Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Classifying radiographic changes of the pubic symphysis in male athletes: Development and reproducibility of a new scoring protocol

Andreas Serner [a,*], Javier Arnaiz [a], Andrea Mosler [a,b], Emad Almusa [a], Zarko Vuckovic [a], Igor Tak [c,d,e], Mario Maas [f], Adam Weir [a,g,h], Per Hölmich [a,i]

[a] *Aspetar Orthopaedic and Sports Medicine Hospital, Doha, Qatar*
[b] *La Trobe Sport and Exercise Medicine Research Centre, School of Allied Health, La Trobe University, Melbourne, VIC, Australia*
[c] *Physiotherapy Utrecht Oost, Bloemstraat 65d, 3581 WD, Utrecht, The Netherlands*
[d] *Amsterdam Collaboration on Health and Safety in Sports (ACHSS), Amsterdam UMC, Vrije Universiteit Amsterdam, Department of Public and Occupational Health, Amsterdam Public Health Research Institute, Amsterdam, The Netherlands*
[e] *Academic Center for Evidence Based Sports Medicine (ACES), Amsterdam University Medical Center, Amsterdam, The Netherlands*
[f] *University of Amsterdam. Amsterdam UMC, Department of Radiology and Nuclear Medicine, Division of Musculoskeletal Radiology, Amsterdam, The Netherlands*
[g] *Erasmus MC Center for Groin Injuries, Department of Orthopaedics, Erasmus University Medical Centre, Rotterdam, The Netherlands*
[h] *Sport Medicine and Exercise Clinic Haarlem (SBK), Haarlem, The Netherlands*
[i] *Sports Orthopaedic Research Center-Copenhagen (SORC-C), Department of Orthopaedic Surgery, Copenhagen University Hospital, Amager-Hvidovre, Denmark*

ARTICLE INFO

ABSTRACT

Purpose: To develop a specified radiographic scoring system for the pubic symphysis and adjacent bones, and to examine the intra- and inter-rater reproducibility of this system.

Method: Development of the scoring protocol was performed in three stages using AP pelvis radiographs of 102 male adult athletes. The final protocol included 5 overall scoring items, which included further specification of locations: 1) bone lucency (erosion-like configuration and cysts), 2) proliferation, 3) fragmentation, 4) sclerosis, and 5) joint space width. Intra- and inter-rater reproducibility were determined using Cohen's kappa statistic (κ) and intraclass correlation coefficient (ICC). The standard error of measurement (SEM) and minimal detectable difference (MDD) were also determined.

Results: We present a radiographic scoring protocol with clear definitions and examples to improve clinical usability. Intra-rater reproducibility was: bone lucency (erosion-like configuration or cysts): $\kappa = 0.67$ (95 %CI 0.56−0.78), proliferation: $\kappa = 0.54$ (95 %CI 0.38−0.70), fragmentation: $\kappa = 0.80$ (95 %CI 0.67−0.93), sclerosis: $\kappa = 0.60$ (95 %CI 0.49−0.71), and joint space width: ICC(2.1) 0.85 (95 %CI 0.78−0.89), SEM 0.4 mm, MDD 1.2 mm. Inter-rater reproducibility was: bone lucency: $\kappa = 0.61$ (95 %CI 0.50−0.72), proliferation: $\kappa = 0.34$ (95 %CI 0.20−0.48), fragmentation: $\kappa = 0.67$ (95 %CI 0.50−0.84), sclerosis: $\kappa = 0.30$ (95 %CI 0.17−0.43), and joint space width: ICC(2.1) 0.72 (95 %CI 0.59−0.81), SEM 0.5 mm., MDD 1.5 mm.

Conclusions: The Aspetar pubic symphysis radiographic scoring protocol contains five overall scoring items, with additional specifications. These five items showed moderate to almost perfect intra-rater reproducibility, and fair to substantial inter-rater reproducibility. This protocol provides the basis for use in clinical practice, and will allow future investigations of the clinical significance of radiographic changes at the pubic symphysis in athletes.

1. Introduction

The pubic symphysis joint is a fibro-cartilaginous joint with a central disc between two hyaline cartilage-covered joint surfaces [1]. These structures are exposed to considerable stress during sports, especially high intensity change of direction movements are considered provocative [2]. Pubic-related groin pain has recently been categorised as a separate defined clinical entity of groin pain in athletes [3], and is diagnosed in the presence of local tenderness of the pubic symphysis and the immediately adjacent bone.

Plain radiographs of the pelvis are a commonly used investigation in athletes who present with long-standing groin pain [4]. They are usually

made to visualize the pelvis as whole, including the hip and pubic symphysis joints. For the pubic symphysis specifically, degenerative changes, including findings, such as sclerosis and joint surface irregularities, have long been considered to be associated with groin pain in athletes. As far back as the 1970's it was reported that athletes often have changes at the pubic symphysis, but also that changes often exist in asymptomatic athletes [5]. Over the past 50 years, a substantial number of pelvic radiographs have been performed, but a well-defined radiographic protocol for the pubic symphysis is still lacking, and scientific knowledge about the normal radiographic appearance of the pubic symphysis is scarce, both in general and athlete populations [6].

