The validation of simple scoring methods for evaluating compartment-specific synovitis detected by MRI in knee osteoarthritis

L. A. Rhodes, A. J. Grainger¹, A.-M. Keenan², C. Thomas², P. Emery² and P. G. Conaghan²

Objectives. Synovitis is common in osteoarthritis (OA) of the knee. In order to evaluate its importance, valid and reliable quantification methods are required. The aim of this study was to compare simple, semiquantitative synovitis scores with detailed volume assessments in subjects with knee OA using magnetic resonance imaging (MRI) as the investigative tool.

Methods. Thirty-five subjects with clinically diagnosed OA of the knee were assessed for synovitis using gadolinium-enhanced MRI. The volume measurement of synovitis was assessed at four sites (medial and lateral parapatellar recesses, intercondylar notch and suprapatellar pouch). The semiquantitative scoring of synovitis was scored on a 0–3 scale.

Results. A total of 140 sites were assessed. There were good correlations between the semiquantitative scores and the volume measurements, ranging from $r = 0.86$ (P<0.00) for the medial parapatellar recess to $r = 0.71$ (P<0.00) for the lateral parapatellar recess. The ordinal regression suggested that, while the synovitis scores predicted between 50 and 70% ($R^2 = 0.53,$ P<0.00 to $R^2 = 0.71,$ P<0.00) of the volume measurements, there was some minor overlap, particularly in the mid-range synovitis scores.

Conclusions. These results suggest that semiquantitative synovitis scores are valid and will enable feasible evaluation of the synovium in OA cohorts.

Key words: Osteoarthritis, Synovitis, MRI, Synovial volume, Knee.

Osteoarthritis (OA) is increasingly common in our ageing society [1, 2]. It is known that the OA process involves the whole joint organ, including the subchondral bone and menisc; synovitis is common from its early stages [1, 3, 4]. The importance of OA synovitis in terms of pain and structural progression has not been clearly determined [5]. The synovium therefore remains a key focus for OA research.

Magnetic resonance imaging (MRI) provides a non-invasive way of studying the synovium throughout the whole joint (compared with arthroscopy, which maybe limited by access). At present the use of the contrast agent gadolinium diethylene-triamine penta-acetic acid (Gd-DTPA) is the most sensitive way of detecting OA synovial hypertrophy, confirmed on biopsy as synovial inflammation [6]. Quantification of this MRI synovitis may be obtained by volumetric methods or by dynamic assessment. Most of the work in this field has been performed in the knee, predominantly in rheumatoid arthritis [7–11]. Detailed volume assessment with manual outlining of enhancement in post-gadolinium images currently represents the gold standard of synovial volume assessment. However, these methods are laborious and not feasible for evaluating large OA cohorts.

Subsequently, semiquantitative or ‘eyeballed’ scores have been reported, such as the Whole-Organ MRI Score (WORMS) [12] and the Knee Osteoarthritis Scoring System (KOSS) [13]. However, there are no reports on the criterion validity of these scores (that is, their comparison with a gold standard). Also, such scores have assessed global knee joint synovial changes. However, given that the importance of synovitis at the level of the individual joint compartment is unknown, it would seem desirable to have a scoring method that includes multicomartment evaluation.

The aim of this study was therefore to compare a semiquantitative synovitis scoring system against more laborious volumetric assessments at a number of distinct anatomical sites throughout the OA knee.

Methods

Thirty-five subjects (20 females and 15 males) with clinically diagnosed OA of the knee were assessed for synovitis using gadolinium-enhanced MRI. All subjects fulfilled the American College of Rheumatology criteria for OA (clinical and radiographic criteria) [14]. The Leeds (West) Research Ethics Committee approved the study protocol and all subjects gave informed written consent to their participation according to the Declaration of Helsinki [15].
Magnetic resonance imaging

