Full length article

Patient-specific finite element computer models improve fracture risk assessments in cancer patients with femoral bone metastases compared to clinical guidelines

Florieke Eggermonta,*, Gerco van der Walb, Paulien Westhoffc, Arjonne Laar, Marianne de Jongd, Tom Rozemae, Herman M. Kroon, Onarisa Ayub, Loes Derikxb, Sander Dijkstrab, Nico Verdonschotta,g, Yvette van der Lindenh, Esther Tancka

a Orthopaedic Research Laboratory, Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, the Netherlands
b Department of Orthopaedic Surgery, Leiden University Medical Center, Leiden, the Netherlands
c Department of Radiation Oncology, Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, the Netherlands
d Radiotherapeutic Institute Friesland, Leeuwarden, the Netherlands
e Bernard Verhees Institute, Tilburg, the Netherlands
f Department of Radiology, Leiden University Medical Center, Leiden, the Netherlands
g Laboratory of Biomechanical Engineering, University of Twente, Enschede, the Netherlands
h Department of Radiotherapy, Leiden University Medical Center, Leiden, the Netherlands

ARTICLE INFO

Keywords:
- Finite element model
- Bone metastases
- Pathological fracture
- Fracture risk
- Clinical guidelines

ABSTRACT

Purpose: To determine whether patient-specific finite element (FE) computer models are better at assessing fracture risk for femoral bone metastases compared to clinical assessments based on axial cortical involvement on conventional radiographs, as described in current clinical guidelines.

Methods: Fifty-four patients with 50 femoral bone metastases, who were treated with palliative radiotherapy for pain, were included (64% single fraction (8 Gy), 36% multiple fractions (5 or 6 x 4 Gy)) and were followed for six months to determine whether they developed a pathological femoral fracture. All plain radiographs available within a two month period prior to radiotherapy were obtained. Patient-specific FE models were constructed based on the geometry and bone density obtained from the baseline quantitative CT scans used for radiotherapy planning. Femoral failure loads normalized for body weight (BW) were calculated. Patients with a failure load of 7.5 x BW or lower were identified as having high fracture risk, whereas patients with a failure load higher than 7.5 x BW were classified as low fracture risk. Experienced assessors measured axial cortical involvement on conventional radiographs. Following clinical guidelines, patients with lesions larger than 30 mm were identified as having a high fracture risk. FE predictions were compared to clinical assessments by means of diagnostic accuracy values (sensitivity, specificity and positive (PPV) and negative predictive values (NPV)).

Results: Seven femurs (14%) fractured during follow-up. Median time to fracture was 8 weeks. FE models were better at assessing fracture risk in comparison to axial cortical involvement (sensitivity 100% vs. 86%, specificity 74% vs. 42%, PPV 39% vs. 19%, and NPV 100% vs. 95%, for the FE computer model vs. axial cortical involvement, respectively).

Conclusions: Patient-specific FE computer models improve fracture risk assessments of femoral bone metastases in advanced cancer patients compared to clinical assessments based on axial cortical involvement, which is currently used in clinical guidelines.

1. Introduction

Patients with bone metastases carry a risk of pathological fractures. [1–3] If a pathological fracture occurs in a weight-bearing bone such as the femur, this leads to an immediate decrease in the patient’s mobility and self-care, and as a result in a reduced quality of life and possibly shortened survival [4,5]. Treatment of bone metastases in intact femurs is, therefore, based on the expected fracture risk [6,7]. Patients with an expected low
Institute, Tilburg). Ethical approval was obtained from all participating radiotherapy institutes in the Netherlands (Radboud university Medical Center, Nijmegen; Leiden University Medical Center, Leiden; Radiotherapeutic Institute Friesland, Leeuwarden; Bernard Verbeeten Institute, Tilburg). Specific inclusion and exclusion criteria have been discussed elsewhere [21,22] and between January 2015 and April 2017, with the aim to investigate fracture risk assessment tool.

