Rheumatology, 2024, **00**, 1–10
https://doi.org/10.1093/rheumatology/keae653
Advance access publication 4 December 2024
Original Article

British Society for Rheumatology RHEUMATOLOGY OXFORD

Basic science

A multivariable prediction model to identify anti-CCP positive people in those with non-specific musculoskeletal symptoms in primary care

Heidi J. Siddle [1,*], Michelle Wilson [1], Jacqueline L. Nam [1], Leticia Garcia-Montoya [1,2], Laurence Duquenne [1,2], Kulveer Mankia [1,2], Paul Emery [1,2], Elizabeth M.A. Hensor [1,2]

[1]Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, UK
[2]NIHR Leeds Biomedical Research Centre, Leeds, UK

*Correspondence to: Heidi J. Siddle, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Chapel Allerton Hospital, 2nd Floor, Chapeltown Road, Leeds LS7 4SA, UK. E-mail: h.siddle@leeds.ac.uk

Abstract

Objectives: We aimed to develop a prediction model identifying people presenting to primary care with musculoskeletal symptoms likely to be anti-CCP positive and therefore at risk of developing RA.

Methods: Participants aged ≥16 years, with new-onset non-specific musculoskeletal symptoms and no history of clinical synovitis, completed a symptom questionnaire and had an anti-CCP test. Model development used LASSO-penalized logistic regression, performance was assessed using area under the receiver operating characteristic curve (AUROC) and decision curve analysis, model over-fit was estimated using bootstrapping and cross-validation. Participants were followed-up at 12 months for RA or seronegative/undifferentiated inflammatory arthritis diagnosis.

Results: Analysis included 6879 participants; 203 (2.95%) of whom were anti-CCP positive. Eleven predictors were retained: male sex, first-degree relative with RA, ever smoked and joint pain in: back, neck, shoulders, wrists, hands/fingers, thumbs, knees, feet/toes. AUROC was 0.65 (95% CI 0.61, 0.69, optimism = 0.03). Using a 4% decision threshold, the model recommended an anti-CCP test in 1288 (18.7%) participants, 78 (6.1%) of whom were anti-CCP positive, compared with 125/5591 (2.2%) below the threshold. Net benefit was 0.0040 (0.0020 corrected). Forty-eight participants were diagnosed with inflammatory arthritis/RA within 12 months. Of those who were above the threshold and anti-CCP positive, 32.1% developed inflammatory arthritis/RA compared with 0.4% of those who were anti-CCP negative. Of those below the threshold, 0.3% were diagnosed with inflammatory arthritis/RA.

Conclusions: Targeted anti-CCP testing in primary care may aid earlier identification of people at risk of RA, prompting specialist referral to rheumatology for earlier diagnosis and initiation of disease-modifying therapy.

Keywords: anti-cyclic citrullinated peptide (anti-CCP) antibodies, at risk of rheumatoid arthritis, inflammatory arthritis, musculoskeletal symptoms, prediction model

Rheumatology key messages

- This novel multivariable prediction model provides the opportunity to support primary care clinicians in identifying people at risk of RA.
- This prediction model identifies individuals with non-specific musculoskeletal symptoms without synovitis who are more likely to test anti-CCP positive and are therefore at risk of RA.
- Targeted anti-CCP testing will aid the implementation of the 'simple score' in anti-CCP positive individuals to stratify risk of developing RA.

Introduction

RA is a chronic, progressive autoimmune inflammatory arthritis characterized by swelling, tenderness and ultimately destruction of synovial joints, leading to disability, associated comorbidities and premature mortality [1–5]. In England, the estimated the point-prevalence of RA diagnoses in 2019 was 0.779% with the annual incidence from 49.1–52.1/100 000 person-years [6]. RA has a substantial socioeconomic burden in terms of direct costs to National Health Service (NHS) healthcare providers (£560 million) and indirect costs related to loss of productivity (£1.8 billion annually in sick leave) [7], and increased morbidity and mortality [8]. Burden is further exacerbated in those with a low socioeconomic status, with increased pain and poorer outcomes in inflammatory arthritis [9].

Early diagnosis and subsequent treatment with DMARDs improves clinical outcomes through preventing joint damage and subsequent disability [10–12], potentially offering the

Received: 13 September 2024. **Accepted:** 13 November 2024

opportunity to alter the natural course of inflammatory arthritis [13, 14]. However, the Rheumatoid Arthritis Classification Criteria and other models that predict RA diagnosis apply to patients who already have clinically apparent joint inflammation (synovitis) [15–17].

A phase of RA-related autoimmunity often precedes the onset of clinical RA by several months or years [18], characterized by the presence of anti-CCP antibodies and RF [19]. Therapeutic intervention with MTX and abatacept in people at-risk of RA has been shown to reduce MRI inflammation, clinical symptoms and rates of RA, with sustained efficacy beyond treatment [20–22]. Models developed for use in secondary care can predict inflammatory arthritis/RA development in people presenting with musculoskeletal symptoms with positive RF and/or anti-CCP antibodies, and positive imaging outcomes, i.e. US power Doppler signal, but without clinical signs of synovitis [23–27]. However, people who develop musculoskeletal symptoms often initially present in primary care rather than secondary care [28].

Those subsequently diagnosed with RA in secondary care visit a general practitioner (GP) on average four times before being referred to a specialist for diagnosis; the median time from symptom onset to diagnosis and first treatment is 9 months [7]. Our research team has developed a simple score to predict risk of inflammatory arthritis in anti-CCP positive individuals, which includes four biomarkers: anti-CCP titre, RF, ESR and early morning stiffness (EMS) [27]. However, anti-CCP testing is not routinely requested or always accessible in primary care. The National Institute for Health and Care Excellence (NICE) only recommends measuring anti-CCP antibodies if the patient is negative for RF when RA is suspected clinically and it should not delay a referral for a specialist opinion [29]. Thus, there is an opportunity to identify people earlier in the pre-clinical phase of inflammatory arthritis through targeted anti-CCP testing in primary care, combined with inflammatory arthritis prediction models that include anti-CCP.

