OP0253

STRUCTURAL DAMAGE IN AXIAL SPONDYLOARTHRITIS: IS THERE A PREFERRED WAY TO ASSESS PROGRESSION OVER TIME?

X. Baraliakos¹, V. Rios Rodriguez², M. Torgutalp², A. Dilbaryan², H. Haibel², M. Verba², J. Sieper², J. Braun¹, M. Rudwaleit³, D. Poddubnyy^{2,4}. ¹Ruhr-University Bochum, Rheumazentrum Ruhrgebiet, Herne, Germany; ²Charité-Universitätsmedizin, Gastroenterology, Infectious Diseases and Rheumatology, Berlin, Germany; ³Klinikum Rosenhöhe, Universitätsklinikum OWL, Bielefeld, Germany; ⁴German Rheumatism Research Centre, Epidemiology Unit, Berlin, Germany

Background: Axial spondyloarthritis (axSpA) is characterized by development of structural damage in the spine, assessed by conventional radiographs (CR). The modified Stokes Ankylosing Spondylitis Scoring System (mSASSS) includes all the SpA-relevant CR findings and is the most frequently used scoring system for quantification of spinal structural damage in patients (pts) with axSpA. Radiographic progression over 2 years using mean mSASSS progression, scored by a blinded 'paired' (sets of CRs of the same patient) approach, is considered as a key treatment outcome reflecting structural progression in clinical trials. However, this approach has been accompanied by questions about the influence of pts showing mSASSS 'improvement', indicating possible disappearance of syndesmophytes over time, something 'naturally unexpected'.

Objectives: To investigate the performance of mSASSS in assessing spinal radiographic damage and progression in axSpA pts using different approaches of CR evaluations.

Methods: Complete sets (cervical and lumbar CRs) from the German SpA Inception Cohort (GESPIC) at baseline and after 2 years were blinded to all clinical and demographic characteristics and then scored using the mSASSS by 5 different experienced readers, 2 blinded and 3 unblinded to the timepoint of CR performance. The final mSASSS score was calculated as a mean of 2 (blinded) or 3 (unblinded) readers. Descriptive statistics, cumulative probability plots and shift analyses (agreement of 2/2 readers in the blinded and 2/3 readers in the unblinded group) were performed for each reader group.

Results: A total of 210 pts (mean age 37.3y, 51% male, 79% HLA-B27+) with non-radiographic (n=115) and radiographic (n=95) axSpA at baseline were included. The mean mSASSS at baseline was 4.3±8.3 vs. 3.4±7.9, while the mean progression was 0.7±2.3 vs. 1.0±1.9 mSASSS units for the blinded vs. the unblinded group, respectively (Figure). On the patient level, progression of ≥2 mSASSS units was found in 30 (14.3%) vs. 37 (17.6%) pts in the blinded vs. the unblinded group, while agreement between groups was seen in 179 (85.2%) pts, 18 (8.9%) pts for progression and 161 (76.7%) for no progression.

Figure 1. Cumulative probability plot for mean mSASSS scores (all scores) in blinded (A) and unblinded (B) and 'definite' scores in blinded (C) and unblinded (D) scoring approach

In the analysis of 'definite' CR findings (only scores of 2=syndesmophytes or 3=ankylosis), the mean mSASSS score at baseline was 3.3±8.0 vs. 2.6±7.2 and the mean radiographic progression was 0.6±2.4 vs. 0.8±2.1 mSASSS units for the blinded vs. the unblinded group, respectively (Figure). On the patient level, progression was found in 37 (17.6%) vs. 33 (15.7%) pts in the blinded vs. the unblinded group, while agreement between groups was seen in 188 (89.5%) pts, 24 (11.3%) for progression and in 164 (78.1%) pts for no progression.

