RHEUMATOLOGY

Rheumatology 2022;00:1–11
https://doi.org/10.1093/rheumatology/keac081
Advance access publication 9 February 2022

Original article

Do patient-reported measures of disease activity in rheumatoid arthritis vary between countries? Results from a Nordic collaboration

Bénédicte Delcoigne [1], Sella Aarrestad Provan [2], Hilde Berner Hammer[2,3], Daniela Di Giuseppe [1], Thomas Frisell[1], Bente Glintborg [4,5], Gerdur Grondal[6], Bjorn Gudbjornson[6], Merete Lund Hetland [4,5], Brigitte Michelsen[2,7], Dan Nordström[8,9], Heikki Relas[8] and Johan Askling[1]

Abstract

Objectives. To investigate whether patient-reported outcomes vary across countries and are influenced by cultural/contextual factors. Specifically, we aimed to assess inter-country differences in tender joint count (TJC), pain and patient's global health assessment (PGA), and their impact on disease activity (DAS28-CRP) in RA patients from five Nordic countries.

Methods. We collected data (baseline, 3- and 12-months) from rheumatology registers in the five countries comprising RA patients starting a first ever MTX or a first ever TNF inhibitor (TNFi). In order to assess the role of context (=country), we separately modelled TJC, pain and PGA as functions of objective variables (CRP, swollen joint count, age, sex, calendar period and disease duration) with linear models. Analyses were performed at each time point and for both treatments. We further assessed the impact of inter-country differences on DAS28-CRP.

Results. A total of 27 645 RA patients started MTX and 19 733 started a TNFi. Crude inter-country differences at MTX start amounted to up to 4 points (28 points scale) for TJC, 10 and 27 points (0–100 scale) for pain and PGA, respectively. Corresponding numbers at TNFi start were 3 (TJC), 27 (pain) and 24 (PGA) points. All differences were reduced at 3- and 12-months, and attenuated when adjusting for the objective variables. The variation in predicted DAS28-CRP across countries amounted to <0.5 units.

Conclusions. Inter-country differences in TJC, pain and PGA are greater than expected based on differences in objective measures, but have a small clinical impact on DAS28-CRP across countries.

Key words: patient-reported outcome (PRO), pain, disease activity, rheumatoid arthritis, inter-country comparison

Rheumatology key messages

- There are inter-country differences in patient-reported outcomes across the five Nordic countries.
- Differences in objective measures (i.e. demographical, clinical and laboratory measures) do not fully explain differences in patient-reported measures.
- Whilst existing, the clinical impact of inter-country differences in patient-reported outcomes is limited.

[1]Department of Medicine Solna, Clinical Epidemiology Division, Karolinska Institutet, Stockholm, Sweden, [2]Division of Rheumatology and Research, Diakonhjemmet Hospital, [3]Faculty of Medicine, University of Oslo, Oslo, Norway, [4]The DANBIO registry and Copenhagen Center for Arthritis Research (COPECARE), Center for Rheumatology and Spine Diseases, Centre of Head and Orthopaedics, Copenhagen University Hospital Rigshospitalet, Glostrup, [5]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, [6]Centre for Rheumatology Research (ICEBIO), Landspitali University Hospital and Faculty of Medicine University of Iceland, Reykjavik, Iceland, [7]Division of Rheumatology, Department of Medicine, Hospital of Southern Norway Trust, Kristiansand, Norway, [8]Division of Medicine and Rheumatology, Helsinki University Hospital and [9]Department of Medicine, University of Helsinki, Helsinki, Finland

Submitted 21 September 2021; accepted 29 January 2022

Correspondence to: Bénédicte Delcoigne, Department of Medicine Solna, Karolinska Institutet, Clinical Epidemiology Division T2, Karolinska University Hospital, SE-171 76 Stockholm, Sweden.
E-mail: benedicte.delcoigne@ki.se

CLINICAL SCIENCE

Introduction

In complex chronic diseases such as RA, disease activity is often assessed using composite scores that combine patient-reported items, assessor-reported items and laboratory markers. For instance, the algorithm to calculate DAS28-CRP (i.e. the disease activity score based on 28 joint count and CRP) combines information on 28 tender joint count (TJC), 28 swollen joint count (SJC), CRP and the patient's global health assessment (PGA). In the algorithm, these items do not contribute equally; for example, TJC contributes almost twice as much as SJC [1].

In the era of treat-to-target strategies [2], RA disease activity scores such as the DAS28-CRP become critical for clinical decision-making. Further, in the era of global rheumatology, inferences from clinical trials and international collaborative studies presume that such composite scores are directly translatable across populations [2, 3]. At the same time, it is widely reported that there are profound cultural and contextual differences in how individuals experience, describe and report symptoms and experiences such as pain [4–7]. It is thus reasonable to assume that the reporting of TJC, and other patient-reported items such as visual analogue scale (VAS) pain measures and PGA, may be influenced by cultural and contextual factors. Such influence may limit the comparability of standard composite metrics such as DAS28-CRP and, in turn, limit inferences from clinical trials performed in (culturally or otherwise) different populations, and the prospects for collaborative research based on pooling of composite scores such as DAS28-CRP from different populations.

In the current study, we used data from rheumatology registers in the five Nordic countries on RA patients at comparable time-points in the course of their RA disease to evaluate whether there are systematic differences in patient-reported measures [here: TJC, pain (VAS) and PGA (VAS)] across the five countries, over and above those that would be predicted by differences in 'objective' measures (here: CRP, ESR and SJC), age, sex and calendar time. We further investigated the impact of any such difference on the predicted DAS28 values, and thus how any inter-country variation in subjective measures might influence the comparability of observed RA disease activity measures across these countries.

Subjects and methods

Data sources

We collected data from DANBIO in Denmark, ROB-FIN in Finland, ICEBIO in Iceland, NOR-DMARD in Norway and SRQ-ARTIS in Sweden [8–16]. These registers are all set up to monitor RA patients in clinical practice, and contain data on patient and disease characteristics, as well as visit data (disease activity, treatment, clinician- and patient-reported measures) [10]. Information regarding the criteria for being included in the registers and for starting TNFi treatment in the different countries are presented in Supplementary Table S1, available at *Rheumatology* online.

Study population

From each register, we assembled two independent, but potentially overlapping study cohorts for which we extracted visit data from 2008 to 2018: all RA patients who started: (i) a first ever treatment with MTX monotherapy that was also the first ever disease-modifying anti-rheumatic (DMARD) treatment (this population was not available in ICEBIO, and was limited to the calendar period 2008–2012 in NOR-DMARD, and 2015–2018 in ROB-FIN); and (ii) a first ever TNF inhibitor (TNFi) as the first ever biologic DMARD treatment during our study period.

