Predictive Value of Active Sacroiliitis in MRI for Flare Among Chinese Patients with Axial Spondyloarthritis in Remission

Qing Zheng · Wen Liu · Yu Huang · Zhenyu Gao · Yuanhui Wu · Xiaohong Wang ·

Meimei Cai · Yan He · Shiju Chen · Bin Wang · Lingyu Liu · Shuqiang Chen ·

Hongjie Huang · Ling Zheng · Rihui Kang · Xiaohong Zeng · Jing Chen ·

Huaning Chen · Junmin Chen · Zhibin Li · Guixiu Shi (iD)

ABSTRACT

Introduction: In recent axSpAx patients with remission lasting at least 3 months and later followed-up monthly for a median of 8 months, we compared the predictive value of baseline MRI of sacroiliac joints and constructed a nomogram model for predicting flare.

1. Qing Zheng, Wen Liu, and Yu Huang are co-first authors and contributed equally to this study. 2. Guixiu Shi, Zhibin Li, and Junmin Chen are corresponding authors. 3. Signature unit of the first author: Department of Rheumatology and Clinical Immunology, The First Affiliated Hospital of Xiamen University.

Q. Zheng · W. Liu · Z. Gao · Y. Wu · X. Wang ·
M. Cai · Y. He · S. Chen · B. Wang · L. Liu ·
G. Shi (✉)
Department of Rheumatology and Clinical
Immunology, The First Affiliated Hospital of
Xiamen University, Xiamen, Fujian, China
e-mail: gshi@xmu.edu.cn

Q. Zheng · L. Zheng · R. Kang · X. Zeng · J. Chen ·
H. Chen · J. Chen (✉)
Department of Rheumatology and Clinical
Immunology, The First Affiliated Hospital of Fujian
Medical University, Fuzhou, Fujian, China
e-mail: drchen_cn@sina.com

Q. Zheng
e-mail: drmichaelzheng@qq.com;
shiguixiuzq2020@163.com

Methods: This study included 251 patients with axial spondyloarthritis, according to the ASAS axSpA classification criteria, who achieved Low Disease Activity (ASDAS) and underwent MRI examination. A total of 144 patients from the First Affiliated Hospital of Xiamen University were used as the nomogram training set; 107 from the First Affiliated Hospital of Fujian Medical University were for external validation.
Results: The median time of relapse was 8.705 months (95% CI 8.215–9.195) and 7.781 months (95% CI 7.075–8.486) for MRI-positive patients and 9.8 months (95% CI 9.273–10.474) for MRI negative patients, respectively. Both active sacroiliitis on MRI (HR 1.792, 95% CI 1.230–2.611) and anti-TNF-α treatments (HR 0.507, 95% CI 0.349–0.736) were significantly associated with disease flares. Gender, disease

Y. Wu · M. Cai · Y. He · S. Chen · B. Wang · L. Liu ·
G. Shi
School of Medicine, Xiamen University, Xiangan
South Road, Xiangan District, Xiamen City, China

Y. Huang
Department of Anesthesiology, Fujian Maternity
and Child Health Hospital, Fuzhou, Fujian, China

Z. Gao
Rheumatology Department, The Affiliated Hospital
of Hubei Minzu University, Enshi, Hubei, China

S. Chen
Department of Ultrasound, The First Affiliated
Hospital of Fujian Medical University, Fuzhou,
Fujian, China

duration, HLA-B27, MRI, and anti-TNF-α treatment were selected as predictors of the nomogram. The areas under the ROC curve (AUROCs) of the 1-year remission probability in the training and validation groups were 0.71 and 0.729, respectively. Nomogram prediction models present better AUROCs, C-indices, and decision curve analysis cure than the clinical experience model.

Conclusions: Active sacroiliitis in MRI requires weighting in order to estimate remission and disease flares, when axSpA patients achieve low disease activity. The simple nomogram might be able to discriminate and calibrate in clinical practice.

Trial registration: ClinicalTrials, NCT03425 812, Registered 8 February 2018, https://clinicaltrials.gov

Keywords: Joint; Magnetic resonance imaging; Predictive value of tests; Sacroiliac spondylitis

---

Key Summary Points

Why carry out this study?

Magnetic resonance imaging (MRI) is a well-accepted technique for detecting bone marrow edema, which is a specific sign of osteitis in patients with axSpA.

The clinical hypothesis of this study is that the MRI status (negative or positive) of remised patients could be used for prediction of disease flare and should be weighted as an appropriate supplement for current clinical remission standards.

What was learned from the study?

MRI and anti-TNF-α treatments were independently related to disease flares. We fit a nomogram predictor including gender, disease duration, HLA-B27, MRI, and anti-TNF-α treatment with ROC curve of the 1-year remission probability in the training and validation groups were 0.71 and 0.729, respectively.

Our study suggested that the MRI status and anti-TNF-α treatment should be well considered among patients with axSpA for predicting risk of flare. Our nomogram predictive model for flare might be well validated before using in practice.

DIGITAL FEATURES

This article is published with digital features, including a summary slide, to facilitate understanding of the article. To view digital features for this article go to https://doi.org/10.6084/m9.figshare.13562414.

INTRODUCTION

Axial spondyloarthritis (axSpA) is a chronic inflammatory rheumatic disease characterized by a diversity of clinical symptomatology, including inflammatory back pain, enthesitis, dactylitis, and the formation of syndesmophytes, all of which significantly affect a patient's healthy functioning and general quality of life [1]. Previous studies reported that nearly 20–70% of patients with axSpA suffered from a flare, including single or repeated flares, as well as flares with no intermittent symptoms [2, 3]. Evidence from the Trial with Human rFSH in Europe and the Rest of the World (ESTHER trial) demonstrated that patients with axSpA can still suffer from flare events, even with the administration of etanercept (ETA) [4]. Constant symptoms of flares are significantly associated with a poorer prognosis and an overall less healthy status [3, 5]. However, little is

H. Huang
Department of Radiology, The First Affiliated Hospital of Fujian Medical University, Fuzhou, Fujian, China

Z. Li (✉)
Epidemiology Research Unit, Center of Translational Medical Research, The First Affiliated Hospital of Xiamen University, Xiamen, Fujian, China
e-mail: zhibinli33@163.com

X. Wang
CT Department, The First Affiliated Hospital of Xiamen University, Xiamen, Fujian, China

known about the influence of the risk factors and the mechanism of a flare. Therefore, it is necessary to explore the potential risk factors of flares, which may provide key information regarding the prediction of flares, as well as the prognosis for axSpA patients.

