Posterior malleolar fractures

Diagnostic accuracy, morphology and clinical outcome

Meijer, D.T.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Download date: 10 Jan 2020
GUESSTIMATION OF POSTERIOR MALLEOLAR FRACTURES ON LATERAL PLAIN RADIOGRAPHS

D.T. Meijer
J.N. Doornberg
I.N. Sierevelt
W.H. Mallee
C.N. van Dijk
G.M.M.J. Kerkhoffs
S.A. Stufkens

Ankle Platform Study Collaborative – Science of Variation Group

Injury, July 2015
ABSTRACT

Background
Accurate assessment of articular involvement of the posterior malleolar fracture fragments in ankle fractures is essential, as this is the leading argument for internal fixation. The purpose of this study is to assess diagnostic accuracy of measurements on plain lateral radiographs.

Methods
Quantification of three-dimensional computed tomography (Q3DCT) was used as a reference standard for true articular involvement (mm²) of posterior malleolar fractures. One-hundred Orthopaedic Trauma surgeons were willing to review 31 trimalleolar ankle fractures to estimate size of posterior malleolus and answer: (1) what is the involved articular surface of the posterior malleolar fracture as a percentage of the tibial plafond? and (2) would you fix the posterior malleolus?

Results
The average posterior malleolar fragment involved 13.5% (SD 10.8) of the tibial plafond articular surface, as quantified using Q3DCT. The average involvement of articular surface of the posterior malleolar fragment, as estimated by 100 observers on plain radiographs was 24.4% (SD 10.0). The factor 1.8 overestimation of articular involvement was statistically significant (p < 0.001). Diagnostic accuracy of measurements on plain lateral radiographs was 22%. Interobserver agreement (ICC) was 0.61. Agreement on operative fixation, showed an ICC of 0.54 (Haraguchi type I = 0.76, Haraguchi type II = 0.40, Haraguchi type III = 0.25).

Conclusions
Diagnostic accuracy of measurements on plain lateral radiographs to assess articular involvement of posterior malleolar fractures is poor. There is a tendency to misjudge posteromedial involvement (Haraguchi type II).

INTRODUCTION

Both size of a posterior malleolar fracture that requires fixation, as well as the reliability of measurements on plain lateral radiographs are subject of ongoing debate. Several studies suggest that posterior malleolar fractures involving 25–33% of the tibial plafond require fixation.1-4 If the size of the posterior fragment is important in decision-making it seems foolish to rely on questionable diagnostics: it has been stated before that reliability of plain radiographs is poor compared to measurements on two-dimensionalComputed Tomography (CT).5 However Ferries’ study was limited by a 2D-CT reference standard, rather than quantification using three-dimensional (3D)-CT.6,7 Moreover, a recent study concluded that plain radiographs allowed for accurate assessment of the size of the posterolateral fragment in terms of interobserver reliability by eight experienced orthopaedic trauma surgeons, as compared to their standard: interpretation of the senior author and experienced musculoskeletal radiologist in a consensus agreement.1 In order to minimise subjectivity, we aimed to compare plain lateral radiographs to a 3D-CT reference standard. Previous research shows that quantification of 3D-CT modelling (Q3DCT) is a reliable technique to calculate articular surface areas.8,9

It has been suggested that morphology of the posterior malleolar fragment might be even more important than fracture size.5,10 Haraguchi and colleagues classified posterior malleolar fractures into three types, based on pathoanatomy of posterior malleolar fragments (Figure 1).11 To the posterolateral fragments usually the posterior syndesmotic ligaments are attached. To the posteromedial fragments the deep deltoid ligament can be attached, which has significant implications for stability.11,12 Weber and colleagues have described the Haraguchi type II fractures (including the posterior colliculus of the medial malleolus as having impacted fragments posteromedially that interfere with spontaneous anatomic reduction. We hypothesise that especially these types of posterior involvement are frequently missed on plain lateral radiographs.

