Comparative effectiveness of certolizumab pegol, abatacept and biosimilar infliximab in patients with rheumatoid arthritis treated in routine care. Observational data from the Danish DANBIO registry emulating a randomized trial

Grøn, Kathrine L; Glintborg, Bente; Nørgaard, Mette; Mehnert, Frank; Østergaard, Mikkel; Dreyer, Lene; Krogh, Niels S; Bjorner, Jakob B; Hetland, Merete L

Published in:
Arthritis & rheumatology (Hoboken, N.J.)

DOI (link to publication from Publisher):
10.1002/art.41031

Publication date:
2019

Document Version
Accepted author manuscript, peer reviewed version

DR. KATHRINE LEDERBALLE GRØN (Orcid ID : 0000-0001-9759-8181)

Article type : Full Length

Comparative effectiveness of certolizumab pegol, abatacept and biosimilar infliximab in patients with rheumatoid arthritis treated in routine care. Observational data from the Danish DANBIO registry emulating a randomized trial

Kathrine L. Grøn, MD[1,2], Bente Glintborg, MD, PhD, DMSci[1,2], Mette Nørgaard, MD, PhD, Professor[3], Frank Mehnert, MSc[3], Mikkel Østergaard, MD, PhD, DMSci, professor[1,2], Lene Dreyer, MD, PhD, professor[1,4], Niels S. Krogh, MSc Econ[5], Jakob B. Bjorner, MD, PhD, professor[6,7] and Merete L. Hetland, MD, PhD, DMSci, professor[1,2]

[1)] DANBIO and Copenhagen Center for Arthritis Research (COPECARE), Center for Rheumatology and Spine Diseases, Centre of Head and Orthopaedics, Rigshospitalet, Copenhagen, Denmark
[2)] Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
[3)] Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark
[4)] Departments of Clinical Medicine and Rheumatology, Aalborg University and Aalborg University Hospital, Aalborg, Denmark
[5)] Zitelab, Copenhagen, Denmark
[6)]Optum Patient Insights, Johnston, RI, US
[7)] Department of Public Health, University of Copenhagen, Copenhagen, Denmark

Conflicts of interest

Kathrine L. Grøn has received research support from BMS, **Bente Glintborg** has received consultancies/speaking fees/honoraria from Biogen, Abbvie, Pfizer (> $10.000), **Mette Nørgaard** has no disclosures, **Frank Mehnert** has no disclosures, **Mikkel Østergaard** has received consultancies/speaking fees/honoraria from Abbvie, BMS, Boehringer-Ingelheim, Celgene, Eli-Lilly, Hospira, Janssen, Merck, Novartis, Orion, Pfizer, Regeneron, Roche, Sanofi, UCB (<$10.000) **Lene Dreyer** has received consultancies/speaking fees/honoraria from UCB, MSD, Eli Lilly, Galderma and Janssen Pharmaceuticals, (< $10.000), **Niels S. Krogh** has no disclosures, **Jakob B. Bjørner:** employee of Optum, **Merete L. Hetland** has received consultancies/speaking fees/honoraria from MSD, Orion(< $10.000), expert testimony: CellTrion, research grants from AbbVie, Biogen, BMS, Novartis, Pfizer, Samsung, UCB.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/art.41031

Corresponding author:

Kathrine L. Grøn

Rigshospitalet Glostrup, Valdemar Hansens Vej 17, 2600 Glostrup

Email: kathrine.l.groen@gmail.com

Tlf: +45 26188891

ABSTRACT

Objectives: Nationwide Danish guidelines regarding rheumatoid arthritis (RA) patients initiating biological treatment(bDMARDs) are issued approximately annually. For bio-naïve patients treated with concomitant methotrexate, the mandatory drugs were: certolizumab pegol (year: 2013-2014, recommended compliance: 80%); abatacept(2014-2015, 80%); biosimilar infliximab(CT-P13) (2015-2016, 50%). We hypothesized that the guidelines could be perceived as a surrogate randomization tool where calendar time rather than patient-specific factors defined choice of bDMARD. Objectives were to 1) assess compliance to guidelines (supporting the assumption of surrogate randomization); 2) compare effectiveness of certolizumab pegol, abatacept and CT-P13 in patients treated according to guidelines.

Methods: Observational cohort study emulating a randomized trial (intention-to-treat). RA patients compliant to the guidelines were identified in DANBIO and information on prior comorbidities were obtained through linkage to national registries. Outcomes were DAS28-remission-rates (at 6/12-months) and one-year treatment retention, compared across treatments (comorbidity/confounder-adjusted multivariable logistic and Cox-regression analyses).