MRI of the knee was performed using a 1.5 T Gyroscan ACS-NT whole-body scanner (Philips Medical Systems, Best, The Netherlands). A Philips quadrature receiver knee coil was used; subjects were placed in the supine position for imaging. T1 pre- and post-gadolinium sagittal and coronal images were acquired with the following parameters: repetition time (TR) 604 ms, echo time (TE) 14 ms, slice thickness/slice gap 3 mm/0.3 mm, field of view (FOV) 160 mm, rectangular field of view (RFOV) 80%, matrix size 256 x 256 and number of signal averages (NSA) 2. For T2W pre-gadolinium fat-suppressed (FS) trans images, TR was 5589 ms, TE 100 ms, slice thickness/slice gap 3 mm/0.3 mm, FOV 160, RFOV 80%, matrix size 256 x 256 and NSA 4. For T1W post-gadolinium FS trans images, TR was 650 ms, TE 15 ms, slice thickness/slice gap 3 mm/0.3 mm, FOV 160, RFOV 80%, matrix size 256 x 256 and NSA 2. Gd-DTPA was administered at 0.2 ml (0.1 mmol)/kg body weight as a 14-s bolus. Post-Gd-DTPA sequences began 4.5 min after injection.

Sites chosen for analysis

The following four anatomical sites were chosen for comparison: the medial and lateral parapatellar recesses, the intercondylar notch and the suprapatellar pouch. The semiquantitative scoring system was compared in areas where formal volume measures could reliably be made. Formal volume measures in the infrapatellar (Hoffa’s) fat pad are difficult, given the multiple infoldings of the synovium and interspersed fat commonly seen in this area, so this area was not included in the comparison.

Quantitative synovitis measurements

Reader 1 (L.A.R.) performed quantitative assessment of synovitis volume within the four anatomical sites. Axial T1-weighted post-Gd-DTPA FS images were chosen in preference to other planes as they allowed the assessment of the four anatomical sites on one set of images.

Image analysis was performed using a commercially available image analysis software package (Analyze; Analyze-Direct, Lenexa, KS, USA). This software allowed the generation of regions of interest (ROIs) that delineated the enhancing synovium by manual outlining in the four anatomical sites on consecutive slices. The number of enhancing pixels within the ROI was calculated in each slice and converted to an area measurement. The volume of synovitis in each anatomical site was calculated using the following formula: Vol_{syn} = \Sigma (A_{syn} \times ST) [16, 17], where ST represents the sum of the slice thickness and the slice gap and A_{syn} represents the area of the synovial volume in each slice.

Synovitis in the medial and lateral parapatellar recesses on the axial sequences was assessed on consecutive slices, with the superior and inferior boundary of these recesses defined as the slice showing the superior and inferior pole of the patella, respectively. Synovial tissue on these slices was assigned to either the medial or lateral parapatellar recess on the basis of a sagittal line dividing the joint space at the crista of the patella (Fig. 1a). The intercondylar notch on the axial sequence was defined as the space lying between the femoral condyles and limited by a line tangential to the posterior surface of the condyles (Fig. 1b). Synovitis in the suprapatellar pouch was assessed on consecutive slices superior to the patella and, where available, to the top of the pouch (Fig. 1c).

Semiquantitative synovitis scoring

Reader 2 (A.J.G.) performed the semiquantitative scoring using all available sequences. Synovitis was scored using grades 0–3, where 0 represents normal synovium, 1 represents diffuse, even thickening, 2 shows nodular thickening and 3 represents gross nodular thickening of the synovium (Fig. 2).

Readers 1 and 2 were blinded to the results of the comparative study at the time of evaluation.

Strategy for statistical analysis

While the level of association between the data from the synovitis score (which is ordinal) and the volume measurement (continuous data) for each anatomical site was initially investigated using Spearman’s \( \rho \), of greater interest was an exploration of the predictive ability of the synovitis scores to predict the volume measurement obtained from the pixel count. Regression is considered a much more powerful tool than simple correlation as it adds prediction capabilities and the strength of a relationship is much more explicit [18, 19]. While regression values are mathematically closely related to correlation, regression values are expected to be lower than reliability estimates. In order to undertake ordinal regression, pixel count was recoded as ordinal data and employed in an ordinal regression model, with the dependent variable as the volume measurement and the covariate as the synovitis score. Problems associated with small sample size are acknowledged. However, confirmation of assumptions and model validation were explored for each anatomical site and are reported. Outliers were included in all analyses.
In order to establish the test–retest reliability of the semiquantitative scoring, 40 images were evaluated on two occasions by the same assessor under blinded conditions. Across all sites, the retest reliability was excellent ($\rho = 0.93$, $P < 0.00$). For the 40 sites, there was perfect agreement on 34 occasions (85%), all six non-agreements differing by only one scoring point.