2.1. Patients

Two multicentre prospective cohort studies were performed between August 2006 and September 2009, [21,22] and between January 2015 and April 2017, with the aim to investigate fracture risk assessment utilizing FE models in patients with femoral bone metastases. Specific inclusion and exclusion criteria have been discussed elsewhere [21,22]. In summary, patients with advanced cancer and referred for radiotherapy of bone metastases in the femur were asked to participate in four radiotherapy institutes in the Netherlands (Radboud university medical center, Nijmegen; Leiden University Medical Center, Leiden; Radiotherapeutic Institute Friesland, Leeuwarden; Bernard Verbeeten Institute, Tilburg). Ethical approval was obtained from all participating centres. Patients were treated with radiotherapy according to the current clinical guidelines [6,7]. Lesions with an axial cortical involvement ≤ 30 mm were treated with 8 Gy SF. Lesions with an axial cortical involvement > 30 mm were considered for prophylactic stabilizing surgery. However, if the patient’s condition was too poor and surgery was undesirable or impossible, the patient received MF radiotherapy (e.g. 5 or 6 fractions of 4 Gy) to induce remineralization of the bone. [23] If a patient was too ill to travel to the radiotherapy department for multiple fractions, it was accepted to deviate from the treatment guidelines and apply 8 Gy SF. Patients who were referred for surgery were not included in this study. In total, 156 patients gave informed consent. Patients were followed for six months or until a fracture occurred or until death, whichever occurred first.

Baseline characteristics (sex, age, primary tumour, radiotherapy schedule, pain (on a scale from 0 to 10), and Karnofsky performance status (KPS, on a scale from 0 to 100) [24]) were recorded prior to radiotherapy. QCT scans used for radiotherapy planning were made at baseline using a standardized protocol (120 kVp, 220 or variable mA, slice thickness 3 mm, pitch 1.5 or < 1, spiral and standard reconstruction, field of view (FOV) 480 mm, in-plane resolution 0.9375 mm). In thirteen patients, it was not possible to create FE models due to a hip or knee prosthesis (n = 5), an incompletely scanned femur (n = 7), or a missing calibration phantom (n = 1) (see Fig. 1 for the flow chart of patient inclusion for the current study). In a previous study it was shown that bone strength of femurs with osteoblastic lesions was overestimated [21], probably due to the fact that the empirically established FE material model is not valid for the highly mineralized (pathological) bone tissue in osteoblastic lesions. As a result, patients with predominantly osteoblastic appearance were excluded (n = 27) from the analysis. Also, patients were excluded if the calibration was affected by an air artefact (n = 43) [21], causing a shading artifact on the upper half of the calibration phantom and therefore resulted in unreliable calibration functions. Additionally, fourteen patients were excluded because their body weight was missing from the clinical research files and three patients were excluded because they underwent preventive stabilizing surgery shortly after inclusion. In total, 66 patients (76 affected femurs) were included in the FE database.

For the current study, we included the patients from the FE database with conventional anteroposterior (AP) and/or lateral radiographs available within a two month period prior to radiotherapy. [12] Patients were excluded from the current study if no radiographs were available (n = 21) [12]. This resulted in inclusion of 45 patients, with 50 affected femurs (Fig. 1). Twelve of these femurs were part of the study group published before [21].

2.2. FE models

Patient-specific femoral FE models were generated as described previously. [16,21] In short, for each irradiated femur, the three-dimensional geometry was segmented from the CT scan (Mimics 11.0 and 14.0, Materialise, Leuven, Belgium), and subsequently converted into a solid mesh of tetrahedral elements (Patran 2005r2 and 2011, MSC Software Corporation, Santa Ana, CA, USA). Additionally, the CT scans were calibrated with the use of a solid calibration phantom containing known calcium equivalent densities (Image Analysis, Columbia, KY, USA) that was scanned along with the patient. With the use of this calibration, Hounsfield units were converted to calcium equivalent values, which were used to calculate non-linear isotropic material behaviour for each tetrahedral element based on the material model of Keyak et al. [25] To correct for inter-scanner differences [26–28], cross-calibration using phantom scans (Gammex 467 phantom, RMI Gammex, Middleton, WI, USA) [27] was performed per scanner.

