treatment is mainly aimed at combating physical affectation and hardly focusses on emotional and social aspects.

Objectives: To identify the situations that produce the greatest negative emotional impact in RA patients and to assess the coping strategies used.

Methods: A previous phase (Project OpinAR) (1) identified situations with negative emotional and social impact. In this project, 2 stages were developed: a) A nominal group of patients, following the McMillan method, selected the most emotionally affected situations from the previous list; b) Voting of patients and rheumatologists to assess the assumed effectiveness of 8 coping strategies for each situation with a Likert scale of 1 to 10 (minimum and maximum): proactive resolution, self-criticism, emotional expression, desiderative thinking, social support, cognitive restructuring, avoidance of problems and social withdrawal. The median, interquartile range (IQR), and statistical significance of the differences between patients and physicians (Student's T test) were calculated with a cut-off value of $p=0.05$. The strategies valued with 8, 9 and 10 were recommended and those valued as 1, 2 or 3 were discouraged. The study was approved by the Ethical Research Committee of the Hospital of La Princesa.

Results: 107 patients from all the Autonomous Communities and 31 expert rheumatologists from 13 Autonomous Communities were recruited, of which 100 (93%) and 17 (55%) participated in the study in a valid way, respectively.

The four situations that most negatively affect them, in the opinion of patients and their doctors are: i) The patient feels that his disease evolves worse than what the doctor says, ii) The patient does not know aspects of the control and monitoring of his disease, iii) The patient feels that he does not really participate in the decision making, and iv) The patient is dissatisfied with the overall treatment received. Regarding coping strategies, doctors recommended using proactive resolution in all situations more frequently than patients ($p<0.05$ for all situations, low variability, with $IQR \leq 2$ among physicians) as well as emotional expression ($p<0.05$ for 3 situations, high variability, with $IQR \geq 3$). Other strategies were less valued by doctors and patients, but with similar responses. There was a lot of dispersion ($IQR > 1$) in most cases.

Conclusion: The situations with the greatest negative emotional impact on RA are those that put the doctor and the patient at odds with regard to the perception of the disease and its control and should be treated during visits. The most recommended strategies are proactive resolution and emotional expression. The most discouraged are self-criticism and the avoidance of the problem. Anyway, there is a great variability of opinions and doctors tend to give more extreme responses. Doctors and patients should have a space and willingness to discuss and exchange opinions on how they emotionally face RA.

REFERENCE:

- [1] Castañeda S, et al. National Barometer to assess the emotional aspect of the patient with Rheumatoid Arthritis (Spanish). Science and Evidence. Bate Scientia Salus (ISSN 2174-8101). Madrid 2016; vol 3: pp 3-11.

Acknowledgement: BMS

Disclosure of Interests: Santos Castañeda Consultant for: Amgen, BMS, Pfizer, Lilly, MSD, Roche, Sanofi, UCB, José Bernardo Negrón: None declared, MAría D. Navarro Rubio: None declared, José L. Baquero Ubeda: None declared, Loreto Carmona Grant/research support from: Abbvie, Actelion, Astellas, BMS, Eisay, Gebro Pharma, Grünenthal, Leo Pharma, Lilly, MSD, Novartis, Pfizer, Roche, Sanofi-Aventis and UCB Pharma, Paid instructor for: Novartis

DOI: 10.1136/annrheumdis-2019-eular.7488

SATURDAY, 15 JUNE 2019

EULAR Projects in musculoskeletal imaging

OP0349

NEURAL NETWORKS FOR AUTOMATED SCORING OF JOINT DISEASE ACTIVITY ON DOPPLER ULTRASOUND IMAGES

Jakob K. H. Andersen¹, Jannik S. Pedersen¹, Martin S. Laursen¹, Kathrine Holtz¹, Jakob Grauslund², Thijsus Rajeeeth Savarimuthu¹, Søren Andreas Just³.

¹University of Southern Denmark, The Maersk Mc-Kinney Møller Institute, Odense, Denmark; ²Odense University Hospital, Department of Ophthalmology, Odense, Denmark; ³Svendborg Hospital, Odense University Hospital, Department of Medicine, Svendborg, Denmark

Background: The development of standardized methods for ultrasound (US) scanning and evaluation of synovitis activity by the OMERACT-EULAR Synovitis Scoring (OESS) system, is a major step forward in the use of US in the diagnosis and monitoring of patients with inflammatory arthritis. The variation in interpretation of disease activity on US images can affect diagnosis, treatment and outcomes in clinical trials.

Objectives: To investigate if we could utilize neural network architecture for the interpretation of disease activity on Doppler US images, using the OESS scoring system.

