

Impact of TNF inhibitor therapy on joint replacement rates in rheumatoid arthritis: a matched cohort analysis of UK registry data

Authors: Samuel Hawley¹, M. Sanni Ali^{1,2}, Rene Cordtz^{3,4}, Lene Dreyer^{5,6}, Christopher J. Edwards⁷, Nigel K. Arden^{8,9}, Cyrus Cooper^{8,9}, Andrew Judge^{1,9,10}, Kimme Hyrich^{11,12}, Daniel Prieto-Alhambra^{1,13}

Affiliations:

1. Pharmaco- and Device-Epidemiology Group, Centre for Statistics in Medicine, NDORMS, University of Oxford, UK
2. Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK
3. Centre for Rheumatology and Spine Diseases, Gentofte, Rigshospitalet, DK
4. The Parker Institute, Copenhagen, DK
5. Department of Rheumatology and Clinical Medicine, Aalborg University Hospital, Aalborg, DK
6. Aalborg University, Aalborg DK
7. NIHR Clinical Research Facility, University Hospital Southampton, UK
8. Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, UK
9. MRC Lifecourse Epidemiology Unit, University of Southampton, UK
10. Translational Health Sciences, University of Bristol, UK
11. NIHR Manchester Biomedical Research Centre, Manchester University NHS Foundation Trust, Manchester, UK
12. Arthritis Research UK Centre for Epidemiology, Division of Musculoskeletal & Dermatological Sciences, University of Manchester, Manchester, UK
13. GREMPAL Research Group, Idiap Jordi Gol and CIBERFes, Unviuersitat Autonomia de Barcelona and Insituto de Salud Carlos III, Barcelona, Spain

Corresponding author: D. Prieto-Alhambra, Botnar Research Centre, Windmill Road, OX3 7LD, Oxford, UK

Abstract:

Objective: Previous ecological data suggest a decline in need for joint replacements in RA patients following the introduction of tumour necrosis factor inhibitor (TNFi) therapy, although patient-level data are lacking. Our primary aim was to estimate the association between TNFi use and subsequent incidence of total hip replacement (THR) and total knee replacement (TKR).

Methods: A propensity score matched cohort was analysed using the British Society for Rheumatology Biologics Registry (2001 – 2016) for Rheumatoid Arthritis (BSRBR-RA) data. Propensity score estimates were used to match TNFi users to similar conventional synthetic DMARD (csDMARD) users (with replacement) using a 1:1 ratio. Weighted multivariable Cox regression was used to estimate the impact of TNFi on study outcomes. Effect modification by baseline age and disease severity were investigated. Joint replacement at other sites was also analysed. An instrumental variable sensitivity analysis was also performed.

Results: The matched analysis contained a total of 19,116 patient records. Overall, there was no significant association between TNFi use versus csDMARD on rates of THR (HRs = 0.86 [95% CI: 0.60-1.22]) although there was significant effect modification by age ($p < 0.001$). TNFi was associated with a reduction in THR among those >60 years old (HR= 0.60 [CI: 0.41 - 0.87]) but not in younger patients. No significant associations were found for TKR or other joint replacement.

Conclusions:

Overall, no association was found between the use of TNFi and subsequent incidence of joint replacement. However, TNFi was associated with a 40% relative reduction in THR rates among older patients.

Keywords: epidemiology, rheumatoid arthritis, biologics, TNFi, joint replacement, total hip replacement, total knee replacement, comparative effectiveness

Key messages:

- Overall, this analysis showed TNFi use was not associated with subsequent rates of joint replacement
- Among elderly patients, TNFi use was associated with a 40% reduction in subsequent THR rates
- Given prior ecological data, future studies are needed to confirm and/or further elucidate this relationship

Background:

Rheumatoid arthritis (RA) is a chronic autoimmune disease driven by pathological inflammatory processes in patients' joints, subsequent hallmarks of which include structural damage to cartilage and bone (1-3). Joint damage is a central feature of RA and has been estimated to account for approximately 25% of disability in established disease (4). The permanent and often progressive nature of joint damage, in conjunction with associated pain, loss of function and failure to adequately respond to therapeutic options are strong indications for eventual joint replacement surgery (5, 6). Outcomes for hip and knee replacement in RA are generally considered good (7), although it has been observed that such patients are at increased risk of various adverse events compared to patients undergoing these procedures for osteoarthritis, including dislocation, infection, myocardial infarction, and revision (8, 9).

Furthermore, there are significant healthcare costs which in the UK are approximately £6,000 – £7,000 per operation (10). The prevention of irreversible joint damage through early and aggressive management using pharmacotherapy has been well demonstrated and is therefore recommended in numerous national guidelines (11-13). First-line therapy options include various conventional synthetic disease modifying anti-rheumatic drugs (csDMARDs).