Study variables

We used a common and harmonized study protocol. For each included subject and cohort, we collected data at three time points: at the start of the DMARD (MTX or TNFi, plus/minus 30 days) (= baseline), at 3 months (defined as the visit closest in time between 90 days and 150 days after the DMARD start) and at 12 months (defined as the visit closest in time between nine and 15 months after the DMARD start), irrespective of any treatment discontinuation. We retrieved baseline data on sex, age (categorized as 18–49, 50–74, 75+ years), birth decade, calendar period of treatment start (2008–2011, 2012–2014, 2015–2018), RA symptom duration (i.e. self-reported time since first RA symptom onset, <1 year, 1–2 years, >2 years) and RA disease duration (i.e. time since the clinical RA diagnosis, <1 year, 1–2 years, >2 years for MTX starters, and <2 years, 2–5 years, 5–10 years, >10 years for TNFi starters). In addition, the following variables were collected at baseline, 3 and 12 months: pain, PGA, fatigue and assessor's global health (AGA) (all four on 0–100 mm VAS scales), SJC, TJC (both on a 28 scale), CRP level, ESR (not available in Denmark and Iceland), DAS28-CRP and a measure of functional status [either the full or the modified HAQ (HAQ or mHAQ)].

Statistical analyses

Baseline demographic and clinical variables (medians and interquartile ranges), and descriptive statistics for TJC, SJC, CRP, pain, PGA and DAS28-CRP at 3 and 12 months were displayed for each country. In the following, we consider TJC, pain and PGA as 'subjective' measures, which were analysed using CRP, SJC, sex, age at treatment start, disease duration (only TNFi data), birth decade and calendar period as 'objective' measures. Percentage of missingness is reported together with the descriptive statistics. We did not impute missing data.

Crude and adjusted inter-country differences

All analyses were performed in each of the two study cohorts (initiators of a first ever MTX and a first ever TNFi, respectively). In unadjusted analyses, baseline TJC, pain and PGA were compared across countries with ANOVA. Next, at each time-point (baseline, 3 and 12 months), we used linear regression to analyse the distribution of the subjective measures: (i) with 'country' as the unique independent variable (equivalent to a one way ANOVA); and (ii) additionally adjusting for sex, birth decade, baseline age, calendar period of DMARD start, CRP and SJC. In the analyses of the TNFi initiator cohort, we additionally adjusted for RA disease duration. For all analyses, we log-transformed TJC, SJC and CRP, which all had a skewed distribution, in order to ensure the achievement of the assumptions required by linear regression models. We used the country with the largest number of included treatment initiators (Sweden) as reference. In addition to the country's coefficient obtained from the linear model, we report the percentage of the total variance explained by the model and the percentage specifically explained by the 'country' variable.

Impact of intercountry differences on predicted DAS28-CRP

The impact of the inter-country differences in subjective markers on DAS28-CRP was assessed as follows: we predicted TJC and PGA values at baseline, 3 and 12 months, using the linear regression models by replacing the value of the 'country' variable with another country (e.g. assuming that Finnish patients would behave as Danish patients given a certain combination of objective variables). For each time point and for each of the country-specific study populations we thus generated four predicted TJC and PGA values for the MTX dataset (as we had four contributing countries) and five predicted TJC and PGA values for the TNFi dataset (as we had five contributing countries). We incorporated these in the algorithm for computing 'predicted' values of DAS28-CRP for each country. These means are displayed graphically for visual comparison.

Statistical programs

SAS (V.9.4) was used for assembling and preparing the data as well as for the descriptive analyses. Linear regressions were performed and graphs made in R (version 4.0.2). The assumptions required by linear regression models (linear relationship between outcome and independent variables, normality of the distribution of the residuals, homoscedasticity and lack of outliers' influence) were graphically tested.

Ethics and data protection

The study data were irreversibly anonymized before pooling, and exported and analysed at Karolinska Institutet (Sweden). Approval from the data protection agencies and registry holders, and/or ethics approvals were provided from the relevant authorities in each country.

Results

MTX initiators

The MTX initiator cohort comprised visit data for 27 645 RA patients (Denmark: 6558; Finland: 607; Norway: 542; and Sweden: 19 938). Data were collected for slightly different calendar time periods in different countries (depending on available data and registration practice). In all countries, around two-thirds were female, most were 50–74 years old (Table 1). Disease duration was on average higher in Finland than in the other countries.

The descriptive statistics at baseline demonstrated statistically significant (ANOVA) and for certain variables also clinically meaningful inter-country differences (Table 1). For instance, we noted large variations for TJC (medians ranging from 2 in Finland to 6 in Denmark), SJC (medians ranging from 2 in Finland to 4 in Denmark, Norway and Sweden), and PGA (medians ranging from 33 in Finland to 57 in Denmark), and pain, though less markedly (medians ranging from 40 in Finland to 50 in Denmark) (Table 1).

Table 2 displays mean crude and adjusted differences in TJC, pain and PGA between countries, at baseline, 3 and 12 months, obtained from linear models. For TJC, the crude mean differences at baseline [up to 3 points (out of 28), between Finland on the one hand and Denmark and Norway on the other hand] were reduced to <2 points following adjustment. For pain VAS, the crude differences were clinically small [up to 8 points (out of 100), between Finland and Denmark at baseline] and generally were little affected by adjustment at any time point. For PGA, the highest crude mean difference was 17 (out of 100) (between Finland and Denmark at baseline) and were generally little affected by adjustment.

Overall, for all time points, whereas our fully adjusted models explained 48%, 36% and 35% of the variance of TJC at baseline, 3 and 12 months, respectively, the contribution of the variable 'country' to the explained variance was below 2%. For pain, the corresponding numbers were 21%, 17%, 16% for the explained variance, and around or <1% at all time points for the variable 'country' contribution. The corresponding numbers for PGA were 22%, 17%, 15%, with the contribution of the variable 'country' being 2%, 3% and 2%, respectively. No sign of deviation from the linear regression assumptions was identified.

