Posterior malleolar fractures

Diagnostic accuracy, morphology and clinical outcome

Meijer, D.T.

Citation for published version (APA):
ABSTRACT
Background
Up to 44% of ankle fractures have involvement of the posterior tibial margin. Fracture size and morphology are important factors to guide treatment of these fragments, but reliability of plain radiography in estimating size is low. The aim of the current study was to evaluate the accuracy of 2-dimensional computed tomography (2DCT) in the assessment of posterior malleolar fractures. Additionally, the diagnostic accuracy of 2DCT and its value in preoperative planning was evaluated.

Methods
Thirty-one patients with 31 ankle fractures including a posterior malleolar fragment were selected. Preoperative CT scans were analyzed by 50 observers from 23 countries. Quantitative 3-dimensional CT (Q3DCT) reconstructions were used as a reference standard.

Results
Articular involvement of the posterior fragment was overestimated on 2DCT by factors 1.6, 1.4, and 2.2 for Haraguchi types I, II, and III, respectively. Interobserver agreement on operative management (“to fix, or not to fix?”) was substantial (κ = 0.69) for Haraguchi type I fractures, fair (κ = 0.23) for type II fractures, and poor (κ = 0.09) for type III fractures. 2DCT images led to a change in treatment of the posterior malleolus in 23% of all fractures. Surgeons would operatively treat type I fractures in 63%, type II fractures in 67%, and type III fractures in 22%.

Conclusion
Surgeons overestimated true articular involvement of posterior malleolar fractures on 2DCT scans. 2DCT showed some additional value in estimating the involved articular surface when compared to plain radiographs; however, this seemed not yet sufficient to accurately read the fractures. Analysis of the CT images showed a significant influence on choice of treatment in 23% with a shift toward operative treatment in 12% of cases compared to evaluating plain lateral radiographs alone.

INTRODUCTION
Approximately 7% to 44% of ankle fractures involve the posterior tibial margin. These fractures tend to have a poorer prognosis than fractures without posterior involvement. The decision whether or not to address the posterior fragment during surgery is a subject of ongoing debate, and practice varies among surgeons. In the current literature, there seems to be a consensus that a posterior fragment that comprises more than 25% to 33% of the tibial plafond requires fixation. However, reliability of these measurements has proven to be questionable. In a previous study, we found that surgeons overestimate articular involvement of the posterior malleolar fracture on plain radiographs. Apart from size, morphology of the posterior malleolar fragment might be more important. Haraguchi and colleagues classified posterior malleolar fractures into 3 types based on their morphology: type I fractures are described as a triangular fragment of the posterolateral corner of the tibial plafond; type II fractures have extension of the fracture line into the medial malleolus; and type III fractures involve smaller shell-shaped fragments at the posterolateral lip of the tibial plafond. Knowledge of the characteristics of posterior malleolar fragments will contribute to the general understanding of ankle fracture patterns. Hence, the relevance to address these posterior fragments operatively should be based on morphology and size instead of mere fragment size alone. Two-dimensional computed tomography (2DCT) is expected to enhance surgeons’ ability to estimate the morphology and size of the posterior malleolar fragment. However, data on the accuracy and reliability of 2DCT in the assessment of posterior malleolar fracture characteristics are scarce. Quantification of 3-dimensional computed tomography (Q3DCT) modeling has proven a useful technique in evaluating fracture morphology.

The aim of the current study was primarily to evaluate the accuracy of 2-dimensional computed tomography (2DCT) in the assessment of posterior malleolar fractures. In addition, we assessed the value of 2DCT imaging in preoperative planning and assessed the variability in surgeons’ management of these fractures and compared the diagnostic accuracy of 2DCT to the accuracy of plain radiographs.

METHODS
This study was approved by our institutional review board.

Subjects
Thirty-one patients with 31 ankle fractures involving a posterior malleolar fragment (OTA type 44) were selected, based on an equal distribution of Haraguchi types I to III. All patients were treated in a level III trauma center (Sint Lucas Andreas Hospital, Amsterdam, the Netherlands) between March 2005 and December 2012 and had both preoperative plain radiographs and CT scans of the injured joint.