Over recent decades the emergence of biologic therapies such as tumour necrosis factor inhibitors (TNFi) has revolutionised the management of RA as these drugs are widely recognized to improve numerous outcomes of the disease, including joint damage (14, 15). Despite this, to our knowledge there are no RCT studies addressing the issue of whether biologic therapies prevent/delay ultimate joint failure, as indicated by the need for a joint replacement. Recent ecological data from the UK and Denmark have indicated a reduction in the incidence of knee replacement among RA patients following the introduction of TNFi in 2002/2003 (16, 17), however concurrent to this has been an increasing emphasis on early and more aggressive usage of csDMARDs. Patient-level data on this topic is required to disentangle these issues, yet such data remain scarce.

Our current aim was to estimate the comparative effectiveness of TNFi vs. csDMARDs on subsequent rates of total hip replacement (THR) and total knee replacement (TKR) among a large cohort of RA patients.

Materials and Methods:

Data sources and exposures

We obtained data from the British Society for Rheumatology Biologics Register for RA (BSRBR-RA). This register contains prospectively collected observational data on over 20,000 RA patients recruited from 2001 onwards, primarily in order to evaluate the real-world safety of TNFi. The initial UK NICE guidance on use of TNFi stipulated clinicians initiating such therapy must register the patient into the BSRBR-RA, and recruitment continued originally until the target of at least 4,000 patients per TNFi cohort (etanercept, infliximab or adalimumab) was reached (last “original” patient recruited 2008) but reopened recruitment to these three originator drugs from 2010 onwards. NICE guidance restricts National Health Service prescribing of TNFi to patients with a sustained 28-joint disease activity score (DAS28) >5.1 who have failed to adequately respond to two csDMARDs, with each treatment lasting ≥ 6 months. The BSRBR-RA contains an additional comparator cohort of non-biologic treated RA patients on csDMARDs, entry into which was dependent on having active disease (guide DAS28 >4.2). Recruitment to this cohort closed in 2008

Participants in all study cohorts are followed up indefinitely using physician questionnaires sent from BSRBR-RA to the patients’ rheumatology clinic. These were sent every six months for a patients’ first three years of follow-up, and annually thereafter. In addition to collecting data on changes to therapy and disease details, physicians were asked what serious/adverse events had occurred since last follow-up date. In addition, patients were asked to complete a health diary every six months for the first three years in the study in which they detailed any hospital admissions. These free text responses were coded by BSRBR-RA staff using the MedDRA hierarchy (Medical Dictionary for Regulatory Activities (18)). Mortality data for participants were obtained from the Health and Social Care Information Centre (HSCIC) (now merged into NHS digital) via alerts generated from Office for National Statistics records. The BSRBR steering committee provided approval for the project, granting data

access in 2016. All patients provided written informed consent, no further approvals were necessary.

Outcomes

The primary outcomes of interest were first occurrence of THR or TKR, analysed separately. Patients were followed up from date of registration into BSRBR-RA until the earliest date of either outcome event, follow-up form date indicating change in TNFi exposure status (stopping among TNFi users or starting biologics among csDMARD users), last follow-up form or death. Other joint replacement (OJR) (a composite outcome consisting of elbow, shoulder, hand or other small joint replacement) was a secondary outcome.

Study population

Our study sample (Supplementary File 1) consisted of all biologic-naïve RA patients, either in the control cohort or those initiating a TNFi (etanercept, infliximab or adalimumab) no more than 6 months prior to registration within BSRBR-RA. Patients with a THR or TKR recorded prior to registration were excluded, as were patients with less than 6 months of follow-up (i.e. those who did not return at least one follow-up questionnaire). In analyses of OJR, further exclusion was made of patients who had undergone an OJR prior to baseline.

Statistical analysis

Owing to confounding by indication, i.e. TNFi users vs. csDMARD users likely having a different baseline risk of THR/TKR, we decided a-priori to match TNFi users to csDMARD users based on their propensity for receiving treatment. Propensity scores (PS), i.e. the probability of receiving treatment conditional on observed baseline characteristics (including those predictive of outcome) were estimated for all patients using logistic regression. The list

of potential confounders included in the propensity score equation (and described in Table 1) consisted of: age, gender, ethnicity, index of multiple deprivation (socio-economic status), body mass index (BMI), smoking status, year of registration (quintiles), time since RA diagnosis, DAS28, health assessment questionnaire (HAQ) score, quality of life (SF36 domains), 1987 ACR criteria, systemic involvement, co-morbidities and co-medications. A full description of these variables are included in Supplementary File 2. We matched each TNFi patient to the csDMARD patient with the most similar PS within a caliper distance of 0.2 standard deviations of the logit of the PS (19). Patients falling outside this common support region remained unmatched and were excluded from further analysis. We used matching with replacement (20) owing to fewer available csDMARD patients than TNFi users in the register. Missing data were imputed using chained equations and 10 imputed datasets were created. Statistical analyses were carried out in Stata 15.1 and R.

