The validity of in vivo ultrasonographic grading of osteoarthritic femoral condylar cartilage: a comparison with in vitro ultrasonographic and histologic gradings

† Department of Physical Medicine and Rehabilitation, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan
‡ Department of Pathology, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan
§ Department of Orthopedic Surgery, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan
∥ Department of Physical Medicine and Rehabilitation, Faculty of Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan
¶ Department of Pathology, Faculty of Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan
# Department of Orthopedic Surgery, Faculty of Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan

Summary

Objectives: To establish an ultrasonographic (US) grading for semiquantitative evaluation of the femoral condylar cartilage of knee osteoarthritis (OA), in vivo, and compare the in vivo US grading with the in vitro US and histologic gradings.

Design: Ninety-five patients going to receive total knee arthroplasty because of OA of the knee were recruited. US examination was performed in vivo in the day before operation using a grading system including parameters of margin sharpness, clarity and thickness. Specimens of the medial and lateral distal femoral condyles taken during the operation were graded with in vitro US and histologic examination. The correlation between the in vivo US and in vitro US as well as between the in vivo US and histologic gradings was analyzed.

Results: In 172 femoral condyles (including medial and lateral ones), the distribution of grading ranged from Grade 1 to 6 in in vivo US and from Grade 1 to 4 in histologic examination. The in vivo US grading was significantly correlated to in vitro US grading over anterior and middle areas (p < 0.001, Rho = 0.35 and 0.45, respectively) and histologic grading over these two areas (p < 0.001, Rho = 0.40 and 0.36, respectively). When the cases with maximal angle of knee flexion less than 120 degree were excluded, the correlation was better.

Conclusions: The significant correlation between in vivo US and histologic gradings might permit semi-quantitative in vivo US assessment of osteoarthritic femoral condylar cartilage.

Introduction

Osteoarthritis (OA) has been a prevalent musculoskeletal disease† usually causing pain and disability‡. The earlier medical treatments largely aimed on symptoms relief using medication like NSAIDs. Recently, slow-acting drugs in OA (SADOAs) have been developed with both symptom modifying effects and structure modifying effects§. The disease-modifying OA drugs (DMOADs) are used to prevent, decelerate, or even reverse degeneration of cartilage. Nonetheless, most studies still employed symptom and disability indices as treatment measurement, such as pain visual analog scale, Lequesne index and questionnaire of Western Ontario and McMaster Universities Osteoarthritis (WOMAC)¶. Radiography is a frequently used image modality#. Cartilage thickness is indirectly evaluated by the joint space width measurement. Yet, variability in knee positioning creates varying estimates of the joint space width#. It was reported that the meniscus accounts for a substantial proportion of the variance explained in joint space narrowing#. Another major limitation of the joint space width measurement is that we can only measure the sum of the two opposing cartilage layers but not the individual cartilage thickness.

Arthroscopy provides a direct view of the cartilage and is considered as gold standard for assessing cartilage lesions§. Its scoring and grading systems were also developed and validated#*. However, the disadvantage of arthroscopy is that it is an invasive procedure.
Non-invasive procedures to directly visualize particular cartilage include computed tomography (CT), magnetic resonance imaging (MRI) and ultrasonography (US)\(^\text{13}\). The CT images can be obtained in the transverse plane, which is perpendicular to the direction of weight bearing, but cartilage over the weight bearing area cannot be visualized. Although the image reconstruction in the sagittal and coronal planes is available, the resultant image quality is inferior to those obtained by a direct imaging technique\(^\text{13}\). MRI has also developed in recent years. Besides the thickness measurement, MRI has also been implemented for molecular imaging of cartilage\(^\text{14}\). Its limitations on clinical application are due to the high cost and limited accessibility.

