Follow us @BoneJointRes

Freely available online OPEN ACCESS

BJR

BIOMATERIALS

Can patient-specific finite element models better predict fractures in metastatic bone disease than experienced clinicians?

TOWARDS COMPUTATIONAL MODELLING IN DAILY CLINICAL PRACTICE

F. Eggermont, L. C. Derikx, N. Verdonschot, I. C. M. van der Geest, M. A. A. de Jong, A. Snyers, Y. M. van der Linden, E. Tanck

Orthopaedic Research Laboratory, Radboud Institute for Health Sciences, Radboud university medical center

- F. Eggermont, MSc, PhD Candidate, Orthopaedic Research Laboratory,
- L. C. Derikx, PhD, Postdoctoral Researcher, Orthopaedic Research Laboratory,
- I. C. M. van der Geest, MD, PhD, Orthopaedic Surgeon, Department of Orthopaedics,
- A. Snyers, MD, Radiation Oncologist, Department of Radiation Oncology,
- E. Tanck, PhD, Associate Professor, Orthopaedic Research Laboratory, Radboud Institute for Health Sciences, Radboud university medical center, Nijmegen, The Netherlands.
- N. Verdonschot, PhD, Professor, Orthopaedic Research Laboratory, Radboud Institute for Health Sciences, Radboud university medical center, Nijmegen, The Netherlands and Laboratory of Biomechanical Engineering, Enschede, The Netherlands.
- M. A. A. de Jong, MD, Radiation Oncologist, Radiotherapeutic Institute Friesland, Leeuwarden, The Netherlands.
- Y. M. van der Linden, MD, PhD, Radiation Oncologist, Department of Radiotherapy, Leiden University Medical Center, Leiden, The Netherlands.

Correspondence should be sent to F. Eggermont; email: Florieke.Eggermont@radboudumc.nl

doi: 10.1302/2046-3758.76.BJR-2017-0325.R2

Bone Joint Res 2018;7:430–439.

Objectives

In this prospective cohort study, we investigated whether patient-specific finite element (FE) models can identify patients at risk of a pathological femoral fracture resulting from metastatic bone disease, and compared these FE predictions with clinical assessments by experienced clinicians.

Methods

A total of 39 patients with non-fractured femoral metastatic lesions who were irradiated for pain were included from three radiotherapy institutes. During follow-up, nine pathological fractures occurred in seven patients. Quantitative CT-based FE models were generated for all patients. Femoral failure load was calculated and compared between the fractured and non-fractured femurs. Due to inter-scanner differences, patients were analyzed separately for the three institutes. In addition, the FE-based predictions were compared with fracture risk assessments by experienced clinicians.

Results

In institute 1, median failure load was significantly lower for patients who sustained a fracture than for patients with no fractures. In institutes 2 and 3, the number of patients with a fracture was too low to make a clear distinction. Fracture locations were well predicted by the FE model when compared with post-fracture radiographs. The FE model was more accurate in identifying patients with a high fracture risk compared with experienced clinicians, with a sensitivity of 89% *versus* 0% to 33% for clinical assessments. Specificity was 79% for the FE models *versus* 84% to 95% for clinical assessments.

Conclusion

FE models can be a valuable tool to improve clinical fracture risk predictions in metastatic bone disease. Future work in a larger patient population should confirm the higher predictive power of FE models compared with current clinical guidelines.

Cite this article: *Bone Joint Res* 2018;7:430–439.

Keywords: Finite element modelling, Fracture prediction, Femur, Metastatic bone disease

Article summary

- Patient-specific finite element (FE) models are evaluated on their ability to identify patients at risk of pathological femoral fracturing resulting from metastatic bone disease.
- These FE results are compared with clinical fracture locations and judgements of experienced clinicians.

Key messages

- The FE model was more accurate at identifying patients with a high fracture risk compared with experienced clinicians, but slightly less accurate in identifying patients with a low fracture risk.
- Fracture locations were well predicted by the FE model when compared with post-fracture radiographs.
- FE models can be a valuable tool to improve clinical fracture risk predictions in metastatic bone disease.