Magnetic resonance imaging (MRI) is a well-accepted technique for detecting bone marrow edema, which is a specific sign of osteitis in patients with axSpA [6]. Moreover, erosions occurring above a certain Spondyloarthritis Research Consortium of Canada (SPARCC) Enthesitis Index score threshold were highly specific for axial SpA, excluding postpartum women [7]. However, MacKay et al. reported that there was no significant correlation between total SPARCC and the Ankylosing Spondylitis Disease Activity Score with CRP (ASDAS-CRP) [8]. Van der Heijde et al. also reported that there was a lack of association between clinical remission and MRI-demonstrated remission [9]. Furthermore, MRI sacroiliitis can be observed in recurrent acute anterior uveitis (rAAU) patients without any musculoskeletal symptoms [10]. One possible reason is that SPARCC from MRI may provide additional imaging information about disease activity, which cannot be provided by ASDAS. This additional information from MRI imaging may be helpful in assessing the activity of axSpA. Therefore, one clinical question is: if a patient achieves clinical remission, could the information of his/her MRI status (negative or positive) be considered as a predictor for the possibility of disease flare, and does it need to be weighted as an appropriate supplement for current clinical remission standards?

In this study, we initially sought to compare the difference in MRI status of active sacroiliitis between patients with and without flare, and then to evaluate whether active sacroiliitis in MRI also needs to be considered once clinical remission has occurred. Previous studies suggested that gender, age, disease duration time, anti-TNF-α treatment, and HLA-B27 more or less affected the disease flare [11–13]. However, a prediction model combining various factors has not yet been reported. Therefore, our second aim is to construct a set of nomogram prediction models to facilitate the forecasting of the possibility of flare for clinical remission in patients with axSpA.

METHODS

Study Design and Participants

Effects of nonsteroidal anti-inflammatory drugs in recurrence of spondyloarthritis patients after remission (NASA study, ClinicalTrials.gov ID: NCT03425812) is an ongoing multicenter, randomized, parallel-controlled study conducted at Fujian, China from 2017 to 2021. The NASA study aimed to investigate the recurrence rate in remission patients with axSpA who withdrew from nonsteroidal anti-inflammatory therapy. As part of the NASA study, eligible participants who fulfilled the following criteria were included in our study: those aged $\geq$ 18 years at screening; those who fulfilled the ASAS (Assessment of Spondyloarthritis International Society) classification criteria for axSpA [14, 15]; were followed up regularly for more than 1 year; achieved a Low Ankylosing Spondylitis Disease Activity Score (ASDAS) [12]; and underwent MRI examination. Follow-up of these patients was continued for 12 months (Fig. 1).

The ASAS classification criteria for axSpA is defined as: the presence of sacroiliitis on plain-film radiography or on MRI, plus at least one SpA feature, or, the presence of HLA-B27 plus at least two SpA features [16]. An ASDAS below 2.1 is defined as low disease activity in a 2018 update of the nomenclature for disease activity states because the majority of patients consider themselves as being in a patient-acceptable symptom state (PASS) in this category [12, 17, 18]. Patients who achieved Low Disease Activity score (ASDAS $\leq$ 2.1) and sustained for 3 months or more were included. Considering the high false-positive diagnosis of axSpA by using MRI among pregnancy women and athletes, patients were not eligible if they had pregnancy-related low back pain [19] or if they participated in intensive physical training, such as military recruits and athletes [20, 21]. Other excluding criteria included: autoimmune diseases other than axSpA, chronic infection, severe organ failure, pregnancy, and mental illness.
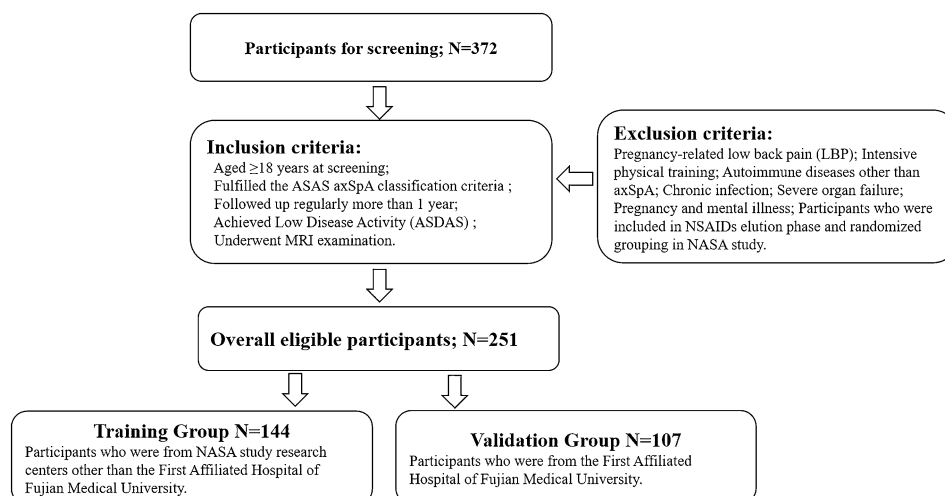
Fig. 1 Flowchart of participants selection

Participants who were included in NSAIDs elution phase and randomized grouping in NASA study were also excluded. We divided participants into a training group and a validation group for model development. Participants who were from the First Affiliated Hospital of Fujian Medical University were included in the validation group. Participants who were from research centers other than the First Affiliated Hospital of Fujian Medical University were included in the training group (Fig. 1).

