteria for axSpA, had an Ankylosing Spondylitis Disease Activity Score (ASDAS) ≥ 2.1 , received non-optimal treatment by non-steroidal anti-inflammatory drugs and were biologic-naïve.

Interventions (1) TC/T2T: visits every 4 weeks and prespecified strategy based on treatment intensification until achieving target (ie, ASDAS < 2.1); (2) UC: visits every 12 weeks and treatment at the rheumatologist's discretion.

Main outcome Percentage of patients with a $\geq 30\%$ improvement on the ASAS-Health Index (ASAS-HI). Other efficacy outcomes and adverse events were recorded. A health economic evaluation was performed.

Statistical analysis Two-level mixed models were used to estimate efficacy outcomes. Cost-effectiveness was assessed by the incremental cost per quality-adjusted life-year (QALY) gained for TC/T2T versus UC.

Results 160 patients were included (80/group). Mean (SD) age was 37.9 (11.0) years and disease duration was 3.7 (6.2) years; 51.2% were men. ASDAS at inclusion was 3.0 (0.7), and ASAS-HI was 8.6 (3.7). ASAS-HI improved by $\geq 30\%$ in 47.3% of the TC/T2T arm and in 36.1% of those receiving UC (non-significant). All secondary efficacy outcomes were more frequent in the TC/T2T arm, although not all statistically significant. Safety was similar in both arms. From a societal perspective, TC/T2T resulted in an additional 0.04 QALY, and saved €472 compared with UC.

Conclusion TC/T2T was not significantly superior to UC for the primary outcome, while many secondary efficacy outcomes favoured it, had a similar safety profile and was favourable from a societal health economic perspective.

Trial registration number NCT03043846.

Key messages**What is already known about this subject?**

- Treat-to-target (T2T) has demonstrated to be an efficacious approach in rheumatic inflammatory diseases such as rheumatoid arthritis and psoriatic arthritis.
- Recommendations for the management of axial spondyloarthritis (axSpA) have been published, including the recommendation to apply a T2T approach in this disease, despite the lack of evidence of the utility of such approach in this disease, compared with usual care (UC).

What does this study add?

- This is the first study evaluating the efficacy of a treat-to-target and tight control (TC) approach in axSpA compared with UC.
- Overall, in this setting of expert centres in spondyloarthritis, UC resulted in very good outcomes for a substantial number of patients.
- Although the primary outcome measure was not achieved, response rates between the two treatment groups differed by 11% in favour of TC/T2T.
- Despite the higher prescription rate of biologics in the TC/T2T arm, safety profiles were similar across arms, and the TC/T2T arm had a favourable outcome from a societal health economic perspective.

How might this impact on clinical practice?

- This trial did not prove that a TC/T2T approach is significantly better than UC for the primary outcome, while many secondary efficacy outcomes favoured it, had a similar safety profile and was favourable from a societal health economic perspective at 1 year.
- This suggests that a TC/T2T approach might be beneficial in axSpA, but other strategy trials for this disease aiming to evaluate the efficacy of TC/T2T in this setting are needed to build on our findings.

INTRODUCTION

Disease activity in axial spondyloarthritis (axSpA) is assessed by measuring inflammation, with composite indices such as the completely patient-reported Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)¹ or the Ankylosing Spondylitis Disease Activity Score (ASDAS).² The ASDAS includes both patient-reported outcomes (PRO) and C-reactive protein (CRP), and is the preferred outcome for axSpA.³ Moreover, several ASDAS thresholds categorising disease activity have been proposed and validated: a score <2.1 means low disease activity and ≥ 2.1 means active disease.^{4,5}

The ASAS and European League Against Rheumatism (EULAR) societies have issued recommendations for the management of axSpA⁶ and indications for the pharmacological interventions depend on disease activity. The two major categories of pharmacological treatments are non-steroidal anti-inflammatory drugs (NSAIDs) and biologic disease-modifying antirheumatic drugs (bDMARDs), such as tumour necrosis factor inhibitors (TNFi) and IL-17 inhibitors (IL-17i). According to these latest recommendations,⁶ TNFi can be prescribed to patients with active disease (evaluated by ASDAS or BASDAI) despite previous exposure to at least two NSAIDs for at least 4 weeks in total (unless these drugs are contraindicated or cause side effects). Moreover, the presence of either objective signs of structural damage on pelvic radiography or inflammation (ie, elevated CRP or abnormal MRI showing subchondral bone oedema at the sacroiliac joint) is required. Finally, the rheumatologist should be convinced that in a particular patient there is a favourable benefit/risk profile. If the first TNFi fails, a switch to another TNFi or IL-17i should be considered.

According to the treat-to-target (T2T) concept, a precise and predefined determination about the target to be reached is defined before treatment starts; more importantly, the patient and treating physician decide in advance to intensify the treatment until the target is reached, unless contraindicated. The concept of tight control (TC) calls for rapid assessment of both efficacy and safety of a new treatment in a patient. For safety, the time frame can be very short, should an adverse event (AE) occur. As part of the TC, the efficacy of NSAIDs for axSpA should be evaluated after 4 weeks of treatment and of TNFi and IL-17i after 12–16 weeks.

In medicine, the combination of TC and T2T (a TC/T2T strategy) has demonstrated benefits in some areas, in particular for hypertension⁷ and diabetes.⁸ In rheumatology, this strategy has proven to be effective in rheumatoid arthritis⁹ and psoriatic arthritis,¹⁰ although these trials defined both the target and the outcome according to disease activity. In contrast with the previous TC/T2T trials in rheumatology, and more in line with what has been published in other disciplines, we decided to differentiate the target (ie, disease activity) and the outcome by choosing as the primary outcome the consequence of disease activity (ie, impact on functioning and health).

Conducting trials in which the new treatment algorithm is actually a complex strategy can be very challenging, as it may be difficult for the staff involved in the study not to apply their recently acquired experience with the new treatment algorithm into their usual standard of care. One approach to overcome this issue is to conduct a cluster-randomised trial. In these studies, not patients but centres/investigators are randomised; therefore, all individuals belonging to a certain centre are assigned to either the new treatment or the usual care (UC). This reduces the likelihood of 'contamination' of the UC treatment, since no centre provides both TC/T2T and UC.¹¹

No trial has yet evaluated the potential benefits of a TC/T2T strategy for patients with axSpA,¹² although experts have already recommended using this strategy in daily practice.¹³ The objective of this trial was thus to compare its potential benefits with UC in patients with axSpA.

METHODS

Study design

This was a pragmatic, prospective, parallel, cluster-randomised (with the centre as the cluster), open, controlled (two arms) trial (TICOSPA—NCT03043846). This study was not considered interventional, as the treatment was either a standard of care approach according to the treating rheumatologist (UC arm) or a TC/T2T algorithm strictly following the current international scientific recommendations for axSpA management. The study duration of 1 year was considered sufficient to demonstrate the benefits of a particular strategy on symptoms. The study was conducted in agreement with local good clinical practice (GCP) and the Declaration of Helsinki.