Computed tomography
All preoperative CT scans were acquired with 1 of 2 different systems (Toshiba Aquilion 4 Slice, Toshiba Medical Systems Cooperation, Tokyo, Japan, or GE Discovery ct750 HD, GE Healthcare,
Reliability of this technique has been assessed in 17, 18, 43. The Diagnostic Accuracy of 2-D Computed Tomography for Articular Involvement and Fracture Pattern of PMF.

Observers
This study was performed as part of the Ankleplatform Study Collaborative—Science of Variation Group. 24, 27 One hundred one independent orthopaedic surgeons were invited to take part in this study. Observers were assigned to review 31 DCT scans of ankle fractures involving the posterior malleolus using an online DICOM viewer. The following 5 questions were asked of each case:

1. What is the involved articular surface of the posterior malleolar fracture as a percentage of the complete tibial plafond articular surface? (in %, open question);
2. According to the Haraguchi classification, which type of posterior malleolar fragment is seen in this patient? (type I, type II or type III);
3. Are the CT images a valuable contribution to preoperative planning? (yes/no);
4. Based on the CT images, would you operatively address the posterior malleolar fragment? (yes/no); and
5. If yes, what would your operative approach be? (multiple-choice: A, anterior; B, posterolateral; C, posteromedial; or D, posterolateral and posteromedial approach)

Statistical analysis
Statistical analysis was performed with SPSS 20.0 for Windows (IBM Corp, Armonk, NY). Data were normally distributed and measurements are presented as means with standard deviations (SDs). For diagnostic accuracy assessment, the average value of the 50 observers was used to describe the difference between the observations (on 2DCT) and the reference standard (Q3DCT). Paired t tests were performed to test the differences for the entire group, and the 3 types of fractures separately. A p value less than .05 was considered to be statistically significant. Assessment of precision of measurements was performed by calculation of the interclass correlation coefficient (ICCagreement), interobserver agreement regarding the Haraguchi classification, and the decision to operate was determined by calculating the kappa value. Both ICC and kappa value were interpreted according to the categorical rating of Landis and Koch: slight agreement, 0.00-0.20; fair agreement, 0.21-0.40; moderate agreement, 0.41-0.60; substantial agreement, 0.61-0.80; and almost perfect agreement, greater than 0.81, with 1.00 being the highest obtainable value. 8 The standard error of measurement (SEM) was used to calculate the smallest detectable difference (SDD) between the observers.

Comparison with plain radiographs
In a previous study, we used the same group of observers and the same cohort of ankle fractures to assess the accuracy of plain radiographs in estimating articular involvement of posterior malleolar fractures. 27 Also, we evaluated management decisions based on plain radiographs. Apart from operative approach and assessment of Haraguchi type, which were not assessed, all questions were identical to the current study. Hence, we were able to compare diagnostic performance characteristics (sensitivity, specificity, and accuracy) and effect on treatment decisions of plain radiographs and 2DCT imaging. To compare these results, we matched identical observers of both study groups.

RESULTS
Of the 201 surgeons invited, 50 surgeons from 23 countries responded and evaluated all images. For characteristics of the participating observers, see Table 1. All participating observers evaluated the complete series of 31 fractures and answered all 5 questions if applicable.

FIGURE 1. On the left, a Haraguchi type I fracture of the posterior malleolus, with a triangular fragment, comprising only the posterolateral corner. In the middle, a Haraguchi type II fracture, with extension of the fracture into the posteromedial corner. Sometimes, there is an extension into the medial malleolar fracture. Mostly type III fractures consist of 2 fragments: posterolateral and posteromedial (posterior colliculus of medial malleolus). On the right, a Haraguchi type III fracture is seen, with small shell-shaped fragments at the posterior rim.