A grading scale combining findings to provide an overall assessment of radiographic changes in the pubic symphysis and adjacent bone has previously been proposed [7]. This scale groups findings of erosions, cysts and proliferation into four grades (0–3), from no changes to advanced changes [7]; however, specific items are not defined in detail. This can limit the generalizability of the scale, as it may result in different understandings of the wording used, and create difficulties in the interpretation [8]. While advanced changes appear to be associated with groin pain in general, associations between specific radiographic findings and the defined clinical entities of groin pain are lacking. Radiographic findings may simply reflect adaptations to load, rather than being an actual source of pain. Skeletal changes in response to load are common in asymptomatic male football players, and may not be associated with the presence or development of groin pain [2,9]. In order to investigate whether specific radiographic findings have an association with pubic-related groin pain in athletes, or any of the other clinical entities of groin pain, a detailed and reproducible scoring system is necessary. A systematic review on radiological findings in athletes with symphyseal and adductor-related groin pain found that only 4 out of 17 included studies described inter-rater reproducibility (1 radiographic grading scale [7] and 3 MRI studies) and no studies reported intra-rater reproducibility [6]. The reproducibility of a scoring system is important for the conduct of clinical studies, as this provides information about the amount of error inherent in the measurement [8]. This will assist in the interpretation of the findings and influence clinical usability.

Our aim was to develop a specified radiographic scoring system for the pubic symphysis and adjacent bones in asymptomatic male athletes, and to examine the intra- and inter-rater agreement of this system.

2. Methods

2.1. Participants

The participants in this study were all included as part of a study on screening and risk factors for groin pain in elite male football players [9]. All study participants were elite male football players <18 years old, who played in the Qatar Stars League (QSL) during the 2013–2014 and 2014–2015 seasons. The QSL is the highest level of professional male football in Qatar. Football teams in the QSL generally train 5 times and play 1 game each week. Football players underwent pre-competition screening in one or both soccer seasons, compliant with the Fédération Internationale de Football Association (FIFA) recommendations, as previously described in detail [10]. This screening was deemed mandatory by the Qatar Football Association for all football players playing in the league, and all screening was performed at Aspetar Orthopaedic and Sports Medicine Hospital, Doha, Qatar. All players were informed of the study and invited to participate during this process. Participation in this study was not a requirement of the mandatory screening. Only players without current groin pain were included in this study. All included players provided written informed consent. Ethical approval was obtained from the Institutional Review Board of Anti-Doping Lab Qatar (approval no. F2013000003).

2.2. Radiographs

We obtained an anteroposterior (AP) pelvic radiograph in standing position with the player having both hips in 15° of internal rotation. A film-focus distance of 115 cm was used, with the beam centred at 2.5 cm superior to the pubic symphysis. For the assessment, we focused only on the pubic bone adjacent to the symphysis joint, not including the pubic rami.

2.3. Procedure

Development of the scoring protocol was performed in incremental stages.

2.3.1. Stage 1

Initially, we assessed a previously published 4-grade general scoring system [7]. Because the scoring criteria were not defined in sufficient detail, both in terms of specific items as well as the grading level, we found difficulty in using this scoring system, and decided to test adjustments of the scoring system. We first tried removing the general severity score and instead rated four items dichotomously as either present or absent, without further definitions. These items were; lucency, sclerosis, osteophytes, and fragmentation. We conducted a pilot inter-rater reproducibility study on these four items between two specialised musculoskeletal (MSK) radiologists, with 13 and 15 years of experience. The radiologists scored the four items separately on 19 cases and found moderate to substantial inter-rater agreement (lucency: kappa (K) = 0.69, sclerosis: K = 0.57, osteophytes: K = 0.63, fragmentation: K = 0.79), indicating this approach could be useful.

2.3.2. Stage 2

We chose to further specify the scoring items and add specific definitions aiming to improve the reproducibility, generalisability, and clinical utility. Five overall radiographic findings were decided upon and defined through a review of 20 randomly selected radiographs. The five defined items were; bone lucency (subcategorised into erosions and cysts), bony proliferation, fragmentation, sclerosis, and joint space narrowing. Once the definitions for these items were considered clear between all authors, the 20 radiographs were scored by one radiologist on two occasions three weeks apart. The intra-rater reproducibility of the scoring in this pilot study varied from fair to almost perfect (lucency: K = 0.29, subcategory of erosions: K = 0.30, subcategory of cysts: K = 0.88, proliferation, K = 0.79, fragmentation: K = 0.61, sclerosis: K = 0.71, joint space narrowing: K = 0.90).

2.3.3. Stage 3

Items with inadequate reproducibility were reviewed and definitions modified to enhance clarity. Based on the pilot studies, we performed a sample size calculation for stage 3. We expected a prevalence of the main findings to be between 30–70 % and the sub-categorisation findings as low as 10 %. With an aim of a kappa value of at least 0.8 with a lower 95 % confidence interval (CI) limit of 0.4, assuming no bias between examiners, the required sample size was determined to be 48 for the main findings and 102 for the less frequent findings [12]. Thus 102 radiographs were included and scored for all items.

The same two MSK radiologists scored 102 radiographs independently, blinded to each other's scoring and to any clinical information. The radiographs were randomly selected from the total number of included players using an online randomization tool [13], and did not include radiographs previously used in the pilot studies. For intra-rater reproducibility, one radiologist reassessed the 102 radiographs in a different sequence after an interval of at least 4 weeks to prevent recognition bias. Assessment of one case took an average of around 2.5 min.