Results: 776 patients were included (certolizumab/abatacept/CT-P13: 336/215/225). Compliance to guidelines was high (70%/65%/59%). Six and 12 months' DAS28 remission rates were: 35%/33%/42% and 35%/31%/35%, respectively. Compared with certolizumab, 6 and 12 months' adjusted odds ratios(OR) for DAS28 remission were for abatacept OR=0.96(95%CI:0.63,1.47) and 0.74(0.47,1.15), for CT-P13 OR=1.38(0.91,2.09) and 0.96(0.62,1.49), respectively. Adjusted hazard ratios(HR) for withdrawal (during 0-90 days/91-365 days) were for abatacept HR=0.70(95% CI 0.39,1.27)/1.16(0.84,1.60) and for CT-P13 HR=0.58(0.33,1.10)/0.83(0.59,1.17), compared with certolizumab.

Conclusion: The surrogate randomization procedure enabled head-to-head comparison of certolizumab pegol, abatacept and CT-P13. Although some differences in the estimated effectiveness were observed across drugs, confidence intervals were wide and statistical significance was not reached.

INTRODUCTION

Rheumatoid arthritis (RA) is a severe inflammatory joint disease associated with pain, disability and joint destruction. An increasing number of biological disease-modifying anti-rheumatic drugs (bDMARDs) is now available. Randomized controlled trials (RCTs) are considered the gold standard for assessing drug effects. [(1)] In RCTs with intention-to-treat analyses (ITT), the different bDMARDs in combination with methotrexate (MTX) appear to have similar efficacy, [(2-4)] but only some of the RCTs have performed direct comparison of different bDMARDs.[(5, 6)] Observational studies in routine care are a valuable supplement to RCTs due to the inclusion of less restricted populations, longer follow-up time and lower costs. However, lack of randomization makes observational effect estimates vulnerable to

confounding, including confounding by indication.[7] Interestingly, it has been shown that a properly designed observational study may produce results very similar to that of RCTs, when observational data emulate a RCT.[8]

DANBIO is a Danish nationwide registry of patients with inflammatory arthritis.[9] Patients are monitored prospectively in routine care with coverage >95% for bDMARD treated patients.[10] Through linkage to other national registries DANBIO data can be enriched with information on e.g. hospitalizations, comorbidities, and vital status.[11] In Denmark, nationwide guidelines on mandatory choice of bDMARD are issued approximately annually based on a combination of expert opinion and costs. The guidelines include the expected proportion of patients treated according to the guideline (=compliance).[12] The guidelines are enforced by the financial incentive of the departments, since the cheapest drug is selected as first choice in the guideline among drugs considered to be equally effective and with comparable safety profile. Also, the treating physician must argue for any deviation from the guideline in each individual case. Between 2013 and 2016, certolizumab pegol (a tumor necrosis factor inhibitor, TNFi), abatacept (an inhibitor of co-stimulation of T-cells) and CT-P13 (biosimilar infliximab) were successively the recommended bDMARDs in combination with MTX in biologics-naive RA patients with insufficient prior response to csDMARDs.

We hypothesized that the guidelines could be perceived as a surrogate randomization tool[13, 14] and that choice of bDMARD would be determined by calendar period rather than by patient-specific factors, thus minimizing confounding by indication. This might provide a unique opportunity for emulating an RCT thus enabling the comparison of real-world effectiveness of three drugs that have never been compared directly in the

same RCT. The "ideal hypothetical trial" that this observational study attempted to emulate would be a pragmatic trial, where eligible patients were biologics-naïve patients with RA who were inadequate responders to MTX/csDMARD and who initiated their first bDMARD between 2013-2016. They would be randomized to either certolizumab pegol, abatacept or CT-P13, and the treatment effect and adherence would be compared after 6 and 12 months.[15, 16]

The aims of the current study were therefore two-fold: 1) to assess compliance to the guidelines to support the assumption of surrogate randomization and 2) to compare the effectiveness of treatment with certolizumab pegol, abatacept and CT-P13 given in combination with concomitant MTX in bDMARD naïve patients with RA in routine care.

PATIENTS AND METHODS

National treatment guidelines and compliance to them

In Denmark, treatment with bDMARDs is tax paid, which facilitates equal access to expensive therapies. For drugs considered equally effective, a tender-process takes place approximately annually, after which it is mandatory to prescribe the cheapest drug as first bDMARD treatment. The expected compliance to the guideline is stated, i.e. the rheumatologists must prescribe the cheapest drug in e.g. 50% or 80% of the patients during the calendar period that the guideline covers.[17] Between 2013 and 2016, certolizumab pegol (Jan 1th 2013-June 30th, 2014), abatacept (July 1th 2014 – June 30th, 2015) and CT-P13 (July 1th 2015-June 30th, 2016) were recommended as first line bDMARDs in biologic-naïve RA patients concomitantly treated with MTX, with expected compliance of 80%, 80% and

50%, respectively. For each of the three calendar periods, the proportion of all biologic-naïve RA patients who started first bDMARD in combination with MTX in accordance with the national guidelines was calculated and compliance was assessed as: (numbers adherent to guidelines/numbers of all bio-naïve pts initiating bDMARD in combination with MTX) * 100%.