Baseline characteristics of the TNFi user and csDMARD cohorts were summarised and differences assessed by way of standardised mean differences (SMD) (22, 23), with smaller values indicative of greater similarity between cohorts. This assessment was carried out for the cohort prior to matching and in the 10th imputed dataset. Incidence rates of each outcome event with 95% confidence intervals (CI) were calculated among matched TNFi users and csDMARD users. Weighted Cox regression was used to compare THR, TKR and OJR rates, taking into account the number of times each individual csDMARD user was included by the matching with replacement, and adjusting the standard error of the estimates accordingly. We censored all patients at 12 years due to the small and unstable size of the csDMARD cohort after this time. Baseline characteristics that were not sufficiently similar post-matching, defined as $SMD > 0.1$ (21), were entered into a final multivariable Cox model for further

adjustment (21). Final models were run for all 10 datasets created in the multiple imputation process and hazard ratios (HRs) were pooled using Rubin's rules.

Age and disease severity were a-priori specified as potential effect modifiers of the association between TNFi use and subsequent need for joint replacement. We tested for these by including interaction terms for approximately median age (\leq / \geq 60 years old) and DAS28 (\leq / $>$ 5.1 NICE cut-off) in the weighted Cox model and used the likelihood ratio test to assess model fit. In the event of a significant interaction ($p < 0.1$), matching and survival models were re-run stratified by the significant effect modifier.

Sensitivity analyses

We addressed the potential of unobserved confounding through the use of an instrumental variable (IV) approach, using physician preference as the instrument, as has been done previously (24, 25). Details of this sensitivity analysis are described in Supplementary File 3. We also repeated the main PS analysis after excluding TNFi users recruited into the registry after the csDMARD cohort had closed, in order to maximise comparability between groups.

Results:

Of 13,126 eligible RA patients identified in BSRBR-RA, 97% (9,558) of the TNFi users and 51% (1,644) of the csDMARD users were retained following PS matching (Supplementary File 1). Given the 1:1 matching with replacement, a total of 19,116 patient records were used in subsequent analyses, with each csDMARD user being used a median of 3 (IQR: 1-6) times.

Baseline characteristics of TNFi users were markedly different in the unmatched study sample compared to the csDMARD cohort (Supplementary File 2), especially in aspects of disease severity. Specifically, the TNFi cohort had on average higher DAS28, HAQ score, proportion fulfilling the 1987 ACR RA criteria, lower health-related quality of life (as per SF36), longer disease duration and a higher prevalence of prior non-major joint replacement. Conversely, baseline characteristics between exposure cohorts were much more similar post-matching (Table 1). The only persisting differences ($SMD > 0.1$) between the matched cohorts were calendar period of registration, low deprivation and the proportion of patients fulfilling ACR criteria.

Total Hip Replacement

A total of 589 THRs were reported during follow-up (median = 4.94 years [IQR: 1.52 to 10.04] for TNFi and 5.97 years [IQR: 2.05 to 9.55] for csDMARD) of the propensity-matched cohorts. Incidence rates (per 1,000 PYs) were 5.22 [95% CI: 4.66 to 5.88] and 6.30 (95% CI: 4.24 to 9.76) among TNFi users and csDMARD users, respectively (Supplementary File 4). Comparing TNFi to csDMARDs yielded a pooled HR = 0.91 [95% CI: 0.64 to 1.31; $p=0.62$], which when adjusted for any remaining post-matching imbalance in baseline covariates was 0.86 [95% CI: 0.60 to 1.22; $p=0.39$] (Figure 1).

Total Knee Replacement

Among the matched sample, a total of 864 TKRs were reported during followup (median = 4.85 years [IQR: 1.50 to 10.01] for TNFi and 5.98 years [IQR: 2.03 to 9.55] for csDMARD) of the propensity-matched cohorts. Incidence rates (per 1,000 PYs) were 8.89 [95% CI: 8.13 to 9.72] and 8.09 [95% CI: 5.32 to 12.89] among TNFi users and csDMARD users, respectively (Supplementary File 4). This yielded a pooled HR = 1.18 [95% CI: 0.90 to 1.56;

$p=0.24$], which when adjusted for any remaining post-matching imbalance in baseline covariates was 1.11 [95% CI: 0.84 to 1.47; $p=0.46$] (Figure 1).

Other Joint Replacement

Among the matched sample, a total of 336 OJRs occurred during follow-up (median = 4.93 years [IQR: 1.52 to 10.02] for TNFi and 5.98 years [IQR: 2.05 to 9.12] for csDMARD) among the propensity-matched cohorts. Incidence rates (per 1,000 PYs) were 4.34 [95% CI: 3.76 to 5.02] and 3.87 [95% CI: 1.97 to 8.73] among TNFi and csDMARD users, respectively (Supplementary File 4). There was no significant difference in OJR rates between the exposure cohorts (Figure 1)

Interactions

Age was found to be a significant ($p<0.001$) effect modifier for both THR and TKR outcomes, although disease severity was not ($p>0.1$). In subsequent stratified analyses (Supplementary File 5), TNFi was associated with an estimated 40% reduction in incidence of THR among older patients (HR = 0.60 [95% CI: 0.41 to 0.87; $p=0.008$]) (Figure 2).

Differences in THR or TKR incidence rates between TNFi and csDMARD cohorts among younger patients were non-significant (Figure 2, Supplementary File 5).