Participants

Centres/Clusters

The first step was to screen for centres in three European countries (France, the Netherlands and Belgium). Each centre had to be willing to apply the treatment to which they would be allocated and thus willing to follow the predefined treatment strategy if allocated to the TC/T2T arm. Thus, in order to minimise the chances of protocol violations in the TC/T2T arm, we selected centres with a particular interest in spondyloarthritis and that were interested in the potential efficacy of T2T and would be willing to follow the TC/T2T algorithm if randomised to this arm. Before randomisation, all selected centres signed a written agreement to adhere to their allocated strategy.

Patients

Patients had to be adults younger than 65 years, with a diagnosis of axSpA according to their rheumatologist and fulfilling the ASAS classification criteria for axSpA.¹⁴ At inclusion, disease had to be active (ie, ASDAS ≥ 2.1). In addition, patients should not have been optimally treated with NSAIDs (ie, they could not have received two full courses of NSAIDs at a daily full dose for at least 2 weeks each) and should not have contraindications to NSAIDs, be biologic-naïve and not have received apremilast in the past 3 months. They also had to have a pelvic radiography and MRI of the sacroiliac joints available, as well as their HLAB27 status. All patients needed to understand the study objectives and to complete questionnaires. They also had to provide written consent.

Treatments

TC/T2T arm

Visits were scheduled every 4 weeks, and the strategy was prespecified, based on the current recommendations for axSpA management,⁶ compiled in an electronic algorithm that guided treatment decisions at each visit, after collection and entry of the ASDAS in the electronic case report form (CRF). The target was an ASDAS <2.1 (ie, low disease activity⁴); although remission has been proposed as the preferred target in recommendations, low disease activity has been proposed as an alternative: an ASDAS <2.1 was selected by the steering committee as <1.3 was considered too stringent. If the target was not met, intensification of treatment was proposed to the investigator until the target was met. In all cases, but in particular with regard to bDMARD

prescriptions, the recommendation had to be consistent with the product labelling. If the recommendation was to initiate a TNFi, the investigator chose which TNFi to use, considering the country's reimbursement criteria. For the specific case of bDMARDs, safety was evaluated every 4 weeks, but, according to recommendations, efficacy (ie, whether ASDAS target was met) was evaluated at the visit occurring after at least 12 weeks after initiation. When the target was met, treatment continuation was recommended; if the patient was considered to have inactive disease (ie, ASDAS <1.3), acceptance of any recommendation to taper NSAIDs was based on a shared decision with the patient. The full algorithm providing the predetermined recommendations at each visit is available in online supplemental file 1 and also in the Protocol.

Usual care

Visits to assess study outcomes were scheduled every 12 weeks, and all treatment decisions, including frequency of follow-up, were left to the investigator's discretion.

Outcomes

All outcomes were assessed at the patient level.

Efficacy outcomes

Primary outcome: The validated ASAS-Health Index (ASAS-HI) is a 17-item questionnaire (range 0–17, with 17 representing the worst health) that represents the most important items of the ASAS Core Set of the International Classification of Functioning as well as five patient-reported outcome (PRO) items.¹⁵ The primary outcome of the trial was the percentage of patients achieving an improvement of at least 30% in the ASAS-HI at the 1 year visit.

Secondary efficacy outcomes can be categorised in two groups: (1) Disease activity outcomes: ASDAS over time, ASDAS states (ie, low disease activity (LDA) and inactive disease (ID)) and changes (ie, major improvement (MI) and clinically important improvement (CII)),^{4–5} ASAS responses (ASAS20, ASAS40),¹⁶ ASAS partial remission, BASDAI, BASDAI 50,¹⁷ Physician Global Assessment and CRP; (2) other PRO measures (Bath Ankylosing Score Global (BASG),¹⁸ Patient' global assessment of disease activity, Visual Analogic Scale (VAS) of Fatigue), functioning and health (ASAS-HI as a continuous score, BASFI¹⁹ and EQ-5D-5L²⁰), work impairment (Work Productivity and Activity Impairment Questionnaire, WPAI),²¹ and treatment (ASAS-NSAID score,²² which aims to quantify the NSAID intake in a defined period of time, and bDMARD initiations).

Safety outcomes

All adverse events (AEs) that were cardiovascular, gastrointestinal or related to infections or allergies during patients' participation in the study were collected and recorded in the case report form, in accordance with EULAR recommendations.²³

Resource utilisation and costs (cost–utility analysis)

At each visit, cost questionnaires were added, assessing health-care resource utilisation (rheumatologist and other specialist visits, nurse, physiotherapist, and emergency department visits, and days in rehabilitation centres and hospitals) and work days (of paid work) missed in the past 12 weeks. Costs per healthcare resource category over the 48-week follow-up were calculated by multiplying the resource used by the Dutch unit costs, taken from the Dutch guideline for economic evaluations, and expressed in 2019 Euros²⁴ (online supplemental file 4). Productivity losses

were valued by using the human capital method in the base-case analysis.

No structural-progression outcomes were evaluated in this study, as the 1-year period was deemed too short to observe any effect.

Sample size estimation

Sample size was calculated in two consecutive steps.

First step: The cluster design was not considered (ie, this calculation used a conventional approach, considering that randomisation would occur at the patient level). It was anticipated that in the UC arm, 25% of the patients would meet the definition of response ($\geq 30\%$ improvement in the ASAS-HI score after 1 year of follow-up). For an α risk of 5% and a β risk of 80%, with a bilateral test, 77 patients per arm were needed to demonstrate a 20-percentage-point difference among responders, that is, at least a 45% responder rate in the TC/T2T arm if the UC responder rate was 25%. For a uniform number of patients per centre, we aimed for a total sample of 160 patients (80 per arm).

Second step: The cluster-randomised design was considered by multiplying the estimated sample needed (ie, 160 patients) by an 'inflation factor' defined as $1 + (m - 1) * \rho$, where m is equal to the size of the cluster (in our study, the number of patients per centre = 10) and ρ is the intracluster correlation, usually set at 0.05.²⁵ The inflation factor for our study was 1.45, and thus the sample needed to take the cluster-randomised design into account was 232 patients (116 patients per arm).

Randomisation and blinding

The 1:1 randomisation was cluster-based: a random number table was used to randomise centres instead of individuals. Centres allocated to the UC arm were blinded for the specific TC/T2T strategy (including treatment and frequency of assessments) in this trial until the end of the study. Separate investigator meetings on different days were organised for each arm.

This was an open study in the sense that centres were not blinded to the study arm to which they were allocated. Within each centre, consecutive patients meeting the inclusion criteria were screened and invited to participate: patients were not blinded but received different information letters, depending on the allocation of the centre in which they were included (one explaining the TC/T2T strategy and visit schedules, and the other the UC visit schedule).