2.4. Scoring items

The final radiographic scoring protocol (see supplementary file) consisted of 5 main radiographic findings, with up to 3 sub-classifications (examples in Figs. 1–5). Four items were scored dichotomously as present or absent for each side separately. Each item was scored as present only if it was considered clearly present, or absent if the radiologist was either uncertain of the presence or if considered clearly absent. This was chosen to avoid overestimation of positive findings. The fifth item, joint space width, was measured and reported as a continuous variable. Joint space width was subsequently categorised as "narrow" if the measure was less than 3 mm, based on previously published mean values and our clinical impressions from athlete populations [1,14].

2.4.1. Bone lucency

Definition: *"a clear area of decreased attenuation compared to the surrounding bone, which corresponds to an erosion-like configuration and/or cysts."*

Bone lucency was sub-classified into erosion-like configuration or cysts.

Erosion-like configuration (ELC) was defined as: *"irregularities of the cortical bone surface, potentially accompanied by loss of the adjacent trabecular bone."* ELC was scored separately according to location;

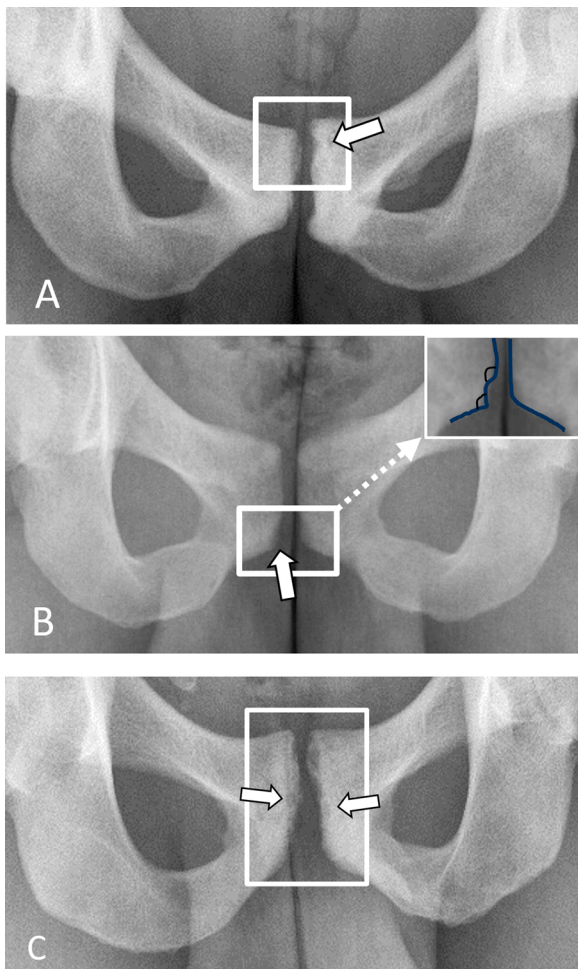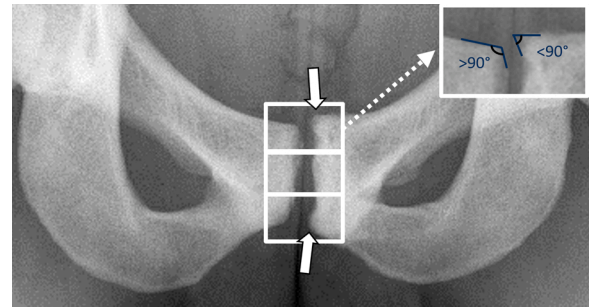a) Superior/central joint surfaces (superior two thirds of the joint surface).

Fig. 2. Proliferation example (superior and inferior).

Fig. 3. Fragmentation examples. A) Central fragmentation, B) Inferior fragmentation.

Fig. 1. Bone lucency examples. A) Superior/central erosion-like configuration, B) Inferior erosion-like configuration, C) Cysts.

Fig. 4. Sclerosis example.

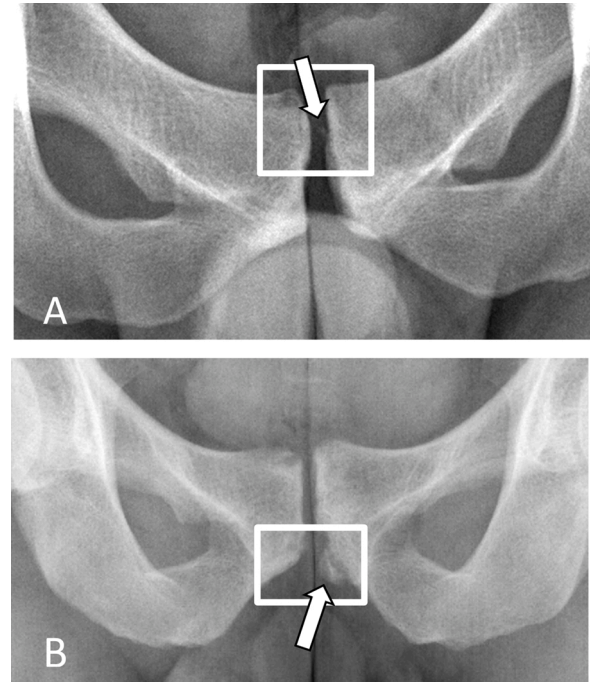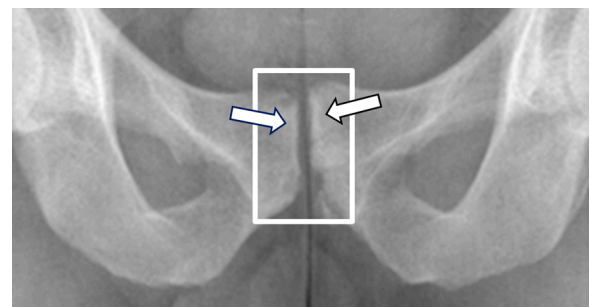b) Inferior margins (lower third of the joint surface)
 - If the entire lower half was considered to have an erosion-like configuration, both of the above were scored as positive.