Fig. 1 displays the mean predicted DAS28-CRP values at baseline, 3 and 12 months, assuming that patients in each country would behave as if they were representative of each of the other countries. The solid shapes in this figure are the mean predicted DAS28-CRP values for the patients in their own country (which, in absence of missing values, will be equal to the observed means). Among these solid shapes at

Table 1 Descriptive statistics of RA patients' demographic and clinical characteristics (first ever MTX)

	Sweden	Denmark[a]	Finland	Norway
n	19 938	6558	607	542
n women (%)	13 799 (69.2)	4370 (66.6)	407 (67.1)	351 (64.8)
Median birth decade	1941–1950	1951–1960	1951–1960	1951–1960
Age at MTX start				
18–49 years	4193 (21.0)	1528 (23.3)	111 (18.3)	154 (28.4)
50–74 years	12 379 (62.1)	4087 (62.3)	374 (61.6)	346 (63.8)
75+ years	3366 (16.9)	943 (14.4)	122 (20.1)	42 (7.7)
MTX start year				
2008–2011	6536 (32.8)	1376 (21.0)	0 0	514 (94.8)
2012–2014	6375 (32.0)	2178 (33.2)	0 0	28 (5.2)
2015–2018	7027 (35.2)	3004 (45.8)	607 (100.0)	0 0
Symptoms duration				
<1 year	961 (48.4)	2177 (44.3)	n/a	n/a
1–2 years	2537 (12.8)	1620 (33.0)	n/a	n/a
>2 years	7725 (38.9)	1112 (22.7)	n/a	n/a
% missing	0	25	100	100
Disease duration				
<1 year	18 831 (94.5)	5602 (92.0)	264 (50.2)	489 (90.4)
1–2 years	460 (2.3)	486 (8.0)	25 (4.8)	17 (3.1)
>2 years	644 (3.2)	0	237 (45.1)	35 (6.5)
% missing	0	7	13	0
Baseline				
Current use of corticosteroids				
Yes (%)	10505 (52.7)	3125 (47.7)	294 (48.4)	353 (65.1)
No (%)	9433 (47.3)	3433 (52.3)	313 (51.6)	174 (32.1)
% missing	0	0	0	2.8
CRP	6.6 [3.0–18.0]	9.0 [3.0–22.0]	5.0 [3.0–14.0]	8.0 [4.0–21.0]
% missing	4	0	8	4
ESR	20.0 [10.0–35.0]	n/a	14.5 [7.0–28.0]	22.0 [12.0–37.0]
% missing	9	–	10	11
28 SJC	4.0 [1.0–8.0]	4.0 [1.0–7.0]	2.0 [0.0–6.0]	4.0 [2.0–8.0]
% missing	2	0	35	0
28 TJC	3.0 [1.0–8.0]	6.0 [2.0–10.0]	2.0 [0.0–5.0]	5.0 [2.0–10.0]
% missing	3	0	35	1
Pain (VAS)	45 [20–67]	50 [28–70]	40 [21–62]	42 [21–63]
% missing	11	4	16	2
PGA (VAS)	45 [20–65]	57 [32–79]	33 [14–60]	44 [25–64]
% missing	10	0	21	2
DAS28-CRP	3.9 [2.7–5.0]	4.4 [3.5–5.3]	3.5 [2.5–4.5]	4.3 [3.4–5.2]
% missing	13	0	31	5
AGA (VAS)	30 [12–50]	29 [16–46]	21 [10–39]	31 [21–47]
% missing	86	9	27	3
Fatigue (VAS)	44 [17–68]	51.0 [25–71]	n/a	38.0 [12–64]
% missing	51	5	100	2
HAQ[b]	0.8 [0.4–1.3]	0.9 [0.4–1.4]	0.9 [0.4–1.4]	0.5 [0.1–0.9]
% missing	16	5	38	2
3 months				
CRP	4.0 [2.0–8.0]	4.0 [2.0–10.0]	3.0 [2.0–7.0]	5.0 [2.0–8.0]
% missing	54	20	55	30
28 SJC	1.0 [0.0–3.0]	0.0 [0.0–2.0]	0.0 [0.0–2.0]	1.0 [0.0–3.0]
% missing	53	20	64	22
28 TJC	1.0 [0.0–4.0]	1.0 [0.0–4.0]	0.0 [0.0–2.0]	2.0 [0.0–5.0]
% missing	53	20	64	23
Pain (VAS)	22 [7–45]	26 [11–48]	23 [5–47]	19 [8–38]
% missing	55	23	58	24
PGA (VAS)	24 [8–48]	32 [13–58]	15 [4–38]	23 [10–42]
% missing	54	21	58	24
DAS28-CRP	2.7 [1.9–3.7]	2.7 [1.9–3.7]	2.4 [1.6–3.2]	2.9 [2.1–3.9]
% missing	56	25	63	30

(continued)

Table 1 Continued

	Sweden	Denmark[a]	Finland	Norway
12 months				
CRP	4.0 [2.0–7.0]	3.3 [1.7–8.0]	3.0 [2.0–7.0]	3.0 [1.0–6.0]
% missing	50	21	57	44
28 SJC	0.0 [0.0–2.0]	0.0 [0.0–1.0]	0.0 [0.0–1.0]	0.0 [0.0–2.0]
% missing	50	22	66	37
28 TJC	1.0 [0.0–3.0]	0.0 [0.0–2.0]	0.0 [0.0–2.0]	1.0 [0.0–4.0]
% missing	50	22	66	36
Pain (VAS)	24 [8–49]	24 [9–49]	28 [8–51]	17 [5–39]
% missing	52	24	59	38
PGA (VAS)	25 [9–49]	29 [10–56]	20 [5–42]	21 [7–42]
% missing	52	22	59	38
DAS28-CRP	2.5 [1.8–3.4]	2.3 [1.7–3.3]	2.2 [1.6–3.0]	2.5 [1.8–3.5]
% missing	53	28	66	44

Variables are measured in four Nordic countries at MTX start (baseline), and at 3 and 12 months. Medians [percentile 25 and 75] or number (%) are displayed. 28 SJC: 28 swollen joint count; 28 TJC: 28 tender joint count; AGA: assessor's global health assessment; DAS28-CRP: disease activity score 28 with CRP; mHAQ: modified HAQ; PGA: patient's global health assessment; VAS: visual analogue scale. [a]modified HAQ in Norway. [b]Patients in Denmark were required to have available baseline visit and at least one visit during 1 year follow-up and <2 years since diagnosis.

baseline, the country associated with the highest DAS28-CRP was Denmark (solid triangle) while Finland was associated to the lowest (solid diamond), which also was the case in Table 1 where observed values are displayed. Comparing similar shapes to each other (i.e. horizontal comparisons), reveals how the DAS28-CRP values of the patients of a given country (for example, Norway looking at squares) would have been assessed should the same patients have been assessed in another country. In this comparison at baseline and 3 months, Denmark was the country where the means of the predicted values were the highest; the lowest were in Finland (i.e. for any given shape at baseline and 3 months, the highest one was in Denmark and the lowest one in Finland). Though in terms of absolute values, the difference between these two extremes was <0.4 DAS28-CRP units at baseline, around 0.3 at 3 months and less than that at 12 months, hence of little clinical relevance. A vertical comparison highlights the differences in covariates distributions (i.e. the ones used in the linear model) between countries. In this comparison, Norwegian patients were showing characteristics that led them to having the highest mean DAS28-CRP and the Finnish patients the lowest, at baseline and 3 months. Twelve-month values displayed a much lower variation.

