Score (from 38.2 % to 47.8%), pulmonary function tests (from 30.9% to 43.5%), measures of dyspnoea (4.4% to 10.9%) and patient global disease (from 10.3% to 21.7%).

Conclusions: There was a wide range of domains reported in trials in systemic sclerosis. The uptake of domains and measure as per the core response set is low in SSc trials compared to other rheumatic diseases with only modest improvement in reporting of 6 out of 31 measures. Improvements in reporting of specific measures align with the recent development of a composite response index in systemic sclerosis (CRISS)2.

REFERENCES:

1. Khanna D, et al. Development of a provisional core set of response measures for clinical trials of systemic sclerosis, *Ann Rheum Dis* 2008;67:703–9.
2. Khanna D, et al. The American College of Rheumatology Provisional Composite Response Index for Clinical Trials in Early Diffuse Cutaneous Systemic Sclerosis. *Arthritis Res Ther* 2016;68:299–311.

Disclosure of Interest: None declared

DOI: 10.1136/annrheumdis-2018-eular.6001

FRI0656 PREDICTIVE VALIDITY OF PRESENTEEISM MEASURES WITH DUAL ANSWER KEYS IN INFLAMMATORY ARTHRITIS

M. Donaldson^{1,2}, A. Kobza^{2,3}, D. E. Beaton⁴, H. Xie^{2,5}, E. C. Sayre², M. A. M. Gignac^{4,6}, D. Lacaille^{1,2}. ¹University of British Columbia, Vancouver, ²Arthritis Research Canada, Richmond, ³University of Ottawa, Ottawa, ⁴Institute of Work and Health, Toronto, ⁵Simon Fraser University, Burnaby, ⁶University of Toronto, Toronto, Canada

Background: Employment studies in arthritis have emphasized the importance of decreased productivity at work, or presenteeism. However, consensus is still lacking on how to best measure this construct. Instruments used include the Work Limitations Questionnaire (WLQ), which measures the amount of time people are limited, and the Workplace Activity Limitations Scale (WALS), which measures the degree of difficulty. We modified the response keys of the WLQ and WALS, creating a dual answer key assessing both degree of difficulty and amount of time with difficulty. Previous work by our group has shown that scores obtained from combining responses to the dual answer keys exhibit good internal consistency, criterion and construct validity.

Objectives: The objective of this study was to evaluate the ability of the combined scores to predict future work cessation and number of work days missed due to arthritis.

Methods: A longitudinal study used baseline and 6-month data from the RCT of an employment intervention, the "Making It Work" Program. Inclusion criteria were: inflammatory arthritis, age 19–59, working at baseline, and concerns about arthritis affecting work. 327 participants were included (RA:173, PsA:48, SLE:42, AS:64). Combined scores were obtained by i) multiplying, and ii) adding, the scores of difficulty and time answer keys at the item level. We assessed the predictive ability of the combined scores, and the original scores, of the WALS and WLQ to predict work cessation due to arthritis within 6 months using binary logistic regression models, or number of work days missed due to arthritis using zero-inflated Poisson models.

Results: WLQ and WALS original and combined scores were significant predictors of future work cessation and number of work days missed. The WLQ combined scores predicted work cessation better than the original score (AIC:117.6 & 118.2 for multiplicative and additive combinations, respectively vs. 125.4 for time) and work days missed (AIC:2179.8 & 2177.8 vs. 2188.7). WALS original score (ie. degree of difficulty) predicted work cessation better than combined scores (AIC:95.8 vs. 109.3 & 109.4). However, WALS combined scores predicted number of work days missed better (AIC:1696.9 & 1673.4 vs. 1795.2).

Conclusions: Combined scores from dual answer keys applied to the WLQ and WALS significantly predicted future work cessation and number of work days missed. Using combined scores, rather than original answer keys, improved the ability of the WLQ to predict both work cessation and work days missed, but only improved the ability of the WALS to predict work days missed.