Cysts were defined as: *"areas of bone lucency with a sclerotic rim inside the trabecular bone compartment, without accompanying cortical bone surface irregularity."*

Examples in Fig. 1.

Fig. 5. Joint space width measurement example.

2.4.2. Bony proliferation
Definition: *"clear osteophyte outgrowths at the joint margins or within the articular space."* Proliferation was further sub-classified into location:

a) Superior joint margin. This can be considered "bone spurs" or classified as "pubic beaking" when bilateral. Well rounded (smooth) bumps at the superior aspect, even if asymmetrical in size, were not considered proliferation. For superior proliferation the "sharpness" of the superior bone corner angle was used for assistance with angles higher than 90 deg. (obtuse angle) considered "rounded" and scored negative/absent, whereas angles lower than 90 deg. (acute angle) were considered "sharp" and scored as positive/present.
b) Central portion of the articular space.
c) Inferior joint margin. Similar considerations as superior proliferation.

Example in Fig. 2.

2.4.3. Fragmentation
Definition: *"clear loose fragment(s) within the symphyseal joint space, or at the inferior medial margin of the pubic bone."*
Examples in Fig. 3.

2.4.5. Sclerosis
Definition: *"a clear area of increased attenuation of the subchondral bone compared to the surrounding bone, corresponding to an area of increased bone density."*
Example in Fig. 4.

2.4.6. Joint space width
Symphyseal joint space was measured in millimetres at the narrowest point of the joint surfaces. Fragmentation within the joint space was ignored in this measurement. The radiograph had to be well centred, i.e. tip of coccyx should be aligned with joint space, as a pelvic rotation could give an impression of more narrow space.
Example in Fig. 5.

2.5. Statistical methods

Intra- and inter-rater reproducibility of the dichotomous scoring was determined using Cohen's kappa statistic (κ). As a low prevalence of certain findings may adversely affect the kappa results [15], the positive, negative, and overall percent agreement were also calculated. Prevalence (P) and bias index (BI) were calculated from the 2×2 tables. For the items scored for both sides (right and left), analyses were performed using 204 sides. For joint space width, analysis was performed for 102 measures. Joint space width was analysed as a continuous variable using intraclass correlation coefficient with a two-way random model using single measures and absolute agreement (ICC 2,1). The standard error of measurement (SEM) and minimal detectable difference (MDD) were also determined, and a potential systematic difference between raters was examined using paired t-tests. Additionally, narrow

joint space, as previously described, was analysed using kappa statistics. For data interpretation, the κ and ICC results were considered poor if <0, slight $0-0.20$, fair $0.21-0.40$, moderate $0.41-0.60$, substantial $0.61-0.80$ and almost perfect if $0.81-1.00$ [16].

3. Results

3.1. Participants

In total 445 players of the 575 male QSL football players screened in these two football seasons had radiographs taken. The demographic data for the 102 players included in this study were: age: median 23 years [interquartile range, IQR, 20–29], range 18–35; height: median 176 cm [IQR 173–181], range 159–197; weight: median 71.3 kg. [IQR 66–77], range 55–105; BMI: mean 23.0 kg/m^2 (SD 2.0), range 18.4–27.1.

3.2. Reproducibility

The intra- and inter-rater reproducibility results are shown in Tables 1–3.

4. Discussion

We have developed a radiographic scoring protocol for the pubic symphysis in several stages to optimize definitions, and analysed the reproducibility hereof. We included five general items, which included further specification of locations. Overall, the final scoring protocol showed substantial intra-rater agreement for most items, whereas inter-rater agreement had higher variation depending on the specific item, generally from fair to substantial agreement.

4.1. Bone lucency

Bone lucency was separated into erosion-like configuration (ELC) and cysts. The intra-rater agreement for these categories were substantial, as well as for the further division into specific locations. In contrast, the inter-rater agreement was moderate for ELC and poor for cysts. The second radiologist explained that he often found this differentiation difficult, and that bowel gas could influence the visualisation of the joint surface line, creating uncertainty (example in Fig. 6A). Joint surface irregularities and cysts determined with magnetic resonance imaging (MRI) have previously been compared separately to groin pain in athletes, however without reproducibility analyses [17,18]. In MRI studies where subchondral cysts and joint surface irregularities were combined, intra- and inter-rater agreement is reported to range from poor to substantial (k = 0.32–0.60 and k = 0.18–0.64, respectively) [19,20], indicating that radiographs may be a better way to determine bone lucency. Alternatively, volumetric interpolated breath-hold examination (VIBE) MRI has shown potential to provide improved visualisation of the pubic symphysis, and may be considered for future studies [21]. Due to the poor to moderate inter-rater agreement of the specification of ELC and cysts in this study, scoring bone lucency as an overall scoring item is more reliable, especially when scoring from different radiologists is compared. Detailed specification into ELC and cysts can be used with more confidence when only one radiologist is involved, considering the substantial intra-rater agreement found in this study. A discussion with the specific radiologist would however be required to ensure clarity.