The existing evidence demonstrates the need for early diagnosis and intervention for inflammatory arthritis in secondary care populations, with the potential for greater gains if people can be identified earlier in primary care. Primary care studies have shown the importance of determining healthcare usage in the 'presentation stage' prior to the diagnosis of RA [30, 31], and ensuring patients receive early screening for known concomitant diseases (e.g. cardiovascular) [32]. We aimed to develop a prediction model to identify which people, among those presenting to primary care with new-onset non-specific musculoskeletal symptoms, are likely to be anti-CCP positive and should be tested for anti-CCP antibodies. Targeted anti-CCP testing could aid early identification of those in the pre-clinical phase of inflammatory arthritis and reduce time to diagnosis.

Methods

The transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) statement has been followed in reporting this prediction model study [33].

Source data

Participants were selected from the Leeds 'Co-ordinated Programme to Prevent Arthritis', a prospective, observational primary care cohort study. Ethics approval was received from the NHS Health Research Authority National Research Ethics Service Committee Yorkshire & the Humber—Leeds West (REC reference number 06/Q1205/169) and the study was registered with ClinicalTrials.gov (ID NCT02012764). Individuals were recruited through the former UK Primary Care Clinical Research Network between June 2007 and March 2020. The cohort recruited individuals aged ≥16 years presenting to primary care with new-onset, non-specific musculoskeletal symptoms, without clinical synovitis who were not being referred to a rheumatology service. Participants provided informed consent before completing a baseline questionnaire, providing their age, gender, cigarette smoking status and participant-reported details of first-degree relatives (FDR) with RA. Participants indicated whether they had current or new pain in their neck, back, shoulders, elbows, hips, wrists, thumbs, hands and/or fingers, knees, ankles, and feet and/or toes. All participants had a second-generation anti-CCP test. Anti-CCP positive participants were invited to attend secondary care follow-up appointments at Chapel Allerton Hospital, Leeds, every 3 months for 12 months and then annually until they received a diagnosis of inflammatory arthritis. Participants who were anti-CCP negative (or who declined follow-up in secondary care) received standard care with their GP and were sent a questionnaire 12 months after enrolment to obtain details of any inflammatory arthritis/RA diagnoses. If questionnaires were not returned or inflammatory arthritis/RA status was unclear, the research team attempted to contact participants and/or their GP by telephone for clarification.

Outcome

The outcome was anti-CCP antibody status (positive/negative). Positivity of the test was determined using machine specific cut-offs, initially using an ImmunoCAP 250 (Phadia, Thermo Fisher Scientific Inc., Massachusetts, USA) (reference range >7 U/ml) and later on a BioPlex 2200 (Bio-Rad, Thermo Fisher Scientific Inc., Massachusetts, USA) machine (reference range >2.99 U/ml).

Predictors

In total, 15 baseline predictors were considered: age at date of blood sample (or date of questionnaire if not available); gender (male/female); smoking status (never/ever); and FDR (parent, sibling or child) with RA. The remaining predictor variables were based on participant-reported pain in specific joints and/or body regions: neck, back, shoulders, elbows, hips, wrists, thumbs, hands and/or fingers, knees, ankles, and feet and/or toes. For bilateral joints/regions, pain in one or both sides was considered an affirmative response.

Statistical analysis methods

Missing data was minimal; complete case analysis was employed. The linearity assumption was assessed by examining the association between the logit probability and age (the only continuous variable). Variables were selected via least absolute shrinkage and selection operator (LASSO) logistic regression using the R function 'cv.glmnet' in the 'glmnet' package [34], with the tuning parameter (lambda) selected to give the minimum mean cross-validated error [34]. Model coefficients were multiplied by 10 and rounded to one decimal place [35], resulting in an anti-CCP prediction score. Discrimination was assessed using the area under the receiver operating characteristic curve

(AUROC) calculated using the 'roc' function in the R package 'pROC' [36] with a corresponding 95% CI estimated via 2000 bootstraps, and slope was estimated by regressing the linear predictor from the model on the outcome. The optimism (as a measurement of model over-fit) around the slope and AUROC were estimated via bootstrap re-sampling with replacement (200 bootstraps).

Sample size for multivariable model development was estimated using the methods described by Riley *et al.* (2019) [37] and implemented using default parameters in the R package & function 'pmsampsize' [38]. Assuming an event rate of 3% [25] and 15 potential predictors (one continuous plus 14 two-level categorical), ≥3724 participants were required.

Decision curve analysis

Decision curve analysis using the 'dca' R function in the 'dcurves' package [39] assessed net benefit of the model when compared with strategies testing all or no participants. Overfit-corrected estimates were obtained using 10-fold cross-validation [40]. Given the high costs of treating RA *vs* the relatively low cost of anti-CCP testing, we chose two cut-offs at relatively low likelihood of being anti-CCP positive: 3% (the overall rate of anti-CCP positivity in the group) and 4% (to achieve higher specificity). An anti-CCP prediction model score at or above the threshold would trigger an anti-CCP test which, if positive, could support referral to secondary care. For those scoring below the threshold, an anti-CCP test is not indicated. A threshold of 3% suggests that failing to identify one participant who is anti-CCP positive is 33 times worse than testing one participant who is anti-CCP negative (24 times worse for the 4% threshold). Twelve-month follow-up data were examined to determine the proportions of participants who received a diagnosis of RA (or undifferentiated inflammatory arthritis/seronegative inflammatory arthritis) within 12 (±3) months. Analyses were performed using R [41] in R Studio [42].

Secondary care data and a 'simple score'

Recently, our group published a simple score to predict risk of inflammatory arthritis [27]. The simple score was developed for use in primary care and consisted of four variables: anti-CCP value, RF, ESR and participant-reported EMS. Referral to secondary care for follow-up was recommended if the simple score met a threshold calculated based on net benefit and clinical impact. We applied the anti-CCP prediction model to the data used to derive this simple score to mimic how the patient pathway could look if both were implemented. We assessed how many participants would: be identified as likely to be anti-CCP positive; be identified as high risk of inflammatory arthritis and referred to secondary care; and receive an inflammatory arthritis diagnosis within 1 year. Since not all of the variables included in the simple score were collected in the primary care cohort, we were unable to calculate it for our entire prediction model development cohort.