Statistical methods

All statistical analyses were performed with the open-source software R, V.3.5.2,²⁶ and Stata SE release V.14.0.

The analysis population was the intention-to-treat population.

Efficacy analysis

A cluster-randomisation design must consider two main limitations. The first is that observations of groups of individuals from the same cluster have a lower variance. To take this into account, all the efficacy analyses first used a two-level mixed model with two random effects to estimate the percentage of responders or the change in outcome over the follow-up (ie, a cluster-adjusted model). The second limitation is that in this type of trial, randomisation occurs at the cluster level, while outcomes are assessed at the individual (patient) level. Therefore, to compensate for the potential imbalance of some covariates across treatment groups, we identified those that were not balanced and included them as adjustment factors in a second model otherwise identical to the first (ie, cluster and imbalance-adjusted model, presented

in detail in online supplemental file 3). The use of mixed-effects models enabled us to deal with the missing data by using maximum likelihood.²⁷

Safety analysis

The number and types of AEs were described both globally and by study arm, as recommended by EULAR²⁸ for cardiovascular, infection-related and gastrointestinal events, as well as allergy (skin reactions and anaphylaxis).

Cost–utility analysis

A health economic evaluation from the perspectives of healthcare and of society (including also costs related to sick leave) was performed. In the base case, cumulative costs and time-averaged health utilities were calculated for each patient after multiple imputation of missing cost categories and EuroQol 5 domains and 5 levels (EQ-5D). Next, the incremental costs (iCosts) and effects (quality-adjusted life-years, iQALY) were used to calculate the incremental cost per QALY gained (incremental cost–utility ratio or ICUR; $ICUR = iCosts/iQALY$) and incremental net monetary benefit (iNMB; $iNMB = iQALY * [willingness-to-pay\ threshold] - iCosts$). To account for baseline differences in costs and QALYs between the study arms, these incremental costs and QALYs were baseline-adjusted by using seemingly unrelated regressions (SUREG). With SUREG, costs and effects can be modelled jointly within the same model, thereby taking their (expected) correlation into account.^{29 30} For this analysis, SUREG was used with a robust variance estimator to reflect the clustering of the data. Finally, the conventional non-parametric bootstrapping method was considered inappropriate for estimating the CIs of the ICUR and cost-effectiveness acceptability curves, because it ignores the clustering of patients inherent to a cluster-randomised clinical trials (RCT). Instead, two-stage bootstrapping, which resamples clusters (first stage) and individuals (second stage), was performed.^{31 32} As the cluster sizes were not balanced, a modified version³³ that allows for this imbalance was used to resample the current study population 5000 times. In addition to the base-case analysis, several sensitivity analyses were conducted. These assessed the impact of reducing bDMARD costs, a different valuation method for productivity losses (the friction cost approach, which considered only productivity losses during the 13-week friction period), the inclusion of costs of presenteeism (at-work productivity loss) and utility based on ASAS-HI (instead of EQ-5D).³⁴ For all analyses, both the societal perspective (including all costs) and the healthcare perspective (including only healthcare costs and excluding productivity losses) are presented. The willingness-to-pay threshold used to interpret the ICUR and to calculate the iNMB was set at €20 000/QALY, based on the level of health reported by the subjects in the current study and recommended Dutch thresholds.³⁵

RESULTS

Participant flow

Eighteen centres (4 in Belgium, 10 in France and 4 in the Netherlands), all of them rheumatology departments with an expertise in SpA, were invited and agreed to participate. All centres signed an agreement to follow the treatment to which they were allocated. After randomisation, nine centres were allocated to the TC/T2T arm and nine to the UC arm.

During the recruitment period (from February 2017 to June 2019; dataset locked on August 2019), centres from the TC/T2T arm screened and included 80 patients, while centres from

the UC arm screened 83 patients and included 80. The analysis population comprised 160 patients. After 1 year, one centre (UC arm) had not included any patient. At the patient level, in the TC/T2T arm, seven patients were lost to follow-up and one refused to continue the study; in the UC arm, three patients were lost to follow-up, four refused to continue and one was excluded because of a non-spondyloarthritis diagnosis during follow-up. In total, 144 patients (72 per arm) completed the last visit (figure 1 presents the study flow diagram, and online supplemental file 2 presents the flow diagram of the study at the cluster level).

Baseline data

The patients' mean age (SD) was 37.9 (11.0) years with a mean disease duration of 3.7 (6.2) years; 51.2% were men. Radiographic damage of the sacroiliac joints was found in 46.9% of patients, 81.9% had had MRI-positive sacroiliitis at least once, and 75.0% were HLA-B27+. Mean ASDAS at inclusion was 3.0 (0.7) and mean ASAS-HI was 8.6 (3.7). The study design resulted in a significant baseline imbalance of some variables at the patient level: patients from the TC/T2T arm had a university education, a history of anterior uveitis and a higher Physician Global Assessment significantly more often; they had a history of gastrointestinal events related to NSAID intake less often and fewer mean days in rehabilitation facilities in the 3 months before the study (see online supplemental file 3). These variables were therefore included as adjustment variables in the 'cluster and imbalance-adjusted model'. Baseline characteristics are presented in table 1.

Efficacy outcomes

The estimated percentages of patients achieving an improvement $\geq 30\%$ on the ASAS-HI at the 1 year visit with the 'cluster and imbalance-adjusted model' was 47.3% in the TC/T2T arm and 36.1% in the UC arm. The estimated difference (11.2%, 95% CI 8.5% to 13.9%) was not statistically significant in either model ($p=0.094$ and $p=0.079$ for the 'cluster-adjusted' and 'cluster and imbalance-adjusted' models, respectively). Some estimated efficacy outcomes were significantly favouring the TC/T2T arm, for example, 76.5% vs 59.5% ($p<0.05$ with both models) for ASDAS low disease activity and 52.3% vs 34.7% ($p<0.05$ with both models) for the ASAS40 at 1 year (figure 2), but the majority did not reach a statistical significance (table 2). The ASAS-NSAID score was not significantly different between treatment arms. The prescription rate for biologics was initially lower in the TC/T2T arm, but was quickly (ie, within 30 days) significantly higher among these patients (56.2% vs 27.2%, $p<0.01$; online supplemental file 5).

Safety

Overall, 55 AEs were reported, 33 in the TC/T2T and 22 in the UC arm. The TC/T2T arm had more allergies (8 vs 1) mainly because of local skin reactions at the injection site. Both groups had a similar number of infections (15 vs 16 in the TC/T2T and UC arms, respectively), with 2 of them severe, both in the UC arm. Table 3 summarises all the AEs.

Cost–utility analysis

During this 48-week follow-up, there were more visits to rheumatologists and more days of bDMARDs used, but fewer visits to physical therapy, fewer days of rehabilitation care and especially fewer days of sick leave in the TC/T2T arm (online supplemental file 4).