US is a non-invasive, widely available, and relatively inexpensive technique. It could be promptly performed and easily accepted by patients, and uses no ionizing radiation\(^\text{15}\). Musculoskeletal US was used to evaluate periarticular and intraarticular structures\(^\text{15,16}\). Naredo \textit{et al.} used US to evaluate articular cartilage. Besides the absolute measurement of thickness, they believed that the ratings of clarity and sharpness correlated better with clinical status\(^\text{18}\). McCune \textit{et al.} also found that the gradings of clarity and sharpness were most reliable quantitative predictors of pathologic change, whereas the thickness measurement became less reliable as the severity of pathology increased\(^\text{19}\). Grassi \textit{et al.} concluded that all the features should be combined to evaluate articular cartilage damage\(^\text{20}\). Our previous study had set up a grading system including these three major features to evaluate the femoral condylar cartilage and found good in vitro validity in comparison with histologic grading\(^\text{21}\). The purpose of this study is to demonstrate in vivo validity by examining the correlation of in vivo US and histologic gradings in a group of severely diseased group.

**Material and method**

**SUBJECTS**

We recruited patients with knee OA who were going to receive total knee arthroplasty (TKA). Clinical symptom and radiograph determined the diagnosis of OA and decision of treatment with TKA. TKA is decided according to the incurrence of intolerable pain and radiographic severity equal to or greater than Ahlbäck stage III\(^\text{22}\). Thus, all the radiographic severity was supposed to be equal to or greater than Ahlbäck stage III. Patients with inflammatory arthritis (e.g., rheumatic arthritis) and infectious disease (e.g., hepatitis B) were excluded.

**IN VIVO US EXAMINATION**

All patients received in vivo US examination in the day before operation of TKA, with a commercially available US device using a real time 5–12 MHz high-resolution linear transducer (HDI 1500; Advanced Technologies Laboratories, Bethell, WA). Patients were supine on an examination bed, with the knee to be replaced flexed as much as possible. The angle of knee flexion was recorded. A horizontal grid just above the patella was drawn. The transducer was placed on the grid and perpendicular to the surface of the knee. The imaging parameters were set properly without change in the evaluation of all the subjects. The US grading over lateral and medial femoral condyles was made by the same experienced investigator. Another investigator repeated the US examination on some of the subjects.

We created the US grading system based on the conclusions of previous studies, including evaluation on sharpness of the superficial margin, and clarity and thickness of the cartilage band\(^\text{18,21}\). Grade 0 indicates normal cartilage with absolute sharpness and clarity and uniform thickness. Grade 1 indicates blurred margin or partial lack of the clarity without thickness change. Grade 2 shows blurred margin and partial lack of the clarity without thickness change, while Grade 3 indicates blurred margin and completed lack of the clarity. If the margin is almost difficult to be defined and the band is complete opaque, Grade 4 is coded. If the thickness of the band is markedly changed, Grade 5 is coded. Grade 6 is coded if the cartilage band could not be visualized (Fig. 1).

Fig. 1. US grading of OA cartilage (indicating the area inside the square mark). Grade 1: blurred margin or partial lack of the clarity, without thickness change. Grade 2: blurred margin and partial lack of the clarity, without thickness change. Grade 3: blurred margin and complete lack of the clarity. Grade 4: difficult-defined margin and the complete-opaque band. Grade 5: marked thickness change. Grade 6: no visualized cartilage band. Gr: Grade.
SPECIMEN AND IN VITRO US EXAMINATION

During the TKA operations, distal femoral condyles were excised with an oscillating saw and rinsed in saline. The distal medial and lateral condyle specimens were chosen. The specimens were examined using US in vitro by the first US investigator. The grading system was the same as our previous report\textsuperscript{21}. Since the investigator saw the specimens but not the subjects, she was blinded to the results of in vivo US performed 1 day early.

HISTOLOGIC EXAMINATION

The specimens were fixed with 4% paraformaldehyde for 24 h, and decalcified in a decalifier for a few days. Then the specimens were processed for embedding in paraffin wax. Standardized 4-μm-thick sections of the specimens were prepared from selected areas. The sections were stained with haematoxylin and eosin (HE) for histologic examination. The characteristics of histologic changes in osteoarthritic cartilage include loss of matrix staining, surface fibrillation, deeper fissure or clefts, reduced thickness, loss and erosion of cartilage, and thickened sclerotic subchondral bone\textsuperscript{22}. Grading or staging systems have been previously described\textsuperscript{24,25}. Two pathologists, who were blinded to US results, evaluated the histologic changes in the articular cartilage under light microscope. They classified the histologic change into four grades according to Huang’s criteria\textsuperscript{24} (Fig. 3).