Results

Participants

In total, the primary care cohort consisted of 9957 records; out of these, 7032 participants were eligible for inclusion in our model development cohort (Fig. 1) and of these, 153 with missing baseline data were excluded (2.2%). Two-hundred and three participants were anti-CCP positive (2.95%). Demographic and clinical characteristics of the 6879 participants with data are shown in Table 1.

Anti-CCP prediction model development and performance

The only continuous predictor, age, did not violate the linearity assumption. LASSO regression resulted in a model containing 13 predictors (age and ankle not selected). Model coefficients were multiplied by 10 and rounded to one decimal place to give the anti-CCP prediction model, which ranged from −6 to 22, shown in Table 2 (coefficients for hips and elbows became near-zero once rounded, leaving 11 predictor variables). Note that it is not possible to interpret the individual coefficients from a prediction model; prediction and causal inference are distinct data science tasks, requiring different methodologies [43]. Therefore, it is not a concern if the direction of the coefficients for certain variables appears counter-intuitive as they are not intended to represent causal effects.

The distribution of the anti-CCP prediction model calculated for all participants is shown in Supplementary Fig. S1, available at *Rheumatology* online. When used to predict anti-CCP status the score gave an AUROC of 0.65 (95% CI 0.61, 0.69, optimism = 0.03) and a slope of 1.26 (95% CI 0.94, 1.58, optimism = 0.21).

When applying the 3% and 4% thresholds (at respective anti-CCP prediction model score cut-offs ≥8 and ≥11), 40.4% and 18.7% of participants, respectively, were above each threshold (Table 3) and would therefore be recommended for anti-CCP testing. Of these participants, 4.4% (3% threshold) and 6.1% (4% threshold) were anti-CCP positive (positive predictive value). Out of those participants below each threshold 2% (3% threshold) and 2.2% (4% threshold) were anti-CCP positive (1 − negative predictive value). The net benefit when applying each threshold was 0.0058 (0.0040 corrected, 3% threshold) and 0.0040 (0.0020 corrected, 4% threshold; Table 3, Fig. 2), these figures are equivalent to the proportion of true positives that would be detected among those tested, without incurring false positives, i.e. 4/1000 and 2/1000. Equivalent figures for strategies where all or no participants are tested are shown in Supplementary Table S1, available at *Rheumatology* online.

High anti-CCP titres (>3 times the upper limit of normal) have previously been shown to be particularly predictive of inflammatory arthritis/RA [27]. Using the 3% risk threshold, 1.0% (39/4101) of those below the threshold had high-titre anti-CCP compared with 3.1% (85/2778) above the threshold. Using the 4% risk threshold, the corresponding results were 1.2% (65/5591) *vs* 4.6% (59/1288).

Anti-CCP prediction model performance in 12-month inflammatory arthritis/RA follow-up data

Twelve-month follow-up data were available for 2480 participants (36.1%); of the 4399 participants without follow-up data, 4306 were anti-CCP negative (97.9%). Since follow-up questionnaires are more likely to be completed and returned in the case of an inflammatory arthritis/RA diagnosis [44], and attempts were made to confirm diagnosis for all anti-CCP positive participants, these participants were presumed to have not developed inflammatory arthritis/RA. The proportion of participants who would/would not be tested according to the anti-CCP prediction model, split by

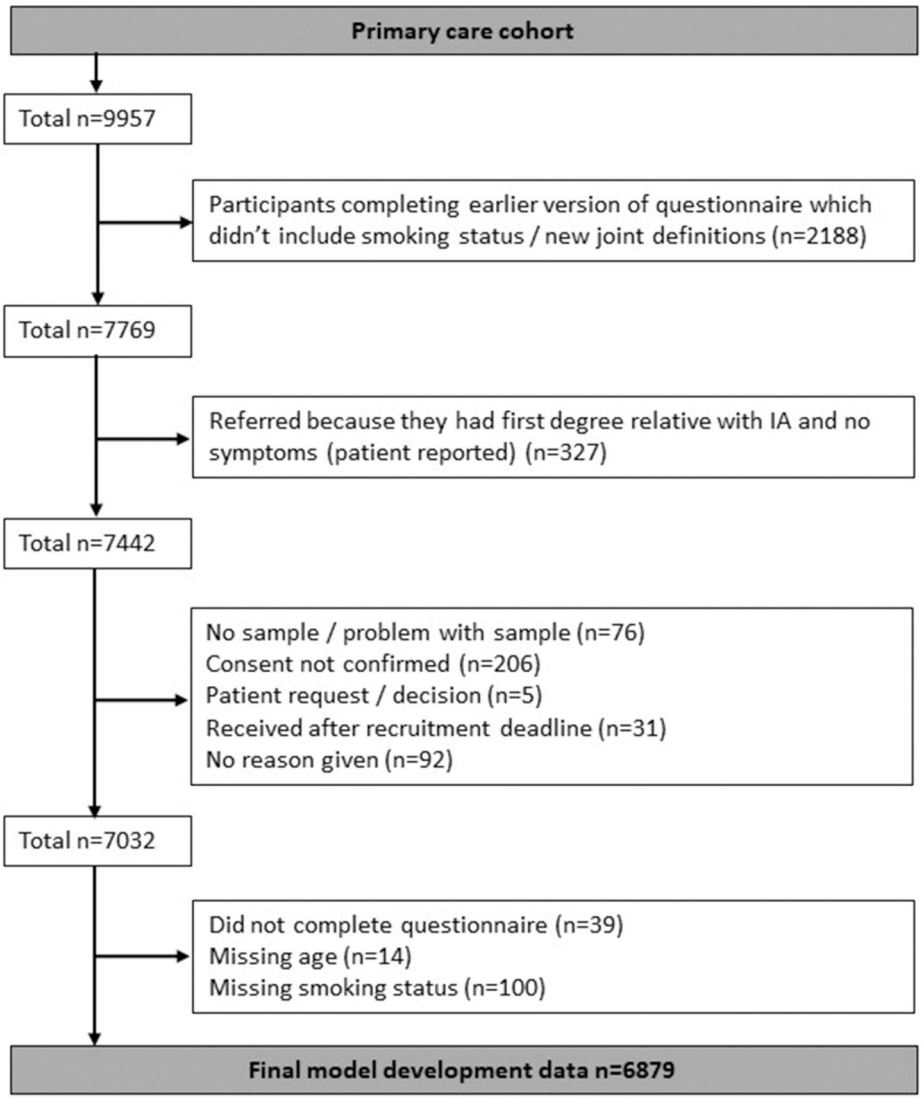

Figure 1. Flow chart detailing selection of participants for model development from the primary care cohort

anti-CCP result and 12-month inflammatory arthritis/RA diagnosis is shown in Table 3.