In the shift analysis, mSASSS worsening was found in 35 (0.8%) and mSASSS 'improvement' in only 4 (0.1%) out of a total of 4.373 vertebral edges (VE) analyzed in the blinded and in 109 (2.2%) and 2 (0.04%), respectively, out of a total of 4.914 VE analyzed in the unblinded group (Table). The majority of progression was found for the development of 'definite' signs of progression in both the blinded (25/35, 71.4%) and the unblinded (61/109, 56%) group, while more VE showing 'minor' (only scores of 1=sclerosis, squaring, erosion) signs of progression were found in the unblinded (48/109, 44%) vs. to the blinded (10/25, 28.6%) group.

Conclusion: Despite lower mean baseline mSASSS, higher mean mSASSS progression was found using the unblinded approach, while in the shift analysis the unblinded approach was also more specific, confirming the absence of 'improvement' of radiographic damage over time. These results were similar in the analysis on the patient's level. The scoring approach and the influence of 'minor' radiographic changes should be taken into account for avoiding the so-called 'background noise' in the evaluation of mSASSS.

Acknowledgements: GESPIC was initially supported by the BMBF. As a consequence of the funding reduction by BMBF according to schedule in 2005 and stopped in 2007, complementary financial support has been obtained also from Abbott, Amgen, Centocor, Schering-Plough, and Wyeth. Starting from 2010, the core GESPIC cohort was supported by AbbVie.

Table 1. Shift analysis of the blinded and unblinded reading approach on the basis of VE.

Blinded reading						
mSASS Scores		2-year follow-up				Total
		0	1	2	3	
Baseline	0	4.166	10	8	9	4.193
	1	0	29	4	0	33
	2	3	1	48	4	56
	3	0	0	0	92	92
Total		4.169	40	60	105	4.374

Unblinded reading						
mSASS Scores		2-year follow-up				Total
		0	1	2	3	
Baseline	0	4.575	48	14	3	4.640
	1	1	75	16	8	100
	2	0	0	69	20	89
	3	0	1	0	84	85
Total		4.576	124	99	115	4.914

Disclosure of Interests: None declared

DOI: 10.1136/annrheumdis-2021-eular.2532

OP0254

AN ARTIFICIAL NEURAL NETWORK FOR THE DETECTION OF DEFINITE RADIOGRAPHIC SACROILIITIS WITH HIGH SPECIFICITY IN THE DIAGNOSTIC SETTING

J. L. Vahldiek¹, K. K. Bressemer¹, S. Niehues¹, L. Adams¹, L. Spiller², M. Protopopov², V. Rios Rodriguez², B. Mueche², J. Rademacher², H. Haibel², M. Torgutalp², F. Proft², D. Poddubnyy^{2,3}. ¹Charité – Universitätsmedizin Berlin, Division of Radiology, Berlin, Germany; ²Charité – Universitätsmedizin Berlin, Division of Gastroenterology, Infectious Diseases and Rheumatology, Berlin, Germany; ³Deutsches Rheuma-Forschungszentrum Berlin (DRFZ), ein Institut der Leibniz-Gemeinschaft, Epidemiology Unit, Berlin, Germany

Background: Radiographs of the sacroiliac joints are commonly used as the first imaging method for the diagnosis of axial spondyloarthritis (axSpA), but the reliability of sacroiliitis detection is usually low. Recently we developed a deep convolutional neural network (CNN) that can detect radiographic sacroiliitis with expert-like accuracy in patients with axSpA, i.e., classification as radiographic or non-radiographic [1]. There is frequent criticism that many artificial intelligence algorithms for diagnostic analysis of medical images lack robust validation in real-world clinical applications.

Objectives: The aim of this study was to evaluate the performance of our deep CNN in detecting definite radiographic sacroiliitis in the diagnostic setting.

Methods: In this study, we included a total of 361 patients with chronic back pain who presented to a rheumatologist in a specialized SpA center with a suspicion of axSpA within the OptiRef project [2]. All patients received a structured rheumatologic diagnostic work-up that resulted in the final diagnosis of axial SpA/no axial SpA. Radiographs of sacroiliac joints were evaluated by a rheumatologist and radiologist according to the modified New York criteria; the consensus judgement of the presence of definite radiographic sacroiliitis (≥grade 2 bilaterally or ≥grade 3 unilaterally) was used as a reference. The predictions of the deep CNN's inference (with a balanced cutoff of 0.724 for the predictions of the model that was derived from the training and validation steps [1]) on all available pelvic radiographs was compared to this reference judgement.

