Objectives: To compare classification accuracy between readings provided by expert readers belonging to three different laboratories and those automatically returned by CAD.

Methods: We acquired 1771 images from 583 consecutive sample with an acquisition unit using HEp-2 cells (MBL) at 1:80 screening dilution. Each image was blindly classified as positive, negative, or weak positive by two experienced physicians, with more than ten years of experience in IIF, for each of the three laboratories. We obtained gold standard on the basis of annotations provided by expert physicians. We described a CAD system for HEp-2 classification that relies upon features provided by a deep neural network architecture, namely an Invariant Scattering Convolutional Network (Scatnet). We therefore compared human readings with automatic classification provided by a CAD.

Results: The dataset contains 215 positive samples, 136 weak positive, 219 negative. The CAD-system classification and experts showed a sensitivity of 93% vs 92.9%, 79.4% vs 74.5%, and 92.2% vs. 92.8% on positive, negative and weak positive, respectively (Fig 1). The CAD-system obtained an accuracy of 89.5%, slightly better than average experts' classification (88.5%) and, interestingly, it better recognised the weak positive samples.

Fig1

Sensitivity	CAD	Experts
Positive	0.930	0.929
Weak positive	0.794	0.745
Negative	0.922	0.928
Positive Predictive Value		
Positive	0.939	0.933
Weak positive	0.777	0.779
Negative	0.927	0.913

Conclusions: Solid gold standard is essential for use CAD systems in routine work lab. The CAD-system classification and the 3 experts have provided comparable results. Laboratories' agreement improves using digital images and comparing each single human evaluation to a potential reference data and for this reason nowadays the CAD system should be considered a reliable tool of standardisation reducing the inter-laboratory variability.

REFERENCE:

[1] Rigon A, Infantino M, et al. Autoimmun Rev 2017.

Disclosure of Interest: None declared

DOI: 10.1136/annrheumdis-2018-eular.3230

SAT0643

ULTRASOUND ABNORMALITIES IN WRIST, MCP2 AND MTP5 ARE MOST DISCRIMINATIVE IN PREDICTING ARTHRITIS DEVELOPMENT IN SEROPOSITIVE ARTHRALGIA PATIENTS

A. Blanken¹, M. van Beers-Tas¹, M. Nielen², M. Meursing Reyniers¹, F. Turkstra¹, C. van der Laken¹, D. van Schaardenburg¹. ¹Amsterdam Rheumatology and immunology Center, Reade, VUmc and AMC, Amsterdam; ²Netherlands Institute for Health Services Research, NIVEL, Utrecht, Netherlands

Background: Seropositive arthralgia patients are at risk of developing rheumatoid arthritis (RA). Ultrasound (US) might be used to further predict which seropositive individuals will progress to RA. However, the value of US in the prediction of RA is still a point of debate, mainly due to the use of different scoring systems and compositions of joints in US protocols in literature.

Objectives: To investigate which joints are most discriminative in predicting arthritis development in seropositive arthralgia patients.

Methods: We included 174 seropositive patients with arthralgia, but without clinical arthritis. US was performed at baseline in 16 joints: bilateral metacarpal phalangeal (MCP) 2 and 3, proximal interphalangeal (PIP) 2 and 3, wrist and metatarsal phalangeal (MTP) joints 2, 3 and 5. Images were scored for grey-scale (GS) synovitis and Power Doppler (PD) on a scale of 0–3. Grades ≥ 2 for GS synovitis and grades ≥ 1 for PD were regarded as abnormal. Clinical arthritis development was assessed in any of 44 joints during yearly follow-up or during an

unscheduled visit in case of progression of symptoms. Patients were followed until clinical arthritis development or for a maximum of 5 years.