TNFi initiators

The TNFi initiator cohort included 19 733 RA patients (Denmark: 5606; Finland: 1422; Iceland: 397; Norway: 1012; and Sweden: 11 296) and their demographic characteristics were similar to the ones of the MTX initiator cohort (Table 3).

Inter-country large variations were observed for TJC, SJC, PGA, pain and AGA at baseline (Table 3).

Crude and adjusted differences in mean TJC, pain and PGA across countries, obtained from linear models are displayed in Table 4 and show a similar pattern to that of MTX initiators in Table 2. In short, the highest inter-country differences were observed at baseline and were somewhat attenuated with adjustment for TJC and pain, but not modified for PGA. At 3 and 12 months, differences in TJC were meaningless (<1 unit). In contrast, while differences also were somewhat attenuated for pain and PGA, the largest differences (between Denmark and Norway) remained above 10 units, adjustments did not substantially reduce them. Overall, the proportion of variance explained by the fully adjusted model at baseline, 3 and 12 months was 39%, 31% and 34% for TJC, with the country variable explaining only 2%, 2% and 1% of the variance, respectively. For pain, the corresponding proportions were 16%, 15%, 17%, and 1%, 2%, 1%, respectively, and for PGA, these proportions were 18%, 16%, 17%, and 4%, 3%, 2%, respectively. The assumptions required for linear regressions were reasonably met.

Fig. 2 displays the five predicted DAS28-CRP values at baseline, 3 and 12 months for each country. Among the solid shapes (i.e. mean predicted values for patients in their own country) at baseline, the country associated with the highest DAS28-CRP was Iceland (solid reversed triangle), but this was no longer true at 3 and 12 months. The horizontal comparison of similar shapes shows that, at baseline, Denmark and Iceland were the countries where the means of the predicted values were the highest while the lowest were in Finland, Norway and Sweden. In absolute values, these differences never exceeded 0.5 DAS28-CRP units, and were thus of little clinical relevance. Finally, the vertical comparison reveals that at baseline, Icelandic patients had characteristics that led them to have the highest mean DAS28-

Table 2 Mean differences in TJC, pain and PGA between countries (first ever MTX)

Baseline	Sweden	Denmark	Finland	Norway
Crude model[a]				
TJC (0–28)	ref	2.2 (2.0, 2.4)***	−0.9 (−1.2, −0.4)***	2.2 (1.6, 2.8)***
Pain (0–100)	ref	4.7 (3.9, 5.5)***	−2.6 (−5.0, −0.2)*	−1.2 (−3.5, 1.2)
PGA (0–100)	ref	11.8 (11.0, 12.5)***	−5.1 (−7.6, −2.6)***	1.3 (−1.0, 3.7)
Adjusted model[b]				
TJC (0–28)	ref	1.7 (1.5, 1.8)***	−0.1 (−0.5, 0.2)	1.0 (0.7, 1.4)***
Pain (0–100)	ref	2.7 (2.0, 3.4)***	−0.3 (−2.9, 2.3)	−3.1 (−5.3, −0.9)**
PGA (0–100)	ref	10.0 (9.3, 10.7)***	−3.3 (−6.0, −0.6)*	−0.9 (−3.1, 1.3)
3 month				
Crude model[a]				
TJC (0–28)	ref	0.0 (−0.1, 0.1)	−0.6 (−0.8, −0.3)***	0.7 (0.4, 1.0)***
Pain (0–100)	ref	3.1 (2.2, 3.9)***	0.6 (−2.5, 3.7)	−2.9 (−5.4, −0.4)*
PGA (0–100)	ref	7.3 (6.4, 8.1)***	−5.9 (−9.1, −2.7)***	−1.3 (−3.9, 1.3)***
Adjusted model[b]				
TJC (0–28)	ref	0.4 (0.3, 0.5)***	−0.1 (−0.3, 0.2)	0.6 (0.3, 0.8)***
Pain (0–100)	ref	4.5 (3.7, 5.3)***	−0.4 (−3.7, 3.0)	−2.3 (−4.7, 0.2)†
PGA (0–100)	ref	7.0 (6.1, 7.9)***	−8.5 (−12.4, −4.6)***	−0.1 (−2.6, 2.8)
12 month				
Crude model[a]				
TJC (0–28)	ref	−0.2 (−0.2, −0.1)***	−0.4 (−0.6, −0.1)**	0.2 (0.0, 0.5)*
Pain (0–100)	ref	0.4 (−0.4, 1.3)	1.8 (−1.4, 5.0)	−4.6 (−7.4, −1.8)**
PGA (0–100)	ref	6.5 (5.6, 7.3)***	−4.7 (−8.0, −1.3)**	−2.9 (−5.8, 0.0)*
Adjusted model[b]				
TJC (0–28)	ref	0.2 (0.1, 0.3)***	−0.2 (−0.4, 0.1)	0.4 (0.2, 0.6)***
Pain (0–100)	ref	2.9 (2.0, 3.8)***	1.4 (−2.6, 5.4)	−1.1 (−3.9, 1.7)
PGA (0–100)	ref	6.3 (5.6, 7.2)***	−6.2 (−10.4, −2.1)**	0.8 (−2.1, 3.7)

The crude and adjusted differences are calculated at MTX start (baseline), and at 3 and 12 months, using the largest country (Sweden) as reference and are displayed with 95% CIs. *P*-values: †<0.10; *<0.05; **<0.01; ***<0.001. [a]Crude models include the variable country only. [b]Adjusted models additionally include sex, birth decade, age at treatment start, calendar period, log (CRP + 1) and log (SJC + 1). PGA: patient's global health assessment; SJC: 28 swollen joint count; TJC: 28 tender joint count.

CRP, but this was no longer the case at 3 and 12 months.