Sensitivity analyses

Results were unchanged when PS matching and subsequent survival analysis was repeated following exclusion of ($n=1,213$) patients recruited into the TNFi cohort after the csDMARD cohort had closed. Comparative effectiveness estimates were HR=0.81 (95% CI: 0.55 to 1.18) for THR and 1.11 (95% CI: 0.83 to 1.50) for TKR. Results of the sensitivity IV analysis (Supplementary Files 6 & 7) confirmed main findings, with a borderline reduction in risk of

THR (absolute risk reduction of -1.88 per 100 patients [95% CI: -3.86 to 0.10; $p=0.063$]) but no association between biologics use and TKR risk (-0.57 per 100 patients [95% CI: -2.69 to 1.54; $p=0.60$]).

Discussion:

Main findings

We've here sought to address the scarcity of patient-level data comparing use of TNFi versus csDMARDs on rates of joint replacement in RA. Using a large UK-based RA biologics register, overall we found no difference in subsequent rates of joint replacement between PS matched TNFi and csDMARD users. When stratified by age, TNFi was associated with a significant 40% reduction ($p=0.008$) in THR incidence among patients ≥ 60 years old (Figure 2), although non-significant increases were observed in TKR for the same age group and in THR for those < 60 years old.

Findings in context

Our overall incidence rates of THR and TKR (results not shown) of 4.95/1,000 PYs and 7.84/1,000 PYs, respectively, align well with previous estimates of joint replacement among RA patients within the UK (16) and elsewhere (17, 26, 27).

Emerging observational data indicate that the number and/or incidence of RA related joint surgery has been in decline across numerous developed countries (16, 17, 27-37), although this has primarily been seen for smaller joints (35, 37-39). Many of these studies have inferred a possible role of biologics in this decline. Indeed, a reduction in need for joint replacement associated with TNFi use would be an expected finding given previous evidence

of TNFi use reducing joint damage as measured radiographically (15). A previous meta-analysis of 70 RCT studies reported that annual radiographic progression was 0.6% less in patients treated with biologics compared to those on a single DMARD (40). Similarly, another meta-analysis has demonstrated that patients on initial combination therapy (methotrexate plus a biologic agent) are 30% more likely to experience non-progression at 1-year than those on methotrexate alone (41). In this context, a 40% reduction in THR rates associated with TNFi use among a more elderly subgroup of patients as found in the present study is quite plausible, and a lack of translation of positive findings on joint erosion from prior RCTs into a “real world” reduction in rates of joint replacement within our main study sample is initially surprising.

However, a more detailed examination of prior RCT findings indicate that an expectation of widespread reduction in joint replacement associated with TNFi use is potentially unwarranted. There is a large degree of variation in the nature of comparator groups used in prior studies (42, 43), and this could be an important factor in considering the lack of effect as described in our main findings. For instance, whilst RCT studies have shown reduced radiographic progression among biologic users versus csDMARD monotherapy, an almost equal reduction has been achieved among combination csDMARD users relative to csDMARD monotherapy (40). Similarly, whilst TNFi has been shown to confer early benefits over combination csDMARD therapy, these benefits have been reported to disappear during the second year of follow-up (44), possibly due to time-to-efficacy and time to achieve maximal dose of csDMARDs. The use of etanercept (vs. oral triple therapy) resulted in only small radiographic benefits in another trial (45), with another showing triple therapy to be non-inferior to biologics in terms of change in DAS28 at 48 weeks (46). Given these previous data, it could be that the reduced THR rates in older patients may reflect a general

improvement in management of RA over the past 20 years and earlier and more aggressive use of csDMARDs rather than solely the effect of TNFi. Intriguingly, a registry-based study on the topic (47) recently found increased rates of major joint replacement associated with use of biologic therapy, although the authors concluded residual confounding was an issue given the small number of confounding factors for which the analysis was adjusted.

We found a reduction in THR incidence among an older patient subgroup, but no significant impact on TKR incidence, which is interesting as one might expect any effect to be expressed approximately equally at hip and knee. It could be that the relatively long disease duration at our baseline meant there was greater potential for prevention of joint destruction at the hip over knee, although details of differential natural history of RA disease at these two joints are not well established. It's also very difficult to disentangle the impact of TNFi on improved function and overall quality of life and how this may have mediated effects on longer-term progression of joint damage, potentially differentially at the knee and hip. Another factor could be the role of trauma related THRs among the older subgroup and whether there may be some pathway to reduced THR rates via TNFi associated improvement in bone quality. It should be emphasised however that the positive impact on THR incidence was only observed in a subgroup of patients that have not been well studied in this regard previously and that prior studies at the population-level have identified different patterns in this regard, some finding reduced rates of TKR (16, 17) and others THR (27, 34) following introduction of TNFi

Limitations

The potential for residual confounding by indication is a key limitation of the current study. Given that prior to matching there was a much higher disease severity among the TNFi group