The purpose of the present study is to find the diagnostic accuracy of measuring articular involvement of posterior fragments in ankle fractures on plain radiographs in a web-based collaborative13,14 using Q3DCT as a reference standard. A second goal is to assess the reliability of lateral radiographs on decision making, whether or not to fix the posterior fragment. We expect that surgeons overestimate true articular involvement on plain radiographs (because of discrepancy of the orientation of the fracture plane and the obliquity of the roentgen beam); but hope that the inter-observer agreement is good to excellent, since estimating fragment size on plain lateral radiographs has been the standard of care for decades.
METHODS

Subjects

A retrospective search for plain radiographs plus preoperative CT-scans of patients with ankle fractures (OTA type 44) involving the posterior malleolar fragment was performed in a Level III Trauma Center (St. Lucas Andreas, Amsterdam) treated between 2005 and 2012. This resulted in a total of 57 patients. After exclusion of 12 tibial pilon fractures (OTA type 43) in a consensus meeting, and 14 because of poor image quality, 31 ankle fractures were included and evaluated using Q3DCT-modelling technique as previously described.\(^3\)\(^,\)\(^9\)

Q3DCT modelling technique as a reference standard

We used quantitative three-dimensional CT modelling (Q3DCT) techniques as previously validated and reported for upper extremity articular anatomy,\(^10\)\(^,\)\(^11\)\(^,\)\(^12\)\(^,\)\(^14\)\(^,\)\(^16\)\(^,\) pathoanatomy of distal radius, coronoid, radial head and distal humeral fractures.\(^16\)\(^,\)\(^17\)\(^,\)\(^20\)\(^,\)\(^21\)\(^,\)\(^24\) Reliability of Q3DCT to determine articular involvement of posterior malleolar fractures as a percentage of the tibial plafond to establish a reference standard has been validated (ICC 0.993), in a separate study.\(^3\)\(^,\)\(^22\) A video to illustrate the methodology in detail is available at http://www.traumaplatform.org. The DICOM files were exported for further processing into MATLAB 8.0 (Natick, Massachusetts, USA). The created images and additional data were then loaded into Rhinoceros 4.0 (Seattle, Washington, USA). A wire model was constructed (Figure 2) to be used to form a mesh that represented the surface of the cortical bone and the individual fragments. The edges of the articular surfaces were manually checked and marked by the investigators (Figure 3). Measurement of surface is a standard feature in the Rhinoceros software. The area of the articular surface was presented as square millimetres (mm\(^2\)).

Diagnostic accuracy – study design

A web-based platform was created (The Science of Variation Group) to facilitate large international interobserver studies.\(^14\)\(^,\)\(^21\)\(^,\)\(^24\) Independent members of the Amsterdam Foot & Ankle Platform from several countries were contacted by email and invited to evaluate the 31 cases. The assessments and measurements were carried out on http://www.ankleplatform.com using a built-in radiology viewer (Figure 4). Observers were asked to review anteroposterior and lateral radiographs of the 31

Statistical analysis

Statistical analysis was performed using IBM SPSS 20.0 (SPSS Inc., Chicago, IL). Data were normally distributed and measurements are presented as means with Standard Deviations (SD). To calculate diagnostic accuracy of plain radiographs, the average value of the 100 observers was used to describe the difference between estimations on plain lateral radiographs and the reference standard (Q3DCT). Paired t-tests were performed to test the differences between the reference standard and described ankle fractures in order to measure the size of the posterior malleolar fragment and answer two questions: (1) what is the involved articular surface of the posterior malleolar fracture as a percentage of the tibial plafond? and (2) when you decide on operative treatment of this ankle fracture, would you fix the posterior malleolar fracture? The study was performed under a protocol approved by the local Institutional Research Board (IRB).
Within the Haraguchi type I fractures, the mean posterior malleolar fragment involved 16.3% (SD 10.0) of the articular surface on Q3DCT, compared to 29.3% (SD 17.4) on plain lateral radiographs. This difference of 13% (95% CI 12.34-13.63) was significant (p < 0.001). Within Haraguchi type II fractures, the reference articular surface found on Q3DCT was 18.1% (SD 10.3), compared to an estimation of 26.7% (SD 7.7) on plain radiographs by our group of 100 observers. This difference of 8.6% (95% CI 7.76 – 9.49) was significant (p = 0.032). Haraguchi type III fractures, had 6.6% (SD 4.7) of the articular surface involved, compared to the estimation of 17.8% (SD 8.1) on plain radiographs. This difference of 11.2% (95% CI 10.28 – 12.03) was significant (p = 0.001). These values relate to a factor 1.8, 1.5, and 2.7 overestimation respectively for Haraguchi types I, II and III (Table 1).