The degeneration of the articular cartilage is usually inhomogeneous. It is difficult to select an area to represent the whole change. We selected the areas for examination just based on the location. Two transverse areas on each specimen, including anterior and middle areas, were chosen (Fig. 2). The middle area of the specimen was located over the middle part of the specimen and represented the weight bearing portion of the femoral condyle. The anterior area of the specimen was located over the middle of the anterior part of the specimen and represented the sub-weight bearing portion. Both in vitro US examination and histologic examination were performed on both these two areas.

STATISTICAL ANALYSIS

Weighted Kappa was used for evaluating the agreement between two observers of in vivo US examination. Spearman correlation was performed to determine the correlation between in vivo US and in vitro US gradings, and between in vivo US and histologic gradings of articular cartilage. A significant correlation was defined as $p$ value less than 0.05. A Rho score of less than 0.40 indicated weak correlation; that of 0.40–0.75, moderate correlation; and that of 0.75–1.00, good correlation.

Results

Ninety-five patients, 80 females and 15 males, were recruited in this study. Mean age was 71.9 years with SD of 5.9 years. Fifty-one patients received TKA of right knee, 44 patients of left knee. Maximal angle of knee flexion ranged from 50 to 145 degrees.

THE DISTRIBUTION OF US AND HISTOLOGIC GRADINGS

After decalcification, only 172 specimens (84 from medial and 88 from lateral femoral condyles) could be analyzed. The distributions of in vivo US and histologic gradings are listed in Tables I and II. The discrepancy of cartilage degeneration severity over lateral and medial femoral condyles was noted in both in vivo US and histologic gradings. Lateral parts were less severe than medial parts. While mixing the data from lateral and medial femoral condyles, the grades included ranged from Grade 1 to 6 in in vivo US and from Grade 1 to 4 in histologic examinations.

INTER-OBSERVER REPEATABILITY

The first 34 patients received a repeated in vivo US examination by a second investigator. Inter-observer repeatability revealed moderate agreement in examining 34 medial femoral condyles, 34 lateral femoral condyles, and 68 mixed medial and lateral femoral condyles (Weighted Kappa = 0.67, 0.61 and 0.68, respectively).

THE CORRELATION OF IN VIVO US GRADING VS IN VITRO US GRADING AND IN VIVO US GRADING VS HISTOLOGIC GRADING

We performed in vitro US examination on 171 specimens. The in vivo US grading was significantly correlated with in vitro US grading over anterior and middle areas ($p < 0.001$, Rho = 0.35 and 0.45, respectively). After excluding the specimens from subjects who could not flex their knee equal to or greater than 120 degrees, the correlation was better over anterior area and similar over middle area ($p < 0.001$, Rho = 0.41 and 0.44, respectively). The relationship between in vivo US and histologic gradings over anterior and middle area is presented, respectively, in Tables III and IV. We also examined the condition where the cases with maximal angle of knee flexion less than 120 degrees were excluded.

The in vivo US grading was significantly correlated to histologic grading over anterior and middle areas ($p < 0.001$, Rho = 0.40 and 0.36, respectively). When the cases with maximal angle of knee flexion less than 120 degrees were excluded, the relationship between in vivo US and histologic gradings over anterior and middle areas was still significant and the correlation coefficient was better (Rho = 0.44 and 0.39 over anterior and middle areas, respectively) (Table V).

Discussion

This study demonstrates the validity of in vivo US grading of osteoarthritic femoral condylar cartilage with comparison with histologic grading. The correlation between these two types of grading was significant. Besides, in vivo US grading was also significantly correlated to in vitro US grading. The correlation was better in cases with maximal angle of knee flexion greater than or equal to 120 degrees.