Ethical approval for the NASA study, including the consent procedure, was granted by the Research Ethics committee of the First Affiliated Hospital of Xiamen University (reference number: KYH2018-007) and the First Affiliated Hospital of Fujian Medical University (reference number: MRCTA, ECFAH of FMU [2018] 198) with the Declaration of Helsinki V and the Danish legislation. All participants gave written informed consent prior to study inclusion.

Clinical Assessment

History and physical examination were performed by rheumatologists at baseline and at 4-week interval visits for a duration of 1 year. The general information collected at the baseline evaluation visit included age, gender, disease duration of axSpA, tobacco smoking status, alcohol drinking status, history of inflammatory

back pain, family history of axSpA, and presence of any extra-articular manifestations (including psoriasis, inflammatory bowel disease, and uveitis).

All selected participants completed outcome measurements. The outcome measurements in spondylarthritis were recorded following the report from Jane Zochling [22], which included ASDAS-CRP and Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) at baseline and at the follow-up visits. Details of past and present antirheumatic therapies were also recorded, such as nonsteroidal anti-inflammatory drugs (NSAID), antitumor necrosis factor-α (anti-TNF-α) treatment, and disease-modifying antirheumatic drugs (DMARDs). These therapies were considered noteworthy only if the patient related a history of regular administration of a particular therapy, and if the total time of the therapy administration (both regular administration and tapering) was greater than 50% of the time from enrollment to the end of study. Peripheral arthritis was defined as presence of pain, swelling, and/or tenderness of any peripheral joint excluding shoulder and hip joints.

Laboratory Assessment

Human leucocyte antigen B27 (HLA-B27; flow cytometry method, Beckman Coulter Inc.) status, erythrocyte sedimentation rate (ESR, mm/

h), and C reactive protein (CRP, mg/l; immune turbidimetry method, Beckman Coulter Inc.) were assessed as well. A limited value of 2 mg/l was used when CRP level was below 5 mg/l [23].

MRI Scoring

An MRI of the sacroiliac (SI) joints was performed on each patient at the time of his/her baseline visit, using the Siemens Magnetom Skyra 3.0 T MRI with Picture Archiving and Communication System. The acquired MRI images included semi-coronal T1-weighted imaging sequences with a repetition time of 700 ms, echo time of 24 ms, slice thickness of 3 mm, gap of 0.45 mm, 200 × 200-mm field of view, number of excitations of two, and 384 × 288–pixel matrix size. A semi-coronal short tau inversion recovery (STIR) sequence was also applied with repetition time of 3500 ms, inversion time of 150 ms, echo time of 45 ms, slice thickness of 3 mm, gap of 0.45 mm, 200 × 200-mm field of view, number of excitations of three, and 320 × 224-pixel matrix size [7, 24, 25].

The MRIs were anonymized and evaluated by two independent, experienced radiologists with extensive, and well-calibrated axSpA MRI reading experience who were blinded to the clinical data and the patient's final diagnosis. In the case of score discrepancies, a third radiologist with 20 years of MRI experience scored the images. According to the ASAS definition, an MRI scan of SI joints was recorded as positive (active sacroiliitis) if BME was clearly present and located in a typical anatomical area (subchondral bone), on a T2-weighted sequence, and on at least two consecutive slices of an MRI scan. Detection of inflammation on a single slice would have to show the presence of more than one inflammatory lesion. Other clinical conditions, such as infection, tumor, and trauma had to be excluded [26]. The STIR of MRI scans was assessed for disease activity using the SPARCC system for sacroiliac joints (0–72 scale) [9, 24, 27]. SPARCC score ≥ 2 is defined as SPARCC positive [9].

Outcome and Censoring

The primary outcome in this study was flare, which was defined as one flare, occurring after Low Disease Activity (ASDAS), if ASDAS-CRP ≥ 2.1 [17] and the change of ASDAS-CRP ≥ 0.9 during the 12-month follow-up [28]. Censoring included patients who did not experience the event duration the follow-up period, patients who were lost to follow-up during the study period, or patients who withdrew from the study [29].

Quality Control

The NASA study followed the principles of Good Clinical Practice [30], the Declaration of Helsinki V, and the Danish legislation. A standard protocol of quality control was established for this study. An investigators' meeting was held prior to the initiation of the study to discuss the scheme, research operations, and the method of sample detection in detail. Standard training programs for data collection and method of sample detection for investigators from all study centers were administered. We set up a multicenter coordinating committee to coordinate research matters. Professional clinical research auditors were employed in order to ensure the quality of collected data and information. Each center established regulations and standard management procedures and Standard Operation Procedure [31] in order to strengthen the training of researchers and to ensure the progress of the trial.

Statistical Analysis

Medians and interquartile ranges (IQR) of continuous variables and frequencies (percentages) of categorical variables were calculated. Mann–Whitney tests were used for comparison of patient characteristics.

A total of 15 potential predictors were selected in this study based upon our literature review [11, 12], expert consensus from a Delphi study [32], and our previous cohort study [13, 25, 33]. We then selected four predictors by multiple Cox regression models, one predictor

selected by literature review, and an experts' consensus in order to construct a nomogram (Figure S1) [34]. Cox regression models were performed to explore the potential confounders among the overall participants. Variables with a p value of less than 0.05 in crude models were selected as predictors of the multiple Cox regression models to identify the independent predictors. The forward likelihood ratio method was used to select independent risk factors. Factors entered the stepwise model at $p < 0.05$ and exited at $p < 0.10$. A Kaplan–Meier curve was used to estimate means for relapse time and assess the proportional hazards assumption (PH) for categorical variables. Goodness of fit testing (Schoenfeld residuals) was used to assess PH assumption for continuous data [29].