Study design

For each of the three calendar periods, biologic-naïve patients were identified in DANBIO and included in this comparative effectiveness study, if they started bDMARD treatment (with concomitant MTX) in accordance with the guideline. Thereby we aimed at establishing a non-blinded observational cohort study with surrogate (calendar time) randomization.

Patient characteristics and disease activity (baseline, 6 months, 12 months) were retrieved from DANBIO. Baseline (=index date) was defined as the start date of the bDMARD. Time windows were applied for baseline and follow-up visits as follows: Baseline was defined as the time window from 30 days before until 6 days after treatment start. For the 6-months visit, the time window was from 8 to 39 weeks and for the 12-months visit from week 40 to 72 weeks after treatment start. If more than one visit occurred within a given time window, the one closest to the given time point was selected for the analysis.

By use of unique civil registration numbers, comorbidities prior to the index date in each patient were identified by linkage to the Danish National Registry of Patients (DNRP).[11] The DNPR has virtually complete data on in- and out-patient contacts (in-patient data since 1977 and out-patient since 1995).

Ethics

According to Danish law, informed consent and ethical approval were not required for the present study.

Clinical variables and outcomes

The following baseline covariates were identified in DANBIO: age (years), gender (female/male), RA disease duration (years), anti-cyclic citrullinated peptide (anti-CCP) status (positive/negative), current smoking status (yes/no/missing), physician global score on a visual analogue scale (VAS) 0-100, patient pain VAS 0-100, fatigue VAS 0-100, patient global VAS 0-100, functional status (Health Assessment Questionnaire, HAQ (continuous, 0-3), C-reactive protein (CRP, continuous) and glucocorticoid use at baseline (injections or oral) (yes/no). Similarly, at baseline and at the 6- and 12-month visits DAS28 (based on four variables and C-reactive protein; continuous),[18, 19] and clinical disease activity index (CDAI; continuous) were calculated. Remission criteria were: DAS28<2.6 and CDAI≤2.8.

Study outcomes were: proportion of patients in DAS28 remission (ITT analysis) after 6 and 12 months, and one-year retention rates. Furthermore, odds for achieving DAS28 remission after 6 and 12 months, and one-year risk of withdrawal (hazard ratio, HR) were compared across treatments. Secondary outcomes were median DAS28 and proportion of patients in CDAI remission at 6 and 12 months.

Comorbidities were identified in DNPR 10 years prior to baseline and the Charlson Comorbidity Index (CCI) score (0/ ≥ 1) was calculated excluding category 7, connective tissue disease **(Supplementary Table S1).**[20]

In the analysis of treatment retention, time on bDMARD was defined as number of days from baseline to the first of: stop date of the bDMARD, emigration, start of treatment with another bDMARD, death, or end of study period. The stop date was recorded in DANBIO by the treating physician as the date of first missed dose. Reasons for withdrawal (lack of effect/adverse events/cancer/pregnancy wish/treatment at another hospital/infection/loss to follow-up/death/surgery/project participation/remission/other/unavailable) were retrieved from DANBIO. End of study period was January 1st, 2018 to ensure 12 months of follow-up on patients in the third calendar period.

Statistical analysis

All analyses were conducted using SAS version 9.4. Descriptive statistics for continuous variables were summarized as medians and interquartile ranges (IQR), categorical variables were assessed as frequencies and percentages.

Treatment responses at 6 and 12 months: For patient with missing DAS28 remission status at 6 months, the following assumptions were made as part of an approximated ITT analysis: 1) patients who had withdrawn from treatment due to "lack of effect" were classified as not being in remission, 2) patients who had withdrawn due to "remission" were classified as being in remission, 3) patients still on treatment who had more than one swollen joint at 6 months were classified as not being in remission, 4) patients who still had missing outcome data were in the main analyses classified as not being in remission. Similar approaches were applied for DAS28 remission status at 12 months and for CDAI remission. When calculating medians of DAS28, patients with missing DAS28 were excluded.

The odds of being in DAS28 remission at 6 months were estimated by logistic regression analyses adjusted for baseline confounders identified *a priori* as clinically relevant: a) age, gender, b) age, gender, DAS28, HAQ, smoking and CCI score (=fully adjusted). Certolizumab was the reference drug, because most patients were treated with this drug. In "best case/worst case" sensitivity analyses, the above-mentioned ITT analysis was expanded further to assess the impact of extreme endpoints: In patients who had missing outcome data for DAS28 remission (yes/no) at 6 months, logistic regression analyses were carried out where one drug at a time was assigned "best case" for all missing data and compared to the two other drugs assigned "worst case" for all missing data. In the "best case", patients with missing data on DAS28 at 6 months were categorized as being in remission. In the "worst case" they were categorized as not being in remission (i.e. non-responder imputation). Similar analyses were applied for DAS28 remission at 12 months.

