

THU0697 FEASIBILITY AND RELIABILITY OF THE SPARCC SACROILIAC JOINT STRUCTURAL SCORE FOR CHILDREN WITH SPONDYLOARTHRITIS

N.A. Chauvin¹, W.P. Maksymowych², R.G. Lambert³, J.L. Jaremko³, D.M. Biko¹, J. Paschke⁴, T.G. Brandon⁵, P.F. Weiss⁵. ¹Radiology, Children's Hospital of Philadelphia, Philadelphia, United States; ²Rheumatology; ³Radiology, University of Alberta; ⁴CARE Arthritis, Alberta, Canada; ⁵Rheumatology, Children's Hospital of Philadelphia, Philadelphia, United States

Background: Clinical trials in juvenile spondyloarthritis (JSpA) and axial disease are lacking. To assess the effectiveness of medications, we need measures to evaluate structural progression in the pediatric sacroiliac joint (SIJ).

Objectives: To evaluate the reliability of the SPARCC sacroiliac joint structural score (SSS) in children with suspected or confirmed JSpA.

Methods: The SSS assesses a spectrum of structural lesions of the SIJ on MRI including fat metaplasia, erosion, backfill, and ankylosis on 5 consecutive slices through the cartilaginous part of the joint. These components are scored 0–20 (backfill and ankylosis) or 0–40 (fat metaplasia, erosion). We developed a pediatric training module that included a detailed description of each SSS component plus sclerosis (0–40), scoring methodology, and numerous examples. After reviewing the module, 6 readers (mix of adult and pediatric radiologists and rheumatologists) scored 30 studies (Exercise 1). All readers then reviewed a second training module based on DICOM images and scored an additional 29 studies (Exercise 2). Inter-observer reliability was assessed using intraclass correlation (ICC).

Results: The SSS had face validity and was feasible to score in the 59 pediatric cases. 35 (59%) were male and median age was 15 years (IQR 12–16). The ICCs for the SSS components from both calibration exercises are shown in the Table. In the 1st exercise, fat metaplasia and sclerosis had good reliability (≥ 0.4) while reliability for erosion, backfill, and ankylosis were low. During the first exercise, 31 (17%), 131 (73%), 56 (31%), 100 (56%), and 21 (12%) of the 180 ratings from 6 readers had a score >0 for fat metaplasia, erosion, backfill, sclerosis, and ankylosis, respectively. In the second exercise, the ICC for erosion, backfill, and sclerosis improved and had good reliability. Of the 174 ratings from 6 readers, 91 (52%), 18 (10%), and 49 (28%) of scores were >0 for erosion, backfill, or sclerosis in exercise 2, respectively. Reliability for fat metaplasia and ankylosis were not calculated due to low frequency of lesions (8 (5%) and 3 (2%), respectively).

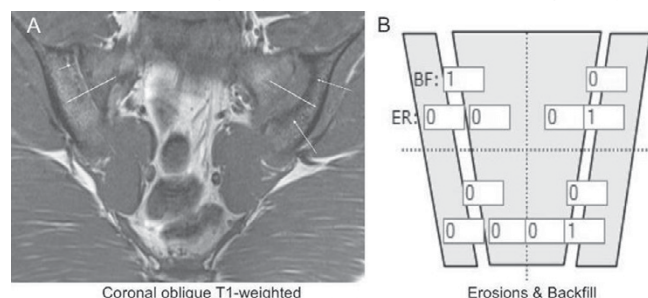


Figure 1. Sample slice and scoring methods for the sacroiliac joint structural score. A. Coronal oblique T1-weighted image of the sacroiliac joints shows each joint divided into 4 quadrants. There is backfill along the superior aspect of the right sacroiliac joint, as demonstrated by increased T1 signal in the joint space (short arrow). There is an extended erosion of the left iliac cortex affecting both upper and lower quadrants of the left iliac bone (long arrows). No sclerosis, ankylosis or fat metaplasia. B. Scoring schematic demonstrates a "1" for backfill (BF) within the R upper SIJ quadrant and "1" in both the upper and lower quadrants of the left SIJ representing erosions (ER).

Conclusions: The SSS was feasible to score and had acceptable reliability for pediatric SIJ MRI evaluation. ICC improved with additional calibration exercises based on DICOM, even for readers with limited experience.

Disclosure of Interest: None declared

DOI: 10.1136/annrheumdis-2017-eular.6146

THU0698 MEASUREMENT PROPERTIES OF PRESENTEEISM MEASURES WITH DUAL ANSWER KEYS IN INFLAMMATORY ARTHRITIS

M. Donaldson¹, A. Kobza¹, D.E. Beaton², M.A. Gignac², D. Lacaille^{1,3}. ¹Arthritis Research Canada, Richmond; ²Institute for Work and Health, Toronto; ³University of British Columbia, Vancouver, Canada

Background: Employment studies in arthritis have emphasised the importance

of decreased productivity at work, or presenteeism. Yet, how to best measure presenteeism remains challenging. The "Work Limitations Questionnaire" (WLQ) is frequently used. A drawback is that it measures the amount of time people are limited, but not the degree to which they are limited. In contrast, the "Workplace Activity Limitations Scale" (WALS) measures the degree of limitation, but not the time. We modified the response keys to the WALS and WLQ to measure both degree of difficulty and amount of time with difficulty.

Objectives: Our objective was to evaluate measurement properties, i.e. internal consistency and construct validity, of the WALS and WLQ with combined scores from dual answer keys.