Disclosure of Interest: None declared

DOI: 10.1136/annrheumdis-2018-eular.5970

FRI0657 CONSTRUCT VALIDITY AND RESPONSIVENESS OF THREE STIFFNESS ITEMS IN RHEUMATOID ARTHRITIS: AN APPLICATION OF THE OMERACT FILTER 2.1

E. T. Craig^{1,*}, C. O. Bingham III¹, S. J. Bartlett², U. Haque¹, A. Leong³, A.-M. Orbai¹. ¹Department of Rheumatology, Johns Hopkins University, Baltimore, United States, ²Center for Mind Body Research, McGill University, Montreal, Canada, ³Patient Research Partner, Santa Barbara, United States

Background: Stiffness is frequent and important to people with rheumatoid arthritis (RA), but measuring it is an unmet need. Qualitative research with people with RA has shown that stiffness duration, severity and interference have face and content validity. These items are being evaluated to develop a patient reported outcome measure (PROM) for stiffness.

Objectives: To evaluate individual stiffness item measurement properties using the OMERACT Filter 2.1.

Methods: Consecutive patients in a longitudinal RA cohort were enrolled. Participants completed 3 items assessing stiffness severity, interference, and duration. PROMs collected included patient global assessment (PGA), pain VAS, self-reported flare (Y/N), PROMIS measures, and patient global impression of change (PGIC). Physician assessments included 28 tender/swollen joint counts, and global (EGA). Disease activity indices were CDAI and DAS28-CRP. Stiffness construct validity was assessed using Spearman correlation coefficients. To assess discrimination in groups expected to change vs stable groups, we calculated standardized response means (SRMs) using multiple anchors.

Results: 196 patients with RA were included. **Construct validity:** all stiffness items correlated highly with pain; moderately with DAS-28 CRP and CDAI; and low with joint counts and EGA. Stiffness severity and interference correlated highly with PGA, pain interference and fatigue. Stiffness interference correlated highly with physical function. **Discrimination:** all stiffness items had moderate effect size for resolution of flare, and for 'much improved' by PGIC. Interference had moderate effect for improved CDAI category and for 'little improved' by PGIC. Stiffness duration had large effect for new RA flare, and 'much worse' by PGIC; severity and interference had moderate effect for RA flare; severity had moderate effect for worsened CDAI and for 'little worse' and 'much worse' by PGIC (TbI).

Table 1 Discrimination of stiffness items by anchor in RA longitudinal sample (n=181). SRM=mean change/pooled SD of change. SRM >0.8 large, 0.5–0.79 moderate, 0.2–0.5 small effect.

Stiffness Anchors	Duration SRM	Severity SRM	Interference SRM
Flare status (n=181)	0.67	0.68	0.54
Flare to No flare (improved) (n=26)	0.01	0.01	0.14
No flare to No flare (unchanged) (n=111)	0	0.19	0.27
Flare to Flare (unchanged) (n=14)	0.81	0.60	0.54
No flare to Flare (worse) (n=30)			
PGIC vs baseline (n=181)	0.52	0.61	0.79
Much improved (n=26)	0.43	0.46	0.60
Little improved (n=25)	0.08	0.08	0.11
No change (n=68)	0.41	0.67	0.42
Little worse (n=49)	1.05	0.59	0.39
Much worse (n=13)			
CDAI vs baseline (n=181)	0.25	0.21	0.61
Better CDAI category (n=45)	0.07	0.12	0.04
Unchanged CDAI category (n=87)	0.48	0.52	0.29
Worse CDAI category (n=49)			

Conclusions: In this cohort of RA patients there is emerging evidence for good construct validity for stiffness duration, severity and interference; moderate discrimination for improvement in RA disease state, and moderate to high discrimination for worsening RA. Additional items are being tested to assemble an instrument for measuring stiffness in RA.