In the training group, in order to construct nomogram prediction models, a crude Cox regression model was created, and selected those factors reaching a significance threshold of $p < 0.25$ to forward stepwise multiple Cox regression models in order to determine predictors (according to each of the response criteria examined). Factors entered the stepwise model at $p < 0.05$ and exited at $p < 0.27$. The coefficients of the multiple Cox regression model were used to generate a nomogram. Then, we evaluated the validity of the nomogram in the validation group. The concordance index (C-index) and AUROC were used to evaluate the discrimination ability of the nomogram. Calibration curves were plotted to assess the calibration ability of the nomogram. Decision curve analysis (DCA) was performed to demonstrate the clinical usefulness of the nomogram [33–35]. DCA is a novel method for evaluating alternative prognostic strategies, which has advantages over AUROC [36, 37].

We further constructed a clinical experience model following the same procedure as was mentioned earlier; the C-index, ATROC and DCA were also calculated to compare whether the nomogram prediction model was better than the clinical experience model.

We used the research data collecting system (RDC, version 1.0, patent application number 202010087680.0) to collect the data. The R (The R Foundation for Statistical Computing, version 3.6.2) statistical packages "rms," "survival,"

"foreign," and "survival ROC" were used to generate the nomogram, calculate the C-index, and plot the calibration curves, the AUROC curves, and Kaplan–Meier curves, respectively. The source file "stdca.r" was obtained from the website www.mskcc.org, which was used to draw the DCA curves. Two-sided p value < 0.05 was considered statistically significant.

RESULTS

Demographic and Clinical Characteristics

Between 2017 and 2019, a total of 372 participants were screened for inclusion in the NASA cohort. A total of 251 patients (67.4%) met the criteria for study inclusion and completed their baseline visit (Fig. 1). Overall, the median age was 31 years (IQR: 27–39 years) and 66.9% were male. Among them, the symptom duration at baseline was 4 years (IQR: 2–6 years) (Table 1). Most of the participants (89.2%, $n = 224$) were receiving NSAIDs therapy, more than half (65.3%, $n = 164$) were receiving DMARDs, and nearly one-half of the participants (49.4%, $n = 124$) were receiving anti-TNF-α treatment. The median ASDAS was 1.31 (IQR: 0.69–2.63) and 55.8% ($n = 140$) of participants were MRI positive. Participants in the flare group had longer symptom duration at their first visit, a lower percentage with anti-TNF-α treatment and MRI negative status, and a higher SPARCC score of the SI joints as compared with those of the non-flare group (all p values < 0.05). The mean relapse-free time for total participants was 8.7 months (95% CI 8.21–9.19); and 7.78 months (95% CI 7.07–8.48) for MRI-positive participants and 9.87 months (95% CI 9.27–10.47) for MRI negative participants, respectively (p for log-rank test < 0.05). Among all patients, peripheral joint involvement was present in 33.4% of all patients (84/251); 40 (32.0%) in non-flare group and 44 (34.9%) in flare group. There was no statistical difference in peripheral joint involvement in two groups ($p = 0.625$). A total of 25.0% patients (63/251) presented nr-axSpA at baseline, 36 (28.8%) in non-flare, and 27 (21.4%) in the flare group.

Table 1 Characteristics of participants including in this study

Variables	Total $n = 251$	Non-flare $n = 125$	Flare $n = 126$	p value
Female, n (%)	83 (33.1)	38 (30.4)	45 (35.7)	0.371
Age at enrollment (years)	31 (27, 39)	33 (28, 42.5)	31 (26, 38)	0.760
Symptom duration at first visit (years)	4 (2, 6)	4 (2, 5)	5 (2, 7)	0.037
Peripheral arthritis, n (%)	84 (33.4)	40 (32.0)	44 (34.9)	0.625
Non-radiographic axial SpA, n (%)	63 (25.0)	36 (28.8)	27 (21.4)	0.178
History of IBP, n (%)	199 (79.3)	99 (79.2)	100 (79.4)	0.974
Family history of SpA, n (%)	37 (14.7)	22 (17.6)	15 (11.9)	0.203
History of IBD, n (%)	10 (4.0)	6 (4.8)	4 (3.2)	0.510
History of psoriasis, n (%)	6 (2.4)	3 (2.4)	3 (2.4)	0.992
History of enthesitis, n (%)	38 (15.1)	17 (13.6)	21 (16.7)	0.498
Alcohol drinkers, n (%)	191 (76.0)	98 (78.4)	93 (73.8)	0.394
Current cigarette smoker, n (%)	56 (22.3)	30 (24.0)	26 (20.6)	0.522
HLA-B27, n (%)	173 (68.9)	81 (64.8)	92 (73.0)	0.160
Drug use information, n (%)				
Anti-TNF-α drugs	124 (49.4)	80 (64)	44 (34.9)	< 0.0001
NSAIDs	224 (89.2)	115 (92.0)	109 (86.5)	0.160
DMARDs	164 (65.3)	82 (65.6)	82 (65.1)	0.931
CRP at baseline (mg/l)	2.4 (0.5, 5.0)	2.3 (0.5, 4.3)	2.6 (0.5, 5.5)	0.225
ESR at baseline (mm/h)	10 (5, 17)	10 (5, 17)	11 (5, 17)	0.999
ASDAS at baseline	1.1 (0.69, 1.63)	1.1 (0.70, 1.42)	1.18 (0.61, 1.73)	0.315
BASDAI at baseline	1.9 (1.30, 2.90)	1.8 (1.26, 2.80)	1.9 (1.29, 3.00)	0.397
With positive MRI, n (%)	140 (55.8)	56 (44.8)	84 (66.7)	< 0.0001
SPARCC score, SI joints	7 (0, 19)	2 (0, 12)	12 (0, 22)	< 0.0001

There was no statistical difference for nr-axSpA in two groups ($p = 0.178$).

Independent Risk Factors of Flares

Crude Cox proportional hazards models among all participants showed that features of active sacroiliitis on MRI and history of anti-TNF-α treatment were significant predictors of flare at the 0.05 level of significance (Table 2). Multiple Cox regression models for the significant variables from crude Cox models found that both MRI status (HR 1.792, 95% CI 1.230–2.611, $p = 0.002$) and anti-TNF-α treatment (HR 0.507, 95% CI 0.349–0.736, $p < 0.0001$) were the significantly associated with disease flares.