One-year treatment retention was investigated with multivariable Cox regression analyses adjusted for the above-mentioned confounders. The proportional hazard assumption was not met for the initial model but was fulfilled when analyses were done with separate effects for the time-intervals 0-90 days and 91-365 days.

To mimic common inclusion and exclusion criteria in RCTs, further sensitivity analyses were done in which we excluded patients with severe comorbidities (myocardial infarction, cerebrovascular disease, dementia, hemiplegia, moderate to severe renal disease, diabetes with end organ damage, any tumor, leukemia, lymphoma, moderate to severe liver disease, metastatic solid tumor and AIDS) and only included patients with a baseline DAS28 > 3.2 who were seropositive for rheumatoid factor or anti-CCP. In these analyses, no adjustment for baseline comorbidities was done.

RESULTS

Compliance to the national guidelines

Table 1 shows the pattern of bDMARD prescriptions in the three calendar periods. The national guidelines were followed in 776 RA patients initiating certolizumab pegol (336 patients), abatacept (215 patients) or CT-P13 (225 patients) (**Table 1**). Compliance regarding choice of drug in each calendar-period was 70%, 65% and 59%, respectively.

Baseline characteristics of the patients included in the comparative effectiveness study

Patients had similar DAS28, CDAI, CRP and patient reported outcomes (HAQ, VAS fatigue and pain) across treatments (**Table 2**). Patients treated with CT-P13 were slightly older and abatacept treated patients had fewer comorbidities and fewer were smokers (**Table 2**).

Visits during follow-up

The number of visits registered in DANBIO during the first year of treatment was similar for all three drugs (median = 3 visits). Furthermore, the median number of days until the 6- and 12-months visits were similar for the three drugs (6 months: 186-199 days, 12 months: 385-404 days).

Treatment responses at 6 and 12 months and one-year retention rates

At 6 months, DAS28 remission had been achieved in 35% of certolizumab pegol, 33% of abatacept, and 42% of CT-P13 treated patients. At 12 months, it was 35%, 31% and 35%, respectively (ITT analyses) (**Table 3**). Results for CDAI showed similar patterns (**Supplementary table S2**). One-year crude retention rates were 60% for certolizumab pegol, 57% for abatacept, and 69% for CT-P13. With certolizumab pegol as the reference drug, the estimated age- and gender adjusted OR for achieving DAS28 remission at 6 months was OR=0.95 (95% CI:0.65 to 1.39) for abatacept and OR=1.38, (0.95 to 1.99) for CT-P13. After 12 months the corresponding results were OR= 0.84 (0.56 to 1.24) and OR=1.04 (0.70 to 1.54), respectively. Similar estimates were found in confounder-adjusted analyses (**Figure 1**). For all analyses the CIs included 1. ”Worst case” and “best case“ imputations changed the estimates so whichever drug that was assigned “best case” was superior to the two other drugs (assigned “worst case”, **Supplementary table S3**).

Comparison of retention rates

The estimated risk of withdrawal for abatacept was 30% lower (HR:0.70 (95% CI: 0.39 to 1.27), fully adjusted model) than certolizumab pegol during the first 90 days of treatment. However, from 91 to 365 days the risk was 16% higher (HR:1.16 (0.84 to 1.60)), (**Table 4**). For CT-P13, the risk was 42% lower (HR:0.58 (0.33 to 1.10)) during the first 90 days and 17% lower from 91 to 365 days (HR:0.83 (0.59 to 1.17). All CIs included 1 (**Table 4**). After 6 months’ follow-up, median DAS28 was 2.9 (IQR: 2.0-3.9) for certolizumab pegol, 3.0 (2.0-3.9) for abatacept and 2.6 (2.0-3.8) for CT-P13 in patients still on treatment. Similarly, after 12 months DAS28 was 2.2 (1.7-3.0) for certolizumab pegol, 2.4 (1.8-3.3) for abatacept and

2.5 (1.7-3.5) for CT-P13. Reasons for discontinuation (lack of effect, adverse events) are registered routinely in DANBIO. The main reason for withdrawal was lack of effect (36-60%) followed by adverse events (15-42%), while other reasons were given in less than 10% of cases.

Sensitivity analyses restricted to patients fulfilling strict RCT in- and exclusion criteria

Adherence to guidelines, baseline characteristics, treatment responses and retention rates were similar in the analyses that excluded patients with severe comorbidities and included only seropositive patients with a DAS28 >3.2. In the restricted cohort, the baseline characteristics, treatment responses and retention rates were largely similar across the three drugs, and the CI were wide (**Supplementary table S4-S8)**.