The rate of inflammatory arthritis/RA at 12 months amongst those the model would not recommend for testing, either at the 3% or 4% threshold, was 0.3%.

Of those above the 3% threshold who were anti-CCP positive, 25.4% were diagnosed with inflammatory arthritis/RA within 12 months, compared with 0.2% of those who were anti-CCP negative. At the 4% percent threshold, corresponding figures were 32.1% inflammatory arthritis/RA in those tested who were anti-CCP positive compared with 0.4% RA in those who were anti-CCP negative.

The same data for only participants with follow-up available is shown in Supplementary Table S2, available at *Rheumatology* online. In 203 anti-CCP positive participants, 38 were diagnosed with inflammatory arthritis/RA within the first 12 months. If we exclude participants without follow-up data available, this equates to an inflammatory arthritis/RA rate of 34.5% (38/110), if we include those without follow-up, assuming no inflammatory arthritis/RA, this figure is 18.7% (38/203) which is more comparable to the 12-month inflammatory arthritis progression rates reported for the anti-CCP positive cohort followed in secondary care at the coordinating centre (15.4%). Note that >90% of those developing inflammatory arthritis in the secondary care dataset met RA criteria [27].

Performance of anti-CCP prediction model combined with 'simple score' in secondary care data

Our anti-CCP prediction model is designed to facilitate targeted anti-CCP testing, aiding the implementation of the simple score developed by our group in anti-CCP positive participants, demonstrated in Supplementary Fig. S2, available at *Rheumatology* online. The simple score was developed in secondary care data consisting of 455 anti-CCP positive participants, 70 (14.4%) of whom were from our full primary care cohort [35 (7.7%) included in the current anti-CCP prediction model development]. Demographics of the 455 participants can be found in Duquenne *et al.* (2023) [27], 70 participants were diagnosed with inflammatory arthritis within 12 months. Four-hundred and one participants had sufficient data to calculate both the anti-CCP prediction model (3% and 4% threshold) and the simple score. The

Table 1. Baseline demographic and clinical characteristics of participants included in the model development cohort

Variable		Total ($N = 6879$)
Age[a] (years)	Mean (s.d.)	54.6 (14.8)
	Median (Min, Q1, Q3, Max)	55.0 (16.0, 45.0, 66.0, 91.0)
Sex	Female	4815 (70.0%)
Smoking status	Ever	3539 (51.4%)
FDR with RA	Yes	1871 (27.2%)
Participant-reported joint pain (left and/or right where applicable)		
Neck	Yes	2151 (31.3%)
Back	Yes	2577 (37.5%)
Shoulders	Yes	2973 (43.2%)
Elbows	Yes	2029 (29.5%)
Hips	Yes	2691 (39.1%)
Wrists	Yes	2719 (39.5%)
Hands/fingers	Yes	3768 (54.8%)
Thumbs	Yes	2582 (37.5%)
Knees	Yes	3909 (56.8%)
Ankles	Yes	2096 (30.5%)
Feet/toes	Yes	2449 (35.6%)

[a] At date of blood sample (or date of questionnaire if not available).
FDR: first-degree relative (parent, sibling or child); Q1: 1st quartile; Q3: 3rd quartile.

Table 2. Anti-CCP prediction model; scores to be added up to obtain one total score, on a scale of –6 to 22 (for bilateral joints, pain in one or both sides should be considered an affirmative response)

Predictor variables	Score
Joint pain: back	–3
Joint pain: neck	–2
Joint pain: knees	–1
Joint pain: wrists	4
Sex: male	3
First-degree relative[a] with RA	3
Joint pain: feet/toes	3
Joint pain: hands/fingers	3
Joint pain: shoulders	3
Smoking status: ever	2
Joint pain: thumbs	1
Total score	___

[a] Mother, father, brother, sister, son or daughter.

anti-CCP prediction model correctly identified 47.9% (3% threshold) and 25.2% (4% threshold) of the anti-CCP positive participants (Table 4). Approximately half of the participants above each anti-CCP prediction model threshold, were also above the threshold for the simple score and would therefore be recommended for follow-up in secondary care. Of these 37% (3% threshold) and 45.5% (4% threshold) received an inflammatory arthritis diagnosis within 12 months, >90% of whom met criteria for RA. Of those below each threshold for anti-CCP testing in our model, 12% developed inflammatory arthritis within 1 year.

Discussion

Prediction of RA can be challenging in primary care as it makes up a small proportion of musculoskeletal conditions that account for a third of all GP appointments. This novel multivariable prediction model provides the opportunity to support primary care clinicians in identifying people at risk of developing RA through targeted anti-CCP testing. More importantly, intervening in those who are at high risk can reverse subclinical inflammation in some patients delaying the onset of RA [20–22] and potentially reducing the long-term disability associated with the disease. There is also evidence that MTX treatment aiming at secondary prevention in the at-risk phase of RA is cost-effective [45]. However, we acknowledge that further research is warranted to confirm that long-term outcomes associated with identifying and treating people at risk of RA are more clinically and cost-effective rather than waiting for clinical synovitis to develop.