Results: In a total of 2784 joints that were imaged, 112 showed GS synovitis and 14 PD. The majority of GS synovitis was present in MTP2 and MTP3 joints (56 (50%) and 32 (29%), respectively), followed by wrists (15 (13%)), MCP3 (4 (4%)), MTP5 (3 (3%)), MCP2 (2 (2%)) and none in PIPs. Out of 14 joints with PD, 7 were wrists, 3 MTP5, 2 MCP2 and 2 PIP3. Fifty-one (29%) of the patients developed clinical arthritis in at least one joint after a median follow-up of 12 (interquartile range 6–23) months. For GS synovitis, the wrist, MCP2 and MTP5 were most discriminative in predicting clinical arthritis development (12/15 (80%) of patients with GS synovitis in wrist developed clinical arthritis, 3/3 (100%) in MTP5 and 2/2 (100%) in MCP2). MTP2 and 3 were least discriminative (<27%). No substantial differences were found between left and right joints. No clear association with clinical arthritis development was found in the limited number of joints that had positive PD.

Conclusions: Wrist, MCP2 and MTP5 joints (although numbers were small) showed the highest predictive value for development of clinical arthritis in any of 44 joints. Although most GS synovitis was observed in MTP2 and 3, predictive value of MTP2 and 3 joints for development of clinical arthritis was low. These results indicate that the choice of joints in the US protocol may influence the predictive value of ultrasound in predicting clinical arthritis development in seropositive arthralgia patients.

Disclosure of Interest: None declared

DOI: 10.1136/annrheumdis-2018-eular.2614

SAT0644

T1RHO MAPPING IN THE ASSESSMENT OF ARTICULAR CARTILAGE INTEGRITY OF THE KNEE IN CHILDREN WITH JUVENILE IDIOPATHIC ARTHRITIS

A.M. Barendregt¹, V. Mazzoli², C. van Gulik², M. van den Berg³, A. Nassar Sheikh Rashid⁴, K.M. Dolman⁵, D. Schonenberg-Meinema⁴, T.W. Kuijpers⁴, M. Maas⁶, A. J. Nederveen⁶, R. Hemke⁶. ¹Department of Radiology and Nuclear Medicine; ²Radiology and Nuclear Medicine, University of Amsterdam; ³Department of Pediatric Hematology, Immunology, Rheumatology and Infectious Disease, Emma Children's Hospital; ⁴Department of Pediatric Hematology, Immunology, Rheumatology and Infectious Disease, Emma Children's Hospital; ⁵Pediatrics, OLVG; ⁶Department of Radiology and Nuclear Medicine, Academic Medical Center, Amsterdam, Netherlands

Background: Early detection of microstructural damage to cartilage of patients with juvenile idiopathic arthritis (JIA) might prevent irreversible cartilage damage by timely treatment and follow-up of cartilage microstructure. New, quantitative MRI sequences such as T1 ρ are possibly able to detect pre-erosive cartilage damage by quantifying proteoglycan (PG) loss of cartilage.¹⁻⁴

Objectives: To study feasibility of T1 ρ mapping for assessment of articular cartilage integrity in children with JIA and study correlation between T1 ρ relaxation time and the Juvenile Arthritis MRI score (JAMRIS)⁵ for disease activity.

Methods: After IRB approval and informed consent, the knee of patients with JIA or suspected JIA was imaged at 3T MRI using T2, pre and post contrast T1 and a sagittal T1 ρ sequence with 400 Hz and spin lock time of 5, 10, 20, 40 and 50 ms. A region of interest (ROI) was drawn in articular cartilage of the knee on the T1 ρ images, resulting in a mean T1 ρ value of articular cartilage per patient. Using regular T2W and T1W pre and post contrast scans, JAMRIS was assigned to discriminate inflamed knees (JAMRIS ≥ 1) from non-inflamed knees (JAMRIS 0). In SPSS, Mann-Whitney U test and Spearman correlation coefficient were used to compare the mean T1 ρ value between patients with and without arthritis on MRI and to correlate T1 ρ values with JAMRIS score. ROI drawing was performed twice in 5 subjects. Intraclass correlation coefficient (ICC) was used to study intra-reader reliability.