### Results

Thirty-five subjects, including 12 males and 23 females with a median age of 63 yr (range 49–77 yr), participated in this study. Participants demonstrated mild to moderate OA, as indicated by the Kellgren–Lawrence scores: the median score for the medial femoral tibial area was 3, for the lateral femoral tibial area it was 1.5, and for the patellofemoral area it was 2. Participants were evaluated at the four different sites, resulting in a total of 140 sites for assessment between the quantitative and the semiquantitative assessment of the synovitis volume. Box plots for each anatomical site are presented in Figs 3–6. The median scores for the medial parapatellar recess was 2, with the corresponding median volume measure 8373.75; for the lateral parapatellar recess this was 2, with a median lateral volume measurement of 8404.69; the suprapatellar pouch median score was 1, with a median suprapatellar pouch volume measurement of 4217.81; and for the intercondylar notch, the median score was once again 2, with a median volume measurement of 5001.56. As can be seen in these figures, there was a positive visual association between the synovitis score and the volume measurement.

Outliers are presented graphically for each scoring category and were included in the analysis. Note there were no zero scores in the intercondylar notch data.

Statistics of agreement for each anatomical site, including Spearman’s $\rho$ and levels of statistical significance, are presented in Table 1. There was strong agreement between the synovitis score and the volume measurements for each of the sites. The level of correlation was excellent for the medial parapatellar recess ($\rho = 0.86$, $P < 0.00$) and good for the lateral parapatellar recess, suprapatellar pouch and intercondylar notch ($\rho = 0.71$, $\rho = 0.72$ and $\rho = 0.78$, respectively; $P < 0.00$).

A summary of the ordinal regression is presented in Table 2. Testing for goodness of fit indicated that the model was supported for each anatomical site ($\chi^2 = 42.75$, $P < 0.00$; $\chi^2 = 26.04$, $P < 0.00$; $\chi^2 = 96.21$, $P < 0.00$; $\chi^2 = 37.43$, $P < 0.00$). The model suggested that the synovitis score predicted 70% of the volume measurement for the medial parapatellar recess ($R^2 = 0.71$, $P < 0.00$), 53% of the volume measurement for the lateral parapatellar recess ($R^2 = 0.53$, $P < 0.00$), almost 60% for the suprapatellar pouch ($R^2 = 0.59$, $P < 0.00$) and 66% for the
intercondylar notch ($R^2 = 0.66, P < 0.00$). However, when explored for discrimination, there was significant overlap in the 95% confidence interval parameters for the volume measurements. This suggests that there is overlap in whether a particular volume measurement received a particular score. This difficulty in discrimination was particularly noted in the mid-range volume measurements, where there was more inconsistency with respect to whether a particular volume measurement was classified as 1 or 2. This is also reflected in the descriptive data in the box plots.

**Discussion**

This study demonstrated that semiquantitative scoring of OA synovitis correlated well with more laborious manual volume assessments, at the level of individual knee compartments. Although semiquantitative synovitis scores have accepted face validity, this study provides important information on their criterion validity; that is, their agreement with the current synovial volume gold standard. Clearly, no longitudinal data have been presented here with which to report on sensitivity to change.

This study employed gadolinium, which is not used commonly in OA-evaluated cohorts. Although gadolinium increases the inconvenience to the patient and the cost of such studies, it is currently the best way of detecting synovitis, and initial studies on the importance of synovitis in OA should include gadolinium enhancement.

There was a good to strong correlation between the synovitis scores and the volume measurements for each of the anatomical sites. The highest correlation was seen for the medial parapatellar recess, with good correlation observed in the lateral parapatellar recess, suprapatellar pouch and the intercondylar notch. The regression modelling suggested that the synovitis scores were a strong predictor of the volume measurement and predicted between 50 and 70% of the volume measurement; such findings are consistent with what have been considered to be ‘strong predictors’ elsewhere in the medical literature [20–24].