The FE models of the proximal femur were distally fixed at the knee joint centre using two high-stiffness springs (200.000.000 N/m) and loaded by incrementally displacing a cup on the femoral head in axial direction (Fig. 2, MSC.MARC 2007r1 and 2013.1, MSC Software...
Fig. 1. Flowchart of patient and femur inclusion.

Fig. 2. The workflow of generating the FE model. The CT scan is used to obtain the geometry, which is converted into a solid mesh. Additionally, the CT is calibrated using the calibration phantom to obtain bone mineral densities, which are used to calculate non-linear isotropic material behaviour for each tetrahedral element. The FE model is distally fixed at the knee joint centre (KJC) using two high-stiffness springs. Load was applied by displacing a cup on the femoral head in line with the hip joint centre (HJC) in axial direction.
2.4. Statistical analysis

Baseline characteristics were compared between the groups of patients who did and did not develop a pathological fracture during follow-up using Mann-Whitney U (age, pain, KPS), Fisher exact (sex, primary tumour) tests. FE predictions were compared to clinical assessments by means of diagnostic accuracy values: sensitivity, specificity, PPV and NPV. We compared sensitivities and specificities between FE and clinical assessments using McNemar’s test [29].

3. Results

In total, 45 patients with 50 affected femurs were included in this study. Six patients developed seven fractures (14%). Median time to fracture was 8 weeks (range 1–18). Thirteen patients died during the 6 month follow-up (33%). There were no significant differences in baseline characteristics between patients who did and did not develop a femoral fracture (Table 1). Five out of the seven fractured femurs had been treated with 8 Gy SF. Examples of CT images of a few patients can be found in Supplementary Fig. 1.

Table 1

<table>
<thead>
<tr>
<th>Patients with a femoral fracture (n = 6)</th>
<th>Patients without a femoral fracture (n = 39)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>2 (33%)</td>
<td>19 (49%)</td>
</tr>
<tr>
<td>Female</td>
<td>4 (67%)</td>
<td>20 (51%)</td>
</tr>
<tr>
<td><strong>Age at inclusion</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD), years</td>
<td>64.2 (6.2)</td>
<td>63.5 (10.9)</td>
</tr>
<tr>
<td><strong>Primary tumour</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>2 (33%)</td>
<td>11 (28%)</td>
</tr>
<tr>
<td>Prostate</td>
<td>0 (0%)</td>
<td>4 (10%)</td>
</tr>
<tr>
<td>Lung</td>
<td>2 (33%)</td>
<td>12 (31%)</td>
</tr>
<tr>
<td>Multiple myeloma</td>
<td>2 (33%)</td>
<td>5 (13%)</td>
</tr>
<tr>
<td>Other</td>
<td>0 (0%)</td>
<td>7 (18%)</td>
</tr>
<tr>
<td><strong>Radiotherapy schedule</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single fraction (1 x 8 Gy)</td>
<td>5 (83%)</td>
<td>24 (62%)</td>
</tr>
<tr>
<td>Multiple fractions (5-6 x 4 Gy)</td>
<td>1 (17%)</td>
<td>15 (38%)</td>
</tr>
<tr>
<td><strong>Pain score</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>7.5 (2.5)</td>
<td>5 (2.5)</td>
</tr>
<tr>
<td>Karnofsky performance status [24]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>68.3 (14.7)</td>
<td>79.3 (10.1)</td>
</tr>
<tr>
<td><strong>Affected femur</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unilateral</td>
<td>5 (83%)</td>
<td>34 (87%)</td>
</tr>
<tr>
<td>Bilateral</td>
<td>1 (17%)</td>
<td>5 (13%)</td>
</tr>
</tbody>
</table>

Pain was assessed at baseline using a pain score ranging from 0 (= no pain) to 10 (= worst imaginable pain). Karnofsky performance status [24] ranging from 0 (= lowest performance) to 100 (= highest performance).