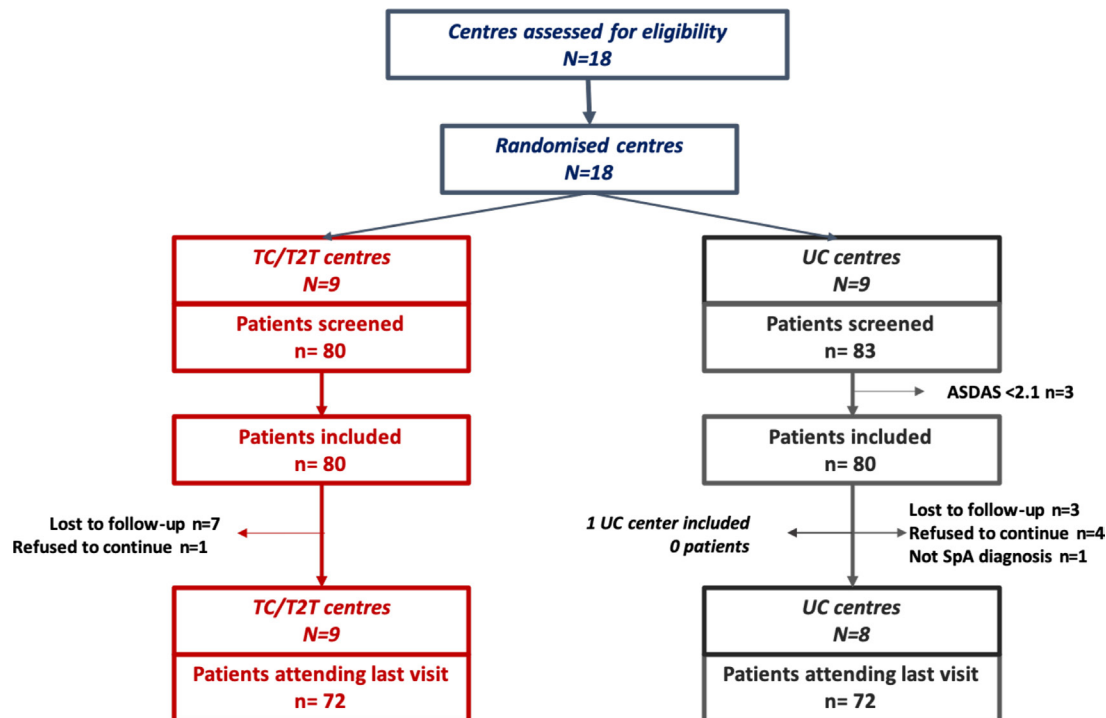


Figure 1 Flow chart of the study. ASDAS, Ankylosing Spondylitis Disease Activity Score; SpA, spondyloarthritis; TC/T2T, tight control and treat-to-target; UC, usual care.

In the unadjusted base-case analysis, TC/T2T dominated UC (iQALY 0.07; iCosts −€1472) from a societal perspective and was cost-effective with an ICUR of €11 538 (iQALY 0.07; iCosts €801) from a healthcare perspective. The baseline-adjusted and cluster-adjusted base-case analysis produced similar results, with TC/T2T dominating UC and being cost-effective with an ICUR

of €19 430 from the societal and healthcare perspective, respectively. In bootstrapped adjusted analyses, the probability that TC/T2T would be cost-effective compared with UC was 72% from the societal perspective and 52% from the healthcare perspective, given a threshold for willingness-to-pay of €20 000/QALY.

In the sensitivity analyses, lower bDMARD costs favoured TC/T2T even more strongly, as expected. The estimate of productivity costs based on the friction cost approach (ie, considering only the first 13 weeks of absence) still favoured TC/T2T, as did inclusion of the presenteeism costs (ie, productivity loss while at work), both somewhat less strongly than the base case. Finally, using disease-specific utilities based on the ASAS-HI instead of the EQ-5D to calculate QALYs resulted in less favourable results for TC/T2T due to the smaller gain in QALYs (0.015–0.018) for TC/T2T relative to UC (table 4 and online supplemental file 4).

Table 1 Baseline characteristics of patients included in the two treatment arms

	TC/T2T (n=80)	UC (n=80)	Total (n=160)
Age (years)*	37.6 (10.8)	38.1 (11.4)	37.9 (11.0)
Sex (Male)	45 (56.2%)	37 (46.2%)	82 (51.2%)
Smoking status (current)	29 (36.2%)	32 (40.0%)	61 (38.1%)
University studies †	57 (71.2%)	42 (52.2%)	99 (61.9%)
Disease duration (years)	4.2 (6.56)	3.3 (5.83)	3.7 (6.20)
Radiographic sacroiliitis	42 (52.5%)	33 (41.2%)	75 (46.9%)
MRI sacroiliitis	63 (78.8%)	68 (85.0%)	131 (81.9%)
HLA B27 positive	62 (77.5%)	58 (72.5%)	120 (75.0%)
History of anterior uveitis †	16 (20.0%)	6 (7.5%)	22 (13.8%)
Good NSAID response	63 (78.8%)	55 (68.8%)	118 (73.8%)
ASDAS	3.0 (0.7)	3.0 (0.6)	3.0 (0.7)
BASDAI (n=80/79)	5.2 (1.7)	5.2 (1.9)	5.2 (1.8)
CRP (mg/L)	9.2 (9.8)	7.4 (10.5)	8.3 (10.1)
ASAS-Health Index (0–17) (n=80/78)	8.2 (3.8)	9.0 (3.6)	8.6 (3.70)
NSAID score (last 3 months)	32.7 (35.4)	37.5 (36.6)	35.1 (36.0)

*Numerical variables are presented as means (SD) and categorical variables as n (percentage).

†Significantly imbalanced variables; other imbalanced variables at baseline were gastrointestinal events in relation to NSAIDs, Physician Global and Rehabilitation days (see online supplemental file 3).

ASDAS, Ankylosing Spondylitis Disease Activity Score; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; CRP, C-reactive protein; NSAID, non-steroidal anti-inflammatory drugs.

DISCUSSION

This is the first strategy trial evaluating the potential benefits of a TC/T2T strategy in patients with axSpA. This was a negative trial, in which the main outcome did not reach a statistically significant difference in both groups. Safety was similar in both arms, even though biologics were prescribed significantly more often in the TC/T2T arm and produced more local injection site allergic reactions in this arm. The health economic analysis also favoured the TC/T2T arm from the societal perspective and (although with a lower level of certainty) from the healthcare perspective as well.