Disclosure of Interest: None declared

DOI: 10.1136/annrheumdis-2018-eular.2142

FRI0658 PROGNOSTIC MARKERS FOR RESPONSE ACCORDING TO THE NEW ACR/EULAR 2016 RESPONSE CRITERIA FOR IDIOPATHIC INFLAMMATORY MYOSITIS

E. Espinosa-Ortega¹, M. Holmqvist², M. Dastmalchi¹, I. Lundberg¹, H. Alexanderson¹. ¹Rheumatology Unit, Department of Medicine, Karolinska Institutet, ²Institutionen för medicin, Solna, Stockholm, Sweden

Background: Idiopathic inflammatory myopathies (IIM) is a heterogeneous group of chronic, inflammatory diseases with varying response to treatment. To

date there is no biomarker to predict such response. Autoantibodies are found in up to 80% of patients with IIM; myositis associated (MAA) and myositis specific antibodies (MSA) are associated with certain clinical features. Whether MSAs can be used as biomarkers to predict response to treatment is unknown.

Objectives: To investigate predictors of response according to the New ACR/EULAR 2016 response criteria for idiopathic inflammatory myositis using a national rheumatology register.

Methods: Patients with IIM who were included in the Swedish Rheumatology Quality Register (SRQ) between January 1st, 2003 and December 31st, 2015 within 12 months of diagnosis were included in this study. Response was assessed at the visit registered closest to one year after inclusion. IIM subgroup (dermatomyositis (DM), polymyositis (PM) and overlap myositis), autoantibody profile, time to diagnosis and treatment were categorized into dichotomous variables. Time to diagnosis (from first symptoms) less than 3 months was considered as early; initial dose of glucocorticoids was considered high when >45 mg. The ACR/EULAR 2016 criteria for Clinical Response were applied to measure response, and it was categorized in minimal (20–39/100), moderate (40–59/100) and major (>60/100) response. The association between response and clinical characteristics was assessed by multinomial logistic regression.

Results: 179 patients were included. Median age at inclusion was 58.9 years (Interquartile range (IQR) 46.7–68.6), 65% were female and 49% had ever smoked. Thirty-three percent had DM, 45% PM and 22% overlap myositis. Lung disease was present in 37% and cancer within ± 3 years from IIM diagnosis was present in 13%. Ninety-one percent were given glucocorticoid treatment, 72% immunosuppressive drugs, 22% cyclophosphamide and 13% a biological drug. **Response rates.** Two thirds (62%) were responders, 23% had minimal, 20% moderate and 18% major response. Minimal response was associated with high initial glucocorticoid dose (OR 3.4, CI95% 1.4–8.1); moderate response to high initial glucocorticoid dose (OR 4.8 CI95% 1.9–11.8) and major response to early diagnosis (OR 3.9 CI95% 1.3–11.9), and high initial glucocorticoids dose (OR 9.5 CI95% 3.0–29.8). No associations between IIM subgroup, autoantibody profile and response rates were observed.

Conclusions: Early and intensive treatment with high doses of glucocorticoids was associated with high rates of clinical response. These data suggest that early intensive immunosuppressive treatment is important in IIM.

REFERENCES:

1. Betteridge Z, McHugh N. Myositis-specific autoantibodies: an important tool to support diagnosis of myositis. *J Intern Med* 2016;280:8–23.
2. Aggarwal R, et al. 2016 American College of Rheumatology/European League Against Rheumatism criteria for minimal, moderate and major clinical response in adult dermatomyositis and polymyositis. *Ann Rheum Dis* 2017;76:792–801.

Acknowledgements: Börje Dahlin Foundation.

Disclosure of Interest: None declared

DOI: 10.1136/annrheumdis-2018-eular.5189

FRI0659

INVESTIGATION OF SELF-REPORTED PAINFUL JOINT COUNT AS OUTCOME MEASURE IN HAND OSTEOARTHRITIS

F. Kroon¹, J. van der Plas¹, S. van Beest¹, W. Damman¹, M. Kloppenburg^{1,2}.

¹Rheumatology, ²Clinical Epidemiology, Leiden University Medical Center, Leiden, Netherlands

Background: Hand osteoarthritis (OA) research is in need of disease-specific validated instruments to measure patient-reported outcomes. Self-reported painful joint count (PJC) could be useful to assess pain, including information about the distribution of painful joints, while being less time-consuming than assessor-reported tender joint count (TJC).