Results: Pelvic radiographs of 340 patients (110 with axSpA including 61 patient with radiographic and 49 with non-radiographic axSpA, and 230 without SpA) were available for the CNN evaluation. The deep CNN achieved a sensitivity of 79% for the diagnosis of radiographic axSpA. The specificity of radiographic sacroiliitis detection was 94% (Table 1). The area under the receiver operating characteristics curve for the prediction of the presence of definite radiographic sacroiliitis was 88%. Figure 1 shows an exemplary class activation map of our CNN.

Table 1. Convolutional neural network predictions of the presence of radiographic sacroiliitis in patients with suspected axSpA according to the final diagnosis by rheumatologist in OptiRef (N=340).

Clinical diagnosis	CNN's prediction on the presence of definite radiographic sacroiliitis	
	Present	Not present
Radiographic axial SpA	48/61 (78.7%)	13/61 (21.3%)
Non-radiographic axial SpA	4/49 (8.2%)	45/49 (91.8%)
No SpA (other diagnosis)	14/230 (6.1%)	216/230 (93.9%)

Conclusion: The artificial neural network showed good generalizability and a high specificity with acceptable sensitivity in the detection of radiographic sacroiliitis when applied in the diagnostic setting of patients with chronic back pain and suspicion of axSpA. This algorithm can therefore be used to aid the detection of radiographic sacroiliitis as a part of the diagnostic approach.

REFERENCES:

- [1] Bressemer KK, et al. medRxiv. 2020:2020.05.19.20105304.
- [2] Proft F, et al. Semin Arthritis Rheum. 2020;50:1015-1021.

Acknowledgements: The OptiRef study was supported by a research grant from Novartis.

Figure 1. Inference of the proposed CNN on a randomly chosen pelvic radiograph from the OptiRef study.

This exemplary class activation map shows that our CNN almost exclusively and correctly focusses on the sacroiliac joints to predict whether definite radiographic sacroiliitis is present or not.

Disclosure of Interests: None declared

DOI: 10.1136/annrheumdis-2021-eular.3512

OP0255 THE ROLE OF PELVIC MORPHOLOGY IN AXIAL SPONDYLOARTHRITIS

E. Zaccagnino¹, R. Patel², L. S. Gensler³. ¹UCSF Medical Center, School of Medicine, San Francisco, United States of America; ²UCSF Medical Center, Radiology & Biomedical Imaging, San Francisco, United States of America; ³UCSF Medical Center, Rheumatology, San Francisco, United States of America

Background: Axial Spondyloarthritis (axSpA) is a chronic inflammatory disease affecting the axial skeleton. It includes non-radiographic axSpA and radiographic axSpA [Ankylosing Spondylitis (AS)]. Male axSpA patients often have greater damage, while women report a higher disease burden. The role of pelvic morphology in the axSpA phenotype has not been explored. There is anatomic sexual dimorphism between the male and female pelvis. Given the phenotypic gender differences in axSpA, the role of pelvic morphometry is of interest.

Objectives: The purpose of this study is to determine whether an association exists between pelvic dimensions and radiographic damage in patients with axSpA, as well as to compare these measurements in axSpA patients and healthy controls.

Methods: This was a cross-sectional analysis comparing axSpA cases from a prospective cohort and non-axSpA controls from the UCSF radiology database. Informed consent was obtained from axSpA cohort patients and this study was approved by the institutional IRB. To be included in the analysis, we limited inclusion to age ≤ 50 with an Anterior Posterior (AP) pelvis radiograph in the system. We excluded non-nulliparity, pelvic fracture history, BMI ≥ 30 kg/m², any prosthetic history and avascular necrosis. We measured the pelvic inlet, pelvic outlet, and subpubic angle (based on validated scoring methods) (Figure 1) and assessed its relation to sacroiliac joint (SIJ) damage (average SIJ score, New York criteria) and modified Stoke Ankylosing Spondylitis Spine Score (mSASSS) in cases. AxSpA patients were also compared to age/gender matched controls. Pelvic measurements were performed by 2 blinded independent-trained readers in randomized, blinded image order. Inter-rater reliability was assessed. When examining the relationship between pelvic measurements and damage, linear regression was used to stratify by gender and adjust for potential confounders.