Nevertheless, in almost all of the estimated efficacy outcomes scores and responses were in favour of the TC/T2T, even if most did not reach a statistical significance. Therefore, it is important to analyse the reasons behind the lack of a statistically significant difference between the groups for the main outcome measure. While some would consider this a negative trial and would argue that no further efficacy endpoints should be interpreted, this approach has been considered by some as overly simplistic in

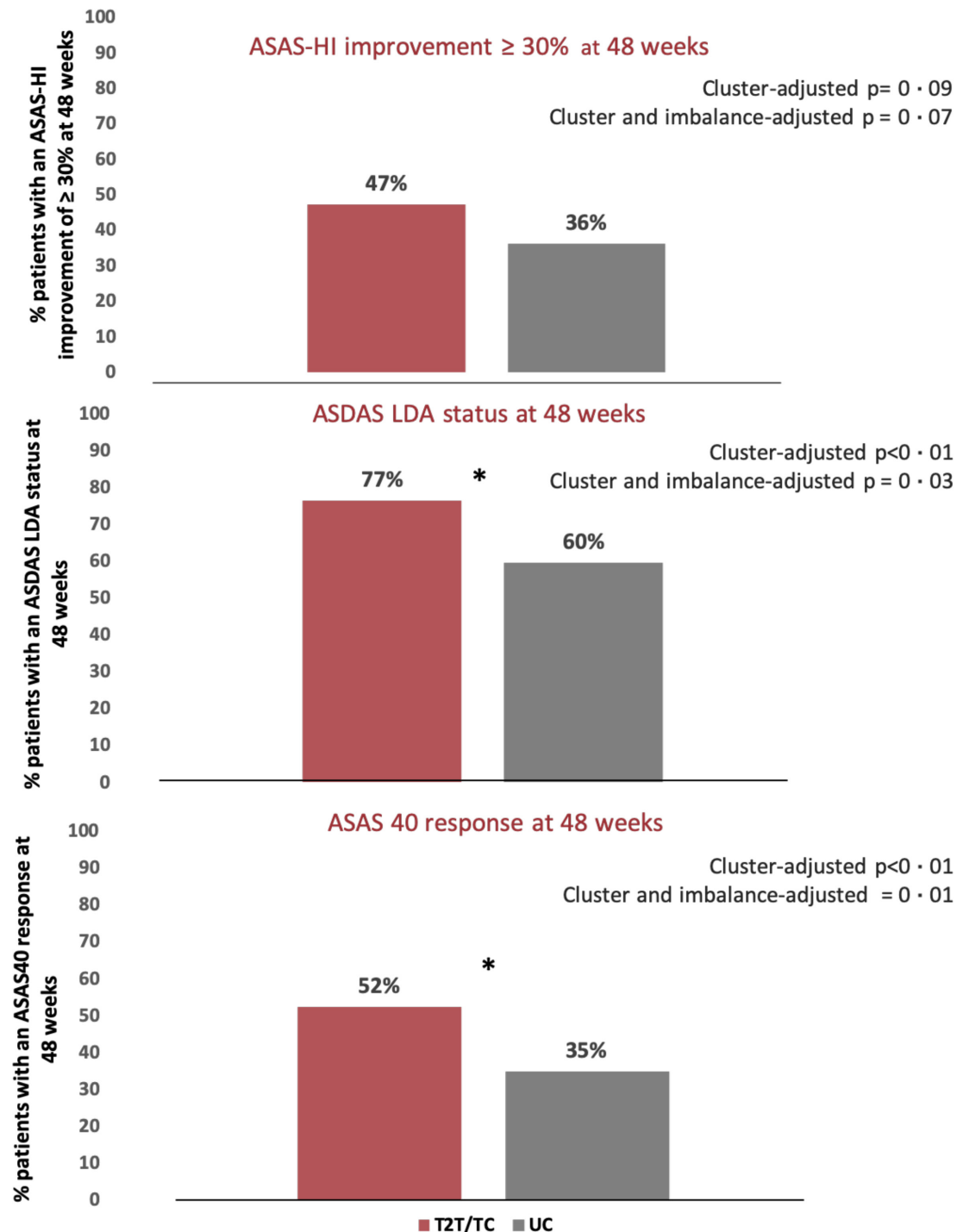


Figure 2 ASAS-HI improvement $\geq 30\%$, ASDAS LDA status and ASAS40 response estimated at 48 weeks. *Statistical significance. ASAS-HI, Axial Spondyloarthritis International Society-Health Index; ASDAS, Ankylosing Spondylitis Disease Activity Score; LDA, low disease activity; T2T/TC, treat-to-target and tight control; UC, usual care.

the literature,³⁶ and in some cases, the presentation of all results is needed to guide appropriate clinical interpretation of the findings.

The first reason behind our not statistically significant results may be related to the choice of the outcome measure: in most strategy trials the target is usually defined by a relevant threshold in a factor that predisposes patients to that outcome, below which the risk of developing this outcome is abolished or significantly reduced. A good example of this can be found in a TC/T2T strategy in diabetes that aimed for a target of

glycosylated haemoglobin below 7% and used a decreased rate of diabetic retinopathy as the primary objective.¹¹ Interestingly, in the two published T2T trials in rheumatology, the TICORA⁹ and TICOPA¹⁰ trials, both the outcome and the target were 'disease activity-related', that is, Disease Activity Score 28 (DAS28) and EULAR response in TICORA and minimal disease activity and American College of Rheumatology (ACR) 20 in TICOPA. To avoid the potential circularity, in TICOSPA the target was disease activity but the main outcome was a consequence of disease activity, that is, disease-specific functioning

Table 2 Estimated efficacy outcomes at the last study visit (week 48)

	Estimated outcomes at week 48		Cluster-adjusted model	Cluster and imbalance-adjusted model
	TC/T2T	UC		
ASAS-HI significant improvement	47.3%	36.1%	NS	NS
ASDAS LDA	76.5%	59.5%	<0.01	0.03
ASDAS ID	25.9%	18.7%	NS	NS
ASDAS CII	61.2%	46.0%	<0.01	0.02
ASDAS MI	16.5%	14.9%	NS	NS
ASAS40	52.3%	34.7%	<0.01	0.01
ASAS20	94.9%	85.9%	<0.01	0.03
BASDAI 50	79.0%	43.8%	0.01	0.03
Physician Global (0–10)	2.0 (0.2)	1.8 (0.2)	NS	NS
CRP (mg/L)	3.9 (1.4)	3.5 (1.5)	NS	NS
BASG (0–10)	2.6 (0.5)	3.4 (0.5)	NS	NS
BASFI (0–10)	1.7 (0.5)	2.4 (0.5)	NS	NS
EQ5D-5L	0.7 (0.1)	0.8 (0.1)	0.02	NS
ASAS-NSAID score	1.5 (2.2)	−4.9 (2.9)	NS	NS

ASAS-HI, ASAS-Health Index; ASDAS, Ankylosing Spondylitis Disease Activity Score; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; CII, clinically important improvement; EQ5D, EuroQol 5 domains and 5 levels; ID, inactive disease; LDA, low disease activity; MI, major improvement; NS, non-significant.

and health, measured by ASAS-HI. Furthermore, others may consider that our outcome was not ambitious enough and we should have aimed at reducing structural damage, like retinopathy for diabetes⁸ or cardiovascular events (such as myocardial infarction or stroke) for hypertension.⁷ However, keeping in mind the lower prevalence of axSpA (compared with diabetes and hypertension) and the slow progression rate in axSpA nowadays, it would have been very difficult to run such TC/T2T trial in axSpA.