According to the FE model, all seven fractures had a failure load below 7.5 x BW, and were accordingly correctly assessed as high risk, resulting in a sensitivity of 100% (Fig. 4, Table 2). Based on the 30 mm threshold, all but one of the fractures were correctly identified (sensitivity of 86%, difference of 14%, 95% confidence interval of -23% to 51%). Of the 43 non-fractured femurs, the FE model correctly predicted 32 femurs as low risk (specificity of 74%), whereas 18 non-fractured femurs were accurately assessed as low risk (specificity of 42%, statistical significant difference of 32%, 95% confidence interval of 14%–47%) by using the 30 mm threshold. The NPV of the FE model was 100%, demonstrating that none of the femurs with a low fracture risk as calculated by the FE model fractured. The NPV of the axial cortical involvement was slightly lower (95%). The PPV of the FE model was 39%, indicating that 39% of the femurs with a high fracture risk assessment in reality fractured. For the axial cortical involvement, the PPV was 19%.

There was no correlation between the femoral failure load corrected for bodyweight and the measured axial cortical involvement measured on conventional radiographs (Fig. 5).

4. Discussion

For assessment of expected fracture risk in patients with femoral bone metastases, this study compared FE computer models with axial cortical involvement on conventional radiographs as described in current clinical guidelines. FE models were better at assessing fracture risk in comparison to the clinical guidelines: the fracture risk of more femurs, either high or low, was correctly assessed. Clinically, fracture risk is estimated based on radiographs or CT scans, and FE models are currently not used. Clinical guidelines, such as the 30 mm threshold of axial cortical involvement, have been constructed to align the method for fracture risk assessment between different medical specialists. However, previous studies have shown that substantial numbers of patients are under- and overtreated based on such clinical guidelines [6,7,12] Although promising, fracture risk assessment by the FE model
stated in the clinical guidelines, which underlines the need for a better, standardized and more reliable fracture risk assessment tool.

Previously, we performed a comparable study, [21] in which we also generated FE models and compared the FE with clinical fracture risk assessments based on the axial cortical involvement measured on DRRs instead of conventional radiographs. Such DRRs are reconstructed radiographs based on the CT images with a rather coarse resolution (0.9375 x 0.9375 x 3 mm). Hence, the DRR image quality was rather poor in comparison to conventional radiographs. In that study, we included 47 femurs of which nine fractured (twelve of these femurs are part of the current study too) and found that the FE model had higher sensitivity, i.e. was more accurate in identifying patients with a high fracture risk, compared to clinical assessments, whereas specificity was lower for the FE models than for the clinical assessments on DRRs [21]. With respect to our previous study, the sensitivity and NPV of the FE model increased (89% for the previous study vs. 100% for the current study, and 97% vs. 100%, respectively), whereas specificity and PPV were slightly lower (79% vs. 74% and 50% vs. 39%, respectively). This can be explained by the fact that part of the CT scans used in the previous study were affected by air artefacts, in addition to possible inter-scanner differences, [26–28] and were, therefore, analyzed for each institute separately. In the current FE database, the CT scans with air artefacts were excluded and inter-scanner differences were corrected for, resulting in more accurate FE fracture predictions. Furthermore, the sensitivity and NPV of the clinical assessments in the previous study were lower (between 0% and 22% vs. 86% for the current study, and between 80% and 84% vs. 95%, respectively), whereas mainly specificity was much higher (between 84% and 97% vs. 42%) in contrast to that of the current study. PPV varied largely in the previous study between the different clinicians (0%–50%). The diagnostic accuracy values of the current study are closer to diagnostic accuracy of previous studies using the 30 mm threshold on conventional radiographs [6,12].

Another study by Goodheart et al. [30] compared fracture risk estimation by FE models with Miresl’s scoring system. Miresl’s scoring system, based on ratings of pain, lesion type, size and location [11], is, however, known to overestimate fracture risk, leading to large numbers of overtreatment [31], which can be reduced when applying the 30 mm threshold [7]. Goodheart et al. [30] showed that the FE model and Miresl’s scoring system had similar sensitivity, whereas specificity was higher for the FE model compared to Miresl’s. They concluded that FE models can improve fracture prediction over clinical assessments based on Miresl’s scoring system [30]. Previously, Van der Linden et al. [7] and recently, Van der Wal et al. [12] showed that fracture risk assessment based on axial cortical involvement of the metastases was more accurate in comparison with Miresl’s scoring system. In the current study, we

Table 2
Diagnostic accuracy values of the FE model and the clinical assessments based on 30 mm threshold for axial cortical involvement of the femoral metastases on radiographs.