DISCUSSION

In this paper we present how data from an observational cohort study of more than 700 bio-naïve patients with RA who initiated treatment with bDMARD according to national guidelines during the years 2013 to 2016 could emulate an RCT study, which allowed a head-to-head comparison of certolizumab pegol, abatacept and CT-P13. Compliance to the guidelines was high. Although some differences in the estimated remission and retention rates were observed between the drugs, confidence intervals were wide and statistical significance was not reached.

RCTs are considered the gold standard for assessing the efficacy of bDMARDs. RCTs comparing the active drug with placebo or with another bDMARD have reported comparable treatment effects across the bDMARDs within each study.[5, 6, 21-23] However, inclusion criteria and definition of treatment responses have varied between RCT studies, which makes comparisons across studies difficult. Studies with indirect comparison of bDMARDs have reported equal effectiveness but with wide CIs and heterogeneity between studies.[24, 25] Observational cohort studies may be a valuable alternative when RCTs are considered too expensive or impossible to conduct.[26] Furthermore, it is of interest to study effectiveness in patients treated in routine care, as most of these patients would not have been eligible in an RCT.[27] Thus, in the present study the included patients represented a non-selected, nationwide cohort that started one of three bDMARDs within a limited time period, which enabled us to compare the three drugs directly. Historically, it has been claimed that the effect size is often higher in observational studies than in RCTs.[26] However, several recent studies comparing RCTs and observational studies could not confirm this.[1, 26, 28] Our study supports that a properly designed cohort study seems to produce results comparable to that of RCTs.

We took several methodological steps to emulate an RCT. Firstly, we took advantage of the national guidelines, which resulted in a surrogate randomization procedure in which the choice of bDMARD was defined by the calendar time, thereby reducing confounding by indication. Compared to the calendar periods prior to and after each of the three drugs being first choice, a steep decline in number of patients treated with the three drugs was observed, demonstrating that compliance to the guidelines was high. Furthermore, similar baseline characteristics in the three cohorts indicated a successful

"randomization". Secondly, to assess the impact of missing follow-up data, we analyzed data by an ITT-methodology, and also conducted best/worst case analyses. Thirdly, linkage of DANBIO to other national registries (e.g. the DNPR) allowed us to adjust for a wide range of potential confounders, including comorbidities. We did this in two ways: by statistical modelling and by restriction of the cohort by applying some common exclusion criteria in RCTs (i.e. excluding patients with severe comorbidities, low disease activity or seronegative status).[29] Overall, the results of the various sensitivity analyses were largely similar, thus confirming the robustness of results.

We took advantage of the fact that the treatment assignment was largely driven by the guidelines. Thus, confounding control started at the time of study design and not at the time of adjusted analyses.[14] As a result, patient characteristics at baseline were very similar across calendar years. Some patients did not adhere to the guidelines. We have previously investigated differences between those who adhered to guidelines and those who did not.[30] Patients adherent to the guidelines had higher DAS28 and patient global score, indicating that some selection of patients took place, which is similar to what is observed in the recruitment of patients for an RCT.[31] Shortly after CT-P13 was marketed in Denmark, the guidelines were revised and CT-P13 became the first choice (expected compliance 50%) whereas certolizumab became the second choice (with an expected compliance of 30%). We decided not to include the latter cohort in our emulated RCT since patient numbers were low and certolizumab at this point in time was only the second choice.

This study has strengths and limitations. High quality clinical prospective data from a nationwide rheumatologic registry, DANBIO, was enriched with data from a virtually complete national registry. Although we used ITT and best/worst case imputation to handle missing data, missingness still occurred. This is an inherent challenge of observational studies due to loss of follow-up and incomplete registrations. Some channeling cannot be ruled out. In contrast to what is the case in an RCT, the physician had other potential treatment alternatives. Since compliance to the guideline was not expected to be 100%, it was possible e.g. to prescribe rituximab rather than the recommended first choice to patients with a prior cancer. We found that CT-P13 patients were slightly older. Subcutaneously administered bDMARDs (i.e. abatacept and certolizumab pegol) may be preferred by younger patients associated with the labor market, because these treatments are less time consuming and require fewer hospital visits compared to intravenous treatments (i.e. CTP-13). A study from the Swedish Rheumatology Quality register (SRQ) showed that patients with poor health-status were channeled to non-TNFI bDMARDs, which made the non-TNFi bDMARDs appear more harmful compared to TNFi.[32] This may, at least in part, explain the slightly lower estimates of treatment responses found for abatacept (a non-TNFi bDMARD) in most of our analyses.