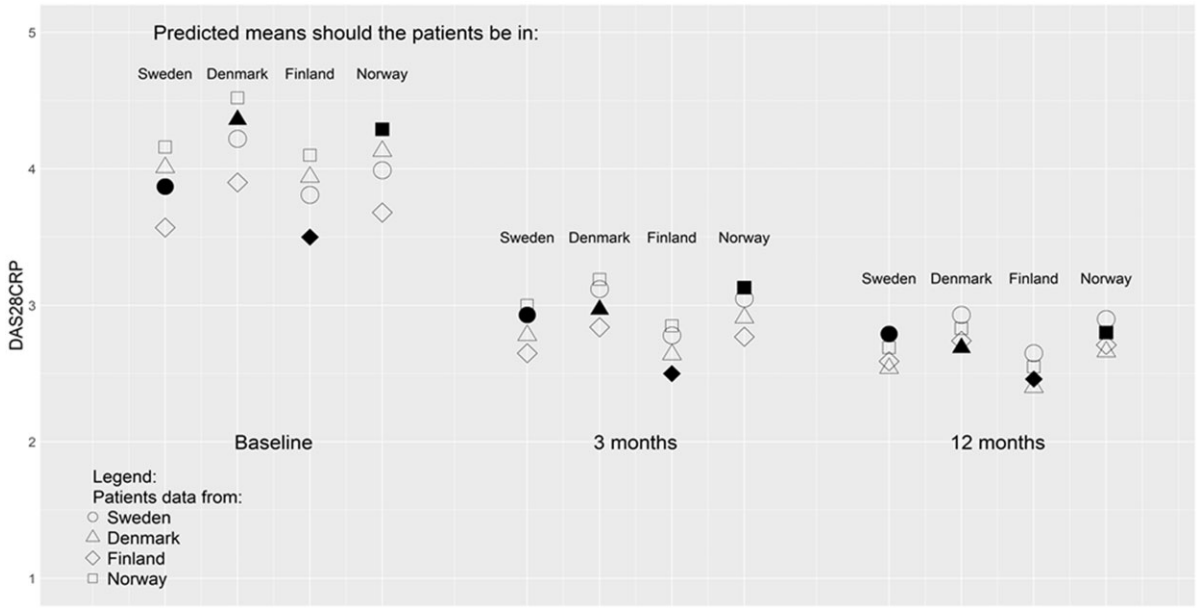
Discussion

Based on visit data from almost 50 000 treatment initiations registered in five different clinical rheumatology registers, we noted that patient-reported data such as TJC, pain and PGA differed across countries. Whereas some of these differences were typically small (e.g. TJC at 12 months for MTX and TNFi starters, pain at all time points for MTX starters), others were more substantial (e.g. baseline TJC, and PGA at almost all time points). Adjustment for objective markers of RA disease activity and context reduced these differences for TJC and pain but had little effect on the inter-country differences observed for PGA. Subsequently, the inter-country differences in TJC and PGA beyond what was explained by objective markers led to a variation in predicted DAS28-CRP values between countries, which reached up to (but not above) 0.5 DAS28-CRP units.

As DAS28-CRP is critical for clinical decision-making in the era of treat-to-target strategy, our study aimed to assess if the TJC and PGA items were as objective as they usually are considered to be. Our results indicate that, in the context of the Nordic countries, the inter-country differences are below the threshold of clinical relevance, and therefore do not threaten the validity of the inter-country comparison of composite metrics such as DAS28-CRP, nor do they limit the prospects for collaborative research based on pooling of these metrics from different Nordic populations.

A substantial body of literature has identified differences in pain reporting between ethnic and national groups, though most studies have been conducted in the US [5]. Some studies have compared pain perception and its report in populations who culturally differed markedly [6], such as Dutch/Egyptian [7], or Japanese/Belgian [17], and found noticeable differences. In our study, the inter-country differences, once accommodating the objective variables, remained limited, and were of modest (if any) clinical relevance. Our results thus indicate that, by and large, Nordic populations are rather similar to each other regarding pain aspects, and therefore also that pooling of individual-level data across these countries does not appear to introduce substantial bias arising from reporting differences.

FIG. 1 Mean predicted values of baseline, 3 and 12 months DAS28-CRP for each country population initiating a first ever MTX treatment



The four solid shapes are the mean predicted DAS28-CRP values for patients in their own country. Similar shapes (i.e. displayed horizontally, not solid), show the means of the DAS28-CRP values that were predicted for the population of a given country, should this population be assessed in another country. Denmark was the country where the means of the predicted values at baseline and 3 months were the highest, as for any given shape, the highest one was in Denmark.

Our study has some limitations. We selected patients defined via a first treatment start (either MTX as the first csDMARD or TNFi as the first bDMARD) aiming to include patients who were similar across countries according to key disease parameters such as disease duration and physical function. However, as the decision to start the treatment is not independent of our outcome variable (i.e. TJC, pain and PGA), and as this dependency may differ by country, the fully objective similarity of the patients at baseline (but perhaps less so during follow-up) is, at least partially, not possible to ensure. Also, whereas the Nordic biologics registers were launched for following all patients treated with bDMARDs, we cannot formally exclude the possibility that the registration systems for patients initiating treatment with MTX may have included subsets of patients not fully representative of the MTX-treated RA population. As an example, the MTX monotherapy cohort of patients in Finland only included mild cases because patients with substantial disease activity levels are treated with combination of csDMARD in this country. While we had information on important variables for running the linear models, we lacked information on variables such NSAIDs, comedication, socio-economic factors and comorbidities that also could play a role in TJC, pain and PGA values. To what extent socio-economic factors and comorbidities (such as depression) could have modified our results is not quantifiable, but we don't suspect the prevalence of these characteristics largely differing between Nordic countries. Another limitation is that some variables such as age were broadly categorized in order to preserve data privacy, preventing a fully adequate adjustment in the models. Some variables (mainly at 3 and 12 months) also were characterized by a substantial amount of missing values. TJC was characterized by a highly skewed distribution. To account for this, and for better ensuring the achievement of assumptions required by linear regression models, we log-transformed TJC + 1 for running the models, and back-transformed the TJC predicted values for calculating the predicted DAS28-CRP. These successive mathematical manipulations, together with missing values for some covariates resulted in some discrepancies between the observed and the predicted means, mainly at the 3- and 12-month assessments. The size of these differences did, however, not impact the interpretation of our results. Another limitation is the large difference in the numbers of patients included from the different countries. This could impact the precision of the estimates and thus the power to detect significant differences, but will not impact the mean values themselves, which guarantees the validity of the above interpretations. As the predicted values are, due to the modelling, usually less spread than the originally observed scores, we abstained from presenting, e.g. the impact of the inter-country differences on (predicted) disease activity categories (remission/low/moderate/high disease activity).

Table 3 Descriptive statistics of RA patients' demographic and clinical characteristics (first ever TNFi)