Normal articular cartilage is a smooth-surface hyaline cartilage. The cartilage contains chondrocytes and an abundant extracellular matrix which is primarily made up of type II collagen fibers, proteoglycans, and water\textsuperscript{26}. The appearance on US is a homogenous anechoic band with a sharp
margin. The lack of echoes is due to uniform transmission of sound wave in cartilage with high water content and densely packed, regularly organized collagen. The sharp margin corresponds to smooth surface of healthy cartilage. In early OA, surface fibrillation develops and results in a loss of sharpness of the margin on US images. Deeper fissure or clefts cause the increased echogenicity. The thickness reduction, which could be showed on US examination, develops as the course of degeneration proceeds.

Previous studies had proposed that the major US features of osteoarthritic cartilage were loss of margin sharpness, loss of clarity of cartilage band, and thickness reduction. The thickness measurement by in vivo US was demonstrated to be highly correlated with that by MR imaging over corresponding area of intercondylar notch (Spearman’s rank correlation coefficient 0.82). However, another study showed that the correlation was significant but only moderate over central portion of the femoral condyles in symptomatic patients (Spearman correlation coefficient 0.44–0.61). Besides, the in vivo US and MR gradings of the clarity and the sharpness of the cartilage were significantly correlated in the knees with OA (Spearman’s rank correlation coefficient 0.27–0.38). In a cadaver study, Mathiesen et al. revealed agreement between real measurement and US examination in cartilage thickness and the extent and depth of induced cartilage defects.

Combined evaluation was thought to be important for diagnosing cartilaginous damage. Spietz et al. set up a grading system using these US features. When they applied this grading system in a standard surgical model of OA in rabbits with anterior cruciate ligament transaction, they revealed highly significant correlation between US and histologic grades. They used high frequency US transducer (40 MHz) and performed the US examination in vitro. Our previous study on human articular cartilage using 5–12 MHz transducer also revealed good correlation between in vitro US and histologic gradings. In the current study, we evaluated the validity of in vivo US grading of OA cartilage.

### Table I

**Distribution of in vivo US grades over various locations**

<table>
<thead>
<tr>
<th>Grade</th>
<th>M</th>
<th>L</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>22</td>
<td>27</td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td>33</td>
<td>49</td>
</tr>
<tr>
<td>4</td>
<td>17</td>
<td>10</td>
<td>27</td>
</tr>
<tr>
<td>5</td>
<td>27</td>
<td>12</td>
<td>39</td>
</tr>
<tr>
<td>6</td>
<td>19</td>
<td>7</td>
<td>26</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>84</strong></td>
<td><strong>88</strong></td>
<td><strong>172</strong></td>
</tr>
</tbody>
</table>

M: medial femoral condyle; L: lateral femoral condyle.

### Table II

**Distribution of histologic grades over various locations**

<table>
<thead>
<tr>
<th>Grade</th>
<th>Anterior area</th>
<th>Middle area</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>L</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>56</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>32</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>38</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>84</strong></td>
<td><strong>88</strong></td>
</tr>
</tbody>
</table>

M: specimen of medial femoral condyle; L: specimen of lateral femoral condyle.
and histologic gradings, which was presented in our previous study\(^\text{21}\). We also found that the correlation was better if we excluded subjects who could not flex their knee well. That means adequate knee flexion may be necessary to expose the weight bearing area of femoral condyle. Since anterior sub-weight bearing portion may be less involved in general condition, we may underestimate the severity of cartilage erosion by using US evaluation. However, US still has its value in serial follow-up treatment measurement by keeping the area evaluated to be the same.

The second problem is that overlying soft tissue may influence the appearance of the underlying cartilage band. Synovial fluid in patients with synovitis may impair the visualization of the synovial–cartilage interface\(^\text{20}\). Varying thickness of the overlying tissue may also verify the image echogenicity\(^\text{19}\). These factors also contributed the inconsistency of the US results on the femoral condyle cartilage. This problem partially explained the less validity of \textit{in vivo} US evaluation than \textit{in vitro} US evaluation.