Methods: Two state-of-the-art neural networks were used to extract information from 1342 Doppler US images from Rheumatoid Arthritis (RA) patients. One neural network that divided images as either healthy (Doppler OESS score 0 or 1) or diseased (Doppler OESS score 2 or 3). The other to score images across all four of the OESS systems Doppler US scores (0-3). The neural networks were hereafter tested on a new set of RA Doppler US images ($n=176$). Agreement between rheumatologist's scores and network scores was measured with the kappa statistic.

Results: For the neural network assessing healthy/diseased score, the highest accuracies compared to an expert rheumatologist were 86.4% and 86.9% with a sensitivity of 0.864 and 0.875 and specificity of 0.864 and 0.864, respectively. The other neural network developed to four class Doppler OESS scoring achieved an average per class accuracy of 75.0%, and a quadratically weighted kappa score of 0.84.

Conclusion: This study is the first to show that neural network technology can be used in the scoring of disease activity on Doppler US images according to the OESS system.

Disclosure of Interests: None declared

DOI: 10.1136/annrheumdis-2019-eular.2174

OP0350

ULTRASOUND POWER DOPPLER SEMI-QUANTITATIVE SCORING, IS IT TIME WE MOVED TO A RED PIXEL DENSITY ALGORITHMIC PROGRAM TO DETERMINE TRUE TREATMENT RESPONSE?

Nirupam Purkayastha¹, Samiha Ismail², Peter C. Taylor³, Costantino Pitzalis¹, Stephen Kelly⁴. ¹Queen Mary University, William Harvey Research Institute, EMR, London, United Kingdom; ²King George Hospital, General internal medicine, London, United Kingdom; ³University of Oxford, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Oxford, United Kingdom; ⁴Barts Health NHS Trust, Rheumatology, London, United Kingdom

Background: The introduction of the semi-quantitative (SQ) grey-scale and power Doppler (PD) 0-3 scoring system, as agreed by the EULAR-OMERACT ultrasound (US) working group (1, 2), has undoubtedly been the most recognised and accepted method in both clinical practice and trials. Studies have suggested that the SQ scoring method is reliable as the quantitative when scoring the grey-scale (3, 4). Additional studies have shown that the use of red pixels in determining the PD signal in tenosynovitis, compared to SQ scoring was not different (5). However, the distribution of the SQ PD score over estimated the percentage compared to the quantitative score, which could reflect treatment responses inaccurately (6).

Objectives:

1. To describe how a quantitative pixel density score maps to a SQ PD score.
2. Can using a 0-4 PD grading scale better distribute the PD percentages, as determined by pixel count, than 0-3 scale?

Methods: Patients were selected from an observational study of 122 DMARD naïve patients (PEAC study REC: 05/Q0703/198; IRAS Project ID:60271) and 60 DMARD failure patients treated with anti-TNF (THERAPIST study, MREC: 15/SC/0045; IRAS Project ID:147240). All patients were classified according to the 1987 ACR or 2010 ACR/EULAR criteria for rheumatoid arthritis. In total 334 images were analysed, from a total of 127 patients. The calculator for the US PD signal/synovial area ratio was created on software MATLAB 2018b. PD was scored according to the SQ definitions for the visual grade 0-3 and 0-4 scoring systems (Figure 1A). Agreement between two readers using 80 US images: ICC SH —0.82 (interval 0.79–0.9), ICC PD —0.93 (interval 0.9–0.95), calculated using SPSS software. The results were analysed using R v3.4.2.

Results: The SQ grade 3 score was shown to over estimate PD signal/synovial area ratio, with a mean ratio of only 40.8%. Figure 2A, illustrates how the spread in PD signal varied within a single grade, in this case SQ grade 3. Figure 1B showed that adding a SQ grade 4 reduced the SD in the SQ grade 3 group but increased the SD overall by causing a larger SD in SQ grade 1, 2 and 4 grades.

A k-means (KM) clustering algorithm was used to find the most efficient ratio cut off points for a SQ grade 3 scoring system. The SQ grades are not taken into account for this scoring system, it is based on the mathematically optimum cut off points (Figure 1C). It is seen that the SD dramatically reduced in the KM score and the 95% confidence interval and grade overlaps.

When comparing the KM score with the SQ score it was shown to be significantly dependent (Pearson's Chi squared p -value $<2.2e-16$) and strongly positively correlated (Spearman's 0.84, p -value $<2.2e-16$), this is illustrated in Figure 2B and C. Furthermore, the KM score was seen to have a stronger linear correlation (Pearson's 0.95, p -value $<2.2e-16$) when compared to the SQ score (0.82, p -value $<2.2e-16$).