	Sweden	Denmark	Finland	Norway	Iceland
n	11296	5606	1422	1012	397
n women (%)	8450 (74.8)	4180 (74.6)	1035 (72.8)	722 (71.3)	288 (72.5)
Median birth decade	1951–1960	1951–1960	1951–1960	1951–1960	1951–1960
Age at TNFi start					
18–49 years	3513 (31.1)	1844 (32.9)	551 (38.7)	347 (34.3)	159 (40.1)
50–74 years	7057 (62.5)	3442 (61.4)	832 (58.5)	626 (61.9)	217 (54.7)
75+ years	726 (6.4)	320 (5.7)	39 (2.7)	39 (3.9)	21 (5.3)
TNFi start year					
2008–2011	4166 (36.9)	3134 (55.9)	751 (52.8)	466 (46.0)	168 (42.3)
2012–2014	2987 (26.4)	1273 (22.7)	396 (27.8)	275 (27.2)	95 (23.9)
2015–2018	4143 (36.7)	1199 (21.4)	275 (19.3)	271 (26.8)	134 (33.8)
Disease duration					
<2 year	7424 (65.7)	303 (5.4)	310 (22.1)	262 (32.5)	25 (6.3)
2–5 years	2075 (18.4)	979 (17.5)	286 (20.4)	170 (21.1)	54 (13.6)
5–10 years	1311 (11.6)	4085 (72.9)	305 (21.8)	146 (18.1)	165 (41.6)
>10 years	483 (4.3)	239 (4.3)	500 (35.7)	228 (28.3)	153 (38.5)
(% missing)	0	0	1	20	0
Current use of corticosteroids					
yes	5104 (45.2)	1493 (26.6)	892 (62.7)	613 (60.6)	82 (20.7)
no	6192 (54.8)	4113 (73.4)	530 (37.3)	399 (39.4)	
% missing	0	0	0	0	79.3
CRP	6.0 [2.3–16.0]	9.0 [3.0–20.0]	8.0 [3.0–20.0]	5.0 [3.0–14.0]	7.5 [3.0–18.5]
% missing	6	21	12	6	62
ESR	16.0 [8.0–31.0]	n/a	14.0 [6.0–28.0]	17.0 [8.0–28.0]	n/a
% missing	12	–	14	17	100
28 SJC	4.0 [1.0–8.0]	3.0 [1.0–7.0]	4.0 [1.0–9.0]	3.0 [1.0–7.0]	6.0 [3.0–10.0]
% missing	6	24	17	5	61
28 TJC	4.0 [1.0–8.0]	6.0 [2.0–11.0]	4.0 [1.0–10.0]	4.0 [1.0–9.0]	7.0 [4.0–12.0]
% missing	6	24	17	5	61
Pain (VAS)	50 [27–70]	59 [37–75]	52 [29–70]	41 [23–65]	68 [50–81]
% missing	10	23	9	2	62
PGA (VAS)	50 [28–70]	66 [46–81]	50 [26–70]	48 [26–68]	72 [55–88]
% missing	9	21	11	3	61
DAS28-CRP	4.2 [3.1–5.1]	4.5 [3.7–5.4]	4.3 [3.2–5.2]	4.0 [3.1–4.9]	4.9 [4.1–5.7]
% missing	14	29	22	10	63
AGA (VAS)	30 [10–50]	31 [19–46]	35 [20–50]	32 [22–45]	60 [44–70]
% missing	86	29	21	13	62
Fatigue (VAS)	50 [24–73]	65 [43–80]	n/a	49 [24–73]	71.5 [52–85]
% missing	49	29	100	49	62
HAQ[a]	0.9 [0.4–1.4]	1.1 [0.6–1.6]	0.9 [0.3–1.4]	0.5 [0.1–0.9]	1.3 [0.8–1.9]
% missing	14	23	15	3	62
TNFi					
Adalimumab	1802 (16.0)	1172 (20.9)	405 (28.5)	94 (9.3)	9 (2.3)
Certolizumab pegol	1067 (9.4)	871 (15.5)	145 (10.2)	264 (26.1)	5 (1.3)
Etanercept	5051 (44.7)	1690 (30.1)	570 (40.1)	396 (39.1)	153 (38.5)
Golimumab	842 (7.5)	198 (3.5)	129 (9.1)	62 (6.1)	45 (11.3)
Infliximab	2534 (22.4)	1675 (29.9)	173 (12.2)	196 (19.4)	185 (46.6)
3 months					
CRP	4.0 [1.3–8.0]	4.0 [2.0–10.0]	4.0 [2.0–8.0]	3.0 [1.0–6.0]	3.0 [1.0–6.0]
%missing	49	31	46	27	66
28 SJC	1.0 [0.0–3.0]	0.0 [0.0–2.0]	1.0 [0.0–3.0]	1.0 [0.0–3.0]	1.0 [0.0–3.0]
%missing	49	31	52	24	63
28 TJC	1.0 [0.0–4.0]	1.0 [0.0–5.0]	1.0 [0.0–4.0]	1.0 [0.0–4.0]	1.0 [0.0–3.0]
%missing	49	31	52	24	63
Pain (VAS)	27 [10–52]	32 [14–60]	26 [10–54]	22 [8–45]	24.5 [10–51]
%missing	50	30	45	25	66
PGA (VAS)	30 [11–53]	41 [18–69]	24 [9–50]	23.5 [8–49]	35 [12–60]
% missing	50	29	48	26	65

(continued)

Table 3 Continued

	Sweden	Denmark	Finland	Norway	Iceland
DAS28-CRP	2.8 [2.0–3.8]	2.9 [2.1–4.1]	2.9 [1.9–3.8]	2.7 [1.9–3.6]	2.7 [1.8–3.8]
%missing	52	36	54	28	68
12 months					
CRP	4.0 [1.3–8.0]	4.0 [1.4–9.0]	4.0 [2.0–8.0]	3.0 [1.0–6.0]	3.0 [2.0–6.0]
%missing	47	30	49	42	51
28 SJC	1.0 [0.0–3.0]	0.0 [0.0–2.0]	0.0 [0.0–2.0]	0.0 [0.0–2.0]	0.0 [0.0–2.0]
%missing	46	30	54	40	51
28 TJC	1.0 [0.0–4.0]	1.0 [0.0–4.0]	1.0 [0.0–3.0]	1.0 [0.0–3.0]	1.0 [0.0–3.0]
%missing	47	30	54	40	50
Pain (VAS)	29.0 [11.0–55.0]	31.0 [14.0–58.0]	26.0 [10.0–55.0]	18.0 [6.0–38.0]	29.0 [11.0–54.0]
%missing	49	29	49	42	54
PGA (VAS)	30.0 [11.0–55.0]	38.0 [16.5–67.0]	24.0 [8.0–50.0]	19.0 [6.0–40.0]	32.0 [13.0–54.0]
% missing	48	28	51	42	51
DAS28-CRP	2.8 [1.9–3.8]	2.7 [1.9–3.8]	2.6 [1.8–3.5]	2.3 [1.7–3.3]	2.6 [1.8–3.7]
%missing	50	36	56	43	53