4.2. Proliferation

Proliferation showed moderate intra-rater agreement and fair inter-rater agreement. The kappa values of this scoring item were however highly influenced by the low prevalence of positive findings, specifically inferior and central proliferation, which were only present in 1–5 % of the cases. Superior proliferation had a higher prevalence of 12–17 %, and showed substantial intra- and moderate inter-rater agreement.

Table 1
Intra-rater agreement results.

	Intra-rater agreement					
Scoring item	Kappa (95 % CI)	Overall agreement	Positive Agreement	Negative Agreement	Prevalence (PI)	Bias index
Lucency	**0.67 (0.56, 0.78)**	85 %	88 %	78 %	65 % (0.30)	0.05
- Erosion-like configuration	**0.65 (0.54, 0.76)**	83 %	87 %	78 %	62 % (0.24)	0.07
• Superior/central	0.67 (0.56, 0.78)	85 %	78 %	88 %	35 % (−0.31)	0.08
• Inferior	0.72 (0.62, 0.82)	86 %	86 %	86 %	50 % (0.00)	−0.02
- Cysts	**0.61 (0.42, 0.80)**	94 %	65 %	96 %	9% (−0.82)	0.00
Proliferation	**0.54 (0.38, 0.70)**	88 %	62 %	93 %	16 % (−0.68)	0.00
• Superior	0.69 (0.53, 0.85)	94 %	72 %	96 %	12 % (−0.77)	−0.03
• Central	0.29 (0.01, 0.57)	94 %	32 %	97 %	5% (−0.91)	0.02
• Inferior	0.40 (−0.14, 0.94)	99 %	40 %	99 %	1% (−0.98)	0.01
Fragmentation	**0.80 (0.67, 0.93)**	96 %	82 %	97 %	13 % (−0.75)	0.01
• Central	.	99 %	0%	99 %	1% (−0.99)	0.01
• Inferior	0.86 (0.75, 0.97)	97 %	88 %	98 %	12 % (−0.76)	0.00
Sclerosis	**0.60 (0.49, 0.71)**	81 %	76 %	84 %	40 % (−0.19)	−0.19
Narrow joint space (3 mm cut-off)	**0.61 (0.44, 0.78)**	84 %	71 %	89 %	27 % (−0.45)	0.04

N = 102 (204 sides), CI = confidence interval, PI = prevalence index, prevalence is reported as a mean of the 2 compared scorings.

Table 2
Inter-rater agreement results.

	Inter-rater agreement					
Scoring item	Kappa (95 % CI)	Overall agreement	Positive Agreement	Negative Agreement	Prevalence (PI)	Bias index
Lucency	**0.61 (0.50, 0.72)**	81 %	85 %	76 %	61 % (0.22)	−0.03
- Erosion-like configuration	**0.57 (0.46, 0.68)**	79 %	82 %	75 %	58 % (0.16)	0.00
• Superior/central	0.56 (0.44, 0.67)	80 %	71 %	85 %	34 % (−0.32)	0.07
• Inferior	0.54 (0.42, 0.65)	77 %	75 %	79 %	45 % (−0.09)	−0.06
- Cysts	**0.16 (-0.06, 0.37)**	91 %	18 %	95 %	5% (−0.89)	−0.07
Proliferation	**0.34 (0.20, 0.48)**	77 %	49 %	86 %	22 % (−0.58)	0.13
• Superior	0.43 (0.27, 0.59)	84 %	52 %	90 %	17 % (−0.66)	0.07
• Central	0.18 (-0.08, 0.43)	93 %	21 %	96 %	5% (−0.91)	0.02
• Inferior	0.28 (-0.15, 0.71)	98 %	29 %	99 %	2% (−0.97)	0.02
Fragmentation	**0.67 (0.50, 0.84)**	94 %	70 %	97 %	10 % (−0.80)	−0.04
• Central	.	100 %	.	100 %	0 (−1)	0
• Inferior	0.67 (0.50, 0.84)	94 %	70 %	97 %	10 % (−0.80)	−0.04
Sclerosis	**0.30 (0.17, 0.43)**	70 %	48 %	79 %	29 % (−0.43)	−0.19
Narrow joint space (3 mm cut-off)	**0.36 (0.16, 0.56)**	75 %	54 %	82 %	27 % (−0.45)	0.04

N = 102 (204 sides), CI = confidence interval, PI = prevalence index, prevalence is reported as a mean of the 2 compared scorings.

Table 3
Intra- and inter-rater reproducibility results of joint space width measurements.

Joint space width (mm)	ICC (95% CI)	Mean (SD) R1a	Mean (SD) R1b/R2	Mean difference (p-value)	SEM (SEM%)	MDD (MDD%)
Intra-rater	0.85 (0.78, 0.89)	3.45 (1.1), range 0.9−8.0	3.33 (1.2), range 0.9−8.1	−0.12 (0.038)	0.4 (13)	1.2 (36)
Inter-rater	0.72 (0.59, 0.81)	3.45 (1.1), range 0.9−8.0	3.22 (1.0), range 0.9−7.4	−0.23 (<0.001)	0.5 (16)	1.5 (45)

N = 102, CI = confidence interval, SD = standard deviation, SEM = standard error or measurement, MDD = minimal detectable difference.