In conclusion, in this observational study of more than 700 patients with RA treated according to the national Danish guidelines, we showed how a surrogate randomization procedure (calendar year), enabled direct comparison of the effectiveness of certolizumab pegol, abatacept and CT-P13. Although some differences in the estimated remission and retention rates across the three drugs were observed, the confidence intervals were wide, and no statistically significant differences were found.

Acknowledgement

Thanks to Laura Köcher from Karolinska Institutet for assistance with the figure. Thanks to the Danish Departments of Rheumatology that contributed data.

Funding

The DANBIO registry has entered into agreements with AbbVie, Biogen, BMS, Eli Lilly, MSD, Novartis, Pfizer, Roche, UCB. They receive post-marketing data and had no influence on the data collection, statistical analyses, manuscript preparation or decision to submit.

REFERENCES

1. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. The New England journal of medicine. 2000;342(25):1887-92.
2. Lipsky PE, van der Heijde DM, St Clair EW, Furst DE, Breedveld FC, Kalden JR, et al. Infliximab and methotrexate in the treatment of rheumatoid arthritis. Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy Study Group. The New England journal of medicine. 2000;343(22):1594-602.
3. Kavanaugh A, Fleischmann RM, Emery P, Kupper H, Redden L, Guerette B, et al. Clinical, functional and radiographic consequences of achieving stable low disease activity and remission with adalimumab plus methotrexate or methotrexate alone in early rheumatoid arthritis: 26-week results from the randomised, controlled OPTIMA study. Annals of the rheumatic diseases. 2013;72(1):64-71.
4. Nam JL, Takase-Minegishi K, Ramiro S, Chatzidionysiou K, Smolen JS, van der Heijde D, et al. Efficacy of biological disease-modifying antirheumatic drugs: a systematic literature review informing the 2016 update of the EULAR recommendations for the management of rheumatoid arthritis. Annals of the rheumatic diseases. 2017;76(6):1113-36.
5. Schiff M, Keiserman M, Codding C, Songcharoen S, Berman A, Nayiager S, et al. Efficacy and safety of abatacept or infliximab vs placebo in ATTEST: a phase III, multi-centre, randomised, double-blind, placebo-controlled study in patients with rheumatoid arthritis and an inadequate response to methotrexate. Annals of the rheumatic diseases. 2008;67(8):1096-103.
6. Schiff M, Weinblatt ME, Valente R, van der Heijde D, Citera G, Elegbe A, et al. Head-to-head comparison of subcutaneous abatacept versus adalimumab for rheumatoid arthritis: two-year efficacy and safety findings from AMPLE trial. Annals of the rheumatic diseases. 2014;73(1):86-94.
7. Psaty BM, Koepsell TD, Lin D, Weiss NS, Siscovick DS, Rosendaal FR, et al. Assessment and control for confounding by indication in observational studies. Journal of the American Geriatrics Society. 1999;47(6):749-54.
8. Hernan MA, Alonso A, Logan R, Grodstein F, Michels KB, Willett WC, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. Epidemiology (Cambridge, Mass). 2008;19(6):766-79.

9. Ibfelt EH, Jensen DV, Hetland ML. The Danish nationwide clinical register for patients with rheumatoid arthritis: DANBIO. Clinical epidemiology. 2016;8:737-42.
10. Hetland ML. DANBIO--powerful research database and electronic patient record. Rheumatology (Oxford, England). 2011;50(1):69-77.
11. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. Scandinavian journal of public health. 2011;39(7 Suppl):30-3.
12. The Danish Council for the Use of Expensive Hospital Medicines http://www.regioner.dk/media/28308/radsfolder-engelsk.pdf.
13. Craig P, Cooper C, Gunnell D, Haw S, Lawson K, Macintyre S, et al. Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. Journal of epidemiology and community health. 2012;66(12):1182-6.
14. Walker A.M. PAR, Hornbrook M. C. Marin M. G., Platt R., Roger V. L., Stang P., Schneeweiss S., . A tool for assessing the feasibility of comparative effectiveness research. Dovepress. 2012;DOI https://doi.org/10.2147/CER.S40357.
15. Hernan MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. American journal of epidemiology. 2016;183(8):758-64.
16. Hernan MA, Sauer BC, Hernandez-Diaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. Journal of clinical epidemiology. 2016;79:70-5.
17. Medicinrådet. https://medicinraadet.dk/anbefalinger-og-vejledninger.
18. Fransen J, Creemers MC, Van Riel PL. Remission in rheumatoid arthritis: agreement of the disease activity score (DAS28) with the ARA preliminary remission criteria. Rheumatology (Oxford, England). 2004;43(10):1252-5.
19. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis and rheumatism. 1995;38(1):44-8.
20. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Journal of chronic diseases. 1987;40(5):373-83.
21. Weinblatt ME, Schiff M, Valente R, van der Heijde D, Citera G, Zhao C, et al. Head-to-head comparison of subcutaneous abatacept versus adalimumab for rheumatoid arthritis: findings of a phase IIIb, multinational, prospective, randomized study. Arthritis and rheumatism. 2013;65(1):28-38.
22. Atsumi T, Yamamoto K, Takeuchi T, Yamanaka H, Ishiguro N, Tanaka Y, et al. The first double-blind, randomised, parallel-group certolizumab pegol study in methotrexate-naive early rheumatoid arthritis patients with poor prognostic factors, C-OPERA, shows inhibition of radiographic progression. Annals of the rheumatic diseases. 2016;75(1):75-83.
23. Smolen JS, Burmester GR, Combe B, Curtis JR, Hall S, Haraoui 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 (London, England). 2016;388(10061):2763-74.
24. Guyot P, Taylor P, Christensen R, Pericleous L, Poncet C, Lebmeier M, et al. Abatacept with methotrexate versus other biologic agents in treatment of patients with active rheumatoid arthritis despite methotrexate: a network meta-analysis. Arthritis research & therapy. 2011;13(6):R204.
25. Singh JA, Hossain A, Tanjong Ghogomu E, Mudano AS, Maxwell LJ, Buchbinder R, et al. Biologics or tofacitinib for people with rheumatoid arthritis unsuccessfully treated with biologics: a systematic review and network meta-analysis. The Cochrane database of systematic reviews. 2017;3:Cd012591.

26. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. The New England journal of medicine. 2000;342(25):1878-86.
27. Vashisht P, Sayles H, Cannella AC, Mikuls TR, Michaud K. Generalizability of Patients With Rheumatoid Arthritis in Biologic Agent Clinical Trials. Arthritis care & research. 2016;68(10):1478-88.
28. Rothman KJ. Six persistent research misconceptions. Journal of general internal medicine. 2014;29(7):1060-4.
29. Schneeweiss S, Patrick AR, Sturmer T, Brookhart MA, Avorn J, Maclure M, et al. Increasing levels of restriction in pharmacoepidemiologic database studies of elderly and comparison with randomized trial results. Medical care. 2007;45(10 Supl 2):S131-42.
30. Gron KL, Glintborg B, Nørgaard M, Mehnert F, Østergaard M, Dreyer L, Krogh N. S., Hetland M.L. . Mandatory choice of first bDMARD in Denmark - An opportunity to study real-life effectiveness? Results from the DANBIO registry. Annals of the Rheumatic Diseases, volume 76, supplement 2, year 2017, page 451. 2017;DOI: 10.1136/annrheumdis-2017-eular.4116

31. Aaltonen KJ, Ylikyla S, Tuulikki Joensuu J, Isomaki P, Pirila L, Kauppi M, et al. Efficacy and effectiveness of tumour necrosis factor inhibitors in the treatment of rheumatoid arthritis in randomized controlled trials and routine clinical practice. Rheumatology (Oxford, England). 2017;56(5):725-35.
32. Frisell T, Baecklund E, Bengtsson K, Di Giuseppe D, Forsblad-d'Elia H, Askling J. Patient characteristics influence the choice of biological drug in RA, and will make non-TNFi biologics appear more harmful than TNFi biologics. Annals of the rheumatic diseases. 2018;77(5):650-7.

Table 1 Prescription patterns of first bDMARD given in combination with MTX in the three calendar periods

	Numbers (%) of RA patients who started bDMARD in combination with MTX		
	Jan 2013- June 2014	**July 2014-June 2015**	**July 2015- June 2016**
Abatacept	6 (1%)	**215 (65%)**	23 (6%)
Adalimumab	29 (6%)	20 (6%)	12 (3%)
Certolizumab pegol	**336 (70%)**	31 (9%)	62 (16%)
Etanercept (originator)	46 (10%)	8 (3%)	1 (0.5%)
Etanercept (biosimilar, SB4)	NA	NA	14 (4%)
Golimumab	10 (2%)	3 (1%)	0
Infliximab (originator)	12 (3%)	4 (1%)	1 (0.5%)
Infliximab (biosimilar, CT-P13)	NA	7 (2%)	**225 (59%)**
Rituximab	22 (5%)	14 (4%)	15 (4%)
Tocilizumab	12 (3%)	27 (8%)	26 (7%)
Total	474 (100%)	329 (100%)	379 (100%)

Bold: first choice according to national guidelines. NA: Not applicable (i.e. the treatment not available in calendar period)

Table 2 Baseline characteristics of RA patients following the Danish national guidelines during one of the three calendar periods and included in the comparative effectiveness study