Another reason behind this lack of statistically significant difference is the unexpected good results of the UC arm: in our initial hypothesis, we expected a 25% response in the UC group, while UC presented a 36% response rate. This can be explained by the open cluster-trial design, as only SpA expert centres, aware of the need for such a trial and of the TC/T2T recommendations could be selected to participate, and were thus probably applying recommendations in their ‘usual care’. Also, we calculated the expected treatment effect as if the study was a trial against placebo (in which a 20% difference is usually used). Here, patients from the control group were not receiving a placebo. In the rheumatology literature, a 12% difference in response between two active-treatment groups has been set as the difference to be observed in non-inferiority head-to-head trials.^{37–39}

Table 3 Adverse events observed during the trial

	Total (n=160)	T2T (n=80)	UC (n=80)
Total number of adverse events	55	33	22
Cardiovascular events	0	0	0
Allergies	9	8	1
Skin, local reaction	7	6	1
Anaphylactic reaction	2	2	0
Infection	31	15	16
Opportunistic infections	2	0	2
Viral	22	12	10
Severe infections	2	0	2
Gastrointestinal events	12	7	5

UC, usual care.

Alternatively, there is also the possibility that this trend observed across efficacy outcomes in favour of TC/T2T is not due to a true-positive effect but is either non-clinically relevant or only reflecting a higher placebo effect on the TC/T2T arm. Indeed, patients and centres were not blinded to the arm of treatment they were allocated to: patients from the TC/T2T arm received an information sheet in which it was stated that they would be receiving the ‘state-of-art’ of treatment, strictly following current recommendations; as all efficacy outcomes are patient-reported, it is also possible that the observed and estimated differences in favour of TC/T2T are only reflecting a placebo effect.

Nevertheless, in an era of cost-containment, the cost-utility analysis favoured the TC/T2T strategy, especially from the societal perspective. Results of the economic evaluation were different from those observed in the TICOPA¹⁰ trial, where the active arm was significantly more expensive than the UC arm, despite iQALYs similar to those in TICOSPA. This might be explained by the partial offset in TICOSPA of the extra cost of bDMARDs and visits by the substantially lower costs due to reduced numbers of days of sick leave and of visits for physiotherapy and rehabilitation facilities. The costs per unit of the different health resources, such as visits to the rheumatologist, varied according to the study setting: the UK for TICOPA, compared with Dutch true costs, adjusted for purchasing power parities for Belgium and France, for TICOSPA. This might have driven differences between the studies and limited comparisons. Furthermore, with the arrival of biosimilars, the costs of bDMARDs decreased substantially compared with TICOPA. Moreover, patients from the UC arm used rehabilitation and physical therapy more often at baseline, and this trend continued during the study. Our cluster-randomisation design of the trial may explain this difference: different habits in different centres may have led to differences in costs.

Overall, in this setting of expert centres in SpA, UC resulted in very good outcomes for a substantial number of patients. Although the primary outcome measure was not achieved, with response rates between the two treatment groups differed by 11% in favour of TC/T2T. Despite the higher prescription rate of biologics in the TC/T2T arm, safety profiles were similar across arms, and the TC/T2T arm had a favourable outcome from a

Table 4 Cost–utility analyses

	Perspective	Adjustment	iCosts	iQALY	iCUR	iNMB*	p20†	p50†
Base-case analyses								
Unadjusted	Societal	None	−€1472	0.069	Dominates UC	€2860	0.85	0.92
Adjusted	Healthcare	None	€801	0.069	€11 538/QALY (NE)	€587	0.69	0.91
	Societal	Baseline costs/effects	−€472	0.041	Dominates UC	€1295	0.72	0.83
	Healthcare	Baseline costs/effects	€789	0.041	€19 430/QALY (NE)	€23	0.52	0.78
Sensitivity analyses								
bDMARD discount (25%)	Societal	Baseline costs/effects	−€856	0.041	Dominates UC	€1678	0.78	0.86
	Healthcare	Baseline costs/effects	€406	0.041	€9888/QALY (NE)	€415	0.68	0.86
bDMARD discount (50%)	Societal	Baseline costs/effects	−€1239	0.041	Dominates UC	€2060	0.82	0.88
	Healthcare	Baseline costs/effects	€23	0.042	€553/QALY (NE)	€809	0.86	0.91
ASAS-HI utility	Societal	Baseline costs/effects	−€466	0.018	Dominates UC	€817	0.64	0.67
	Healthcare	Baseline costs/effects	€793	0.015	€51 938/QALY (NE)	−€487	0.29	0.48
Friction cost approach‡	Societal	Baseline costs/effects	€183	0.042	€4400/QALY (NE)	€648	0.63	0.79
Costs of presenteeism included	Societal	Baseline costs/effects	−€876	0.040	Dominates UC	€1679	0.67	0.76

*For lambda (willingness-to-pay threshold)=€20 000/QALY.

†Probability that TC/T2T is cost-effective for willingness-to-pay thresholds (lambda) of €20 000/QALY or €50 000/QALY (p20 or p50, respectively).

‡For the friction cost approach, absenteeism (sick leave) that lasts longer than the friction period (13 weeks) is not included in the costs.

iCosts, incremental costs; iQALY, incremental QALY; NE, northeast (position of intervention in cost-effectiveness plane, comparator at centre [0,0]).

societal health economic perspective. This suggests that a TC/T2T approach might be beneficial in axSpA, but other strategy trials for this disease aiming to evaluate the efficacy of TC/T2T in this setting are needed to build on our findings.