Fairfield, CT, United States) with a maximal slice thickness of 1 mm. Reconstructions in 3 (sagittal, coronal, and axial) planes were available for review. There were no 3D reconstructions available for review.

Quantitative 3-Dimensional computed tomographic modeling
We used Q3DCT modeling techniques as a reference standard to quantify fragment size and morphology. 22, 24-27 Reliability of this technique has been assessed in a separate study. 25 To create reconstructions, sagittal images of CT scans (DICOM files; Digital Imaging and Communications in Medicine) were analyzed with an algorithm that identified the outer margin of highest density (cortical or subchondral) bone using Matlab (version 8.0; Mathworks, Natick, MA). These outlines were then stacked using Rhinoceros (version 4.0; McNeel North America, Seattle, WA), creating a wire mesh representing the outer margin of the bone. This wire mesh model was then transformed into a polygon mesh, a hollow 3-D model of the outer surface of the bone. Fracture fragments with articular surface attached were then identified and isolated for analysis. Articular surface area of the posterior fragments was reported as a percentage of articular surface of the complete tibial plafond. In case of multiple posterior fragments, the surface areas of the separate fragments were combined. Fracture patterns were analyzed at the level of the tibial plafond and categorized according to the Haraguchi classification. 18

<table>
<thead>
<tr>
<th>Stage</th>
<th>Event</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Review</td>
<td>31 images of ankle fractures involving the posterior malleolus were reviewed.</td>
</tr>
<tr>
<td>2</td>
<td>Question</td>
<td>Observers were asked to determine the involved articular surface of the posterior malleolar fracture.</td>
</tr>
<tr>
<td>3</td>
<td>Agreement</td>
<td>The observers’ responses were compared using the kappa coefficient.</td>
</tr>
<tr>
<td>4</td>
<td>Decision</td>
<td>The decision to operate was determined by calculating the kappa value.</td>
</tr>
<tr>
<td>5</td>
<td>Optimization</td>
<td>The smallest detectable difference (SDD) between the observers was calculated.</td>
</tr>
</tbody>
</table>
According to the reference standard Q3DCT, the mean posterior malleolar fragment involved 14% (SD = 10.8) of the tibial plafond articular surface. The mean articular involvement of the posterior malleolar fracture as measured by 50 observers on the 2DCT images was found to be 22% (SD = 10.39). This difference of 9% (95% CI = 6.4, 10.8) was statistically significant (p < .001). Haraguchi type 1 fractures involved 16% (SD = 13.0) of the articular surface according to the reference standard Q3DCT, compared to 27% (SD = 13.1) estimated by the observers on the 2DCT images. This difference of 11% (95% CI = 8.1, 12.9) was significant (p < .001). Haraguchi type 2 fractures involved 18% (SD = 10.1) on Q3DCT, compared to 26% (SD = 6.49) on the 2DCT images. This difference of 8% (95% CI = 1.3, 14.7) was significant (p = .024). Haraguchi type 3 fractures involved 7% (SD = 4.7) of articular surface, compared to 14% (SD = 4.6) according to observers evaluating 2DCT images. This difference of 8% (95% CI = 4.8, 9.97) was significant (p < .001). These overestimations compare to a factor 1.6, 1.4, and 2.2 for Haraguchi types I, II, and III, respectively.

Diagnostic performance characteristics of 2DCT

2DCT showed an accuracy of 0.74 with a sensitivity of 0.77 and a specificity of 0.77 for Haraguchi type I fractures. For type II fractures, accuracy was 0.79 with a sensitivity of 0.65 and a specificity of 0.86. For type III fractures, accuracy was 0.68 with a sensitivity of 0.44 and a specificity of 0.80. See Table 2 for 95% confidence intervals.

The diagnostic accuracy of 2DCT for posterior malleolar fragment size depended on cut-off values chosen. Within limits ranging 5% below and above the reference standard value, accuracy was 30%.