Results: Of all 13 patients (median age 13.7 years), 7 patients had inflammation in the knee. No cartilage lesions were observed on standard MRI sequences. Acquisition of the T1 ρ was successful and without artefacts in 100% of the children. Patients with inflammation in the knee showed a significantly longer T1 ρ value than patients without inflammation in the knee: 36.3 ms (IQR 29.0–40.4) versus 27.7 ms (IQR 25.8–30.2), $p=0.02$. Correlation between T1 ρ value and JAMRIS score was 0.76 ($p=0.003$). Repeatability of ROI drawing was characterised by ICCs >0.99 , $p<0.05$.

Conclusions: This pilot study indicates that, even in the absence of cartilage erosions, significant differences are seen in cartilage integrity with higher $T1\rho$ relaxation times in patients with knee arthritis. This might indicate loss of glycosaminoglycan content in active synovitis of the knee. Thus, our pilot data suggest that $T1\rho$ could serve as an imaging biomarker for cartilage integrity, ultimately aiming to prevent irreversible cartilage damage and long-term disability in this young JIA population.

REFERENCES:

- [1] Tsushima, et al. *Rheumatol Int* 2012;32:2867–2875.
- [2] Takayama, et al. *Eur J Radiol* 2013;82:e192–198.
- [3] Duvvuri, et al. *Osteoarthritis Cartilage* 2002;10:838–844.
- [4] Regatte, et al. *Acad Radiol* 2002;9:1388–1394.
- [5] Hemke, et al. *Eur Radiol* 2013;23:1075–1083.

Disclosure of Interest: None declared

DOI: 10.1136/annrheumdis-2018-eular.2177

SAT0645

DIAGNOSTIC ACCURACY OF SERUM AUTOANTIBODIES MULTI-TESTING IN RHEUMATOID ARTHRITIS AND ECONOMIC CONSEQUENCES ACROSS EUROPE

B. Mascialino¹, S. Swiniarski², I. Gehring², M. Poorafshar¹. ¹*ImmunoDiagnostics, Thermo Fisher Scientific, Uppsala, Sweden*; ²*ImmunoDiagnostics, Thermo Fisher Scientific, Freiburg, Germany*

Background: Rheumatoid arthritis (RA) diagnosis requires interpretation of a combination of clinical, laboratory and imaging investigations. In RA, the consequences of incorrect serology test results are particularly important: patients with False Positive results (FPs) are initially managed as RA patients, bringing about extra costs until a correct diagnosis is made.

Objectives: The first aim of the study was to evaluate the diagnostic performance of RF-IgA, RF-IgM, and CCP, used alone or in multi-testing parallel or sequential combinations. The secondary goal focused on the economic consequences of serology FPs in selected European countries.

Methods: 190 established RA patients and 197 controls (either affected by other conditions or healthy donors) were used to assess the diagnostic performance of mono- and multi-serology testing in PC; both testing in Primary Care (PC) and in Secondary Care (SC) settings were considered.

For the secondary objective, a 12 month Markov model simulated, from the National Health Services perspective, 1.000 RA-suspected individuals tested in PC and SC with mono- or multi-testing. Costs come from the published literature. The European countries included in the analysis were the United Kingdom, Austria, Belgium, Denmark, France, Germany, Italy, the Netherlands, Portugal, Spain and Sweden. Uncertainty was addressed with sensitivity analysis.

Results: The mono-testing diagnostic performance was:

- RF-IgA: sensitivity [95% CI]=40.5% [33.5%, 47.9%], specificity=92.4% [87.8%, 95.7%];
- RF-IgM: sensitivity=59.0% [51.6%, 66.0%], specificity=89.3% [84.2%, 93.3%];
- CCP: sensitivity=59.5% [52.1%, 66.5%], specificity=96.5% [92.8%, 98.6%].

In multi-testing, defining a “positive result” as “positivity to at least one test” increased sensitivity; “positivity to all the tests” increased specificity. In a PC scenario:

- parallel testing:
 - o using CCP and RF-IgM increased specificity to 99.5% [97.2%, 100.0%];
 - o the three tests used simultaneously maximised specificity (100.0% [98.1%, 100.0%]), but reduced sensitivity to 35.8% [29.0%, 43.1%];
- sequential testing: testing positive to both RF-IgA and RF-IgM followed by CCP testing led to 90.7% [81.7%, 96.2%] sensitivity and to 100.0% [54.1%, 100.0%] specificity.