Author affiliations

¹Rheumatology Department, Hospital Cochin, Paris, France

²ECAMO team, INSERM U1153, Paris, France

³Rheumatology Department, Reina Sofia University Hospital, Cordoba, Andalucía, Spain

⁴Rheumatology Department, University Hospital Ghent, Gent, Oost-Vlaanderen, Belgium

⁵Department of Internal Medicine, Division of Rheumatology, Maastricht University, Maastricht, The Netherlands

⁶Care and Public Health Research Institute (CAPHRI), Maastricht University, Maastricht, The Netherlands

⁷Rheumatology Department, Le Mans Hospital, Le Mans, France

⁸Rheumatology, Leiden University Medical Center, Leiden, The Netherlands

⁹Rheumatology, CHU, Clermont-Ferrand, France

¹⁰Rheumatology, Henri Mondor Hospital Rheumatology Service, Creteil, France

¹¹EpiDermE, Université Paris Est Créteil, Créteil, France

¹²Rheumatology, Hôpital Sud, Echirolles, France

¹³GREPI EA7408, Université Grenoble Alpes, Saint-Martin-d'Herès, France

¹⁴Rheumatology, Zuyderland Medical Centre Heerlen, Heerlen, The Netherlands

¹⁵Rheumatology & Clinical Immunology, University Medical Centre Groningen, Groningen, The Netherlands

¹⁶Rheumatology, VIB-UGent Center for Inflammation Research, Zwijnaarde, Belgium

¹⁷Rheumatology, University Hospital Ghent, Gent, Belgium

¹⁸Rheumatology, ZNA UKJA, Antwerpen, Belgium

¹⁹Rheumatology, Reumainstituut Hasselt, Hasselt, Belgium

²⁰Rheumatology Department, University Hospital Pitié Salpêtrière, Paris, France

²¹Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, Paris, France

²²Rheumatology, Bois-Guillaume CHU, Rouen, France

²³Rheumatology, University Hospital Centre Toulouse, Toulouse, France

²⁴Paul Sabatier University, Toulouse, France

²⁵Rheumatology department, Hospital Centre Mulhouse, Mulhouse, France

²⁶Department of Medicine, Division of Rheumatology, Maastricht University Medical Center, Maastricht, The Netherlands

²⁷Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands

Twitter Anna Molto @annamolto and Philippe Carron @PhilippeCarron

Acknowledgements The authors would like to thank PNSResearch for organising the logistics of implementing the study in the three countries.

Contributors All authors contributed to the manuscript as follows: AM, CL-M, MD, FEVdB, DvdH AB: conception, design and interpretation of data. AM, CL-M, CW: analysis of the data. AM drafted the manuscript. All authors critically reviewed the

results and contributed to the interpretation of data. All authors read and approved the final manuscript, as well as the answer to the reviewers' comments.

Funding UCB provided an unrestricted grant for this study, but had no role in providing any drug, or in the study design, data collection, data analysis, data interpretation or writing of the report.

Competing interests Dr van Tubergen reports grants and personal fees from Novartis, grants from UCB, grants from Biogen, grants from AbbVie, outside the submitted work. Dr van der Heijde reports personal fees from AbbVie, Amgen, Astellas, AstraZeneca, BMS, Boehringer Ingelheim, Celgene, Cytone, Daiichi, Eisai, Eli-Lilly, Galapagos, Gilead, Glaxo-Smith-Kline, Janssen, Merck, Novartis, Pfizer, Regeneron, Roche, Sanofi, Takeda, UCB Pharma, outside the submitted work; and Director of Imaging Rheumatology bv. Dr van Gaalen reports grants from Stichting vrienden van Sole Mio, grants from Stichting ASAS, grants and personal fees from Novartis, grants from UCB, personal fees from MSD, personal fees from AbbVie, personal fees from Bristol Myers Squibb, outside the submitted work. AB received a research grant to her department from AbbVie, consultation fees from Eli Lilly and Galapagos and a speakers fee from UCB, all paid to her department. Dr Van den Bosch reports personal fees from AbbVie, personal fees from Celgene, personal fees from Eli Lilly, personal fees from Galapagos, personal fees from Janssen, personal fees from Novartis, personal fees from Pfizer, personal fees from UCB, outside the submitted work. Dr Claudepierre reports personal fees from Roche Chugai, Novartis, Pfizer, MSD, grants from Roche Chugai, Novartis, Pfizer, UCB, MSD, AbbVie, Lilly, Celgene, Janssen, BMS, outside the submitted work. Dr Molto reports grants from UCB during the conduct of the study; personal fees from AbbVie, grants and personal fees from UCB, personal fees from BMS, grants and personal fees from Pfizer, grants and personal fees from MSD, personal fees from Novartis, personal fees from Gilead, personal fees from Lilly, outside the submitted work. Dr Gossec reports grants from Amgen, Lilly, Janssen, Pfizer, Sandoz, Sanofi, Galapagos, personal fees from AbbVie, Amgen, BMS, Biogen, Celgene, Gilead, Janssen, Lilly, Novartis, Pfizer, Samsung Bioepis, Sanofi-Aventis, UCB, outside the submitted work. Dr Dougados reports grants from UCB, during the conduct of the study; grants and personal fees from AbbVie, grants and personal fees from Pfizer, grants and personal fees from Lilly, grants and personal fees from Novartis, grants and personal fees from Roche, grants and personal fees from BMS, grants and personal fees from Merck, outside the submitted work. Dr Derris, Dr Ruysen-Witrand, Dr Lenaerts, Dr Lopez-Medina, Dr Sparsa, Dr Starmans-Kool, Dr C Webers, Dr Pouplin, Dr Soubrier and Dr Joos have nothing to declare.

Patient consent for publication Not required.

Ethics approval The study was approved by the ethics committee CPP Ile de France III (Réf. Am7156-1-S.C.3394). All patients gave their informed consent.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are

solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Anna Molto <http://orcid.org/0000-0003-2246-1986>
 Clementina López-Medina <http://orcid.org/0000-0002-2309-5837>
 Filip E Van den Bosch <http://orcid.org/0000-0002-3561-5932>
 Annelies Boonen <http://orcid.org/0000-0003-0682-9533>
 Athan Baillet <http://orcid.org/0000-0001-6006-2519>
 Peggy Jacques <http://orcid.org/0000-0001-9227-257X>
 Philippe Carron <http://orcid.org/0000-0001-9254-6171>
 Laure Gossec <http://orcid.org/0000-0002-4528-310X>
 Astrid van Tubergen <http://orcid.org/0000-0001-8477-0683>
 Désirée van der Heijde <http://orcid.org/0000-0002-5781-158X>