There are limitations to this work. With only 35 patients included, there cannot be a complete spread of scores for each of the sites evaluated. This cohort had a greater proportion of subjects with mid-range scores and fewer extreme scores. This may have resulted in higher correlations with the minimal number of extreme values and must be considered in the evaluation of this study. While the regression modelling suggests that the synovitis scoring predicts a reasonable proportion of the volume scoring, it is acknowledged that the numbers involved are at the lower end of what is recommended when undertaking ordinal regression modelling. Furthermore, while the volume measurement has been used as the gold standard in this study, this should not imply that the quantity of synovitis alone may be the most important factor to measure in OA studies, when other features, such as location and relationship to other structural abnormalities, may also be important.

In the absence of well-defined anatomical markers, this study employed pragmatic definitions for anatomical localization. Another limitation is the use of a single plane for assessing a given anatomical compartment. This study employed pragmatic definitions for anatomical localization. Another limitation is the use of a single plane for assessing a given anatomical compartment.

**Table 2. Single-factor ordinal regression modelling for volume measurement and synovitis scores, where B is the parameter estimate and the Wald statistic is an estimation of the parameter estimate to the standard error (SE)**

<table>
<thead>
<tr>
<th>Volume measure</th>
<th>Goodness of fit $\chi^2$ (significance)</th>
<th>Cox and Snell pseudo $R^2$</th>
<th>n</th>
<th>Parameter estimate</th>
<th>s.e. (B)</th>
<th>Wald</th>
<th>P value</th>
<th>95% CI (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medial parapatellar recess</td>
<td>42.75 ($P &lt; 0.00$)</td>
<td>0.71</td>
<td>35</td>
<td>3.81</td>
<td>0.70</td>
<td>29.41</td>
<td>&lt;0.00</td>
<td>2.43</td>
</tr>
<tr>
<td>Lateral parapatellar recess</td>
<td>26.04 ($P &lt; 0.00$)</td>
<td>0.53</td>
<td>35</td>
<td>2.24</td>
<td>0.49</td>
<td>20.94</td>
<td>&lt;0.00</td>
<td>1.28</td>
</tr>
<tr>
<td>Suprapatellar pouch</td>
<td>96.21 ($P &lt; 0.00$)</td>
<td>0.59</td>
<td>35</td>
<td>3.09</td>
<td>0.62</td>
<td>24.13</td>
<td>&lt;0.00</td>
<td>1.85</td>
</tr>
<tr>
<td>Intercondylar notch</td>
<td>37.43 ($P &lt; 0.00$)</td>
<td>0.66</td>
<td>35</td>
<td>4.16</td>
<td>0.82</td>
<td>25.62</td>
<td>&lt;0.00</td>
<td>2.55</td>
</tr>
</tbody>
</table>

Synovitis scores for each anatomical site were included as the covariate in the model.
region using formal volumes; in the present study semiquantitative scoring therefore utilized all available views. For formal volume measures it would have been possible to measure the volume on multiple views and combine the result (e.g. the mean) but, given the time-consuming nature of the measurements, this was not felt feasible. While the axial view may be inaccurate at the apex of the suprapatellar pouch due to partial voluming, the whole of the pouch usually has an anterior convex configuration, making partial voluming an equal if not greater problem in the sagittal plane. The axial plane seemed to be the best compromise and it allowed all regions to be measured in the same plane.

In summary, this study has demonstrated that semiquantitative assessment of compartment-specific synovial volumes in knee OA is valid and provides a reliable, feasible method for evaluating large OA cohorts.

### Key messages

- This study used gadolinium-enhanced MRI to assess synovitis in the OA knee.
- This study demonstrated a valid semiquantitative assessment of compartment-specific synovial volumes in knee OA.

### Acknowledgements

We are grateful to MSD UK for partial sponsorship of this study through an unrestricted educational grant. P.G.C. is partially funded through a grant from the Arthritis Research study through an unrestricted educational grant. P.G.C. is We are grateful to MSD UK for partial sponsorship of this study through an unrestricted educational grant. P.G.C. is partially funded through a grant from the Arthritis Research study through an unrestricted educational grant. P.G.C. is

### References