Existing prediction models developed by our group [27] that identify people who are likely to develop inflammatory arthritis include anti-CCP testing. However anti-CCP testing is not routinely undertaken in primary care and often with limitations imposed on access to testing. Furthermore, in the UK, current NICE guidance only exists for those where RA is suspected clinically (presence of synovitis) and anti-CCP testing is only recommended if a patient is negative for RF. Anti-CCP testing everyone in primary care who presents with musculoskeletal symptoms would cause a significant burden on an already over-stretched healthcare system; the cost of the tests, an increase in phlebotomy services and time to review results by healthcare professionals will be required. Therefore, changing the diagnostic paradigm to detect RA prior to the onset of clinical synovitis requires targeted anti-CCP testing in people who present with new-onset non-specific musculoskeletal symptoms in primary care to identify those at high risk, before they present with clinical synovitis. Additionally, it will support the implementation of existing tools, developed in anti-CCP positive participants, to refer high-risk individuals to secondary care for more timely review and intervention, potentially altering the course of their disease. In addition to delivering potential prevention intervention, identifying high-risk individuals will enable an earlier diagnosis and treatment of RA should it develop. There are clear benefits to treating RA within 12 weeks of symptom onset [46]. Treatment-related costs have been reported to be lower in patients who were identified earlier (symptom duration <12 weeks) compared with those with longer symptom duration (>12 weeks) [47]. Importantly, in those who have autoantibody-positive RA, treatment initiation in those who had symptoms for >12 weeks resulted in more intensive biologics use with overall higher costs within those requiring biologics, suggesting that earlier identification and treatment initiation is likely cost-effective.

Limitations

Although the prediction model was developed using data from a large, multicentre prospective cohort, we acknowledge limitations of the data. The majority of the variables comprising the prediction model are subjective and self-reported (i.e. joint pain) meaning that they are subject to a certain amount of bias. The accurate reporting of RA in FDRs relies on participants having knowledge of the medical history of their immediate family and the inability to differentiate between RA and other types of inflammatory arthritis or even OA may cause false reporting of RA. Furthermore, the percentage of participants with a FDR with RA (27%) included in the prediction model appears high, which may in turn influence the predictor variable score in the model. However, FDRs with RA perceive themselves to be at high absolute risk of developing RA [48] and were therefore more

Table 3. Performance of the anti-CCP prediction model and 12-month follow-up of primary care participants: sensitivity, specificity and predictive values of the score used to predict anti-CCP status at key decision thresholds

	Anti-CCP prediction model		Anti-CCP result		IA/RA diagnosis[a]		Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)	Net benefit	Corrected net benefit
3% threshold/anti-CCP prediction model score ≥8	No test	4101/6879 (59.6%)	–ve	4020/4101 (98.0%)	Yes	4/4020 (0.1%)	60.1 (53.0, 66.9)	60.2 (59.0, 61.4)	4.4 (3.7, 5.2)	98.0 (97.6, 98.4)	0.0058	0.0040
			+ve	81/4101 (2.0%)	Yes	7/81 (8.6%)						
	Test	2778/6879 (40.4%)	–ve	2656/2778 (95.6%)	Yes	6/2656 (0.2%)						
			+ve	122/2778 (4.4%)	Yes	31/122 (25.4%)						
4% threshold/anti-CCP prediction model score ≥11	No test	5591/6879 (81.3%)	–ve	5466/5591 (97.8%)	Yes	5/5466 (0.1%)	38.4 (31.7, 45.5)	81.9 (80.9, 82.8)	6.1 (4.8, 7.5)	97.8 (97.3, 98.1)	0.0040	0.0020
			+ve	125/5591 (2.2%)	Yes	13/125 (10.4%)						
	Test	1288/6879 (18.7%)	–ve	1210/1288 (93.9%)	Yes	5/1210 (0.4%)						
			+ve	78/1288 (6.1%)	Yes	25/78 (32.1%)						

[a] Within 1 year ± 3 months, participants with follow-up unavailable were presumed to have not received an inflammatory arthritis/RA diagnosis. CI: exact binomial confidence interval; IA: undifferentiated or seronegative inflammatory arthritis; PPV: positive predictive value; NPV: negative predictive value.

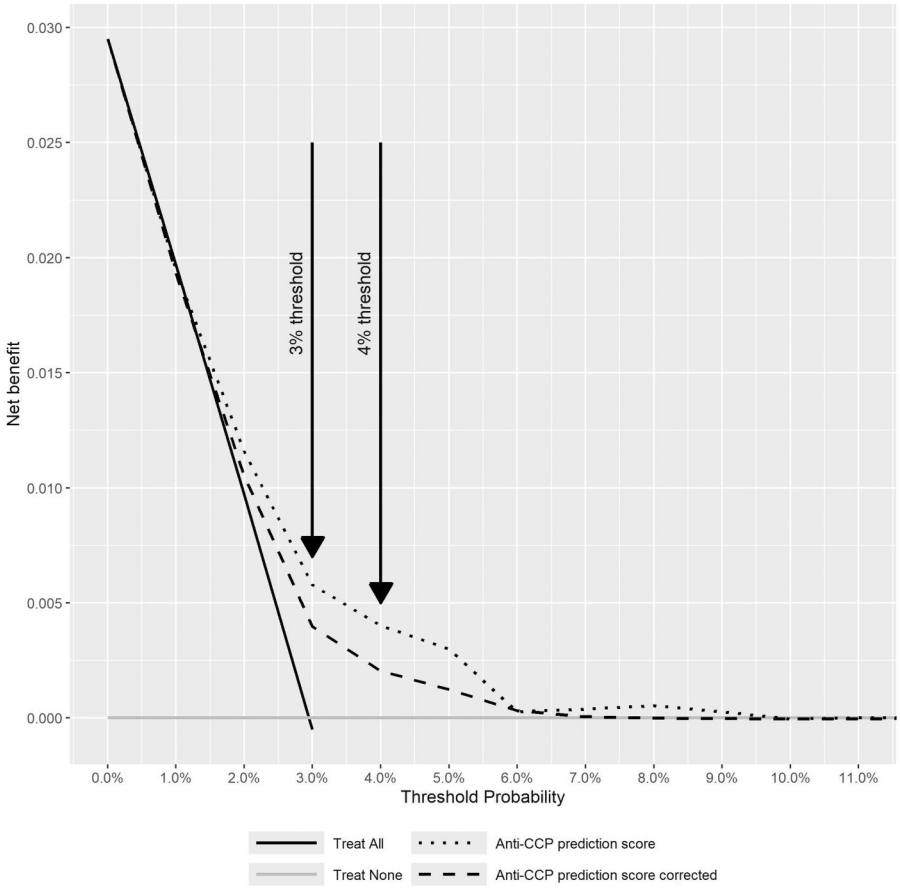

Figure 2. Decision curve showing net benefit for 'treat all', 'treat none', anti-CCP prediction model score and corrected (for over-fit) anti-CCP prediction model score with 3% and 4% decision thresholds