Methods: A cross-sectional study used baseline data from the RCT of an employment intervention, the "Making It Work" Program. Participants were recruited from BC, Alberta and Ontario. Inclusion criteria included: having inflammatory arthritis, currently employed, age 19–59, and having concerns about arthritis affecting ability to work. 364 participants were included (RA:195, PsA:54, SLE:46, AS:69; 77% female, mean (SD) age: 45.9 (9.8) yrs). Combined scores were obtained by i) multiplying, and ii) adding, the scores of difficulty and time answer keys at the item level. No significant differences were observed between the additive and multiplicative models. Hence, we report on the multiplicative model, which reflects consumers' preference. Internal consistency was analyzed using Cronbach's alphas; construct validity by measuring correlation (Spearman coefficients) between WALS or WLQ subscales and constructs such as work productivity activity impairment (WPAl), risk of impending work loss (work instability, RA-WIS), disease measures, and job characteristics.

Results: Analyses at the item level revealed a strong floor effect (WALS: 16% to 56%; WLQ: 27% to 81% of answers for all items except one) but no ceiling effect, likely reflecting the relatively low limitation expected in a working sample. High (i.e. ≥ 0.7) internal consistency (α 0.70–0.82) was found for WALS and all WLQ subscales except WLQ Physical Demands (0.67). As a priori hypothesized, moderate correlation was observed between the time (0.33–0.57), or combined (0.44–0.61), scores of WLQ subscales and WPAl, or WIS; and correlation was consistently higher for combined than time scores. Moderate correlation was also observed between degree of difficulty (0.70–0.77), or combined (0.60–0.69), scores of the WALS and WPAl, or WIS. Contrary to our hypothesis, the WALS combined score did not have a higher correlation with WPAl, or WIS, than degree of difficulty score. Correlations with disease (fatigue, pain, physical function, depression) and job characteristics (job demand, autonomy, social support at work, commuting difficulty) with time, difficulty, or combined scores of the WALS and WLQ met the a priori hypothesized correlation levels.

Conclusions: Our previous research confirmed the value of measuring both degree of difficulty and amount of time, showing they measure different concepts, which can serve different purposes and are both important to patients. This research demonstrates that applying a dual answer key to two validated instruments shows good initial indicators of internal consistency and construct validity.

Disclosure of Interest: None declared

DOI: 10.1136/annrheumdis-2017-eular.3031

THU0699 CHARACTERIZING AND VALIDATING THE PHENOTYPE OF KNEE PAIN: A LATENT CLASS ANALYSIS

F. Pan¹, J. Tian², D. Aitken¹, F. Cicuttini³, C. Ding¹, G. Jones¹. ¹Musculoskeletal Unit; ²Public Health, Menzies Institute for Medical Research, University of Tasmania; ³Department of Epidemiology and Preventive Medicine, Monash University Medical School, Hobart, Australia

Background: Pain in osteoarthritis (OA) is very common and often involves multiple joints. It is multifactorial and individualised with multiple factors involved in the genesis and pain experience, such as structural pathology, psychological factors and pain coping strategies. Thus it may be possible to group people together based on specific factors which are linked to experiencing pain.

Objectives: To identify and validate the phenotype of knee pain over 10.7 years.

Methods: 1099 participants (mean age 63 years; range 51–81 years) from the population-based Tasmanian Older Adult Cohort study were recruited at baseline. 875, 768 and 563 participants attended years 2.6, 5.1 and 10.7 follow-up, respectively. Demographic, psychological, lifestyle and comorbidities data were obtained at baseline. T1-weighted or T2-weighted fat saturated MRI of the right knee was performed to measure knee structural pathology—cartilage defects, bone marrow lesions (BMLs) and effusion-synovitis at baseline. Knee pain was assessed using Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) at each time-point. Presence of pain (yes/no) at the neck, back, hands, shoulders, hips, knees and feet was also assessed by questionnaire at each time-point. Latent class analysis, which can identify unmeasured class membership

Abstract THU0697 – Table 1

| | Exercise 1 ICC (95% CI) | | | | Exercise 2 ICC (95% CI) | | | |
|----------------|-------------------------|------------------------|------------------|-------------------|-------------------------|------------------------|------------------|------------------|
| | All | Pediatric Radiologists | SSS developers | Rheumatologists | All | Pediatric Radiologists | SSS developers | Rheumatologists |
| Fat metaplasia | 0.40 (0.25–0.58) | 0.46 (0.25–0.67) | 0.89 (0.77–0.95) | 0.82 (0.65–0.91) | | | | |
| Erosion | 0.37 (0.22–0.55) | 0.39 (0.15–0.61) | 0.72 (0.48–0.86) | 0.16 (-0.17–0.47) | 0.54 (0.34–0.72) | 0.51 (0.16–0.74) | 0.96 (0.89–0.98) | 0.61 (0.32–0.79) |
| Backfill | 0.39 (0.25–0.58) | 0.36 (0.14–0.58) | 0.82 (0.56–0.92) | 0.45 (0.12–0.69) | 0.47 (0.31–0.64) | 0.12 (-0.08–0.38) | 0.90 (0.80–0.95) | 0.98 (0.95–0.99) |
| Sclerosis | 0.42 (0.25–0.60) | 0.42 (0.17–0.65) | 0.61 (0.08–0.83) | 0.72 (0.49–0.86) | 0.47 (0.30–0.65) | 0.35 (0.13–0.58) | 0.88 (0.75–0.94) | 0.81 (0.63–0.90) |
| Ankylosis | 0.31 (0.17–0.49) | 0.46 (0.24–0.66) | 0.72 (0.49–0.85) | 0.19 (-0.16–0.51) | | | | |

Legend: ICC <0.40 is poor, $0.40 \leq \text{ICC} < 0.75$ is good, $\text{ICC} \geq 0.75$ is excellent.