Table 2 Predictors associated with disease flare using univariate and multiple Cox regression models in training group

Variables	Crude models			Multivariate models		
	p value	HR	95% CI	*p* value	HR	95% CI
Females vs. males	0.603	1.160	(0.662, 2.032)			
Age at enrollment (years)	0.709	0.994	(0.966, 1.024)			
Symptom duration at first visit (years)	0.224	1.081	(0.953, 1.227)	0.088	1.127	(0982, 1.292)
History of IBP	0.806	0.932	(0.532, 1.633)			
Family history of SpA	0.200	0.617	(0.295, 1.291)			
History of IBD	0.448	0.580	(0.142, 2.368)			
History of psoriasis	0.965	1.045	(0.145, 7.532)			
History of enthesitis	0.927	1.031	(0.540, 1.966)			
Alcohol drinkers	0.996	1.002	(0.537, 1.869)			
Current cigarettes smoker	0.519	0.824	(0.458, 1.484)			
HLA-B27	0.196	1.437	(0.830, 2.489)	0.266	1.368	(0.788, 2.374)
Anti-TNF-α drugs use	0.010	0.522	(0.319, 0.856)	0.027	0.568	(0.344, 0.938)
NSAIDs use	0.156	0.636	(0.341, 1.189)			
DMARDs use	0.664	0.897	(0.548, 1.468)			
With positive MRI	0.014	1.991	(1.149, 3.451)	0.023	1.941	(1.097, 3.435)

Construction and Validation of the Nomogram Prediction Models

Among the overall 251 eligible participants, 144 participants, and 107 participants were included in the training and validation groups, respectively. The characteristics of participants in the training and validation groups are shown in Table 3.

Initially, we selected four predictors (symptom duration first visit, HLA-B27, MRI, and anti-TNF-α treatment) using the univariate and multiple Cox regression models in the training group (Table 4). Combined with the other 11 predictors selected by literature review [11, 12], expert consensus from a Delphi study [32], our previous cohort study [13, 25], and for the purpose of simplified clinical application [33], a total of 15 potential predictors were included for the nomogram predictive model. Finally, five predictors, including gender, symptom duration at first visit, HLA-B27, MRI, and anti-

TNF-α treatment, were incorporated into the nomogram model (Fig. 2).

The calibration curves of the nomogram for the training set and validation set showed a good probability consistency between the prediction and observation (Fig. 3a, b). The C-indices for the prediction of overall relapse in the training and validation groups were 0.66 (95% CI 0.646–0.67) and 0.667 (95% CI 0.653–0.681), respectively. AUROCs of the 1-year remission probability in the training and validation groups were 0.71 and 0.729, respectively (Fig. 3c, d).

Compared with Clinical Experience Model

We further constructed a clinical experience model including gender, disease duration, HLA-B27, based on clinical experience and literature review [11, 12] to compare with the nomogram model. Compared with the clinical experience model, the calibration curves of the nomogram
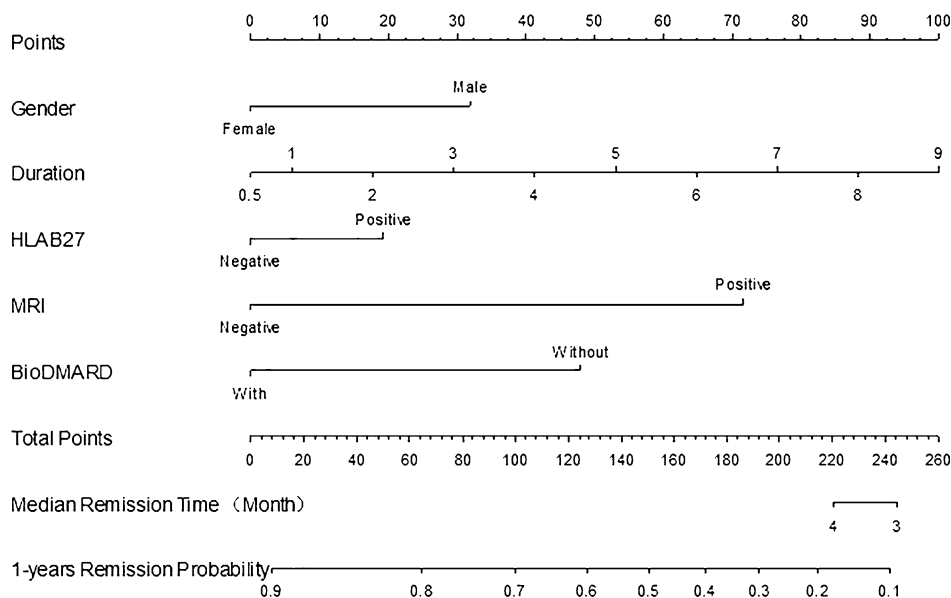
Fig. 2 Nomogram model for predicting flare in axSpA patients achieving low disease activity

showed a good probability consistency between the prediction and observation (Fig. 4a, b). The C-indices of the clinical experience model in the training and validation groups were 0.547 (95% CI 0.0129 0.534–0.559) and 0.576 (95% CI 0.561–0.591), respectively. AUROCs of the clinical experience model was 0.589 and 0.554, respectively, for predicting 1-year remission probability in the training and validation groups (Fig. 4c, d). Compared with the clinical experience model, the nomogram had higher net benefits according to the DCA curves, which indicated that it had better clinical utility (Figure S2).

DISCUSSION

Based on the data from a multicenter randomized clinical trial, we found that MRI status and anti-TNF-α treatment were independent predictive factors of flare among patients with axSpA. We further constructed a nomogram predictive model for flare which provided adequate ability of discrimination and calibration, and good clinical utility.