(supplementary file 2), we cannot rule out that the overall lack of reduction in joint replacement rates among TNFi users may be due in part to a greater prevalence of unmeasured aspects of disease severity and unresponsiveness to therapy in this group, thereby maintaining a baseline ‘disadvantage’ even after PS matching. On-the-other-hand, estimates of impact of TNFi exposure may be subject to a general healthy user bias in that a clinician perceives sufficient patient ability to tolerate and benefit from more intensive therapy regimens, which may have here contributed to the reduced rate of THR in the older TNFi cohort. The use of an IV approach as a sensitivity analysis sought to address the issue of unmeasured confounding, in which the treatment effect was estimated using clinician preference for biologics as an IV, assuming this to be a strong predictor of exposure but unrelated/weakly-related to confounders of the TNFi - joint replacement relationship. However, the findings of this IV approach should be interpreted with caution given the instrument was here associated with several measured confounders (Supplementary File 3) which may undermine its validity as a means to obtain unbiased estimates in the presence of unmeasured confounding. The PS analytical approach taken, in which comparable csDMARD matches were found for each TNFi user - while much improving the internal validity - does mean our findings are not average treatment effects generalizable to the entire RA population but an estimate of the “average treatment effect on the treated” (ATT) (48). This is evident given that the csDMARD users included in the matched sample had more established and severe disease (thereby making them comparable to the TNFi sample) than the unmatched csDMARD sample (Supplementary File 2). The study findings should also be interpreted in the context of a relatively long median disease duration at baseline (Table 1), which means they are not necessarily generalisable to the context of early RA. Finally, we relied on a combination of physician-reported and self-reported incidence of THR, TKR and OJR for our study outcomes as per BSRBR-RA follow-up questionnaires. This may have introduced bias

if events were under-reported, although this would likely act non-differentially in regard to TNFi status and minimally during the early years of follow-up, during which time study participants were sent questionnaires every six months. These various limitations to the present analysis may partly explain differences in the results obtained here compared to previous ecological data on wide-spread reductions in joint replacement rates in RA during the biologic era. While a reduction in THR amongst older TNFi users offers some support for biologics playing a role in reducing need for joint replacement, it must also be noted that the lack of an overall protective effect is suggestive that other factors apart from TNFi are likely to be involved in the aforementioned downward population trends in joint replacement rates in RA.

Strengths

Our study's key strengths are that BSRBR-RA is one of the largest RA registers in the world, which made it possible for us to adjust for many potential confounders and stratify analyses where there was significant effect modification. We were also able to accurately censor follow-up given the linkage to HSCIC mortality data. The use of PS matching is a strong method for dealing with bias (49) arising from likely confounding by TNFi indication, and we were able to reach good balance between the two exposure groups in terms of baseline characteristics across most variables.

Conclusion

In this large prospective study, we found no overall association between TNFi versus csDMARD therapy and subsequent incidence of joint replacement among RA patients, although a 40% relative reduction in THR rates was found among older patients. Future

studies are needed to confirm and/or further elucidate the relationship between TNFi use and joint replacement

Figure Legends:

Figure 1: Estimated impact of TNFi on subsequent joint replacement rates among matched TNFi and csDMARD patients

Figure 2: Estimated impact of TNFi on subsequent joint replacement rates among matched TNFi and csDMARD patients: stratified by age

Author Contributions:

Study Concept: all authors; Data Acquisition: all authors; Data Analysis: SH, SA, DPA;

Drafting Manuscript: SH; Revising and approving final manuscript: all authors

Acknowledgements:

DPA is funded by a National Institute for Health Research (NIHR) Clinician Scientist Award (project number CS-2013-13-012). AJ is supported by the NIHR Biomedical Research Centre at the University Hospitals Bristol NHS Foundation Trust and the University of Bristol. The views and opinions expressed in the manuscript are those of the authors and do not necessarily reflect those of the NIHR, NHS or the Department of Health. The data that support the findings of this study are available from the British Society for Rheumatology Biologics Registry but restrictions apply to the availability of these

data, which were used under license for the current study, and so are not publicly available.

Competing Interests:

S.H., R.C., C.C., S.A. and K.H have nothing to disclose. L.D. reports personal fees from UCB Pharma. and from MSD. C.J.E. has been a speaker for, received honoraria or research support from Abbvie, BMS, Celgene, Pfizer, Biogen, Mundipharma, UCB Pharma., Roche, MSD. N.K.A. reports grants from Bioberica and personal fees from Bioventus, from Regeneration, and from Smith&Nephew. A.J. has received consultancy fees from Freshfields Bruckhaus Deringer, and is a member of the Data Safety and Monitoring Board (which involved receipt of fees) from Anthera Pharmaceuticals outside the submitted work. DPA's group have received unrestricted research grants from Servier Laboratoires, Amgen, and UCB Pharma.