<table>
<thead>
<tr>
<th>Axial cortical involvement [6]</th>
<th>Femurs that fractured during follow-up (n = 7)</th>
<th>Femurs that did not fracture during follow-up (n = 43)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 30 mm (high risk)</td>
<td>6</td>
<td>25</td>
<td>86%</td>
<td>42%</td>
<td>19%</td>
<td>95%</td>
</tr>
<tr>
<td>≤30 mm (low risk)</td>
<td>1</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FE model [21]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤7.5 (high risk)</td>
<td>7</td>
<td>11</td>
<td>100%</td>
<td>74%</td>
<td>39%</td>
<td>100%</td>
</tr>
<tr>
<td>≥ &gt; 7.5 (low risk)</td>
<td>0</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PPV = positive predictive value, NPV = negative predictive value.
showed that FE models can further improve these fracture risk assessments.

Other biomechanical methods, such as computed tomography rigidity analysis (CTRA) [32,33] and high resolution MRI-based FE models [34] also show promising results to be used to assess fracture risk. However, to our knowledge, the MRI-based FE models have not yet been tested in patients with actual metastatic lesions. Also, both CTRA as MRI-based FE models are not yet clinically being implemented.

This study had some limitations. Firstly, conventional radiographs were not available for all 66 patients (representing 76 femurs with bone metastases) in the FE database, resulting in exclusion of 21 patients (26 femurs). Secondly, patients with large lesions and an expected high fracture risk who were surgically treated were not included. Additionally, we only included patients already referred for palliative radiotherapy, so patients without symptoms were not included in these analyses. Consequently, we cannot conclude on whether the FE model would prevent any unnecessary surgeries in those patient groups. Thirdly, some patients died shortly after inclusion or became immobile and, consequently, did not develop femoral fractures during follow-up. They might have developed fractures if they had lived longer or had engaged in activities which would increase load bearing onto their legs.

FE models are probably better at assessing fracture risk in comparison to simple measurements on conventional radiographs due to the fact that they consider for example location and 3D geometry of the lesion, general bone quality, or geometry of the bone, whereas measuring the lesion in the cortex on a two dimensional radiograph does not take these factors into account. To this date, the FE model has only been used in research settings, but we are currently working towards clinical implementation. To enable implementation, we have developed a phantomless calibration method, [35] which facilitates widespread use of QCT and FE by avoiding the requirement of a calibration phantom during the CT scan session. For this phantomless calibration, CT densities of air, fat and muscle tissue are used, which results in FE failure loads comparable to those calculated using the conventional phantom calibration. [35] A limitation of the FE models is that they are currently not applicable for patients with predominantly osteoblastic lesions. This is due to the fact that the material model used is apparently not valid for the high CT density of such lesions and, hence, results in aberrant material properties [21]. None of the patients affected with predominantly osteoblastic lesions developed a fracture during follow-up. In the near future, if patients with femoral bone metastases visit their medical specialist, an FE fracture risk assessment can be ordered, just like ordering a lab-test. The treating physician can then discuss the results of the FE model with the patient, taking into account clinical factors such as the patient’s clinical condition, life expectancy, and activity level.

In conclusion, this study showed that patient-specific FE models improve femoral fracture risk assessments in comparison to measuring axial cortical involvement on conventional radiographs as described in the current clinical guidelines on bone metastases. The FE models could prevent unnecessary surgical procedures and, therefore, improve quality of life of those patients. Therefore, clinical implementation of the FE models is supported.

Funding statement

This work was supported by the Dutch Science Foundation NWO-STW (NPG.06778), Fonds NutsOhra (1102-071), Furlong Research Charitable Foundation, the Dutch Cancer Society (KUN 2012-5591), the Dutch Cancer Society/Alpe d’HuZes (UL2013-6286).

Acknowledgements

The authors thank Leon Driessen and Max Bakker for their help in generating the patient-specific FE models.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.bone.2019.115101.

References