Many previous studies report that MRI is good at detecting BME for the early diagnosis of axSpA [6], however, using MRI in monitoring the patient's response to therapy and managing treatment is still not widely accepted [27, 32]. In our previous study, we measured the capability of MRI to predict the possibility of flare in a small cohort [25]. In this study, we further use a prospective study of 251 participants to evaluate the predictive value of MRI and construct the predictive model of flare among axSpA patients. We used Cox proportional hazards regression models to detect the independent risk factors associated with disease flares when patients had reached a patient-acceptable symptom state (PASS, ASDAS < 2.1). Our results showed that MRI status at the baseline visit was an independent risk factor for flares; which suggested that MRI status also needs to be considered for patient management following clinical remission. We found that anti-TNF-α treatment was another independent predictor significantly related to disease flares. This conclusion is consistent with that from previous studies, which also suggested that anti-TNF-α treatment might directly affect the disease flare [11, 12]. Our previous study found that anti-TNF-α treatment could significantly improve the clinical symptoms and enhance joint activity among axSpA patients [13], which might partially explain the connection. We reconfirmed the previous conclusion that anti-
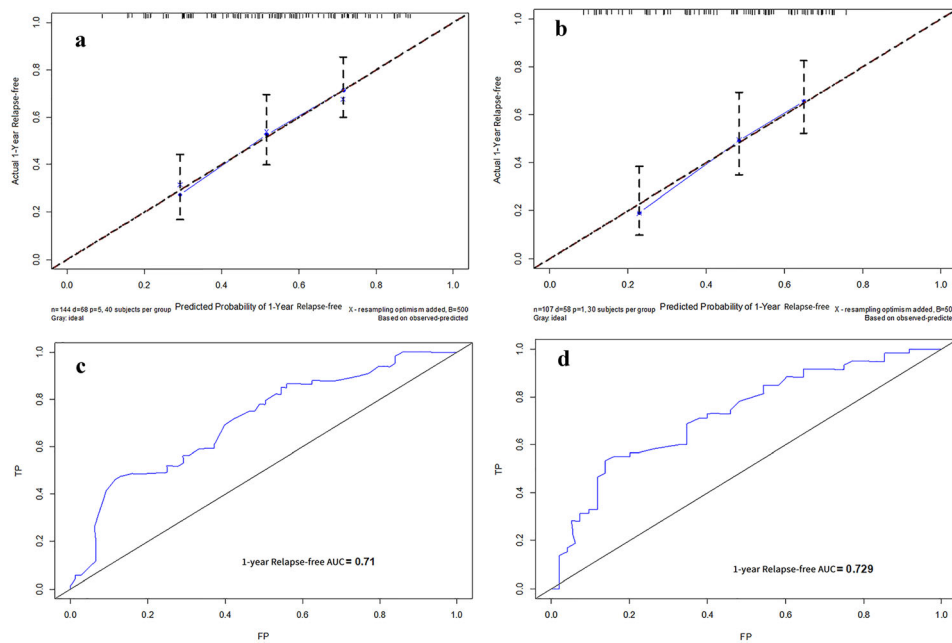
Fig. 3 Calibration curves of the nomogram for the training set (**a**) and validation set (**b**). The areas under the AUROC curve of the 1-year remission probability in the training (**c**) and validation (**d**) groups, respectively
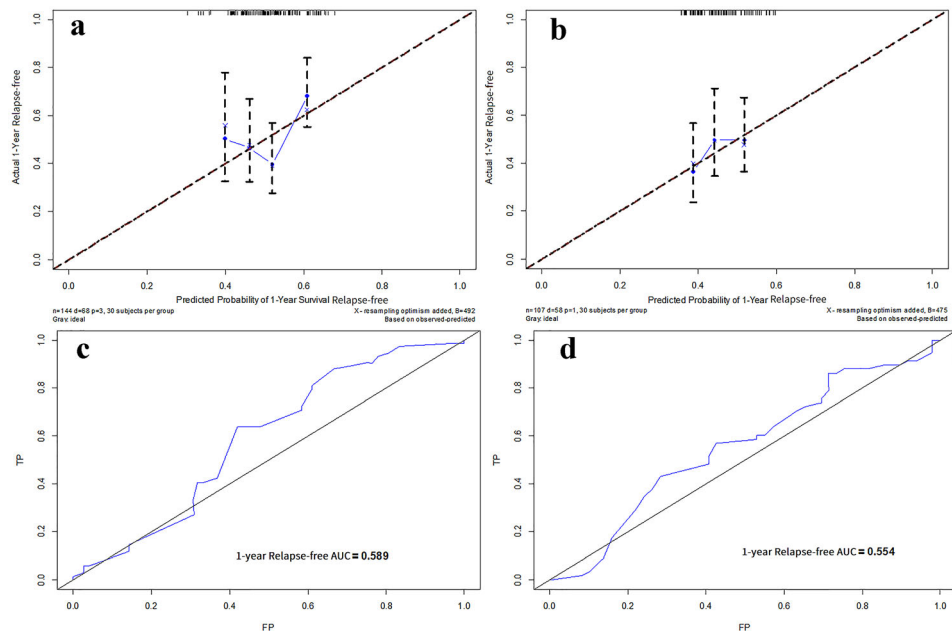

Fig. 4 The C-indices of the clinical experience model in the training (**a**) and validation groups (**b**), respectively. AUROCs of the clinical experience model for predicting 1-year remission probability in the training (**c**) and validation groups (**d**)

TNF-α treatment can significantly improve the therapeutic effects of axSpA.

Clinically, various factors and clinical conditions should be considered and balanced when planning to modify patient management

after clinical remission, such as, gender, age, duration of disease, HLA-B27, MRI result, and anti-TNF-α treatment [11–13]. However, some of these factors are primarily used for research purposes. For example, the MRI SPARCC scoring system provides an objective measure of disease activity for axial SpA, but it is barely used in clinical practice [32]. This is primarily because the SPARCC system requires MRI and scoring training, and it is a time-consuming procedure, which is not easy for clinicians to use. The ASDAS score in practice is also faced with the same problem. Thus, we attempted to transform SPARCC scores into a simple dichotomous variable, a positive or negative MRI status. Then we constructed a nomogram prediction model, which may help rheumatologists to predict and evaluate the 1-year probability of remission by a very simple scoring and visualizing mode. Our nomogram model contained only five simple variables, representing a straightforward visual tool for the prediction in clinical practice. To the best of our knowledge, this is the first nomogram model predicting axSpA maintaining remission and disease flares, in order to modify therapy methods. The AUROCs of nomogram prediction models could be range from 0.71 to 0.729, suggesting a relative acceptably sensitivity and specificity for the prediction. It could be helpful for doctors to quickly identify patients with high- and low-risk of flares, and then to decide whether a careful follow-up should be performed in higher-risk patients.