Conclusion: The 0-4 SQ grade was not equivalent to that of the 0-3 SQ grade used PD scoring and it increased the variance in the majority of SQ grades. This study does show that the PD pixel algorithm has a stronger linear correlation when compared to the SQ grading, suggesting a more accurate method at interpreting treatment response, as the intervals between grades are more equal. Therefore, built in algorithms to detect PD signal in US machines are the future for drug monitoring in inflammatory arthritis.

REFERENCES:

- [1] D'Agostino et al. RMD 2017.

Figure 1. A) Definitions of US PD scoring scale 0-3 using the OMERACT grading system and a revised 0-4 score. B) Comparison between SQ 0-3 and 0-4 PD score, with the mean and standard deviation and 95% confidence intervals of PD signal/synovial area ratio. C) Scoring system found by K-means clustering algorithm showing the optimum ratio boundaries and the mean standard deviation and 95% confidence intervals of the ratio in each quantitative grade 0-3.

A Synovitis grade	Power Doppler OMERACT score	Power Doppler revised score
Grade 0 (normal)	No Doppler signal	No Doppler signal
Grade 1 (minimal)	Up to three single Doppler spots OR up to one confluent spot and two single spots OR up to two confluent spots	1-2 Doppler spots OR up to one confluent spot
Grade 2 (mild)	>Grade 1 but ≤50% Doppler signals in the total greyscale background	More than 2 Doppler spots or scattered single spots or <25% Doppler signals in the total greyscale background
Grade 3 (moderate)	>Grade 2 (>50% of the total greyscale background)	25-50% Doppler signals in the total greyscale background
Grade 4 (severe)		>50% of the total greyscale background

B SQ PD Grade	0	1	2	3	4
3-Score					
Mean	0.02%	4.14%	16.02%	40.8%	
SD	0.09%	3.71%	8.14%	11.74%	
95% CI Upper	0.20%	11.56%	32.30%	64.28%	
95% CI Lower	-0.16%	-3.28%	-0.26%	17.32%	
4-Score					
Mean	0.02%	3.92%	15.82%	38.72%	54.25%
SD	0.09%	3.69%	8.19%	9.01%	14.63%
95% CI Upper	0.20%	11.30%	32.20%	56.74%	83.51%
95% CI Lower	-0.16%	-3.46%	-0.56%	20.70%	24.99%

C K-means Grade	0	1	2	3
Ratio Boundaries	[0,5.30]	[5.30, 15.97]	[15.97, 32.79]	[32.79, 100]
Mean	0.7%	9.9%	22.03%	43.55%
SD	1.33%	3.10%	4.38%	9.47%
95% CI Upper	-1.96%	3.70%	13.27%	24.61%
95% CI Lower	3.36%	16.10%	30.79%	62.49%

Figure 2. A) US images demonstrating the spread in PD signal/synovial area ratio with the same SQ PD grade 3 score. B) Comparison between SQ 0-3 and 0-4 PD grading, illustrating the spread of the PD signal/synovial area percentage ratio in each SQ grade category. C) Density graph using K-means statistical test demonstrating how the spread between the individual grades can be minimised compared to the 3 score PD signal/synovial area ratio in Figure 2B.

- [2] Terslev et al. RMD 2017.
 [3] Terslev et al. Rheum 2012.
 [4] Dougados et al. ARD 2010.
 [5] Ammitzboll-Danielsen et al. Rheum 2016.
 [6] Schmidt et al. Rheum 2015.

Disclosure of Interests: Nirupam Purkayastha: None declared, Samiha Ismail: None declared, Peter C. Taylor Grant/research support from: Celgene, Galapagos, Eli Lilly, UCB, Consultant for: AbbVie, Galapagos, Gilead, Eli Lilly, Pfizer Inc, Costantino Pitzalis Grant/research support from: Celgene, Stephen Kelly: None declared

DOI: 10.1136/annrheumdis-2019-eular.1527

Late breaking abstract session

LB0003

EFFICACY AND SAFETY OF FILGOTINIB FOR PATIENTS WITH RHEUMATOID ARTHRITIS NAÏVE TO METHOTREXATE THERAPY: FINCH3 PRIMARY OUTCOME RESULTS