Table 4. Performance of the anti-CCP prediction model combined with simple score in secondary care data (401 anti-CCP positive participants)

	Anti-CCP prediction model		Simple model score ≥18		IA/RA diagnosis[a]	
3% threshold/anti-CCP prediction model score ≥8	No test	209/401 (52.1%)	No	125/209 (59.8%)	Yes	7/125 (5.6%)
			Yes	84/209 (40.2%)	Yes	18/84 (21.4%)
	Test	192/401 (47.9%)	No	100/192 (52.1%)	Yes	6/100 (6%)
			Yes	92/192 (47.9%)	Yes	34/92 (37%)
4% threshold/anti-CCP prediction model score ≥11	No test	300/401 (74.8%)	No	179/300 (59.7%)	Yes	10/179 (5.6%)
			Yes	121/300 (40.3%)	Yes	27/121 (22.3%)
	Test	101/401 (25.2%)	No	46/101 (45.5%)	Yes	3/46 (6.5%)
			Yes	55/101 (54.5%)	Yes	25/55 (45.5%)

[a] Within 1 year. IA: inflammatory arthritis.

likely to present to primary care with new-onset musculoskeletal symptoms. It is important to acknowledge that model scores for individual variables, for example male sex, which at face value is counterintuitive to the RA population, cannot be interpreted clinically; only the total score can be used to indicate predicted likelihood of being anti-CCP positive. One of the hallmark features of RA, EMS, was not collected in this cohort and may have proven to be an important predictor of anti-CCP positivity.

Although the aim of this study was to develop a prediction model to aid targeted anti-CCP testing, the addition of 12-month follow-up data is important as it confirms the higher rate of reported inflammatory arthritis/RA in anti-CCP positive participants (18.7% *vs* <1% in anti-CCP negative participants), highlighting the need to focus on this group of individuals. All anti-CCP positive participants were invited to be followed-up in secondary care, however only 93 (45.8%) accepted. Along with the participants who were anti-CCP negative, these participants were sent questionnaires 12 months after enrolment asking about their RA status. In addition to the 93 participants followed-up in secondary care, 12-month RA status was available for 2387 participants (36.1% in total). Failure to return the questionnaire (change of address, death, not received) was suspected to be the main reason for limited 12-month follow-up data, and we assumed that these participants had not received a diagnosis of RA in order to avoid the over-inflation of positive RA diagnosis. However, seronegative RA represents an important subset who would not be targeted by our prediction model and our assumption of no RA diagnosis amongst those

who were anti-CCP negative and did not return a questionnaire may have underestimated the rate in this group.

External validation of our predication model is required, however there are no comparable existing external cohorts which have collected sufficient data. Following development of an intervention to implement the model in primary care we are proposing an intervention evaluation study, which will support external validation of the prediction score.

Interpretation

Given the high costs of treating someone who develops RA *vs* the relatively low cost of an anti-CCP test, we chose two prediction model thresholds at low levels of likelihood of being anti-CCP positive: 3% (score ≥8) and 4% (score ≥11). These represent a belief that missing an anti-CCP positive person is between 24 and 33 times worse than performing an anti-CCP test in someone who is anti-CCP negative.

Using our prediction model with a threshold of ≥8 to determine who should undergo anti-CCP testing, we estimate that we would test ∼40% of those presenting with new-onset non-specific musculoskeletal symptoms, 4.4% of whom would be anti-CCP positive, and 65% of all individuals diagnosed with inflammatory arthritis/RA within 12 months would be identified by a positive test. Using the ≥11 threshold, 19% of those presenting with new-onset non-specific musculoskeletal symptoms would be tested, 6.1% of whom would be positive, and 52% of those who developed inflammatory arthritis/RA within 12 months would be identified. Such targeted early anti-CCP testing could be combined with existing inflammatory arthritis risk scores [27] to prioritize high-risk individuals for observation and intervention.

The clinical utility of the two score thresholds will be explored with primary care clinicians in qualitative interviews to determine acceptability and to develop an intervention to support clinicians using this prediction model.

Implications

Interviews with primary care clinicians have been undertaken to understand the key behaviours required to support clinicians using this prediction score and to develop an intervention to support implementation in primary care, alongside heath economic modelling to explore potential cost-effectiveness.

There are currently no formal clinical recommendations for monitoring or managing people who are anti-CCP positive and at risk of developing RA. Further research is needed to understand who should be monitored, when and for how long, and determine the most appropriate place (primary or secondary care) for these people to be monitored. Understanding the perspective of rheumatologists who will be receiving referrals from primary care for those who are anti-CCP positive with symptoms will aid the implementation of this prediction model in primary care and improve patient pathways.

Conclusion

This prediction model offers an opportunity to identify people at risk of RA among the rising number of musculoskeletal conditions that are seen in primary care, through targeted anti-CCP testing before they develop synovitis, the hallmark feature of RA. Ongoing work to implement this prediction model in primary care comes at a time when evidence is emerging to support intervention in those who are at high risk of developing RA. Early identification in primary care prompting referral to rheumatologists in secondary care can facilitate earlier diagnosis. It could also facilitate RA prevention strategies, aimed at reversing subclinical inflammation and potentially delaying or preventing the onset of RA. Such approaches have the potential to reduce the long-term disability and comorbidities typically associated with RA.