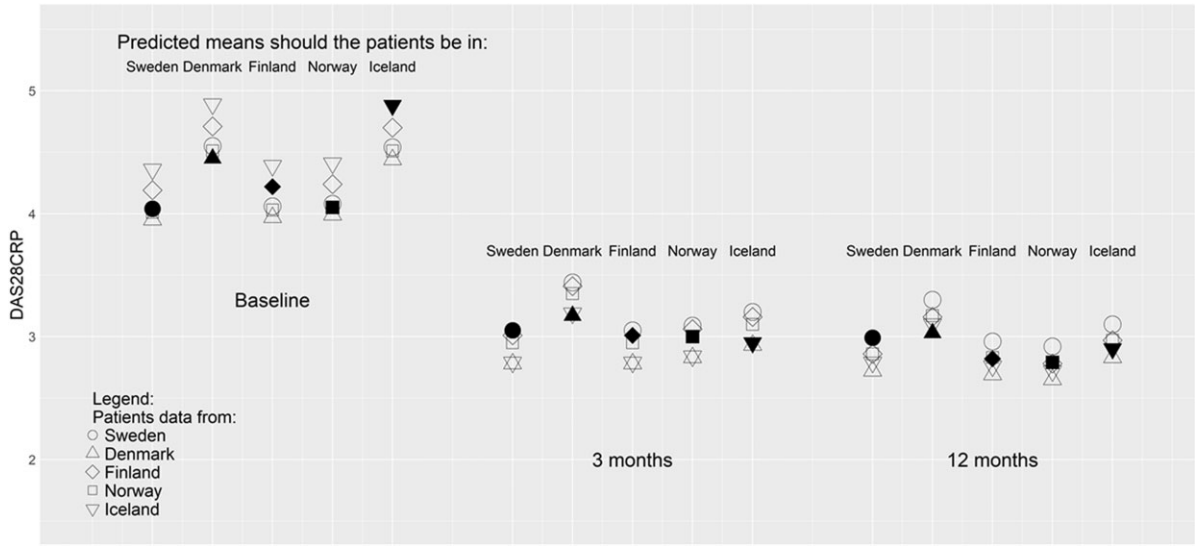
Variables are measured in five Nordic countries at TNFi start (baseline), and at 3 and 12 months. Medians [percentile 25 and 75] or number (%) are displayed. 28 SJC: 28 swollen joint count; 28 TJC: 28 tender joint count; AGA: assessor's global health assessment; DAS28-CRP: disease activity score 28 with CRP; mHAQ: modified HAQ; PGA: patient's global health assessment; TNFi: TNF inhibitor; VAS: visual analogue scale. [a]modified HAQ in Norway.

Table 4 Mean differences in TJC, pain and PGA between countries (first ever TNFi)

Baseline	Sweden	Denmark	Finland	Norway	Iceland
Crude model[a]					
TJC (0–28)	ref	1.7 (1.4, 1.9)***	0.3 (0.0, 0.7)	−0.1 (−0.5, 0.2)	3.2 (1.9, 4.6)***
Pain (0–100)	ref	7.2 (6.2, 8.1)***	0.3 (−1.2, 1.8)	−5.1 (−6.8, −3.4)***	14.3 (10.1, 18.5)***
PGA (0–100)	ref	13.0 (12.0, 13.9)***	−0.8 (−2.3, 0.7)	−1.8 (−3.5, −0.1)*	20.2 (16.1, 24.1)***
Adjusted model[b]					
TJC (0–28)	ref	2.8 (2.5, 3.1)***	0.3 (0, 0.6)†	0.4 (0.1, 0.8)*	2.1 (1.2, 3.1)***
Pain (0–100)	ref	8.4 (7.2, 9.6)***	0.7 (−1.0, 2.3)	−3.6 (−5.5, −1.7)***	12.8 (8.8, 16.9)***
PGA (0–100)	ref	14.8 (13.7, 16.0)***	−0.6 (−2.2, 1.1)	−0.5 (−2.4, 1.4)	19.6 (15.6, 23.5)***
3 month					
Crude model[a]					
TJC (0–28)	ref	0.2 (0.1, 0.4)***	−0.1 (−0.3, 0.1)	0.1 (−0.2, 0.3)	−0.2 (−0.6, 0.3)
Pain (0–100)	ref	4.6 (3.5, 5.7)***	−0.3 (−2.3, 1.7)	−4.5 (−6.5, −2.5)***	0.4 (−4.1, 4.9)
PGA (0–100)	ref	9.6 (8.5, 10.7)***	−2.8 (−4.9, −0.8)**	−3.8 (−5.8, −1.7)***	4.4 (−0.1, 9.0)†
Adjusted model[b]					
TJC (0–28)	ref	1.1 (0.9, 1.2)***	0.1 (−0.1, 0.3)	0.4 (0.2, 0.6)	0.2 (−0.2, 0.6)***
Pain (0–100)	ref	8.7 (7.3, 10.1)***	1.0 (−1.1, 3.1)	−2.9 (−5, −0.7)**	5.0 (0.6, 9.5)*
PGA (0-100)	ref	14.3 (12.9, 15.8)***	−1.3 (−3.6, 0.9)	−1.9 (-4.2, 0.3)†	8.6 (4.1, 13.1)***
12 month					
Crude model[a]					
TJC (0–28)	ref	0.1 (0.0, 0.2)*	−0.3 (−0.5, −0.1)**	−0.3 (−0.5, −0.1)**	−0.2 (−0.5, 0.2)
Pain (0–100)	ref	1.9 (0.8, 3.0)***	−2.2 (−4.3, −0.2)*	−9.2 (−11.4, −6.9)***	0.4 (−3.6, 4.4)
PGA (0–100)	ref	7.1 (6.0, 8.2)***	−5.0 (−7.2, −2.8)***	−8.7 (−11.0, −6.3)***	2.0 (−1.9, 5.9)
Adjusted model[b]					
TJC (0–28)	ref	0.7 (0.6, 0.9)***	0.0 (−0.2, 0.2)	0.0 (−0.2, 0.2)	0.2 (−0.1, 0.5)
Pain (0–100)	ref	6.2 (4.8, 7.6)***	0.8 (−1.3, 3)	−6.3 (−8.6, −3.9)***	4.4 (0.5, 8.2)*
PGA (0–100)	ref	11.5 (10.1, 13.0)***	−2.6 (−4.9, −0.3)*	−5.4 (−7.9, −2.9)***	5.4 (1.7, 9.2)**

The crude and adjusted differences are calculated at TNFi start (baseline), and at 3 and 12 months, using the largest country (Sweden) as reference and are displayed with 95% CIs. †*P*-values: <0.10; *<0.05; **<0.01; ***<0.001. [a]Crude models include the variable country only. [b]Adjusted models additionally include sex, birth decade, age at treatment start, calendar period, disease duration, log (CRP + 1) and log (SJC + 1). PGA: patient's global health assessment; SJC: 28 swollen joint count; TJC: 28 tender joint count; TNFi: TNF inhibitor.

Fig. 2 Mean predicted values of baseline, 3 and 12 months DAS28-CRP for each country population initiating a first ever TNFi treatment



The five solid shapes are the mean predicted DAS28-CRP values for the patients in their own country. Similar shapes (i.e. displayed horizontally, not solid), show the means of the DAS28-CRP values that were predicted for a given country's population, should this population be assessed in another country. At baseline, Denmark and Iceland were the countries where the means of the predicted values were the highest, as for any given shape, the two highest ones were in Denmark and Iceland.