Superior proliferation (also known as "beaking") may be the most relevant location in relation to pubic-related groin pain, as this is possibly associated with central disc protrusion. On MRI, this is described as a cranial bulging on coronal images, and has been shown to be associated with longstanding adductor-related groin pain in male football players [2]. This scoring item, therefore appears specifically relevant to include in future studies on the association between radiographic findings and pubic-related groin pain.

4.3. Fragmentation

Fragmentation agreement values were also influenced by a low prevalence of positive findings. Initially, we chose not to include superior fragmentation in the scoring, as this was not an expected finding based on our experience. There were no cases in this cohort, which led us to change this decision, but it should be noted that we cannot guarantee that this finding may not be present and relevant in rare cases. Our results show that central fragmentation may also be less relevant, with either one or both raters scoring all cases negative. Inferior fragmentation was present in 10–12 % of cases, and showed almost perfect intra- and substantial inter-rater agreement. This inferior fragmentation is likely representing a secondary ossification centre, that could be related to pubic apophysitis, which is receiving increased attention as a potential cause of groin pain in younger athletes [21,22]. Both computed tomography (CT) and MRI scans have been proposed to provide the best assessment of the pubic apophysis [21,22]. Plain radiographs have not been directly compared to these imaging techniques, and may still serve as an initial screening assessment, prior to more expensive and time-consuming investigations or to avoid the radiation dose related to CT scans, as an incomplete apophyseal fusion may be a cause of persistent or recurrent symptoms [22]. The agreement between radiographic findings of inferior fragmentation and CT/MRI findings should however be further explored prior to recommendation on the clinical usability.

4.4. Sclerosis

Sclerosis had substantial intra-rater agreement, but only fair inter-
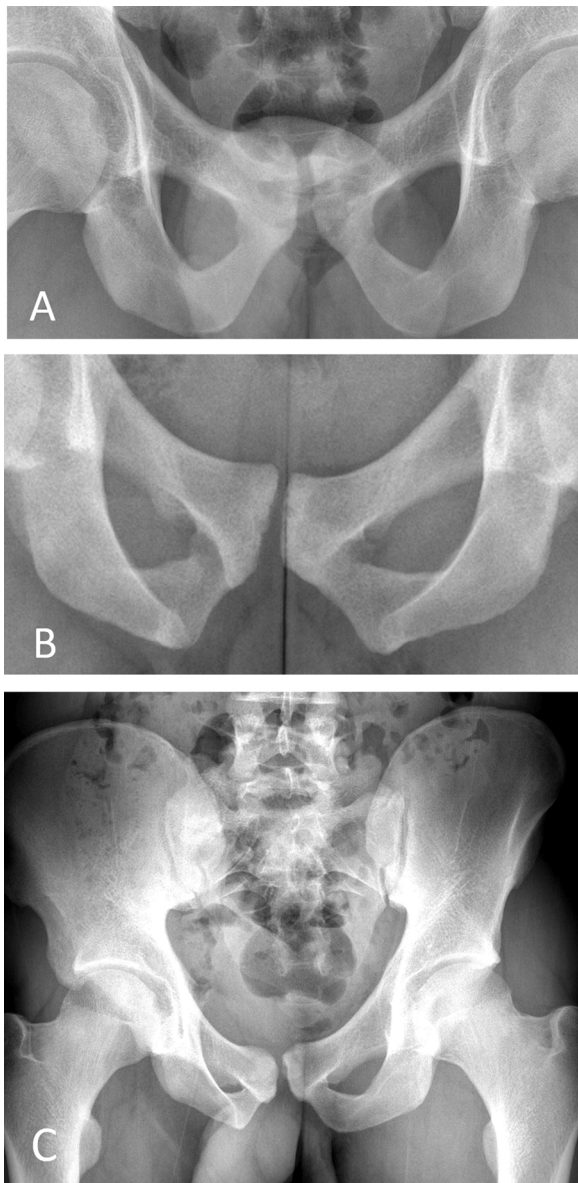
Fig. 6. Examples of how the joint surface visualization can be affected by; A: bowel gas, the coccyx, and the penis, B: Line between the buttocks, C: Example of an image with both excessive pelvic tilt & rotation.

rater agreement. This scoring item had the highest bias index (−0.19) in both analyses, as one radiologist scored this finding higher in the first scoring round than in the second round, and higher than the second radiologist. This indicates that some individual calibration training may be required to determine a cut-off for the presence of sclerosis. Additionally, this item may be difficult to determine in cases where the coccyx or the line between the buttocks may influence the visualisation of sclerosis (examples in Fig. 6A & B). In comparison, intra- and inter-rater agreement MRI scoring of sclerosis in other studies vary from poor to almost perfect (k = 0.34−0.87 and k = 0.19−0.67, respectively) [19,20].