Strengths and limitations

- We performed a prospective multicentre patient study.

- Patient-specific FE analysis is an objective method for the prediction of fracture risk.
- In future studies, the number of patients should be increased, and a solution for inter-scanner differences needs to be found.

Introduction

Cancers of the breast, prostate, lung, kidney, and thyroid can metastasize to bone.[1-3] These metastases can cause pain and, when left untreated, carry a risk of developing complications such as pathological fractures or, in case of vertebral metastases, spinal cord compression.[1,2,4] Pathological fractures in extremities affect quality of life as they hamper the patient's mobility and self-care.

Femoral metastases with a low risk of fracture can be treated conservatively with local radiotherapy. Metastases with a high risk of fracture require prophylactic surgery to retain stability of the bone.[5] This is an invasive procedure requiring anaesthesia, which is generally complex in cancer patients with limited life expectancy and deteriorating condition. Thus, the decision to proceed with either a non-invasive treatment or a prophylactic surgical treatment should be carefully made.

However, current clinical practice lacks an accurate tool to guide clinicians to the correct treatment decision. Numerous studies have evaluated lesion or patient factors on the probability of impending fractures; however, none has shown a sufficient predictive power.[5] A potential tool to improve clinical fracture risk assessments is finite element (FE) modelling, which has been shown to predict human femoral bone strength fairly accurately.[6-10] Our group have shown that the FE model accurately calculated failure load and fairly predicted fracture locations in cadaver femurs with and without artificial lesions compared with mechanical experiments.[11-13] Moreover, we demonstrated that ranking on FE failure load better resembled the experimentally measured failure loads than rankings by experienced clinicians.[11]

In this prospective cohort study, we investigated whether our subject-specific FE models are able to identify patients at risk of pathological femoral fractures resulting from metastatic bone disease. For this purpose, we included patients referred for radiotherapy to treat painful femoral metastases. Against expectations, some of these patients sustained pathological fractures in the femur during follow-up. We calculated the femoral failure loads and compared those between patients who did or did not sustain a fracture. In addition, we compared the FE predictions with assessments by experienced clinicians. We hypothesized that the FE models more accurately identify patients with a high fracture risk than experienced clinicians.

Patients and Methods

Study design. Between August 2006 and September 2009, all patients referred for palliative radiotherapy of the femur to three participating radiotherapy institutes in the Netherlands (Radboud university medical center (institute 1), Radiotherapeutic Institute Friesland (institute 2), and Leiden University Medical Center (institute 3) were asked to participate in this prospective cohort study. Globally, 20% to 25% of the eligible patients participated. Ethical approval was obtained from all participating centres. These patients received palliative radiotherapy following Dutch clinical guidelines. Lesions with an axial cortical involvement < 30 mm have an expected low risk of fracture ($< 5\%$) and were treated with a single dose of 8 Gy.[5] If the axial cortical involvement was > 30 mm, the risk of fracture is estimated at 23%.[14] These patients were referred for prophylactic surgery and therefore excluded from this study.[5] If the patient's condition was such

Table I. Inclusion criteria

Inclusion criteria
Proven malignancy
Karnofsky[16] performance status ≥ 60
No clinical or radiological evidence of pathological fracturing of the femur
No prior palliative surgery for the current treatment site of the femur
No planned surgical intervention of the femoral bone
No systemic radiotherapy 30 days prior to entry into the study
No previous radiotherapy to the current treatment site of the femur
Patient is able and willing to fill out baseline and follow-up forms on pain and quality of life
Patient is willing to undergo additional CT scans for the femoral region

Fig. 1

Boundary conditions for the finite element model. The model was distally fixed by springs with a very high stiffness and the load was applied by means of a cup on the head of the femur, which incrementally displaced in a distal direction.