Supplementary material

Supplementary material is available at *Rheumatology* online.

Data availability

Data and analytical code are available upon reasonable request by contacting the Leeds 'Co-ordinated Programme to Prevent Arthritis' chief investigator (P.E.; E-mail: p.emery@leeds.ac.uk).

Funding

Professor Heidi Siddle, Senior Clinical Lecturer, ICA-SCL-2018-04-ST2-004, is funded by Health Education England (HEE)/National Institute for Health and Care Research (NIHR) for this research project. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR, University of Leeds, NHS or the UK Department of Health and Social Care. The study was supported by the National Institute for Health and Care Research (NIHR) Leeds Biomedical Research Centre (BRC) (grant number 2017–2022: IS-BRC-1215-20015; 2022–2027: NIHR203331); L.G.-Montoya . and L.D. are NIHR BRC Fellows. M.W. and E.M.A.H. are supported in part by the NIHR BRC.

Disclosure statement: K.M. reports grants from Gilead, Eli Lilly, Serac Healthcare, AstraZeneca and Deepcure, and consultancy/honoraria fees from Abbvie, UCB, Galapogos, Lilly, Serac Healthcare, Zura Bio and Deepcure, outside the submitted work. P.E. reports institution grants from Lilly and Samsung, personal consulting fees from Lilly and Galapogos, personal payment or honoraria fees Novartis, Lilly, Abbvie, Gilead and Galapagos, support for attending meetings and/or travel from Lilly, and participation on advisory boards for Lilly, AstraZeneca, BMS, outside the submitted work. H.J.S. reports personal payment or honoraria fees from Janssen, outside the submitted work. M.W., J.L.N., L.G.-M., L.D. and E.M.A.H. have no competing interests.

Acknowledgements

We would like to thank the GPs, health professionals and the UK National Institute for Health Research Clinical Research Network (formerly the UK Primary Care Research Network) for their help with recruitment of the study. The authors would like to thank Diane Corscadden and Katie Mbara for processing the blood samples, Harkee Singh and Collette Hartley for coordinating the study, and Craig Biggs for his IT support in managing the database.

References

1. Scott DL, Symmons DP, Coulton BL, Popert AJ. Long-term outcome of treating rheumatoid arthritis: results after 20 years. Lancet 1987;1:1108–11.
2. Mitchell DM, Spitz PW, Young DY *et al.* Survival, prognosis, and causes of death in rheumatoid arthritis. Arthritis Rheum 1986; 29:706–14.
3. Pincus T, Callahan LF, Sale WG *et al.* Severe functional declines, work disability, and increased mortality in seventy-five rheumatoid arthritis patients studied over nine years. Arthritis Rheum 1984;27:864–72.
4. Berkanovic E, Hurwicz ML. Rheumatoid arthritis and comorbidity. J Rheumatol 1990;17:888–92.
5. Wolfe F, Mitchell DM, Sibley JT *et al.* The mortality of rheumatoid arthritis. Arthritis Rheum 1994;37:481–94.
6. Scott IC, Whittle R, Bailey J *et al.* Rheumatoid arthritis, psoriatic arthritis, and axial spondyloarthritis epidemiology in England from 2004 to 2020: an observational study using primary care electronic health record data. Lancet Reg Health Eur 2022; 23:100519.
7. National Audit Office. Services for people with rheumatoid arthritis. Report by the comptroller and auditor general, 2009. https://www.nao.org.uk/reports/services-for-people-with-rheumatoid-arthritis/ (23 May 2024, date last accessed).
8. Michaud K, Messer J, Choi HK, Wolfe F. Direct medical costs and their predictors in patients with rheumatoid arthritis: a three-year study of 7,527 patients. Arthritis Rheum 2003;48:2750–62.
9. Versus Arthritis. The State of Muscoloskeletal Health 2023. Arthritis and other musculoskeletal conditions in numbers. 2023. https://www.versusarthritis.org/about-arthritis/data-and-statistics/the-state-of-musculoskeletal-health/ (23 May 2024, date last accessed).
10. van der Heide A, Jacobs JW, Bijlsma JW *et al.* The effectiveness of early treatment with "second-line" antirheumatic drugs. A randomized, controlled trial. Ann Intern Med 1996;124:699–707.
11. Bukhari MAS, Wiles NJ, Lunt M *et al.* Influence of disease-modifying therapy on radiographic outcome in inflammatory polyarthritis at five years: results from a large observational inception study. Arthritis Rheum 2003;48:46–53.
12. van Dongen H, van Aken J, Lard LR *et al.* Efficacy of methotrexate treatment in patients with probable rheumatoid arthritis: a double-blind, randomized, placebo-controlled trial. Arthritis Rheum 2007;56:1424–32.
13. Finckh A, Liang MH, van Herckenrode CM, de Pablo P. Long-term impact of early treatment on radiographic progression in rheumatoid arthritis: a meta-analysis. Arthritis Rheum 2006;55:864–72.
14. Lard LR, Visser H, Speyer I *et al.* Early versus delayed treatment in patients with recent-onset rheumatoid arthritis: comparison of two cohorts who received different treatment strategies. Am J Med 2001;111:446–51.
15. Visser H, le Cessie S, Vos K, Breedveld FC, Hazes JM. How to diagnose rheumatoid arthritis early: a prediction model for persistent (erosive) arthritis. Arthritis Rheum 2002;46:357–65.
16. van der Helm-van Mil AH, le Cessie S, van Dongen H *et al.* A prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: how to guide individual treatment decisions. Arthritis Rheum 2007;56:433–40.
17. Aletaha D, Neogi T, Silman AJ 3rd *et al.* 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis 2010;69:1580–8.
18. Sokolove J, Bromberg R, Deane KD *et al.* Autoantibody epitope spreading in the pre-clinical phase predicts progression to rheumatoid arthritis. PLoS One 2012;7:e35296.
19. Rantapää-Dahlqvist S, de Jong BAW, Berglin E *et al.* Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. Arthritis Rheum 2003;48:2741–9.
20. Krijbolder DI, Verstappen M, van Dijk BT *et al.* Intervention with methotrexate in patients with arthralgia at risk of rheumatoid arthritis to reduce the development of persistent arthritis and its disease burden (TREAT EARLIER): a randomised, double-blind, placebo-controlled, proof-of-concept trial. Lancet 2022; 400:283–94.
21. Cope AP, Jasenecova M, Vasconcelos JC *et al.*; APIPPRA Study Investigators. Abatacept in individuals at high risk of rheumatoid arthritis (APIPPRA): a randomised, double-blind, multicentre, parallel, placebo-controlled, phase 2b clinical trial. Lancet 2024; 403:838–49.
22. Rech J, Tascilar K, Hagen M *et al.* Abatacept inhibits inflammation and onset of rheumatoid arthritis in individuals at high risk (ARIAA): a randomised, international, multicentre, double-blind, placebo-controlled trial. Lancet 2024;403:850–9.
23. van de Stadt LA, Witte BI, Bos WH, van Schaardenburg D. A prediction rule for the development of arthritis in seropositive arthralgia patients. Ann Rheum Dis 2013;72:1920–6.
24. Rakieh C, Nam JL, Hunt L *et al.* Predicting the development of clinical arthritis in anti-CCP positive individuals with non-specific musculoskeletal symptoms: a prospective observational cohort study. Ann Rheum Dis 2015;74:1659–66.
25. Nam JL, Hunt L, Hensor EM, Emery P. Enriching case selection for imminent RA: the use of anti-CCP antibodies in individuals with new non-specific musculoskeletal symptoms—a cohort study. Ann Rheum Dis 2016;75:1452–6.
26. Nam JL, Hensor EM, Hunt L *et al.* Ultrasound findings predict progression to inflammatory arthritis in anti-CCP antibody-positive patients without clinical synovitis. Ann Rheum Dis 2016; 75:2060–7.
27. Duquenne L, Hensor EM, Wilson M *et al.* Predicting inflammatory arthritis in at-risk persons: development of scores for risk stratification. Ann Intern Med 2023;176:1027–36.
28. Jordan KP, Kadam UT, Hayward R *et al.* Annual consultation prevalence of regional musculoskeletal problems in primary care: an observational study. BMC Musculoskelet Disord 2010;11:144.
29. NICE. Rheumatoid Arthritis in adults: management. NICE guideline [NG100]. 2018. https://www.nice.org.uk/guidance/ng100 (23 May 2024, date last accessed).
30. Nicholson A, Ford E, Davies KA *et al.* Optimising use of electronic health records to describe the pesentation of rheumatoid arthritis in primary care: a strategy for developing code lists. PLoS ONE 2013;8:e54878.
31. Muller S, Hider SL, Raza K *et al.* An algorithm to identify rheumatoid arthritis in primary care: a Clinical Practice Research Datalink study. BMJ Open 2015;5:e009309.
32. Monk HL, Muller S, Mallen CD, Hider SL. Cardiovascular screening in rheumatoid arthritis: a cross-sectional primary care database study. BMC Fam Pract 2013;14:150.
33. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD). Ann Intern Med 2015; 162:735–6.
34. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw 2010; 33:1–22.
35. Steyerberg E. Chapter 18: Presentation formats. In: Gail M, Krickeberg K, Samet J, Tsiatis A, Wong W, eds. Clinical prediction models: a practical approach to development, validation, and updating. Switzerland: Springer Nature, 2009:319–20.
36. Robin X, Turck N, Hainard A *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12:77.
37. Riley RD, Snell KI, Ensor J *et al.* Minimum sample size for developing a multivariable prediction model: PART II—binary and time-to-event outcomes. Stat Med 2019;38:1276–96.
38. Ensor J, Martin EC, Riley RD. pmsampsize: calculates the minimum sample size required for developing a multivariable prediction model. R package version 1.1.0. 2021. https://cran.r-project.