Furthermore, we also compared the validity of the nomogram mode with a clinical experience model that included gender, disease duration, and HLA-B27 value. The C-indices of the clinical experience model in the training and validation groups were no larger than 0.60, and the AUROCs of the clinical experience model were 0.589 to 0.554, which is far lower than for the nomogram prediction models. We also applied the DCA, a novel method for evaluating alternative prognostic strategies [36, 37], and found an advantage in the nomogram model.

This study has some limitations. Due to the limited number of participants in this cohort, further large-scaled studies are warranted in order to verify the sensitivity and specificity of the model. Also, the participants were from a clinical trial, which might limit the extent of the conclusions; studies in other countries and areas of the world are needed in order to reconfirm our results. Furthermore, additional possible and practical predictors may also be considered.

CONCLUSIONS

In conclusion, using the data from a well-designed clinical trial, this study found that active sacroiliitis in MRI and anti-TNF-α treatment need to be weighted in order to estimate remission and disease flares, when axSpA patients achieve low disease activity. We developed and validated a nomogram predictive model for predicting the remission and flare probability among axSpA patients who have achieved low disease activity. This simple nomogram model had an adequate ability to discriminate and calibrate and appropriate clinical utility. It may be a useful and helpful tool for both patients and doctors conducting a post-remission evaluation.

ACKNOWLEDGEMENTS

Disclosures. Qing Zheng's previous Institution: Health Science Center, University of Sydney, Camperdown/Darlington NSW 2006, Sydney, Australia. Wen Liu, Yu Huang, Zhenyu Gao, Yuanhui Wu, Xiaohong Wang, Meimei Cai, Yan He, Shiju Chen, Bin Wang, Lingyu Liu, Shuqiang Chen, Hongjie Huang, Ling Zheng, Rihui Kang, Xiaohong Zeng, Jing Chen, Huaning Chen, Junmin Chen, Zhibin Li, and Guixiu Shi have nothing to disclose.

Compliance with Ethics Guidelines. Ethical approval for the NASA study, including the consent procedure, was granted by the Research Ethics committee of the First Affiliated Hospital of Xiamen University (reference number: KYH2018-007) and the First Affiliated Hospital of Fujian Medical University (reference number: MRCTA,ECFAH of FMU [2018]198) in compliance with the Declaration of Helsinki V and the Danish legislation. All participants gave written informed consent prior to study inclusion.

Data Availability. The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Open Access. This article is licensed under a Creative Commons Attribution-Non-Commercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES

1. van der Heijde D, Ramiro S, Landewe R, Baraliakos X, Van den Bosch F, Sepriano A, et al. 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. Ann Rheum Dis. 2017;76:978–91.

2. Cooksey R, Brophy S, Gravenor MB, Brooks CJ, Burrows CL, Siebert S. Frequency and characteristics of disease flares in ankylosing spondylitis. Rheumatology. 2010;49:929–32.

3. Stone MA, Pomeroy E, Keat A, Sengupta R, Hickey S, Dieppe P, et al. Assessment of the impact of flares in ankylosing spondylitis disease activity using the flare illustration. Rheumatology (Oxford). 2008;47:1213–8.

4. Song IH, Althoff CE, Haibel H, Hermann KG, Poddubnyy D, Listing J, et al. Frequency and duration of drug-free remission after 1 year of treatment with etanercept versus sulfasalazine in early axial spondyloarthritis: 2 year data of the ESTHER trial. Ann Rheum Dis. 2012;71:1212–5.

5. Cooksey R, Brophy S, Dennis M, Davies H, Atkinson M, Irvine E, et al. Severe flare as a predictor of poor outcome in ankylosing spondylitis: a cohort study using questionnaire and routine data linkage. Rheumatology (Oxford). 2015;54:1563–72.

6. Baraliakos X. Imaging in axial spondyloarthritis. Israel Med Assoc J. 2017;19:712–8.

7. Seven S, Ostergaard M, Morsel-Carlsen L, Sorensen IJ, Bonde B, Thamsborg G, et al. Magnetic resonance imaging of lesions in the sacroiliac joints for differentiation of patients with axial spondyloarthritis from control subjects with or without pelvic or buttock pain: a prospective, cross-sectional study of 204 participants. Arthritis Rheumatol. 2019;71:2034–46.

8. MacKay JW, Aboelmagd S, Gaffney JK. Correlation between clinical and MRI disease activity scores in axial spondyloarthritis. Clin Rheumatol. 2015;34:1633–8.

9. Heijde D, Baraliokos X, Hermann KA, Landewé RBM, Machado PM, Maksymowych WP, et al. Limited radiographic progression and sustained reductions in MRI inflammation in patients with axial spondyloarthritis: 4-year imaging outcomes from the RAPID-axSpA phase III randomised trial. Ann Rheum Dis. 2018;77:699–705.

10. Oliveira TL, Maksymowych WP, Lambert RGW, Muccioli C, Fernandes ARC, Pinheiro MM. Sacroiliac joint magnetic resonance imaging in asymptomatic patients with recurrent acute anterior

uveitis: a proof-of-concept study. J Rheumatol. 2017;44:1833–40.

11. Sieper J, Rudwaleit M, Lenaerts J, Wollenhaupt J, Myasoutova L, Park SH, et al. Partial remission in ankylosing spondylitis and non-radiographic axial spondyloarthritis in treatment with infliximab plus naproxen or naproxen alone: associations between partial remission and baseline disease characteristics. Rheumatology (Oxford). 2016;55:1946–53.