References:

1. Scott DL, Wolfe F, Huizinga TW. Rheumatoid arthritis. *Lancet*. 2010;376(9746):1094-108.
2. Feldmann M, Brennan FM, Maini RN. Rheumatoid arthritis. *Cell*. 1996;85(3):307-10.
3. Schett G, Gravallese E. Bone erosion in rheumatoid arthritis: mechanisms, diagnosis and treatment. *Nat Rev Rheumatol*. 2012;8(11):656-64.
4. Scott DL, Pugner K, Kaarela K, Doyle DV, Woolf A, Holmes J, et al. The links between joint damage and disability in rheumatoid arthritis. *Rheumatology (Oxford)*. 2000;39(2):122-32.
5. Gademan MG, Hofstede SN, Vliet Vlieland TP, Nelissen RG, Marang-van de Mheen PJ. Indication criteria for total hip or knee arthroplasty in osteoarthritis: a state-of-the-science overview. *BMC musculoskeletal disorders*. 2016;17(1):463.
6. Nikiphorou E, Norton S, Young A, Carpenter L, Dixey J, Walsh DA, et al. Association between rheumatoid arthritis disease activity, progression of functional limitation and long-term risk of orthopaedic surgery: combined analysis of two

- prospective cohorts supports EULAR treat to target DAS thresholds. *Ann Rheum Dis*. 2016;75(12):2080-6.
7. Nelissen RG. The impact of total joint replacement in rheumatoid arthritis. *Best Pract Res Clin Rheumatol*. 2003;17(5):831-46.
 8. Burn E, Edwards CJ, Murray DW, Silman A, Cooper C, Arden NK, et al. The impact of rheumatoid arthritis on the risk of adverse events following joint replacement: a real-world cohort study. *Clin Epidemiol*. 2018;10:697-704.
 9. Ravi B, Escott B, Shah PS, Jenkinson R, Chahal J, Bogoch E, et al. A systematic review and meta-analysis comparing complications following total joint arthroplasty for rheumatoid arthritis versus for osteoarthritis. *Arthritis and rheumatism*. 2012;64(12):3839-49.
 10. Burn E, Edwards CJ, Murray DW, Silman A, Cooper C, Arden NK, et al. Trends and determinants of length of stay and hospital reimbursement following knee and hip replacement: evidence from linked primary care and NHS hospital records from 1997 to 2014. *BMJ Open*. 2018;8(1):e019146.
 11. American College of Rheumatology Subcommittee on Rheumatoid Arthritis G. Guidelines for the management of rheumatoid arthritis: 2002 Update. *Arthritis and rheumatism*. 2002;46(2):328-46.
 12. Luqmani R, Hennell S, Estrach C, Birrell F, Bosworth A, Davenport G, et al. British society for rheumatology and british health professionals in rheumatology guideline for the management of rheumatoid arthritis (the first two years). *Rheumatology*. 2006;45(9):1167-9.
 13. Kwok CK, Simms RW, Anderson LG, Erlandson DM, Greene JM, Moncur C, et al. Guidelines for the management of rheumatoid arthritis. *Arthritis and rheumatism*. 1996;39(5):713-22.
 14. Curtis JR, Singh JA. Use of biologics in rheumatoid arthritis: current and emerging paradigms of care. *Clin Ther*. 2011;33(6):679-707.
 15. Scott DL. Biologics-based therapy for the treatment of rheumatoid arthritis. *Clin Pharmacol Ther*. 2012;91(1):30-43.
 16. Hawley S, Cordtz R, Lene D, Edwards CJ, Arden NK, Delmestri A, et al. Association between NICE guidance on biologic therapies with rates of hip and knee replacement among rheumatoid arthritis patients in England and Wales: An interrupted time-series analysis. *Semin Arthritis Rheu*. 2018;47(5):605-10.
 17. Cordtz RL, Hawley S, Prieto-Alhambra D, Hojgaard P, Zobbe K, Overgaard S, et al. Incidence of hip and knee replacement in patients with rheumatoid arthritis following the introduction of biological DMARDs: an interrupted time-series analysis using nationwide Danish healthcare registers. *Ann Rheum Dis*. 2018;77(5):684-9.
 18. MedDRA. Welcome to MedDRA <https://www.meddra.org> [
 19. Austin PC. The performance of different propensity-score methods for estimating relative risks. *J Clin Epidemiol*. 2008;61(6):537-45.
 20. Stuart EA. Matching methods for causal inference: A review and a look forward. *Stat Sci*. 2010;25(1):1-21.
 21. Nguyen TL, Collins GS, Spence J, Daures JP, Devereaux PJ, Landais P, et al. Double-adjustment in propensity score matching analysis: choosing a threshold for considering residual imbalance. *BMC Med Res Methodol*. 2017;17(1):78.
 22. Austin PC. Using the Standardized Difference to Compare the Prevalence of a Binary Variable Between Two Groups in Observational Research. *Communications in Statistics - Simulation and Computation*. 2009;38(6).

23. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in medicine*. 2009;28(25):3083-107.
24. Ertefaie A, Small DS, Flory JH, Hennessy S. A tutorial on the use of instrumental variables in pharmacoepidemiology. *Pharmacoepidemiol Drug Saf*. 2017;26(4):357-67.
25. Brookhart MA, Wang PS, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. *Epidemiology*. 2006;17(3):268-75.
26. Rodriguez-Rodriguez L, Ivorra-Cortes J, Leon L, Fernandez-Gutierrez B, Abasolo L, Gonzalez-Alvaro I, editors. *Temporal Trends of Total Joint Replacement in Rheumatoid Arthritis Patients: A Survival Study 2016: Arthritis Rheumatol*.
27. Hekmat K, Jacobsson L, Nilsson JA, Petersson IF, Robertsson O, Garellick G, et al. Decrease in the incidence of total hip arthroplasties in patients with rheumatoid arthritis--results from a well defined population in south Sweden. *Arthritis Res Ther*. 2011;13(2):R67.
28. National Joint Registry for England and Wales: 1st Annual Report, (2004).
29. National Joint Registry for England and Wales: 5th Annual Report, (2008).
30. Louie GH, Ward MM. Changes in the rates of joint surgery among patients with rheumatoid arthritis in California, 1983-2007. *Ann Rheum Dis*. 2010;69(5):868-71.
31. Jansen E, Virta LJ, Hakala M, Kauppi MJ, Malmivaara A, Lehto MU. The decline in joint replacement surgery in rheumatoid arthritis is associated with a concomitant increase in the intensity of anti-rheumatic therapy: a nationwide register-based study from 1995 through 2010. *Acta Orthop*. 2013;84(4):331-7.
32. Mertelsmann-Voss C, Lyman S, Pan TJ, Goodman SM, Figgie MP, Mandl LA. US trends in rates of arthroplasty for inflammatory arthritis including rheumatoid arthritis, juvenile idiopathic arthritis, and spondyloarthritis. *Arthritis Rheumatol*. 2014;66(6):1432-9.
33. Shourt CA, Crowson CS, Gabriel SE, Matteson EL. Orthopedic surgery among patients with rheumatoid arthritis 1980-2007: a population-based study focused on surgery rates, sex, and mortality. *J Rheumatol*. 2012;39(3):481-5.
34. David G, Tandon N, Waters H, Gunnarsson C, Kavanaugh A. Rheumatoid Arthritis and Joint Replacement: Impact of Biologics. *Americal Journal of Pharmaceutical Benefits*. 2014;6(6):256-64.
35. Nystad TW, Fenstad AM, Furnes O, Havelin LI, Skredderstuen AK, Fevang BT. Reduction in orthopaedic surgery in patients with rheumatoid arthritis: a Norwegian register-based study. *Scand J Rheumatol*. 2015:1-7.
36. Harty L, O'Toole G, FitzGerald O. Profound reduction in hospital admissions and musculoskeletal surgical procedures for rheumatoid arthritis with concurrent changes in clinical practice (1995-2010). *Rheumatology (Oxford)*. 2015;54(4):666-71.
37. Singh JY, B; Watson, S; Perez, J; McGwin, G; Ponce, B. Trends in Joint Replacements Surgery in Patients with Rheumatoid Arthritis. *Arthritis Rheumatol*. 2016;68 (suppl 10).
38. Nikiphorou E, Carpenter L, Morris S, MacGregor AJ, Dixey J, Kiely P, et al. Hand and Foot Surgery Rates in Rheumatoid Arthritis Have Declined From 1986 to 2011, but Large-Joint Replacement Rates Remain Unchanged Results From Two UK Inception Cohorts. *Arthritis Rheumatol*. 2014;66(5):1081-9.
39. Cordtz RH, S.; Prieto-Alhambra, D.; Hojgaard, P.; Zobbe, K.; Odgaard, A.; Overgaard, S.; Kristensen, L.E.; Dreyer, L. Changes in incidence of shoulder, elbow, wrist and finger replacement surgery among rheumatoid arthritis patients following the

introduction of biological dmards: an interrupted time series analysis using danish health care registers. *Annals of the Rheumatic Diseases*. 2018;77.

40. Graudal N, Jurgens G. Similar effects of disease-modifying antirheumatic drugs, glucocorticoids, and biologic agents on radiographic progression in rheumatoid arthritis: meta-analysis of 70 randomized placebo-controlled or drug-controlled studies, including 112 comparisons. *Arthritis and rheumatism*. 2010;62(10):2852-63.
41. Kuriya B, Arkema EV, Bykerk VP, Keystone EC. Efficacy of initial methotrexate monotherapy versus combination therapy with a biological agent in early rheumatoid arthritis: a meta-analysis of clinical and radiographic remission. *Ann Rheum Dis*. 2010;69(7):1298-304.
42. Thorlund K, Druyts E, Avina-Zubieta JA, Wu P, Mills EJ. Why the findings of published multiple treatment comparison meta-analyses of biologic treatments for rheumatoid arthritis are different: an overview of recurrent methodological shortcomings. *Ann Rheum Dis*. 2013;72(9):1524-35.
43. Estellat C, Ravaud P. Lack of head-to-head trials and fair control arms: randomized controlled trials of biologic treatment for rheumatoid arthritis. *Arch Intern Med*. 2012;172(3):237-44.
44. Graudal N, Hubeck-Graudal T, Faurschou M, Baslund B, Jurgens G. Combination Therapy With and Without Tumor Necrosis Factor Inhibitors in Rheumatoid Arthritis: A Meta-Analysis of Randomized Trials. *Arthritis care & research*. 2015;67(11):1487-95.
45. Moreland LW, O'Dell JR, Paulus HE, Curtis JR, Bathon JM, St Clair EW, et al. A randomized comparative effectiveness study of oral triple therapy versus etanercept plus methotrexate in early aggressive rheumatoid arthritis: the treatment of Early Aggressive Rheumatoid Arthritis Trial. *Arthritis and rheumatism*. 2012;64(9):2824-35.
46. O'Dell JR, Mikuls TR, Taylor TH, Ahluwalia V, Brophy M, Warren SR, et al. Therapies for active rheumatoid arthritis after methotrexate failure. *The New England journal of medicine*. 2013;369(4):307-18.
47. Aaltonen KJ, Virkki LM, Jansen E, Sokka T, Konttinen YT, Peltomaa R, et al. Do biologic drugs affect the need for and outcome of joint replacements in patients with rheumatoid arthritis? A register-based study. *Semin Arthritis Rheum*. 2013;43(1):55-62.
48. Pirracchio R, Carone M, Rigon MR, Caruana E, Mebazaa A, Chevret S. Propensity score estimators for the average treatment effect and the average treatment effect on the treated may yield very different estimates. *Statistical methods in medical research*. 2016;25(5):1938-54.
49. Austin PC. The performance of different propensity-score methods for estimating relative risks. *J Clin Epidemiol*. 2008;61(6):537-45.