org/web/packages/pmsampsize/ (18 December 2024, date last accessed).
39. Sjoberg DD. dcurves: Decision Curve Analysis for Model Evaluation. R package version 0.4.0. 2022. https://cran.r-project.org/web/packages/dcurves/index.html (18 December 2024, date last accessed).
40. Vickers AJ, Cronin AM, Elkin EB, Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. BMC Med Inform Decis Mak 2008;8:53. DOI: 10.1186/1472-6947-8-53.
41. R Core Team. _R: A Language and Environment for Statistical Computing_. Vienna, Austria: R Foundation for Statistical Computing, 2024. https://www.R-project.org/ (18 December 2024, date last accessed).
42. Posit team. RStudio: Integrated Development Environment for R. Boston, MA: Posit Software, PBC, 2023. http://www.posit.co/ (18 December 2024, date last accessed).
43. Hernán MA, Hsu J, Healy B. A second chance to get causal inference right: a classification of data science tasks. Chance 2019; 32:42–9.
44. Sackett DL. Bias in analytic research. J Chronic Dis 1979; 32:51–63.
45. van Mulligen E, Bour SS, Goossens L *et al.* Is a 1-year course of methotrexate in patients with arthralgia at-risk for rheumatoid arthritis cost-effective? A cost-effectiveness analysis of the randomised, placebo-controlled TREAT EARLIER trial. Ann Rheum Dis. Published Online First: 18 September 2024. doi: 10.1136/ard-2024-226286.
46. Nell VP, Machold KP, Eberl G *et al.* Benefit of very early referral and very early therapy with disease-modifying anti-rheumatic drugs in patients with early rheumatoid arthritis. Rheumatology (Oxford) 2004;43:906–14.
47. van Mulligen E, Rutten-van Molken M, van der Helm-van Mil A. Early identification of rheumatoid arthritis: does it induce treatment-related cost savings? Ann Rheum Dis 2024; 83:1647–56.
48. Bunnewell S, Wells I, Zemedikun D *et al.* Predictors of perceived risk in first-degree relatives of patients with rheumatoid arthritis. RMD Open 2022;8:e002606. doi: 10.1136/rmdopen-2022-002606.

Rheumatology, 2024, 00, 1–10
https://doi.org/10.1093/rheumatology/keae653
Original Article