4.5. Joint space

Joint space width had substantial to almost perfect ICC values. Although there was a statistically significant difference between the two ratings and between the raters, we consider the standard error of measurement of 0.4−0.5 mm acceptable. This error is better than reported for ultrasound measurements which vary between 0.3–5.4 mm [23]. The measurement error should be considered if a cut-off value for joint space is implemented. For our agreement analysis, we used a cut-off of 3 mm, which showed moderate intra- and fair inter-rater agreement. Using this cut-off, 27 % of the cases where considered to have a narrow joint space. When assessing joint space, age should however also be considered, as the joint space width decreases from childhood into adolescence [24, 25]. Mean joint space width in adult males is usually reported to be above 4 mm [1]. The mean joint space width in this cohort was 3.2–3.5 mm, indicating that joint width may be influenced by the level of football play. On the other end of the spectrum, an upper limit of normal joint space has been suggested to be around 7−10 mm [5,25]. Our results indicate that a wide joint space is not prevalent in male football players without groin pain. For future use, we recommend reporting joint space width as a continuous measure, as an appropriate cut-off value for this population is uncertain.

Other notable considerations on the assessment of pubic symphysis radiographs is the influence of pelvic tilt and rotation (example in Fig. 6C). Excessive tilt or rotation may influence the impression and accuracy of several of the scoring items. Pelvic rotation may for instance give the impression of a lower joint space width. We do not consider this to affect the reproducibility results in this study, as images with excessive tilt and/or rotation were excluded, however, when higher accuracy is required for clinical or other research purposes, we recommend that assessment of tilt and rotation is standardized, e.g. ensuring that the tip of the coccyx is aligned with the midpoint of the joint space, and using a certain distance measure between the superior border of the symphysis and the sacro-coccygeal joint [26]. We have to emphasize that the AP pelvis radiographs were obtained in the standing position, as not in supine as often performed. There is a possibility that this can influence the visualisation of the pubic symphysis, which may affect the assessment and potentially the reproducibility results presented int his study.

It is well known that the prevalence of positive imaging findings can be higher in athletes, and that the presence of some abnormal imaging findings of the pubic symphysis is often not related to symptoms [2,27]. With this scoring protocol, we provide an indication of the prevalence of positive findings in asymptomatic male football players. The clinically relevance will depend on future studies involving patients with groin pain. With this study, we provide a way to differentiate the radiographic changes found in the symphysis joint in a reliable way. Correlations with symptoms, clinical examination, MRI, CT, and possibly others, requires further scientific evaluation. We acknowledge that there may be specific considerations for different populations, such as female athletes, females post-pregnancy, and non-athletes. Additionally, we recognize that we included two specialized MSK radiologists in this study. This should be considered if results are to be extrapolated to assessors with different specialties, professions or experience.

5. Conclusion

We present a radiographic scoring protocol for the pubic symphysis developed through a 3-staged process using clear definitions and examples. The Aspetar pubic symphysis radiographic scoring protocol contains five overall scoring items, with additional specifications. These five items showed moderate to almost perfect intra-rater agreement, and fair to substantial inter-rater agreement. This protocol provides the basis for use in clinical practice and will allow future investigations into the clinical significance of radiographic findings in athletes.

Funding

CRediT authorship contribution statement

Declaration of Competing Interest

None

Acknowledgements

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ejrad.2020.109452.

References

[1] I. Becker, S.J. Woodley, M.D. Stringer, The adult human pubic symphysis: a systematic review, J. Anat. 217 (2010) 475–487, https://doi.org/10.1111/j.1469-7580.2010.01300.x.

[2] S. Branci, K. Thorborg, B.H. Bech, M. Boesen, M.B. Nielsen, P. Hölmich, MRI findings in soccer players with long-standing adductor-related groin pain and asymptomatic controls, Br. J. Sports Med. 49 (2015) 681–691, https://doi.org/10.1136/bjsports-2014-093710.

[3] A. Weir, P. Brukner, E. Delahunt, J. Ekstrand, D. Griffin, K.M. Khan, G. Lovell, W. C. Meyers, U. Muschaweck, J. Orchard, H. Paajanen, M. Philippon, G. Reboul, P. Robinson, A.G. Schache, E. Schilders, A. Serner, H. Silvers, K. Thorborg, T. Tyler, G. Verrall, R.-J. de Vos, Z. Vuckovic, P. Holmich, Doha agreement meeting on terminology and definitions in groin pain in athletes, Br. J. Sports Med. 49 (2015) 768–774, https://doi.org/10.1136/bjsports-2015-094869.

[4] P. Hölmich, Groin injuries in athletes – new stepping stones, Sport. - Sport. - Sport. Orthop. Traumatol. 33 (2017) 106–112, https://doi.org/10.1016/j.orthtr.2017.03.080.

[5] N.H. Harris, R.O. Murray, Lesions of the symphysis in athletes, Br. Med. J. 4 (1974) 211–214.

[6] S. Branci, K. Thorborg, M.B. Nielsen, P. Hölmich, Radiological findings in symphyseal and adductor-related groin pain in athletes: a critical review of the literature, Br. J. Sports Med. 47 (2013) 611–619, https://doi.org/10.1136/bjsports-2012-091905.

[7] J. Besjakov, C. von Scheele, O. Ekberg, C.F. Gentz, N.E. Westlin, Grading scale of radiographic findings in the pubic bone and symphysis in athletes, Acta radiol. 44 (2003) 79–83.