	Certolizumab pegol N=336		**Abatacept N=215**		**CT-P13 N=225**	
Calendar periods	Jan 2013 – June 2014		July 2014-June 2015		July 2015-June 2016	
Expected guideline compliance (%)	80%		80%		50%	
Actual guideline compliance (%)	70%		65%		59%	
Baseline characteristics	N		N		N	
Age, years	336	57 (48-65)	215	57 (48-65)	225	59 (50-66)
Female, n (%)	336	240 (71%)	215	155 (72%)	225	153 (68%)
Anti-CCP positive, n (%)	248	184 (74%)	168	129 (77%)	174	123 (71%)
Disease duration, years	318	3 (1-10)	207	3 (1-9)	212	4 (2-10)
Smoking current, n (%)	279	55 (20%)	191	21 (11%)	204	36 (18%)
Glucocorticoids, n (%)	336	148 (44%)	215	118 (55%)	225	123 (55%)
Physician global, mm	309	31 (18-46)	194	26 (17-41)	211	26 (16-42)
Pain VAS, mm	320	58 (38-75)	197	63 (44-75)	214	61 (34-75)
Fatigue VAS, mm	320	69 (44-82)	195	69 (40-84)	214	64 (40-82)
DAS28	312	4.6 (3.8-5.3)	193	4.5 (3.9-5.3)	206	4.5 (3.7-5.3)
CDAI	307	20 (14-28)	191	19 (13-26)	205	19 (13-26)
HAQ	318	1.1 (0.6-1.8)	194	1.0 (0.6-1.6)	211	1.0 (0.6-1.6)
Patient global VAS, mm	324	70 (49-84)	197	73 (51-84)	214	68 (48-85)
CRP, mg/L	319	8 (3-18)	197	9 (3-20)	214	9 (4-20)
CCI Score ≥ 1, n (%)	334	72 (22%)	213	35 (16%)	222	53 (24%)

Variables are medians (IQR) unless otherwise stated. Abbreviations: CCI: Charlson Comorbidity Index, CRP: C-reactive protein, DAS28: Disease activity score in 28 joints; HAQ health assessment questionnaire; n: number of persons on which estimates are based; VAS: visual analogue scale

Table 3 DAS28 remission rates at 6 and 12 months stratified by treatment, ITT analyses

	Certolizumab pegol	Abatacept	CT-P13
Number of patients	336	215	225
DAS28 remission, 6 months, n* (%)	116 (35%)	72 (33%)	94 (42%)
DAS28 remission, 12 months, n* (%)	118 (35%)	67 (31%)	78 (35%)

*Missing data on DAS28 remission: 6 months: certolizumab pegol: 40 patients, abatacept: 25, CT-P13:29, 12 months: certolizumab pegol: 74, abatacept: 51, CT-P13: 58.
Abbreviations: DAS28: disease activity score in 28 joints, ITT: intention to treat.

Table 4 Withdrawal during the first year of treatment. Results of adjusted Cox regression analyses

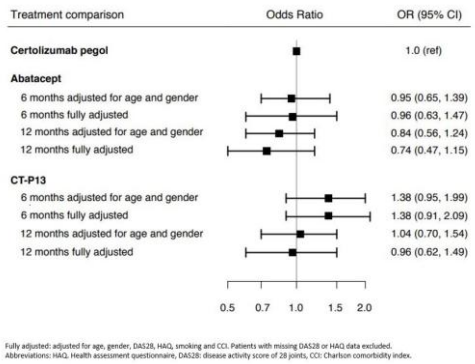
	N	HR (95% CI)		
		Certolizumab pegol	Abatacept	CT-P13
0-90 days				
Adjusted for age and gender	776	1.0 (ref)	0.78 (0.45 to 1.36)	0.63 (0.35 to 1.13)
Fully adjusted*	720	1.0 (ref)	0.70 (0.39 to 1.27)	0.58 (0.33 to 1.10)
91-365 days				
Adjusted for age and gender	702	1.0 (ref)	1.15 (0.85 to 1.56)	0.74 (0.53 to 1.04)
Fully adjusted*	652	1.0 (ref)	1.16 (0.84 to 1.60)	0.83 (0.59 to 1.17)

*age, gender, DAS28, HAQ, smoking, CCI
Abbreviations: CCI: Charlson Comorbidity Index: CI: confidence intervals, DAS28: Disease Activity Score of 28 joints, HAQ: Health assessment questionnaire, HR; hazard ratio

Legends to figures

Figure 1 Adjusted odds ratios (OR) for achieving DAS28 remission at 6 and 12 months.

Results of logistic regression analyses



Fully adjusted: adjusted for age, gender, DAS28, HAQ, smoking and CCI. Patients with missing DAS28 or HAQ data excluded.
Abbreviations: HAQ, Health assessment questionnaire; DAS28: disease activity score of 28 joints, CCI: Charlson comorbidity index.