Reliability of measurements

Within the group of 50 observers, the intraclass correlation coefficient (ICC) of the fragment size measurement for all fractures was 0.57 (95% CI = 0.45, 0.70). For Haraguchi type I fractures the ICC was 0.69 (95% CI = 0.50, 0.88), for Haraguchi type II fractures the ICC was 0.26 (95% CI = 0.13, 0.35) and for type III fractures the ICC 0.27 (95% CI = 0.15, 0.54). The standard error of measurement (SEM) for all fracture types was 9% with a smallest detectable difference (SDD) of 25%. The SEM for Haraguchi type I fractures was 9% with an SDD of 24%. The SEM for Haraguchi type II fractures was 11% with an SDD of 30% and Haraguchi type III fractures had a SEM of 7% with an SDD of 21%. Kappa value for the Haraguchi classification was 0.27 (95% CI = 0.0, 0.54) with an absolute agreement between the observers of 53%.

Operative management of posterior malleolar fractures

None of the observers would operatively address all fractures; neither would an observer treat all fractures conservatively.

Table 2. Sensitivity, specificity and accuracy of 2DCT

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity</th>
<th>95% CI</th>
<th>Specificity</th>
<th>95% CI</th>
<th>Accuracy</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haraguchi type I</td>
<td>0.77</td>
<td>0.71 – 0.83</td>
<td>0.72</td>
<td>0.66 – 0.75</td>
<td>0.74</td>
<td>0.71 – 0.76</td>
</tr>
<tr>
<td>Haraguchi type II</td>
<td>0.65</td>
<td>0.58 – 0.71</td>
<td>0.86</td>
<td>0.83 – 0.88</td>
<td>0.79</td>
<td>0.76 – 0.83</td>
</tr>
<tr>
<td>Haraguchi type III</td>
<td>0.44</td>
<td>0.37 – 0.50</td>
<td>0.80</td>
<td>0.76 – 0.84</td>
<td>0.68</td>
<td>0.64 – 0.71</td>
</tr>
</tbody>
</table>
From all the fractures presented, the majority of observers (consensus agreement) would operatively address 50%. The observer most leaning toward operative treatment would fix 87% of posterior fragments, the most conservative observer would fix 19%. The fracture least operated on (1 observer) was one of the 3 fractures that comprised 0% of the articular surface, a Haraguchi type III avulsion. There were 2 fractures that 100% of the observers would fix, involving 11% and 34% of the articular surface, respectively (according to the reference standard Q3DCT; Figure 2). The fractures most observers would fix were not the largest. Overall, 88% of surgeons would operatively address fractures involving >25% of the articular surface. Fractures involving >15% would be addressed by 85%, fractures involving >10% by 74%, and 14% of the observers would operate fractures of <5% of the articular surface (as measured on Q3DCT). See Table 3 for the management of fractures per Haraguchi subtype.

Out of the 50 observers, 85% found the 2DCT scan to be of added value in the preoperative planning of Haraguchi type II fractures. For Haraguchi type I and III fractures, the 2DCT scan was considered valuable by 62% and 54% of the observers, respectively.

When the surgeon aimed to address the posterior malleolus, 68% and 65% of observers preferred the posterolateral approach for Haraguchi types I and III fracture, respectively. For Haraguchi type II fractures, there was less consensus about the approach; 37% of observers preferred the posterolateral approach, 26% the posteromedial approach, 21% the posterolateral and posteromedial approaches, and 16% the anterior approach.

Comparison with plain radiographs

The answers of the current selection of 50 observers were extracted from the original plain radiography database and compared to the current results. All 50 observers had answered all questions in the evaluation of plain radiography.

The mean articular surface measured on plain radiography was 2.3% higher than on 2DCT (95% CI = −0.1, 4.8, p = .06). The differences in surface area of fracture fragment between plain radiography with 2DCT per Haraguchi type were 2.2% (95% CI = −1.9, 6.4, p = .26), 0.6% (95% CI = −4.0, 5.2, p = .79), and 4.1% (95% CI = −11, 9.2, p = .11) for Haraguchi types I, II, and III, respectively.