Rene Westhovens¹, William Rigby², Désirée van der Heijde³, Daniel Ching⁴, Beatrix Bartok⁵, Franziska Matzki⁵, Zhaoyu Yin⁵, Ying Guo⁵, Chantal Tasset⁶, John Sundy⁷, Neelufar Mozaffarian⁷, Osvaldo Messina^{7,8}, Robert B.M. Landewe⁹, Tatsuya Atsumi¹⁰, Gerd Rüdiger Burmester¹¹. ¹University Hospitals, Leuven, Belgium; ²Dartmouth College, Lebanon, United States of America; ³Leiden University Medical Center, Leiden, Netherlands; ⁴Timaru Hospital, Timaru, New Zealand; ⁵Gilead Sciences, Inc., Foster City, United States of America; ⁶Galapagos NV, Mechelen, Belgium; ⁷Cosme Argerich Hospital, Buenos Aires, Argentina; ⁸IRO Medical Center, Buenos Aires, Argentina; ⁹Amsterdam University Medical Center, Amsterdam, Netherlands; ¹⁰Hokkaido University, Sapporo, Japan; ¹¹Charité-University Medicine Berlin, Berlin, Germany

Background: Filgotinib (FIL), an orally administered, potent, selective inhibitor of Janus kinase 1 (JAK1), has shown good efficacy and was well tolerated for treatment of rheumatoid arthritis (RA).

Objectives: To compare efficacy and safety of FIL with and without methotrexate (MTX) in patients with RA who were naïve to MTX therapy.

Methods: This phase 3, double-blind, active-controlled study randomized patients with moderately to severely active RA (2:1:1:2) to FIL 200mg daily + MTX weekly (up to 20mg), FIL 100mg + MTX, FIL 200mg (+placebo [PBO]), or MTX (+PBO) for up to 52 weeks; results through week 24 are presented. Primary efficacy endpoint was proportion of patients achieving ACR20 response at week 24; additional assessments included ACR50 and ACR70 responses; DAS28-CRP score ≤3.2 and <2.6, and changes in van der Heijde mTSS, HAQ-DI, SF-36 PCS, and FACIT-Fatigue. Safety endpoints included types and rates of adverse events (AEs). Logistic regression adjusting for stratification factors with nonresponder imputation was used for treatment comparisons for ACR response and other binary endpoints. Mixed-effect model adjusting for baseline value, stratification factors, treatment, visit, and treatment by visit interaction as fixed effects with observed cases was used for continuous endpoints.

Results: Of 1,252 randomized patients, 1,249 received study drug (416 FIL 200mg+MTX; 207 FIL 100mg+MTX; 210 FIL 200mg monotherapy; 416 MTX monotherapy) and were analyzed; 1,130 completed week 24. Most (76.9%) were female; mean time since RA diagnosis was 2.2 years (median 0.4 years); mean (standard deviation [SD]) DAS28-CRP was 5.7 (1.0); and 35.9% were using oral steroids at baseline. At week 24, significantly more patients in the FIL 200mg+MTX (81.0%; $P<0.001$) and FIL 100mg+MTX (80.2%; $P<0.05$) arms achieved an ACR20 response compared to MTX monotherapy (71.4%) (Table 1). Compared to MTX monotherapy, more patients receiving FIL with or without MTX achieved

Table 1. Efficacy Outcomes at Week 24*

	FIL 200 mg + MTX (N = 416)	FIL 100 mg + MTX (N = 207)	FIL 200 mg Monotherapy (N = 210)	MTX Monotherapy (N = 416)
ACR20, %	81.0***	80.2*	78.1	71.4
ACR50, %	61.5***	57.0**	58.1***	45.7
ACR70, %	43.8***	40.1***	40.0***	26.0
DAS28-CRP ≤ 3.2, %	68.8***	62.8***	60.0***	46.2
DAS28-CRP < 2.6, %	54.1***	42.5***	42.4***	29.1
mTSS, mean change from BL	0.20	0.22	-0.04**	0.52
HAQ-DI, mean change from BL	-0.94***	-0.90**	-0.89**	-0.79
SF-36 PCS, mean change from BL	12.3***	11.1**	10.4	9.7
FACIT-Fatigue, mean change from BL	10.6	11.4	10.2	10.1

*All patients who were randomized and received at least 1 dose of study drug were included in efficacy analyses.
 ** $P<0.05$ vs MTX monotherapy; *** $P<0.01$ vs MTX monotherapy; **** $P<0.001$ vs MTX monotherapy; *Comparison not adjusted for multiplicity.
 ACR20/50/70, 20%/50%/70% improvement in American College of Rheumatology criteria; BL, baseline; DAS28-CRP, Disease Activity Score based in 28 joints with C-reactive protein; FACIT, Functional Assessment of Chronic Illness Therapy; FIL, filgotinib; HAQ-DI, Health Assessment Questionnaire Disability Index; mTSS, modified total Sharp score; MTX, methotrexate; PBO, placebo; SF-36 PCS, Short-Form 36 Physical Component Summary.