Objectives: To investigate the metric properties of self-reported PJC compared to assessor-reported TJC to measure pain.

Methods: Symptomatic hand OA patients from the Hand OSTeoArthritis in Secondary care (HOSTAS) study marked which joints were painful on standardized hand diagrams. Pain upon palpation was scored on similar diagrams by trained nurses, who were unaware of patients' scores. Self- and assessor-reported joint counts were calculated as the sum of painful/tender joints of both hands (range 0–30). Patients also rated pain on visual analogue scale and completed Australian/Canadian Hand Osteoarthritis index pain subscale, Short Form-36 Mental/Physical Component Summary scales, and Hospital Anxiety and Depression Scale. Radiographs were scored according to Kellgren-Lawrence. MRIs of the right distal and proximal interphalangeal joints (DIPs/PIPs) were evaluated for bone marrow lesions and synovitis. Validity was investigated by assessing correlations between self-reported PJC and assessor-reported TJC, and of both joint counts with other outcome measures. Absolute agreement on joint level, and patient level intra-class correlation coefficients (ICCs) were calculated.

Results: Of 524 hand OA patients (86% women, mean age 61, 90% fulfilling ACR criteria), 506 (96.9%) reported ≥ 1 painful joint (median PJC 8 [interquartile range (IQR) 4–13]), while nurses reported ≥ 1 tender joint in 426 (81.3%) patients (TJC 3 [1–7]). Patients and assessors both most frequently reported pain/tenderness in DIP 2–3, PIP 2–3 and the thumb base, and least often in metacarpophalangeal joints (MCPs) 2–5. Correlation between the joint counts was 0.38. Correlations of both joint counts with other measures were comparable, ranging from 0.11 to 0.37 (table 1). Absolute agreement was highest in MCPs 2–5 and lowest for thumb base, DIP 2–3 and PIP 2–3 (range 61%–89%). ICC between joint counts was 0.28 (95% confidence interval 0.08–0.14).

Table 1 Correlation coefficients for painful joint count (PJC), tender joint count (TJC), and selected clinical and imaging outcomes.

	Self-reported PJC	Assessor-reported TJC
Self-reported PJC	-	
Assessor-reported TJC	0.38**	-
VAS pain	0.37**	0.37**
AUSCAN pain	0.36**	0.36**
SF-36 PCS	-0.30**	-0.27**
SF-36 MCS	-0.12**	-0.11**
HADS depression scale	0.14**	0.15**
Kellgren-Lawrence score	0.08	0.04
MRI synovitis†	0.22*	0.21*
MRI BML†	0.14	0.22*

AUSCAN, Australian/Canadian osteoarthritis hand index; BML, bone marrow lesion; HADS, Hospital Anxiety and Depression Scale; MCS, mental component summary; MRI, magnetic resonance imaging; PCS, physical component summary; SF-36, Short Form 36; VAS, visual analogue scale.

*p<0.05; **p<0.01; †Joint count of DIP and PIP joints of right hand (n=92) was used (range 0–8).

Conclusions: Hand OA patients reported more painful joints than assessors reported tender joints. Self-reported PJC correlated weakly with assessor-reported TJC, and other instruments measuring pain and physical health. Divergent correlations with mental health and imaging scores were even lower. Agreement between patients and assessors was highest in joints with low prevalence of pain/tenderness, but overall agreement was low. Despite apparent advantages, these results suggest that self-reported PJC and assessor-reported TJC cannot be used interchangeably, and further study of PJC as an outcome measure in hand OA is warranted.

Disclosure of Interest: None declared

DOI: 10.1136/annrheumdis-2018-eular.1923