The diagnostic accuracy of plain radiographs for posterior malleolar fragment size depends on cut-off value chosen. Within limits ranging 5% below and above the reference standard value, only 22% of the observers measured accurately. 84% of the observers overestimates true fragment size on plain lateral radiographs. With regard to fragments <15% of the joint surface, the sensitivity is 0.44, the specificity is 0.97. When evaluating fragments between 15 and 25% of the joint surface, the sensitivity of plain radiographs is 0.03 and the specificity is 0.48. For fragments >25% of the joint surface, the sensitivity of measuring on plain radiographs is 0.86, the specificity is 0.60 (Table 2).

**TABLE 1.** Measurements by Q3DCT and 100 observers

<table>
<thead>
<tr>
<th>Fracture type</th>
<th>Mean size of the posterior malleolus as seen on Q3DCT in % (SD)</th>
<th>Mean size of the posterior malleolus as seen on plain lateral radiographs in % (SD)</th>
<th>Mean difference between Q3DCT and plain lateral radiographs in % (95% confidence intervals)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All types</td>
<td>13.5 (10.8)</td>
<td>24.4 (10.0)</td>
<td>10.9 (7.8-14.0)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Haraguchi I</td>
<td>16.3 (11.0)</td>
<td>29.3 (8.8)</td>
<td>13.0 (12.3-13.6)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Haraguchi II</td>
<td>18.1 (10.1)</td>
<td>26.7 (9.4)</td>
<td>8.6 (7.8-9.5)</td>
<td>0.032</td>
</tr>
<tr>
<td>Haraguchi III</td>
<td>6.6 (4.7)</td>
<td>17.8 (11.7)</td>
<td>11.2 (10.3-12.0)</td>
<td>0.001</td>
</tr>
</tbody>
</table>

**TABLE 2.** Diagnostic accuracy of measurements on plain lateral radiographs with Q3DCT as reference standard

<table>
<thead>
<tr>
<th></th>
<th>Fragments &lt;15%</th>
<th>Fragments 15-25%</th>
<th>Fragments &gt;25%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>0.44</td>
<td>0.03</td>
<td>0.86</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.97</td>
<td>0.48</td>
<td>0.60</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>0.96</td>
<td>0.15</td>
<td>0.33</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>0.47</td>
<td>0.16</td>
<td>0.95</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.62</td>
<td>0.16</td>
<td>0.65</td>
</tr>
<tr>
<td>Balanced accuracy</td>
<td>0.70</td>
<td>0.26</td>
<td>0.73</td>
</tr>
</tbody>
</table>
Reliability of measurements

For 100 observers the Intraclass Correlation Coefficient (ICC) for all fractures was 0.61 (95% CI 0.49 – 0.73). However, the values for the respective Haraguchi categories separately showed that the ICC was 0.79 (95% CI 0.64 – 0.93) for Type I fractures; 0.42 (95% CI 0.25 – 0.77) for Type II fractures and 0.31 (95% CI 0.18 – 0.59) for Type III fractures.

The standard error of measurement (SEM) for all fracture types was 10.1% articular surface area with a Smallest Detectable Difference (SDD) of 28.1% articular surface area. The SEM agreement for Haraguchi type I fractures was 9.0% with a SDD of 25.0%. The SEM agreement for Haraguchi type II fractures was 9.1% with a SDD of 25.3% and Haraguchi type III fractures had a SEM of 11.8% with a SDD of 32.7% (Table 3).