The equipment and technique are also possible causes of error\(^\text{25}\). Throughout our study, we used the same US equipment with the same imaging parameter set. We standardized the scanning position and probe angulation. Operator dependence is always a problem of US examination. For validity evaluation in this study, only one investigator was used to perform the US examination to avoid confounding effect. We cautiously prevented these drawbacks.

The US grading is semiquantitative and relatively subjective, however, the inter-observer agreement was shown to be moderate in this study. In other series of studies, quantitative US was used which may provide more objective information\(^\text{31,32}\) and prevent the drawback of qualitative US. However, invasive arthroscopic ultrasound imaging instrument should be used and needs to be developed.

<table>
<thead>
<tr>
<th>Histologic grade</th>
<th>US grade</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0 (0)</td>
<td>18 (15)</td>
</tr>
<tr>
<td>2</td>
<td>2 (2)</td>
<td>2 (1)</td>
</tr>
<tr>
<td>3</td>
<td>2 (1)</td>
<td>6 (4)</td>
</tr>
<tr>
<td>4</td>
<td>0 (0)</td>
<td>1 (0)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>4 (3)</td>
<td>27 (20)</td>
</tr>
</tbody>
</table>

Numbers in parenthesis mean the numbers in the condition when only cases with knee flexion ROM equal to or more than 120 degrees were included.

Table IV

Relationship of \textit{in vivo} US grades of distal femoral condyle and histologic grades over middle area

<table>
<thead>
<tr>
<th>Histologic grade</th>
<th>US grade</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0 (0)</td>
<td>17 (13)</td>
</tr>
<tr>
<td>2</td>
<td>2 (1)</td>
<td>3 (3)</td>
</tr>
<tr>
<td>3</td>
<td>2 (2)</td>
<td>4 (3)</td>
</tr>
<tr>
<td>4</td>
<td>0 (0)</td>
<td>3 (1)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>4 (3)</td>
<td>27 (20)</td>
</tr>
</tbody>
</table>

Numbers in parenthesis mean the numbers in the condition when only cases with knee flexion ROM equal to or more than 120 degrees were included.
be done on animal cartilage. To investigate the histologic changes in the degeneration of human cartilage, data are commonly acquired from specimens from patients undergoing TKA. The possible drawback is that the overall severity would be great. Fortunately, the involvement of degeneration change over two compartments of tibiofemoral joint is usually not identical as shown in our study. In a patient with medial compartment severely involved, the lateral compartment may be only mildly involved. In our study, the specimens from lateral compartment constituted the samples with mild severity. Thus, we were able to investigate cartilage samples with mild to severe severity.

Another problem is that limited flexion angle is the common feature in patients with knee OA. Limited knee flexion causes poor exposure of weight bearing area of articular cartilage. The maximal angle of knee flexion ranged from 50 to 145 degrees in our study. About half of our subjects preserved the flexion angle. When we excluded the subjects with poor knee flexion (maximal angle of knee flexion less than 120 degrees), the correlation between in vivo US and histologic gradings became better. It implied that the angle of knee flexion did influence the correlation. However, the optimal angle for US examination is not yet known. We recommend to flex the knee at the same angle during serial follow-up treatment measurement.

In conclusion, US grading may provide clinical evaluation of osteoarthritic femoral condylar cartilage. We believe that some technical advancement in the US technology would in the future enable even better inspection of the cartilage percutaneously. Thus we could use US examination to detect early change in cartilage degeneration and take action earlier in clinical practice. We also could use US examination to assess the efficacy of treatment in both clinical practice and research field. Further studies should focus on the relationship with clinical presentation and also aim to improve the validity.

Acknowledgments

The authors wish to thank the Statistical Analysis Laboratory, Department of Clinical Research, Kaohsiung Medical University Hospital for all their help. The authors also thank Wan-Ting Huang, M.D., Department of Pathology, for evaluating histologic changes of the cartilage, and Ming-Wei, Chang, M.D., Department of Orthopedic Surgery, for specimen collection.

References