In 23% of cases, treatment of the posterior fragment was changed after reviewing the 2DCT images (interobserver agreement; kappa = 0.54). For shifts in treatment, see Table 4.

DISCUSSION

Our study suggests that surgeons systematically overestimate true articular involvement of posterior malleolar fractures on 2DCT scans. Involved articular surface was overestimated in all Haraguchi subtypes with factor 1.6 (type I), factor 1.4 (type II), and factor 2.2 (type III), respectively. Although 2DCT seemed insufficient to accurately read the fractures, this technique would lead toward a more appropriate treatment when compared to plain radiography alone.

There was higher agreement in estimating the fracture surface of type I fractures (substantial agreement) than type II and III fractures (fair agreement). The fair agreement and high overestimation of type III suggests it is difficult to estimate the involved surface correctly.

<table>
<thead>
<tr>
<th>Fractures surgically addressed (%)</th>
<th>Kappa</th>
<th>95% CI</th>
<th>Absolute Agreement (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haraguchi type I</td>
<td>63</td>
<td>0.69</td>
<td>0.43 – 0.96</td>
</tr>
<tr>
<td>Haraguchi type II</td>
<td>67</td>
<td>0.23</td>
<td>−0.24 – 0.60</td>
</tr>
<tr>
<td>Haraguchi type III</td>
<td>22</td>
<td>0.09</td>
<td>−0.40 – 0.58</td>
</tr>
<tr>
<td>All types</td>
<td>50</td>
<td>0.47</td>
<td>0.16 – 0.78</td>
</tr>
</tbody>
</table>

FIGURE 2. (A) Posterior malleolar fracture with 11% articular involvement that 100% of observers would fix. (B) Posterior malleolar fracture with 34% articular involvement that 100% of observers would fix.
in these shell-shaped fragments of the posterior lip. This is exacerbated by the fact that in general, the smaller the proportion of involved articular surface, the larger the relative over- or underestimation is. Thus, the clinical importance of involved articular surface in type III fractures is likely to be limited, keeping in mind that these fragments are thought to be avulsion fractures that indicate associated posterior syndesmotic injury.

CT evaluation was found to be valuable in 62%, 85%, and 54% for Haraguchi types I, II, and III, respectively. Analysis of the CT images showed a significant influence on choice of treatment in 23%; a shift towards operative treatment was seen in 12% of cases compared to evaluating plain lateral radiographs alone. This suggests CT evaluation would not merely economize on operative treatment; it would enable us to do a better job at selecting the right patients. In line with Büchler and colleagues, we recommend preoperative CT evaluation in all patients with trimalleolar fractures.21

Fragments most observers would operatively address were not the largest. This confirms there are more factors that guide treatment than mere fracture size, as mentioned by Büchler and Gardner et al.16,17 There were 2 fractures that 100% of the observers would fix, involving 11% and 34% of the articular surface, respectively (Figure 2). The first case involving 11% of articular surface was a type II fracture that involved multiple severely displaced fragments creating a large intra-articular gap. The case involving 34% of articular surface was a typical displaced type I fracture that left a large gap in the tibiofibular joint. If surgeons decided to address the posterior fragment operatively, the posterolateral approach was preferred when confronted with a Haraguchi type I (68%) and type III fracture (65%). In recent literature, interest in this posterolateral approach of posterior fragments has seemingly been growing among orthopaedic surgeons, because of its easy visualization and excellent outcomes with a low complication rate.22-24 Additionally, this approach offers more direct access to the posterolateral corner of the tibial plafond, which is specifically affected in Haraguchi type I fractures.25

For Haraguchi type II fractures, there was less agreement about the approach, but the medial involvement guided the surgeon toward a (partial) medial approach in almost half the cases (47%). However, 37% of observers preferred the posterolateral approach and 16% the anterior approach, through which posteromedial fragments cannot be accurately reduced and fixed. A significant amount of the Haraguchi type II fractures are indeed recognized with 2DCT, but would be inadequately treated.