Operative management of posterior malleolar fractures

The decision of the evaluators to recommend surgery or not was based on plain films. The ICC for all fractures on the question ‘When you decide on operative treatment of this ankle fracture, would you fix the posterior malleolar fracture?’ was 0.54 (95% CI 0.24 – 0.83). For the Haraguchi type I fractures the ICC was 0.76 (95% CI 0.53 – 0.99). Haraguchi type II had an ICC of 0.40 (95% CI 0.07 – 0.74) and Haraguchi type III fractures had an ICC of 0.25 (95% CI 0.00 – 0.67) (see Table 4).

None of the observers would operate on all fractures, neither would any observer treat all fractures conservatively. The most aggressive observer would operate on 98.6% of the posterior malleolar fractures. The most conservative observer suggested to fix 25.5% of the fragments. The fracture least operated on, was also one of the three fractures that comprised 0.0% of the articular surface (Haraguchi Type III avulsion fracture). Though, 3 of 100 observers would still fix this fracture. There were two fractures that 100% of the observers would fix, involving 34.1% and 24.1% of the articular surface respectively according to the reference standard Q3DCT (Haraguchi Type I oblique fractures). The fractures most observers would fix were not the largest. Fractures involving >25% of the articular surface would be fixed by 85.3% of the observers, fractures involving >15% by 80.7%, fractures involving >10% by 67.9%, and 8.9% of the observers would operate fractures of <5% of the articular surface. On average, Haraguchi type I posterior malleolar fractures would be fixed by 59.9% of observers, Haraguchi type II by 60.7%, and Haraguchi type III fractures by 25.9% of observers. Interestingly when the fragments of >25% (based on Q3DCT) of the involved articular surface were evaluated separately, 99.3% of Haraguchi type I would have been fixed, while Haraguchi type II fractures would have been fixed in 71.3% of cases (p = 0.23). However for fragments with >15% of the involved articular surface: still 98% of Haraguchi type I fractures were chosen to be fixed, while only 66.2% of Haraguchi type II fractures would be fixed (p = 0.02).

DISCUSSION

Diagnostic accuracy of measuring on plain lateral radiographs to assess articular involvement of posterior malleolar fractures is poor. Surgeons should no longer solely rely on plain lateral radiographs to judge the pathoanatomy of posterior fragments in ankle fractures. Articular involvement of posterior malleolar fractures is overestimated on radiographs in this study with 100 observers evaluating 31 cases using quantification of 3D-CT measurement techniques as the reference standard.

Accurate assessment of articular involvement of the posterior malleolar fracture fragment in ankle fractures is essential, as well as comminution and impaction, as this is the leading argument for internal fixation of these ankle fractures. However overestimation does not mean over-treatment, since size of the fragment is not the main factor influencing outcome, but residual incongruence and instability.

Previous studies concluded that plain lateral radiographs poorly assess the size of the posterior malleolar fractures based on interobserver agreement; however they lack an objective reference standard. As 3D-morphology of the posterior malleolar fragment might be more important than fracture size, we found it very interesting that observers reached substantial agreement on articular involvement of oblique Haraguchi type I fractures (ICC = 0.79) as well as the decision whether or not to fix these fragment (ICC = 0.76), although type I fractures were overestimated in size (factor 1.8 for Haraguchi types I; estimation on radiographs 29.3% versus quantification of 16.3% on 3D-CT). In the classic article by Ferries and colleagues the larger fragments (i.e. Haraguchi type I) were less well assessed.

In our study, the smallest detectable difference between observers for all fractures was 28.1% of the articular surface. Even the large Haraguchi type I fractures had a smallest detectable-difference of 25.05. Hence the rationale behind operative fixation of posterior malleolar fragments involving 25–33% of the articulating surface is highly debatable, when the fracture fragment size is equal to the smallest detectable change between two random observers on plain lateral radiographs.