With respect to mono-testing, multi-testing options reduced the number of FPs. Therefore, in each of the countries considered, multi-testing allowed for important cost savings due to reduced clinical procedures and resource utilisation of FPs.

Conclusions: Multi-serology testing improve the diagnostic accuracy of the individual RF-IgA, RF-IgM and CCP tests. Optimal multi-testing combinations minimise the number of FPs, thus reducing avoidable costs to the National Health Services. Consequently, multi-testing for RA demonstrates superior value from patient and payer perspective.

REFERENCES:

- [1] Hallert E, et al. *Rheumatology* 2006;45(3):325–331.
- [2] Lajas C, et al. *Arthritis Rheum* 2003;49(1):64–70.
- [3] Lundkvist J, et al. *Eur J Health Econ* 2008;(Suppl 2):S49–60.
- [4] Verstappen SM, et al. *Ann Rheum Dis* 2004;63(7):817–24.

Disclosure of Interest: B. Mascialino Employee of: Thermo Fisher Scientific, S. Swiniarski Employee of: Thermo Fisher Scientific, I. Gehring Employee of: Thermo Fisher Scientific, M. Poorafshar Employee of: Thermo Fisher Scientific
DOI: 10.1136/annrheumdis-2018-eular.7275

SAT0646

DEVELOPMENT AND VALIDATION OF A DEEP LEARNING ALGORITHM FOR CLASSIFYING ANTI-NUCLEAR ANTIBODY PATTERNS IN INDIRECT IMMUNOFLOUORESCENCE IMAGES

C.-H. Lin¹, C.-F. Kuo^{1,2}, Y.-F. Fang³, P.-H. Tsai³, Y.-F. Chen³. ¹*Division of Rheumatology, Allergy and Immunology, Chang Gung Memorial Hospital, Taoyuan, Taiwan, Province of China*; ²*Academic Rheumatology, University of Nottingham, Nottingham, UK*; ³*No 5, Fu-Hsing Road, Chang Gung Memorial Hospital, Taoyuan, Taiwan, Province of China*

Background: Antinuclear antibody (ANA) is a valuable test in the evaluate for a variety of autoimmune diseases. It can be detected using immunofluorescence (IIF) which yield both serum titers and also staining patterns. The manual titration of ANA is laborious and evaluation of staining patterns by a microscope requires experts to make correct classification. Deep learning is a collection of computational methods that enable an algorithm to ‘learn’ from a large set of data to acquire ‘skills’ that are characteristic of human intelligence, for example, complex pattern recognition. Application of these methods to ANA pattern classification requires further assessment and validation.

Objectives: To train a deep convoluted neural network to create an algorithm for automated classification of ANA patterns in images of indirect immunofluorescence.

Methods: We construct a deep convoluted neural network based on inception V3. The model adopted a deep learning architecture with 21 layers and ANA images were used to train to detect speckled, homogenous, nucleolar, centromere, dense fine speckled, cytoplasmic and mitochondrial patterns. The data were collected from rheumatology patients from June 2016 to October 2017. We validate the performance of the model using a separate set of images.

Results: Our data set consisted of 81822 ANA images from 8839 patients. The overall accuracy was 99.46% after 200 epochs of iteration using the training dataset (65 458 images) and the accuracy was 96.33% using the validation dataset (16.64 images). The sensitivity was 97.1% for speckled pattern, 91.4% for homogenous pattern, 97.1% for nucleolar pattern, 97.1% for centromere pattern, 97.2% for cytoplasmic pattern, and 100% for both dense fine speckled and mitochondrial patterns. The specificity was also good, ranged between 97.1% and 100%.

Conclusions: In this study, an algorithm based on deep convoluted neural network had high sensitivity and specific for detecting specific ANA patterns. Further research is needed to determine the feasibility to deploy the technology in routine clinical works to assist rheumatologists to interpret ANA.

Disclosure of Interest: None declared

DOI: 10.1136/annrheumdis-2018-eular.6635