Table 1: baseline characteristics of propensity score matched cohorts*: stratified by use of TNFi vs. Conventional Synthetic DMARDs

| Characteristic | <u>CS-DMARD (N=9,558)</u> (1,644 unique patients) | | <u>TNFi (N=9,558)</u> | | SMD |
|--------------------------------------|--|------|-----------------------|------|-------|
| | N | % | N | % | |
| Age: Mean (S.D.) | 55.2 (12.1) | | 55.2 (12.3) | | 0 |
| Gender: % Female | 7,289 | 76.3 | 7,259 | 75.9 | -0.01 |
| ethnicity: % White/Caucasian | 9,114 | 95.4 | 9,118 | 95.4 | 0.00 |
| Index Multiple Deprivation | | | | | |
| Quintile 1 | 1282 | 13.4 | 1322 | 13.8 | 0.01 |
| Quintile 2 | 1420 | 14.9 | 1485 | 15.5 | 0.02 |
| Quintile 3 | 1413 | 14.8 | 1640 | 17.2 | 0.07 |
| Quintile 4 | 1544 | 16.2 | 1710 | 17.9 | 0.05 |
| Quintile 5 | 1284 | 13.4 | 1650 | 17.3 | 0.11 |
| Unknown | 2615 | 27.4 | 1751 | 18.3 | -0.22 |
| BMI | 26.8 (5.9) | | 27.1 (6.3) | | 0.06 |
| Smoking? | | | | | |
| % Current | 2448 | 25.6 | 2259 | 23.6 | -0.05 |
| % Ex | 3272 | 34.2 | 3577 | 37.4 | 0.07 |
| Calendar period of registration | | | | | |
| Oct 01 - Oct 03 | 1481 | 15.5 | 2262 | 23.7 | 0.21 |
| Nov 03 - Aug 04 | 2037 | 21.3 | 2037 | 21.3 | 0 |
| Sept 04 - Aug 05 | 1823 | 19.1 | 1802 | 18.9 | -0.01 |
| Aug 05 - May 07 | 1664 | 17.4 | 1462 | 15.3 | -0.06 |
| May 07 - May 16 | 2553 | 26.7 | 1995 | 20.9 | -0.14 |
| Years since diagnosis: Median (IQR) | 10.8 (10.7) | | 11.0 (8.8) | | 0.02 |
| DAS 28: Mean (S.D.) | 6.47 (1.09) | | 6.43 (0.98) | | -0.04 |
| Overall HAQ score | 1.91 (0.63) | | 1.91 (0.62) | | -0.01 |
| ACR: Ever rheumatoid positive: % | 5,532 | 57.9 | 6,055 | 63.4 | 0.11 |
| ACR: Deformity of ≥3 joint areas?: % | 7,425 | 77.7 | 8,118 | 84.9 | 0.19 |
| ACR: Erosions on hands/feet: % | 4,651 | 48.7 | 5,282 | 55.3 | 0.13 |
| ACR: Ever had nodules % | 4,034 | 42.2 | 3,988 | 41.7 | -0.01 |
| ACR: Symmetry % | 7,468 | 78.1 | 7,883 | 82.5 | 0.11 |
| ACR: Deformity of hand joint % | 6,771 | 70.8 | 7,602 | 79.5 | 0.20 |
| ACR: Morning stiffness >1 hour: % | 8,901 | 93.1 | 8,966 | 93.8 | 0.03 |
| Non-major prior joint replacement^ | 1742 | 18.2 | 1989 | 20.8 | 0.07 |

* Results shown are for the 10th imputed dataset. Matching was performed using replacement of the cs-DMARD users. 9,558 biologic users were each matched to one of the 3,229 cs-DMARD users (with replacement). Number of csDMARD patients represented in final matched sample was 1,644

^ Composite variable consisting of: shoulder, elbow, neck or other small joint replacement (e.g. hand)
SMD: standardised mean difference (smaller values indicative of better balance)