Validity of Hepple Classification Used in the Diagnosis of Talus Osteochondral Lesions

Gokhan Ilyas, MD*
Fikri Burak IPCI, MD, PhD†

*Department of Orthopedics and Traumatology, Usak University Faculty of Medicine, Usak, Turkey.
†Department of Orthopedics and Traumatology, Banaz State Hospital, Usak, Turkey. Dr. Burak is now with Department of Orthopedics and Traumatology, Usak University Faculty of Medicine, Usak, Turkey.

Corresponding author: Gokhan Ilyas, MD, Department of Orthopedics and Traumatology, Usak University Faculty of Medicine, Uşak Eğitim ve Araştırma Hastanesi, Usak, Turkey. (E-mail: gokhan.ilyas@usak.edu.tr)

Background: It was aimed to investigate the intra-observer and inter-observer validity of the Hepple classification used in talus osteochondral lesions.

Methods: This study included 32 patients with osteochondral lesions in the talus after exclusion criteria. A PowerPoint presentation was prepared from the MRI views of the patients. Six observers, divided into two groups according to their experience, were asked to categorize the cases according to Hepple classification. The slices were shuffled and the observers were asked
to reevaluate after 6 weeks. Fleiss kappa (κ) coefficient was used for the inter-observer validity and Cohen’s kappa (κ) coefficient for the intra-observer validity.

**Results:** In the overall inter-observer reliability was at a moderate level of agreement (set one κ=0.511, set two κ=0.406). In the intra-observer evaluation, one observer from the experienced group showed almost perfect agreement (κ=0.809), one observer from the less experienced group had moderate agreement (κ=0.556), and all other observers had substantial agreement (κ=0.556 – 0.730). When all observers were examined, it was seen that there was a substantial agreement in the mean intra-observer evaluation (κ=0.661).

**Conclusions:** While the intra-observer results showed substantial agreement, the inter-observer results showed moderate agreement. Although the Hepple classification system is frequently used, the need for a more reliable classification system for osteochondral lesions of the talus remains.

Osteochondral lesions may result from a traumatic cause or repetitive microtrauma. It starts with cartilage lesions and continues with variable involvement of the subchondral bone. It has been reported that the developing pain is not caused by a cartilage lesion but develops because of repetitive high fluid pressure during walking, stimulating the subchondral bone.¹ Conventional x-ray is preferred in the first stage in the diagnosis of talus cartilage lesions. Magnetic resonance imaging (MRI) is used to make a definitive diagnosis and to determine the size of the lesion, the level of bone marrow edema, and whether it is accompanied by a bone lesion. Accurate determination of the stage of the disease is important in the selection and success of treatment.

Many imaging methods are used in the classification of talus osteochondral lesions; Berndt and Harty’s classification² is used in the radiological evaluation, and Ferkel and Sgaglione’s classification³ is used in the evaluation with computed tomography (CT). Hepple classification⁴ is used in evaluation with MRI. To our knowledge, the validity of the Hepple classification, which is frequently used in the clinic, has not been evaluated. It was hypothesized that the Hepple classification would be reproducible with high confidence in osteochondral lesions of the talus.

The purpose of this study was to investigate the intra-observer and inter-observer validity of the Hepple classification used in talus osteochondral lesions.

Materials and Methods

The MRI images of thirty-two patients with talus osteochondral lesions were included in this study after the approval of the local ethical committee (06.04.2022-64/64/08).

Between January 2021 and January 2022 247 patients who applied to our clinic and underwent ankle MRI were scanned retrospectively. All MRI scans were done with 1.5 tesla (Siemens Magnetom Aera, Erlangen, Germany). Exclusion criteria were; patients without talus chondral lesion (n=195), without optimal MRI (n=7), age under 18 years old (n=2), and over 65 years old (n=11). MRI images (all axial, coronal, and sagittal sections) of 32 (13%) patients whose osteochondral lesions were confirmed and included in the study were obtained from Picture Archiving and Communication Systems (PACS) recordings.

A PowerPoint presentation was prepared from MRI views containing sagittal, coronal, and axial sections of 32 skeletally mature patients of both genders who did not meet the exclusion
criteria (Fig 1). Hepple classification was presented to the observers after MRI views of each patient (Table 1). They were asked to classify the MRI views according to Hepple classification in one week.

Six orthopedic specialists with at least three years of experience participated in the study as observers. The observers were divided into two groups according to their experience, in the first group there were three people with at least six years of orthopedic specialists (7, 8, 14 years respectively), and in the second group, there were three people with less experience (4, 5, 5 years respectively). Each observer made separate evaluations from the others. Observers evaluated the MR images of 32 patients who were relocated at 6-week intervals, without accessing their previous responses. The order of the slides was changed in the second presentation by an orthopedic surgeon who was not included in the study. Intra-observer and inter-observer validity was statistically evaluated.