Strengths of our study include the fact that we used 50 observers from 23 countries to assess posterior malleolar fractures. This group of observers provided us with the best possible representation of the current treatment standards worldwide. Furthermore, to our knowledge, we are the first to examine the accuracy of 2DCT in the evaluation of posterior malleolar fractures by comparing it to Q3DCT as a reference standard.26 The overestimation of involved articular surface is clinically important and adds to our understanding of posterior malleolar fractures.

Limitations of this study include that we have included the articular surface of the medial malleolus in the calculation of articular surface on Q3DCT. Although this technique is indispensable in Haraguchi type II fractures as the medial malleolus itself is part of the posterior fracture, this might partially explain the overestimation in type I and III fractures. Nonetheless, even in type II fractures, there was an overestimation with factor 1.4 which cannot be accounted for by our choice of technique.

Additionally, we acknowledge the possible discrepancy between what CT images found as articular surface and what was true articular surface. As CT images do not show cartilage, the articular surface measured by 2DCT or Q3DCT reconstructions might differ from the actual articular surface. Using proportions to describe the involved surface might have minimized the impact of this effect. Nonetheless, additional studies using magnetic resonance imaging (MRI) or cadaveric bone could point out the degree of over- or underestimation.27,28,29

Finally, as mentioned by Gardner and colleagues, there are other factors such as comminution and impaction of fragments that are important in guiding operative management.30 To optimize the feasibility of this study and limit the workload for our observers, we focused on the estimation of fracture size and pattern and did not address these other factors. This makes it difficult to assess treatment consensus based on Haraguchi type alone. Future studies should take all fracture characteristics into account.

This study showed that observers inaccurately interpret 2DCT data. Studies have shown that extracting 3-dimensional data from 2-dimensional images is a difficult task for the human brain.31 Although quantification of 3DCT data is yet too laborious for use in clinical practice, 3DCT images might assist in further improving the accuracy of human estimation of the degree of articular involvement.

As the attention for operative fixation of posterior fragments grows, we would recommend a prospective follow-up study to further elucidate the clinical relevance of 3D pathoanatomy of posterior malleolar fractures. With more focus on fracture morphology, we may eventually focus less on fracture size. Further understanding of clinical behavior of these fracture types might eventually have a positive effect on the consensus for treatment of posterior malleolar fractures.32

### TABLE 4. Direction of treatment shift after evaluation of 2DCT images

<table>
<thead>
<tr>
<th>Shift in treatment (%)</th>
<th>Kappa</th>
<th>Nonoperative (\rightarrow) operative (%)</th>
<th>Operative (\rightarrow) Nonoperative (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haraguchi type 1</td>
<td>11</td>
<td>0.77</td>
<td>8</td>
</tr>
<tr>
<td>Haraguchi type 2</td>
<td>33</td>
<td>0.29</td>
<td>20</td>
</tr>
<tr>
<td>Haraguchi type 3</td>
<td>25</td>
<td>0.32</td>
<td>10</td>
</tr>
<tr>
<td>All types</td>
<td>23</td>
<td>0.54</td>
<td>12</td>
</tr>
</tbody>
</table>

2

DIAGNOSTIC ACCURACY OF 2-D COMPUTED TOMOGRAPHY FOR ARTICULAR INVOLVEMENT AND FRACTURE PATTERN OF PMF
CONCLUSION

Surgeons overestimated true articular involvement of posterior malleolar fractures on 2DCT scans. There was a wide variety in treatment decisions to manage posterior malleolar fractures, and interobserver agreement on management varied greatly per fracture type. Although 2DCT seemed insufficient to accurately assess posterior malleolar fractures, this technique would possibly lead to more appropriate treatment when compared to plain radiography alone. Analysis of the CT images showed a significant influence on choice of treatment in 23% with a shift toward operative treatment in 12% of cases compared to evaluating plain lateral radiographs alone.

REFERENCES