REFERENCES

- Garrett S, Jenkinson T, Kennedy LG, *et al*. A new approach to defining disease status in ankylosing spondylitis: the Bath ankylosing spondylitis disease activity index. *J Rheumatol* 1994;21:2286–91.
- Lukas C, Landewé R, Sieper J, *et al*. Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. *Ann Rheum Dis* 2009;68:18–24.
- van der Heijde D, Braun J, Dougados M, *et al*. Sensitivity and discriminatory ability of the ankylosing spondylitis disease activity score in patients treated with etanercept or sulphasalazine in the ASCEND trial. *Rheumatology* 2012;51:1894–905.
- Machado PM, Landewé R, Heijde D, *et al*. Ankylosing spondylitis disease activity score (ASDAS): 2018 update of the nomenclature for disease activity states. *Ann Rheum Dis* 2018;77:1539–40.
- Machado P, Landewé R, Lie E, *et al*. Ankylosing spondylitis disease activity score (ASDAS): defining cut-off values for disease activity states and improvement scores. *Ann Rheum Dis* 2011;70:47–53.
- van der Heijde D, Ramiro S, Landewé R, *et al*. 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. *Ann Rheum Dis* 2017;76:978–91.
- SPRINT Research Group, Wright JT, Williamson JD, *et al*. A randomized trial of intensive versus standard blood-pressure control. *N Engl J Med* 2015;373:2103–16.
- ACCORD Study Group, ACCORD Eye Study Group, Chew EY, *et al*. Effects of medical therapies on retinopathy progression in type 2 diabetes. *N Engl J Med* 2010;363:233–44.
- Grigor C, Capell H, Stirling A, *et al*. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. *Lancet* 2004;364:263–9.
- Coates LC, Moverley AR, McParland L, *et al*. Effect of tight control of inflammation in early psoriatic arthritis (TICOPA): a UK multicentre, open-label, randomised controlled trial. *Lancet* 2015;386:2489–98.
- Meurer WJ, Lewis RJ. Cluster randomized trials: evaluating treatments applied to groups. *JAMA* 2015;313:2068–9.
- Schoels MM, Braun J, Dougados M, *et al*. Treating axial and peripheral spondyloarthritis, including psoriatic arthritis, to target: results of a systematic literature search to support an international treat-to-target recommendation in spondyloarthritis. *Ann Rheum Dis* 2014;73:238–42.
- Smolen JS, Braun J, Dougados M, *et al*. Treating spondyloarthritis, including ankylosing spondylitis and psoriatic arthritis, to target: recommendations of an international Task force. *Ann Rheum Dis* 2014;73:6–16.
- Rudwaleit M, van der Heijde D, Landewé R, *et al*. The development of assessment of spondyloarthritis International Society classification criteria for axial spondyloarthritis (Part II): validation and final selection. *Ann Rheum Dis* 2009;68:777–83.
- Kiltz U, van der Heijde D, Boonen A, *et al*. Development of a health index in patients with ankylosing spondylitis (ASAS HI): final result of a global initiative based on the ICF guided by ASAS. *Ann Rheum Dis* 2015;74:830–5.
- Brandt J, Listing J, Sieper J, *et al*. Development and preselection of criteria for short term improvement after anti-TNF alpha treatment in ankylosing spondylitis. *Ann Rheum Dis* 2004;63:1438–44.
- Rudwaleit M, Listing J, Brandt J, *et al*. Prediction of a major clinical response (BASDAI 50) to tumour necrosis factor alpha blockers in ankylosing spondylitis. *Ann Rheum Dis* 2004;63:665–70.
- Jones SD, Steiner A, Garrett SL, *et al*. The Bath ankylosing spondylitis patient global score (BAS-G). *Br J Rheumatol* 1996;35:66–71.
- Calin A, Garrett S, Whitelock H, *et al*. A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath ankylosing spondylitis functional index. *J Rheumatol* 1994;21:2281–5.
- Tsang HHL, Cheung JPY, Wong CKH, *et al*. Psychometric validation of the EuroQoL 5-dimension (EQ-5D) questionnaire in patients with spondyloarthritis. *Arthritis Res Ther* 2019;21:41.
- Reilly MC, Gooch KL, Wong RL, *et al*. Validity, reliability and responsiveness of the work productivity and activity impairment questionnaire in ankylosing spondylitis. *Rheumatology* 2010;49:812–9.
- Dougados M, Simon P, Braun J, *et al*. ASAS recommendations for collecting, analysing and reporting NSAID intake in clinical trials/epidemiological studies in axial spondyloarthritis. *Ann Rheum Dis* 2011;70:249–51.
- Baillet A, Gossec L, Carmona L, *et al*. Points to consider for reporting, screening for and preventing selected comorbidities in chronic inflammatory rheumatic diseases in daily practice: a EULAR initiative. *Ann Rheum Dis* 2016;75:965–73.
- National Health Care Institute. Ministerie van Volksgezondheid W en S. Guideline for economic evaluations in healthcare, 2016. Available: <https://english.zorginstituutnederland.nl/publications/reports/2016/06/16/guideline-for-economic-evaluations-in-healthcare> [Accessed 29 Jun 2020].
- Hemming K, Girling AJ, Sitch AJ, *et al*. Sample size calculations for cluster randomised controlled trials with a fixed number of clusters. *BMC Med Res Methodol* 2011;11:102.
- R: the R project for statistical computing. Available: <https://www.r-project.org/> [Accessed 14 May 2020].
- Allison PD. 312-2012: handling missing data by maximum likelihood 2012;21.
- Baillet A, Gossec L, Carmona L, *et al*. Points to consider for reporting, screening for and preventing selected comorbidities in chronic inflammatory rheumatic diseases in daily practice: a EULAR initiative. *Ann Rheum Dis* 2016;75:965–73.
- Willan AR, Briggs AH, Hoch JS. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. *Health Econ* 2004;13:461–75.
- Pearson. Greene, Econometric analysis. Available: <https://www.pearson.com/us/higher-education/product/Greene-Econometric-Analysis-7th-Edition/9780131395381.html> [Accessed 14 May 2020].
- Davison AC, Hinkley DV. *Bootstrap methods and their application*. Cambridge Core, 1997.
- Flynn TN, Peters TJ. Use of the bootstrap in analysing cost data from cluster randomised trials: some simulation results. *BMC Health Serv Res* 2004;4:33.
- Gomes M, Ng ES-W, Grieve R, *et al*. Developing appropriate methods for cost-effectiveness analysis of cluster randomized trials. *Med Decis Making* 2012;32:350–61.
- Essers I, Hilgsmann M, Kiltz U, *et al*. Development of one general and six country-specific algorithms to assess societal health utilities based on ASAS HI. *RMD Open* 2019;5:e000872.
- Zorginstituut Nederland. Ministerie van Volksgezondheid W en S. Ziektelast in de praktijk - De theorie en praktijk van het berekenen van ziektelast bij pakketbeoordelingen, 2018. Available: <https://www.zorginstituutnederland.nl/publicaties/rapport/2018/05/07/ziektelast-in-de-praktijk> [Accessed 29 Jun 2020].
- Pocock SJ, Stone GW. The primary outcome fails — what next? *N Engl J Med Overseas Ed* 2016;375:861–70.
- Smolen JS, Burmester G-R, Combe B, *et al*. Head-To-Head comparison of certolizumab pegol versus adalimumab in rheumatoid arthritis: 2-year efficacy and safety results from the randomised EXXELERATE study. *Lancet* 2016;388:2763–74.
- Weinblatt ME, Schiff M, Valente R, *et al*. Head-To-Head comparison of subcutaneous abatacept versus adalimumab for rheumatoid arthritis: findings of a phase IIIb, multinational, prospective, randomized study. *Arthritis Rheum* 2013;65:28–38.
- McInnes IB, Behrens F, Mease PJ, *et al*. Secukinumab versus adalimumab for treatment of active psoriatic arthritis (EXCEED): a double-blind, parallel-group, randomised, active-controlled, phase 3b trial. *Lancet* 2020;395:1496–505.