12. Jacquemin C, Maksymowych PW, Boonen A, Gossec L. Patient-reported flares in ankylosing spondylitis: a cross-sectional analysis of 234 patients. J Rheumatol. 2017;44:425–30.

13. Qing Z, Jun XM. Clinical efficacy of recombinant human II tumor necrosis factor receptor-antibody fusion protein in treatment of ankylosing spondylitis. Chin J Clin Pharmacol. 2016;32:27–30.

14. van der Linden S, Akkoc N, Brown MA, Robinson PC, Khan MA. The ASAS criteria for axial spondyloarthritis: strengths, weaknesses, and proposals for a way forward. Curr Rheumatol Rep. 2015;17:62.

15. Sieper J, Poddubnyy D. Axial spondyloarthritis. Lancet. 2017;390:73–84.

16. Rudwaleit M, van der Heijde D, Landewe R, Listing J, Akkoc N, Brandt J, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis. 2009;68:777–83.

17. Machado PM, Landewe R, Heijde DV. Assessment of SpondyloArthritis international S Ankylosing Spondylitis Disease Activity Score (ASDAS): 2018 update of the nomenclature for disease activity states. Ann Rheum Dis. 2018;77:1539–40.

18. Rodriguez-Lozano C, Gantes MA, Gonzalez B, Hernandez-Beriain JA, Naranjo A, Hernandez V, et al. Patient-acceptable symptom state as an outcome measure in the daily care of patients with ankylosing spondylitis. J Rheumatol. 2012;39:1424–32.

19. Arnbak B, Jensen TS, Egund N, Zejden A, Horslev-Petersen K, Manniche C, et al. Prevalence of degenerative and spondyloarthritis-related magnetic resonance imaging findings in the spine and sacroiliac joints in patients with persistent low back pain. Eur Radiol. 2016;26:1191–203.

20. Varkas G, de Hooge M, Renson T, De Mits S, Carron P, Jacques P, et al. Effect of mechanical stress on magnetic resonance imaging of the sacroiliac joints: assessment of military recruits by magnetic resonance imaging study. Rheumatology (Oxford). 2018;57:508–13.

21. Weber U, Jurik AG, Zejden A, Larsen E, Jorgensen SH, Rufibach K, et al. Frequency and anatomic distribution of magnetic resonance imaging features in the sacroiliac joints of young athletes: exploring "Background Noise" toward a data-driven definition of sacroiliitis in early spondyloarthritis. Arthritis Rheumatol. 2018;70:736–45.

22. Zochling J. Measures of symptoms and disease status in ankylosing spondylitis. Arthritis Care Res. 2011;63:S47–58.

23. Machado P, Navarro-Compan V, Landewe R, van Gaalen FA, Roux C, van der Heijde D. Calculating the ankylosing spondylitis disease activity score if the conventional c-reactive protein level is below the limit of detection or if high-sensitivity c-reactive protein is used: an analysis in the DESIR cohort. Arthritis Rheumatol. 2015;67:408–13.

24. Maksymowych WP, Inman RD, Salonen D, Dhillon SS, Williams M, Stone M, et al. Spondyloarthritis research Consortium of Canada magnetic resonance imaging index for assessment of sacroiliac joint inflammation in ankylosing spondylitis. Arthritis Rheum. 2005;53:703–9.

25. Huang H, Zhang Y, Zhang H, Chen J, Zheng Q, Cao D, et al. Qualitative and quantitative assessment of sacroiliitis in axial spondyloarthropathy: can a single T2-weighted Dixon sequence replace the standard protocol? Clin Radiol. 2020;75(4):321. e313–321.e320.

26. Lambert RG, Bakker PA, van der Heijde D, Weber U, Rudwaleit M, Hermann KG, et al. Defining active sacroiliitis on MRI for classification of axial spondyloarthritis: update by the ASAS MRI working group. Ann Rheum Dis. 2016;75:1958–63.

27. Bakker PAC, Ramiro S, Ez-Zaitouni Z, van Lunteren M, Berg IJ, Landewe R, et al. Is it useful to repeat magnetic resonance imaging of the sacroiliac joints after three months or one year in the diagnosis of patients with chronic back pain and suspected axial spondyloarthritis? Arthritis Rheumatol. 2019;71:382–91.

28. Gossec L, Portier A, Landewe R, Etcheto A, Navarro-Compan V, Kroon F, et al. Preliminary definitions of "flare" in axial spondyloarthritis, based on pain, BASDAI and ASDAS-CRP: an ASAS initiative. Ann Rheum Dis. 2016;75:991–6.

29. Kleinbaum DG, Klein M. Survival analysis a self-learning text. 3rd ed. Berlin: Springer; 2012.

30. Pieterse H, Diamant Z. Good clinical practice in clinical interventional studies. Eur Clin Respir J. 2014;12(1):26422.

31. Singh J. International conference on harmonization of technical requirements for registration of pharmaceuticals for human use. J Pharmacol Pharmacother. 2015;6:185–7.

32. Marchesoni A, D'Angelo S, Anzidei M, Bortolotti R, Cantini F, Caramella D, et al. Radiologist-rheumatologist multidisciplinary approach in the management of axial spondyloarthritis: a Delphi consensus statement. Clin Exp Rheumatol. 2019;37:575–84.

33. Lee YH, Bang H, Kim DJ. How to establish clinical prediction models. Endocrinol Metab (Seoul). 2016;31:38–44.

34. Han K, Song K, Choi BW. How to develop, validate, and compare clinical prediction models involving radiological parameters: study design and statistical methods. Korean J Radiol. 2016;17:339–50.

35. Chen S-H, Wan Q-S, Zhou D, Wang T, Hu J, He Y-T, et al. A simple-to-use nomogram for predicting the survival of early hepatocellular carcinoma patients. Front Oncol. 2019;9:1–11.

36. Fitzgerald M, Saville BR, Lewis RJ. Decision curve analysis. JAMA. 2015;313:409–10.

37. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Mak. 2006;26:565–74.