Our study has a number of important strengths. Firstly, we used data from registers that have many similarities in their data collection. Secondly, by using a harmonized study protocol across these registers, together with the choice of comparable time-points in the RA trajectories, we ensured comparing RA patients as similar as possible, thus minimizing contextual differences related to the way these data sources operate. In addition, access to registered data ensured large amounts of patient data and information, and thus high statistical power. We used linear models to analyse the data; the covariates used for adjustment explained a large part of the variance when modelling TJC, which means that important factors for explaining this variable were considered, though we cannot rule out that including other factors (e.g. socio-economic or comorbidities) would have improved our modelling of the inter-country differences.

In conclusion, our study confirmed the existence of inter-country differences in the reporting of TJC, pain and PGA beyond what could be reasonably explained by objective measures, but also demonstrated that the clinical impact of these differences was relatively small, and too small to question the direct comparability of DAS28-CRP values across the studied countries.

Acknowledgements

Jonas Söderling, who assembled the data for Sweden; Niels Steen Krogh, who assembled data for Denmark and Iceland; Joseph Sexton, who assembled data for Norway; and Nina Trokovic, who assembled data for Finland.

Funding: This work was supported by Nordforsk and the Foundation for Research in Rheumatology (Foreum), the research infrastructure was supported by funds from the Swedish Research Council, the Swedish heart Lung Foundation, The Swedish Cancer Society, and funds from Region Stockholm-Karolinska Institutet (ALF).

Disclosure statement: B.Gl.: Rearch grants: Pfizer, AbbVie, BMS. H.B.H.: Speaker for Novartis, Lilly and AbbVie. B.Gu.: Speaker fee form Novartis and Amgen. M.L.H.: Research grants from Abbvie, Biogen, BMS, Celltrion, Eli Lilly, Janssen Biologics B.V, Lundbeck Fonden, MSD, Pfizer, Roche, Samsung Biopies, Sandoz, Novartis. M.L.H.: Chair of the steering committee of the Danish Rheumatology Quality Registry (DANBIO), which receives public funding from the hospital owners and funding from pharmaceutical companies. Co-chair of EuroSpA, which generates real-world evidence of treatment of psoriatic arthritis and axial spondylorthritis based on secondary data and is partly funded by Novartis. B.M.: Consultancy fees and reasearch grant (paid to employer) from Novartis. D.N.: Consultancies Abbvie, BMS, Lilly, Novartis, Pfizer, Roche, UCB. H.R.: Consultancies Abbvie, Celgene, Pfizer, Viatris, UCB. J.A.: Acted or acts as PI for agreements between the following entities and Karolinska Institutet, mainly

regarding safety monitoring of immunomodulator therapies in chronic inflammatory diseases: Abbvie, BMS, Eli Lilly, MSD, Pfizer, Roche, Samsung bioepis, Sanofi and UCB. The other authors have declared no conflicts of interest.

Data availability statement

The data underlying this article cannot be shared publicly due to patients' privacy. The data will be shared on reasonable request by any qualified researchers who engage in rigorous and independent scientific research, addressed to the last author, Johan Askling (johan.askling@ki.se).

References

1 Wells G, Becker JC, Teng J *et al.* Validation of the 28-joint Disease Activity Score (DAS28) and European League Against Rheumatism response criteria based on C-reactive protein against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on erythrocyte sedimentation rate. Ann Rheum Dis 2009;68:954–60.

2 Smolen JS, Landewe RBM, Bijlsma JWJ *et al.* EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. Ann Rheum Dis 2020;79:685–99.

3 Agca R, Heslinga SC, Rollefstad S *et al.* EULAR recommendations for cardiovascular disease risk management in patients with rheumatoid arthritis and other forms of inflammatory joint disorders: 2015/2016 update. Ann Rheum Dis 2017;76:17–28.

4 Orhan C, Van Looveren E, Cagnie B *et al.* Are pain beliefs, cognitions, and behaviors influenced by race, ethnicity, and culture in patients with chronic musculoskeletal pain: a systematic review. Pain Physician 2018;21:541–58.

5 Peacock S, Patel S. Cultural influences on pain. Rev Pain 2008;1:6–9.

6 Rahim-Williams B, Riley JL 3rd, Williams AK, Fillingim RB. A quantitative review of ethnic group differences in experimental pain response: do biology, psychology, and culture matter? Pain Med 2012;13:522–40.

7 Vlaar AP, ten Klooster PM, Taal E *et al.* A cross-cultural study of pain intensity in Egyptian and Dutch women with rheumatoid arthritis. J Pain 2007;8:730–6.

8 Aaltonen KJ, Joensuu JT, Pirila L *et al.* Drug survival on tumour necrosis factor inhibitors in patients with rheumatoid arthritis in Finland. Scand J Rheumatol 2017; 46:359–63.

9 Aaltonen KJ, Ylikyla S, Tuulikki Joensuu J *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 2017;56:725–35.

10 Chatzidionysiou K, Hetland ML, Frisell T *et al.* Opportunities and challenges for real-world studies on chronic inflammatory joint diseases through data enrichment and collaboration between national registers: the Nordic example. RMD Open 2018;4:e000655.

11 Eriksson JK, Askling J, Arkema EV. The Swedish Rheumatology Quality Register: optimisation of rheumatic disease assessments using register-enriched data. Clin Exp Rheumatol 2014;32:S147–9.

12 Ibfelt EH, Jensen DV, Hetland ML. The Danish nationwide clinical register for patients with rheumatoid arthritis: DANBIO. Clin Epidemiol 2016;8:737–42.

13 Ibfelt EH, Sorensen J, Jensen DV *et al.* Validity and completeness of rheumatoid arthritis diagnoses in the nationwide DANBIO clinical register and the Danish National Patient Registry. Clin Epidemiol 2017;9: 627–32.

14 Kvien TK, Heiberg Lie E, Kaufmann C *et al.* A Norwegian DMARD register: prescriptions of DMARDs and biological agents to patients with inflammatory rheumatic diseases. Clin Exp Rheumatol 2005;23(5 Suppl 39):S188–194.

15 Neovius M, Simard J, Sundstrom A *et al.* Generalisability of clinical registers used for drug safety and comparative effectiveness research: coverage of the Swedish Biologics Register. Ann Rheum Dis 2011;70:516–9.

16 Thorsteinsson B, Geirsson AJ, Krogh NS, Gudbjornsson B; ICEBIO. Outcomes and safety of tumor necrosis factor inhibitors in reactive arthritis: a nationwide experience from Iceland. J Rheumatol 2020;47:1575–81.

17 Komiyama O, Kawara M, De Laat A. Ethnic differences regarding tactile and pain thresholds in the trigeminal region. J Pain 2007;8:363–9.