Table II. Characteristics of the patients who sustained a fracture during follow-up

Patient	Gender	Age at inclusion, yrs	Femur	Treatment dose, (fractions, n)	Time to fracture, days	Type of fracture	Activity while fracture occurred
Patient 1	Male	70	Right femur (F1)	24 Gy (6)	123	Neck of femur fracture	Walking
			Left femur (F2)	24 Gy (6)	123	Neck of femur fracture	Walking
Patient 2	Female	53	Right femur (F3)	8 Gy (1)	92	Pertrochanteric fracture	Spontaneously
			Left femur (F4)	N/A	92	Neck of femur fracture	Spontaneously
Patient 3	Male	64	Left femur (F5)	24 Gy (6)	7	Subtrochanteric fracture	Spontaneously
Patient 4	Female	66	Left femur (F6)	24 Gy (6)	13	Neck of femur fracture	Spontaneously
Patient 5	Female	62	Right femur (F7)	24 Gy (6)	3	Unknown	Unknown
Patient 6	Male	89	Right femur (F8)	8 Gy (1)	237	Diaphyseal fracture	Spontaneously
Patient 7	Female	80	Left femur (F9)	N/A	133	Neck of femur fracture	Unknown

N/A, not applicable

Table III. Femurs and patients included in each of the institutes

	Institute 1*		Institute 2†		Institute 3‡	
	Femurs	Patients	Femurs	Patients	Femurs	Patients
Fracture group (F)	5	3	3	3	1	1
Non-fracture group (NF)	11	8	23	22	4	4

*One patient in institute 1 fractured both femurs, but was only radiated on the right side
†One patient in institute 2 was radiated on both sides, but only fractured her right femur, leaving her left femur in the NF group
‡One patient in institute 3 fractured her non-treated femur, leaving her treated femur in the NF group

that surgery was undesirable or impossible, the patient was referred for multiple fraction radiotherapy (e.g. 5 or 6 × 4 Gy) to induce remineralization of the bone.[15] These patients were included in this study. Further inclusion criteria are depicted in Table I.[16] During the study period, 62 patients gave their consent. The patients were grouped according to the predominant appearance of their bone metastases (i.e. lytic or blastic). We excluded patients who had predominant blastic lesions (n=16). Although blastic lesions generally lead to a decreased structural bone strength,[17] in this study the femoral bone strength was overestimated, probably due to the high degree of mineralization which resulted in unrealistically strong material properties in the FE model (see supplementary information). Additionally, patients who had no body weight (BW) recorded (n=5) or sustained a femoral fracture more than a year after inclusion (n=1) were excluded. One patient sustained a femoral fracture following a fall and was therefore excluded. This led to inclusion of 39 patients with predominant lytic bone lesions in this study.

Baseline characteristics of all patients were recorded before radiotherapy. Furthermore, quantitative computed tomography (QCT) scans of the femoral region were retrieved prior to, at 28, and at 70 days after radiotherapy. Patients referred for multiple fraction radiotherapy underwent an additional QCT scan on the final day of their radiation schedule to capture the potential short-term effect of multiple fraction radiotherapy.[18] Through follow-up questionnaires and hospital records, patients were actively followed for six months or until a fracture occurred or until death, as competing risk, whichever occurred first. Based on having sustained a fracture, the patients were divided into either the fracture (F) group or the non-fracture (NF) group. Additionally, after two years, data on fractures and death were updated with the use of hospital records.