SPSS version 24 (IBM Corp., Armonk, New York, USA) was used for statistical analysis. Fleiss kappa (κ) coefficient was used for the inter-observer validity and Cohen kappa coefficient for the intra-observer validity. A κ value is always between 0 and 1; the κ values were graded as slight (0–0.2), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80), and almost perfect (0.81–1).

Results

Seventeen (53%) of the patients whose MRI views were evaluated were female and their mean age was 43.5±14 (22-62) years. The mean age of men was 39.3±11.2 (24-60) years. Nineteen (59%) of the injuries were at the left ankle.
In the inter-observer evaluation, all results, except the second evaluation of the less experienced group, resulted in a moderate agreement ($\kappa=0.440 - 0.590$). The second assessment in the less experienced group showed fair agreement ($\kappa=0.239$). When all observers were evaluated, it was seen that there was a moderate agreement (set one $\kappa=0.511$, set two $\kappa=0.406$) (Table 2).

In the intra-observer reliability assessment, one observer from the experienced group showed almost perfect agreement ($\kappa=0.809$), one observer from the less experienced group showed moderate agreement ($\kappa=0.556$), and all other observers showed substantial agreement ($\kappa=0.556 - 0.730$) (Table 3).

In two assessments, six-week intervals, the intra-observer mean values resulted in a moderate agreement in the less experienced group ($\kappa=0.580$) and substantial agreement in the experienced group ($\kappa=0.742$). When all observers were examined, it was seen that there was a substantial agreement ($\kappa=0.661$) (Table 4).

Discussion

A classification system must be valid, reliable, and reproducible. It should be expressed alike by both surgeons and researchers. Cartilage covers up almost 70% of the circumference of the talus bone. Since the articular cartilage does not have a vascular and nerve structure, the healing potential of developing cartilage lesions is low. Various classification systems are used in talus osteochondral lesions. The classification system based on MRI, which is frequently used in clinical diagnosis, was defined by Hepple. Accurate staging of cartilage lesions is important for correct treatment. Our study showed that the average intra-observer reliability of
the Hepple classification in the diagnosis of talus cartilage lesions was substantial agreement ($\kappa=0.661$), while the inter-observer reliability was moderate ($\kappa=0.511 - 0.406$).

Bexkens et al.\textsuperscript{9} in a study of patients with capitellar osteochondritis dissecans, asked observers to evaluate MRI images according to the Hepple, Dipaola/Nelson, Itsubo classifications, and the DeSmet/Kijowski and Satake criteria for lesion instability. Although the inter-observer evaluation of the Hepple classification showed higher agreement ($\kappa=0.23$) compared to other classifications ($\kappa=0.12 - 0.19$), it was found to be fair agreement ($\kappa=0.20 - 0.40$). However, other classifications appear to be in slight agreement ($\kappa=0.0 - 0.20$). In our study in which talus osteochondritis dissecans were evaluated, a moderate agreement was observed in the inter-observer results ($\kappa=0.406 - 0.511$). However, the second evaluation in the less experienced group resulted in fair agreement, similar to this study ($\kappa=0.239$). In our study, substantial agreement results were obtained in the intra-observer evaluation performed in addition to this study ($\kappa=0.661$).

Tamam et al.\textsuperscript{10} compared MRI with Single-photon emission computed tomography combined with computed tomography (SPECT/CT) in image interpretation and treatment decisions in talus osteochondral lesions. In the study, which included arthroscopically validated cases, the Hart score system was used for SPECT/CT evaluation and Hepple classification was used for MRI evaluation. It was reported that there was no statistically significant difference result between MR and SPECT/CT ($p=0.16$). However, a statistically significant relationship was determined between MRI and SPECT/CT ($r=0.47$, $p=0.018$). According to the study’s conclusions, although MRI is the standard in diagnosing osteochondral lesions of the talus, it is suggested that the combined use of SPECT/CT and MRI is appropriate. However, considering the negative effects of tomography, we recommend using MRI as the first choice in diagnosis. We believe that it is a method that can be preferred in cases where MRI cannot be used.
Many methods have been described in the treatment of osteochondral lesions of the talus, from immobilization to osteochondral autograft transplantation.\textsuperscript{11-18} As in other cartilage lesions, treatment options in talus osteochondral lesions vary according to the degree of the lesion. Accurate knowledge of the degree of cartilage injury will guide the treatment. Therefore, there is a need for classification systems that are evaluated similarly by each observer. MRI is frequently used in the diagnosis and grading of cartilage lesions. For this reason, we planned the validity analysis of the Hepple classification used in the evaluation of talus cartilage lesions with MRI. According to the results of our study, which showed substantial agreement in the intra-observer evaluation and moderate agreement in the inter-observer evaluation, we concluded that the Hepple classification is reproducible and reliable.