Results: The axSpA cohort included 481 patients, of which 210 men and 89 women were included in this analysis and gender/age matched controls. Rater inter-class correlation was above 0.70 for pelvic outlet and above 0.80 for other measures. Cases and controls were similar (Table 1). The regression analysis showed a significant relationship between the sub-pubic angle and damage in the spine (coeff=-0.342, p=0.003) in men with axSpA. A sensitivity analysis, excluding mSASSS outliers (mSASSS ≥ 16) upheld the relationship (coeff=-1.40, p=0.002).

Conclusion: In men with axSpA, there appears to be a relationship between sub-pubic angle and spinal radiographic damage. This is consistent with our finding that women have larger sub-pubic angles and lower spinal radiographic damage than men. A greater sub-pubic angle may protect against spinal involvement or associate with other protective factors. Further work should be performed to understand the contribution of pelvic anatomy to damage in axSpA.

Disclosure of Interests: Ethan Zaccagnino: None declared, Rina Patel: None declared, Lianne S. Gensler Consultant of: AbbVie, Eli Lilly, Gilead, Novartis, Pfizer and UCB., Grant/research support from: Pfizer and UCB.

DOI: 10.1136/annrheumdis-2021-eular.188

OP0256 CHOOSE WISELY: IMAGING FOR DIAGNOSIS OF AXIAL SPONDYLOARTHRITIS

T. Diekhoff¹, I. Eshed², F. Radny¹, K. Ziegeler¹, F. Proft³, J. Greese¹, D. Deppe¹, R. Biesen⁴, K. G. Hermann¹, D. Poddubny^{3,5}. ¹Charité Universitätsmedizin Berlin, Division of Radiology, Berlin, Germany; ²Chaim Sheba Medical Center, Division of Radiology, Tel Aviv, Israel; ³Charité Universitätsmedizin Berlin, Division of Gastroenterology, Infectious Diseases and Rheumatology, Berlin, Germany; ⁴Charité Universitätsmedizin Berlin, Division of Rheumatology and Clinical Immunology, Berlin, Germany; ⁵German Rheumatism Research Centre, Epidemiology Unit, Berlin, Germany

Background: To date, the European Alliance of Associations for Rheumatology (EULAR) guidelines recommend X-ray (XR) as first line imaging in axial Spondyloarthritis (axSpA) and magnetic resonance imaging (MR) if the diagnosis cannot be established by XR and clinical features. However, much knowledge has been gained recently strengthening the applicability of MR for the detection of structural lesions and raising the question, whether XR is still necessary. Also, several publications used low-dose computed tomography (CT) as reference standard and imaging test.

Objectives: In light of this complex diagnostic situation, the aim of this study was to compare the three major modalities, XR, MR and CT of SIJ, in their diagnostic performance of axSpA and differential diagnosis in a cohort of patients with low back pain using the final judgment of the rheumatologist as standard of reference.

Methods: 163 patients (89 with axSpA; 74 with degenerative diseases) underwent XR, CT and MR. Three blinded experts categorized the imaging into axSpA, other diseases or normal in 5 separate reading rounds (XR, CT, MR, XR+MR, CT+MR, respectively). The results were compared to the clinical diagnosis. Sensitivity and specificity values for axSpA and interrater reliability were compared.

Results: XR showed lower sensitivity and specificity (66.3%/67.6% respectively) compared to MR (82.0%/86.5%) and CT (77.5%/97.3%). Sensitivity and specificity of XR+MR was similar to MR alone (77.5% / 87.8%). However, CT+MR was