[8] J. Kottner, L. Audigé, S. Brorson, A. Donner, B.J. Gajewski, A. Hróbjartsson, C. Roberts, M. Shoukri, D.L. Streiner, Guidelines for reporting reliability and agreement studies (GRRAS) were proposed, J. Clin. Epidemiol. 64 (2011) 96–106, https://doi.org/10.1016/j.jclinepi.2010.03.002.

[9] A.B. Mosler, A. Weir, A. Serner, R. Agricola, C. Eirale, A. Farooq, A. Bakken, K. Thorborg, R.J. Whiteley, P. Hölmich, R. Bahr, K.M. Crossley, Musculoskeletal screening tests and bony hip morphology cannot identify male professional soccer players at risk of groin injuries: a 2-Year prospective cohort study, Am. J. Sports Med. 46 (2018) 1294–1305, https://doi.org/10.1177/0363546518763373.

[10] A. Bakken, S. Targett, T. Bere, M.-C. Adamuz, J.L. Tol, R. Whiteley, M.G. Wilson, E. Witvrouw, K.M. Khan, R. Bahr, Health conditions detected in a comprehensive periodic health evaluation of 558 professional football players, Br. J. Sports Med. 50 (2016) 1142–1150, https://doi.org/10.1136/bjsports-2015-095829.

[12] J. Sim, C.C. Wright, The kappa statistic in reliability studies: use, interpretation, and sample size requirements, Phys. Ther. 85 (2005) 257–268.

[13] G.C. Urbaniak, S. Plous, Research Randomizer (Version 4.0) [Computer software], 2013. https://www.randomizer.org/.

[14] B. Alicioglu, O. Kartal, H. Gurbuz, N. Sut, Symphysis pubis distance in adults: a retrospective computed tomography study, Surg. Radiol. Anat. 30 (2008) 153–157, https://doi.org/10.1007/s00276-007-0295-0.

[15] G. Chen, P. Faris, B. Hemmelgarn, R.L. Walker, H. Quan, Measuring agreement of administrative data with chart data using prevalence unadjusted and adjusted kappa, BMC Med. Res. Methodol. 9 (2009) 5, https://doi.org/10.1186/1471-2288-9-5.

[16] J.R. Landis, G.G. Koch, The measurement of observer agreement for categorical data, Biometrics. 33 (1977) 159–174, https://doi.org/10.2307/2529310.

[17] G.M. Verrall, Incidence of pubic bone marrow oedema in Australian rules football players: relation to groin pain, Br. J. Sports Med. 35 (2001) 28–33, https://doi.org/10.1136/bjsm.35.1.28.

[18] H. Paajanen, H. Hermunen, J. Karonen, Pubic magnetic resonance imaging findings in surgically and conservatively treated athletes with osteitis pubis compared to asymptomatic athletes during heavy training, Am. J. Sports Med. 36 (2008) 117–121, https://doi.org/10.1177/0363546507305454.

[19] S. Branci, K. Thorborg, B.H. Bech, M. Boesen, E. Magnussen, M. Court-Payen, M. B. Nielsen, P. Hölmich, The Copenhagen Standardised MRI protocol to assess the pubic symphysis and adductor regions of athletes: outline and intratester and intertester reliability, Br. J. Sports Med. 49 (2015) 692–699, https://doi.org/10.1136/bjsports-2014-094239.

[20] A. Serner, F.W. Roemer, P. Hölmich, K. Thorborg, J. Niu, A. Weir, J.L. Tol, A. Guermazi, Reliability of MRI assessment of acute musculotendinous groin injuries in athletes, Eur. Radiol. 27 (2017) 1486–1495, https://doi.org/10.1007/s00330-016-4487-z.

[21] E. Koh, J. Boyle, Pubic apophysitis in elite Australian Rules football players: MRI findings and the utility of VIBE sequences in evaluating athletes with groin pain, Clin. Radiol. 75 (2020) 293–301, https://doi.org/10.1016/j.crad.2019.12.022.

[22] M. Sailly, R. Whiteley, J.W. Read, B. Giuffre, A. Johnson, P. Hölmich, Pubic apophysitis: a previously undescribed clinical entity of groin pain in athletes, Br. J. Sports Med. 49 (2015) 828–834.

[23] I. Becker, M.D. Stringer, R. Jeffery, S.J. Woodley, Sonographic anatomy of the pubic symphysis in healthy nulliparous women, Clin. Anat. 27 (2014) 1058–1067, https://doi.org/10.1002/ca.22423.

[24] K. Patel, S. Chapman, Normal symphysis pubis width in children, Clin. Radiol. 47 (1993) 56–57, https://doi.org/10.1016/s0009-9260(05)81216-4.

[25] M.J. Budak, T.B. Oliver, There's a hole in my symphysis – a review of disorders causing widening, erosion, and destruction of the symphysis pubis, Clin. Radiol. 68 (2013) 173–180, https://doi.org/10.1016/j.crad.2012.03.021.

[26] M. Tannast, K.A. Siebenrock, S.E. Anderson, Femoroacetabular impingement: radiographic diagnosis–what the radiologist should know, AJR Am. J. Roentgenol. 188 (2007) 1540–1552, https://doi.org/10.2214/AJR.06.0921.

[27] M.A. Martens, L. Hansen, J.C. Mulier, Adductor tendinitis and musculus rectus abdominis tendopathy, Am. J. Sports Med. 15 (1987) 353–356, https://doi.org/10.1177/036354658701500410.