Staats et al.\textsuperscript{19} demonstrated in an arthroscopically validated study in patients with chronic ankle instability that magnetic resonance imaging was helpful in preoperative diagnosis, but not definitively sufficient. Although there is no definite consistency, we think that MRI is the most appropriate imaging method in clinical practice.

Claessen et al.\textsuperscript{20} evaluated the interobserver agreement of Berndt and Harty’s classification and Ferkel and Sgaglione with other classifications in capitellar osteochondritis dissecans lesions. According to the results of the study, a poor agreement was concluded for Berndt and Harty’s classification ($\kappa=0.20$). In the same study, a fair agreement was concluded for Ferkel and Sgaglione classification ($\kappa=0.22$). Inter-observer agreement of Hepple classification was moderate in our study, which we applied to talus cartilage lesions ($\kappa=0.511 - 0.40G$).

It was stated that any $k$ value between 0.40-0.60 indicates moderate agreement among the raters.\textsuperscript{21} Our study showed that interobserver reliability was moderate ($\kappa=0.511 - 0.406$).
The limitation of this study was that all classifications used in talus osteochondral lesions could be evaluated together. However, since MRI is used in the diagnosis of cartilage lesions in our clinical practice, CT images of the same patients were not available.

Conclusion

While the intra-observer results showed substantial agreement, the inter-observer results showed moderate agreement. Although the Hepple classification system is frequently used, the need for a more reliable classification system for osteochondral lesions of the talus remains. In addition, it was determined that surgical experience gave better reliability results.

Financial Disclosure: None reported.
Conflict of Interest: None reported.

References


7 Altman DG. Practical Statistics for Medical Research. Taylor & Francis; 1990.


10 Tamam C, Tamam MO, Yildirim D, Mulazimoglu M. Diagnostic value of single-photon emission computed tomography combined with computed tomography in relation to MRI on


Table 1. Hepple Classification

<table>
<thead>
<tr>
<th>Stage</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Articular cartilage injury</td>
</tr>
<tr>
<td>IIA</td>
<td>Cartilage injury, bony fracture with edema</td>
</tr>
<tr>
<td>IIB</td>
<td>Cartilage injury, bony fracture without edema</td>
</tr>
<tr>
<td>III</td>
<td>Detached, nondisplaced bony fragment</td>
</tr>
<tr>
<td>IV</td>
<td>Displaced fragment</td>
</tr>
<tr>
<td>V</td>
<td>Subchondral cyst</td>
</tr>
</tbody>
</table>

Table 2: Inter-observer reliability values

<table>
<thead>
<tr>
<th></th>
<th>Set 1</th>
<th>Set 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fleiss $\kappa$ (95% CI)</td>
<td>Interpretation</td>
</tr>
<tr>
<td>Experienced (n=3)</td>
<td>0.563</td>
<td>Moderate agr.</td>
</tr>
<tr>
<td>Less experienced (n=3)</td>
<td>0.440</td>
<td>Moderate agr.</td>
</tr>
<tr>
<td>Total (n=6)</td>
<td>0.511</td>
<td>Moderate agr.</td>
</tr>
</tbody>
</table>

*agr: agreement, CI: confidence interval

Table 3: Intra-observer reliability values

<table>
<thead>
<tr>
<th>Observer</th>
<th>Cohen $\kappa$ (95% CI)</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.809</td>
<td>Almost perfect agr.</td>
</tr>
<tr>
<td>2</td>
<td>0.730</td>
<td>Substantial agr.</td>
</tr>
<tr>
<td>3</td>
<td>0.687</td>
<td>Substantial agr.</td>
</tr>
<tr>
<td>4</td>
<td>0.556</td>
<td>Moderate agr.</td>
</tr>
<tr>
<td>5</td>
<td>0.620</td>
<td>Substantial agr.</td>
</tr>
<tr>
<td>6</td>
<td>0.564</td>
<td>Moderate agr.</td>
</tr>
</tbody>
</table>

*agr: agreement, CI: confidence interval

Table 4: Mean intra-observer reliability values

<table>
<thead>
<tr>
<th></th>
<th>Cohen $\kappa$ (95% CI)</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experienced (n=3)</td>
<td>0.742</td>
<td>Substantial agr.</td>
</tr>
<tr>
<td>Less experienced (n=3)</td>
<td>0.580</td>
<td>Moderate agr.</td>
</tr>
</tbody>
</table>
This Original Article is a preprint. It has been reviewed, accepted for publication, and approved by the author but has not been copyedited, proofread, or typeset.

<table>
<thead>
<tr>
<th>Total (n=6)</th>
<th>0.661</th>
<th>Substantial agr.</th>
</tr>
</thead>
</table>

*agr: agreement, CI: confidence interval

Figure 1: An example of a slide