In contrast to Haraguchi type I posterior malleolar fractures, agreement for Haraguchi type II and III fractures on articular involvement and treatment was only fair to moderate. Ferries and colleagues found that plain radiographic interpretations erred in most cases by overrating the size of the involved articular surface.
Interestingly the case presented for underestimation was a Haraguchi type II. Also Weber and colleagues emphasized Haraguchi type II fractures to be a category apart: posteromedial extension of the fracture, if left untreated, leads to instability of the talus. Our results show that the Haraguchi type II fractures are less likely to be operated on. In patients with posteromedial involvement, the fragment often contains the posterior colliculus. Since the deep portion of the deltoid ligament attaches to this colliculus, malunion may lead to medial instability. Finally, in terms of articular involvement and size, 3D morphology of Haraguchi type I fractures results in a fracture plane that is almost parallel to the current standardized orientation of the Rontgen beam in lateral radiographs leading to enhanced visualization and substantial agreement – though overestimated – on articular involvement and treatment in this study. In contrast to Haraguchi type II fractures with conflicting 3D morphology with respect to the rontgen beam orientation resulting in only fair to moderate agreement on size and treatment. Gardner and colleagues also found that significant variation existed, regarding most aspects of posterior malleolar ankle fracture treatment in their survey study of Orthopaedic Trauma and Foot & Ankle surgeons. Most notably, factors other than fragment size guided operative indications. Also other authors conclude that decision on operative treatment is erroneously based on fragment size alone, instead of incorporating other important aspects, such as posteromedial lesions of the tibial plafond, additional osteochondral fragments and impaction.

Recommendations for future studies include a prospective (long-term) follow up of ankle fractures with posterior malleolar fragments including pre- and postoperative CT quantification. The influence of comminution, true fracture fragment size (mm³), 3D fracture morphology, articular involvement (mm²), residual gap, residual step-off and other patient related factors could then be analysed to discover the most important predictors of functional outcome. We philosophize that fracture morphology and associated ligament injury is more important than current emphasis on fragment size and articular involvement. Large undisplaced posterolateral oblique type I fractures could be left unfixed provided there is medial integrity, while small posteromedial type II fractures with associated deep deltoid avulsion could remain unstable even after fixation of the (anterior) medial malleolus.

Of course this study has to be interpreted in the light of its strengths and limitations. We consider the Q3DCT measurement technique a strength, and more reliable in objective quantification of the articulating surface area than subjective methods. However, we acknowledge the limitation of CT in that it does not account for true articular surface, since cartilage is not made visible. Our measurements will probably differ from those made on MRI or on cadaveric bone. Additional studies are needed to compare quantification of articular surface areas on 3D-CT models to true articular surface. A strength of CT however, is that it does account for fracture patterns and fragment morphology. A final limitation of Q3DCT-modeling is the time intensity of this technique. For now, it is primarily used for research purposes.

In conclusion and clinically highly relevant: this study shows that posterior malleolar articular involvement is severely misjudged on plain lateral radiographs. Overall, only 22% of measurements on plain radiographs were accurate (within 5% below and above reference values). The observers showed a remarkable preference to fix Haraguchi type I fractures with substantial agreement. These larger posterolateral fragments are best visible on plain lateral radiographs. Posteromedial fragments are at risk of being overlooked and undertreated, and may lead to persisting medial instability in cases of malunion. We argue that it might not be articular involvement, but 3D pathoanatomy of posterior malleolar fractures, that is most important in decision-making, operative approach, and functional outcome. In order to prevent ‘guesstimation’ of posterior malleolar articular involvement on plain lateral radiographs and subsequent treatment, we recommend additional CT in all trimalleolar ankle fractures.
GUESSTIMATION OF POSTERIOR MALLEOLAR FRACTURES ON LATERAL PLAIN RADIOGRAPHS

REFERENCES