Different CT scanners. Recent work by Carpenter et al[19] has shown that the use of different CT scanners can have a significant effect on bone mineral density measurements and subsequent calculated failure loads, which is difficult to correct for. In the current study, the three institutes used two different types of CT scanner, Philips Big Bore Brilliance (institute 1) and Philips AcQSim CT (institute 2 and 3), both manufactured by Philips Medical Systems, Eindhoven, The Netherlands. Although QCT scan settings were protocolized as far as possible, inter-scanner effects may have been present in the input to our FE models, which could potentially lead to incorrect or at least incomparable FE failure loads. Therefore, apart from a group analysis, we also analyzed the data individually for the three institutes to circumvent such inter-scanner differences. It should be noted that our previous *ex vivo* validation study was conducted using the scanning equipment of institute 1.[11-13]

FE modelling. Patient-specific femoral FE models were generated, for the greater part, using the workflow reported previously.[11] Summarizing, QCT images were generated using a standard protocol (as far as allowed by clinical practice), with the following settings: 120 kVp, 220 mA, slice thickness 3 mm, pitch 1.5, spiral and standard reconstruction, in-plane resolution 0.9375 mm. The patient-specific femoral geometry was segmented from the most recent CT images available and converted to a 3D surface mesh (Mimics 11.0 and 14.0, Materialise, Leuven, Belgium) and a solid mesh consisting of

Table IV. Baseline characteristics

	Fracture group (F) (n = 9)*	Non-fracture group (NF) (n = 38)*	p-value
Gender, n (%)			
Male	4 (*44*)	20 (*53*)	0.7†
Female	5 (*56*)	18 (*47*)	
Median age, yrs (IQR)	66.0 (57.5 to 75.0)	62.5 (52.8 to 76.5)	0.5‡
Median body weight, kg (IQR)	73.0 (63.0 to 76.5)	76.0 (57.8 to 87.3)	0.6‡
Radiation schedule, n (%)§			
Single fraction	4 (*44*)	15 (*39*)	1†
Multiple fractions	5 (*56*)	23 (*61*)	
Median Karnofsky performance status (IQR)	80.0 (70.0 to 80.0)	80.0 (70.0 to 90.0)	0.6‡
Median time since primary tumour, yrs (IQR)	3.6 (1.7 to 6.5)	3.3 (0.8 to 5.5)	0.8‡
Median time since first metastasis, yrs (IQR)	3.2 (0.1 to 3.6)	1.1 (0.1 to 2.8)	0.5‡
Primary cancer site, n (%)			
Breast	2 (*22*)	11 (*29*)	0.3¶
Lung	2 (*22*)	5 (*13*)	
Prostate	2 (*22*)	12 (*32*)	
Kidney	0 (*0*)	4 (*11*)	
Rectum	0 (*0*)	2 (*5*)	
Multiple myeloma	3 (*33*)	2 (*5*)	
Urethra	0 (*0*)	1 (*3*)	
aCUP**	0 (*0*)	1 (*3*)	
Median time to death since inclusion, mths (IQR)††	11.0 (4.0 to 13.0)	8.0 (3.0 to 17.0)	0.9‡
Institute, n (%)			
1	5 (*56*)	11 (*29*)	0.3¶
2	3 (*33*)	23 (*61*)	
3	1 (*11*)	4 (*11*)	
Lesion type, n (%)			
Not visible	1 (*11*)	3 (*8*)	0.4¶
Lytic	1 (*11*)	13 (*34*)	
Mixed	7 (*78*)	22 (*58*)	

*Fracture group: nine femurs in seven patients. Non-fracture group: 38 femurs in 34 patients. Two patients had one fractured and one non-fractured femur
†Fisher's exact test
‡Mann–Whitney U test
§Two femurs in the fracture group were not treated with radiotherapy
¶Chi-squared test
**Cancer of unknown primary origin
††Date of death missing for three non-fracture patients
IQR, interquartile range

tetrahedral elements (average element volume 1.0 mm^3; Patran 2005r2, MSC Software Corporation, Santa Ana, California), subsequently. A solid calibration phantom containing known calcium equivalent densities (Image Analysis, Columbia, Kentucky) was scanned along with the patient at the level of the proximal femur. Since pilot tests showed that calibration in diaphyseal slices was most accurate due to beam hardening in more proximal slices, we performed a mean diaphyseal slice calibration to convert the grey values to calcium equivalent densities, ash densities, and non-linear isotropic material behaviour, respectively, based on the material model of Keyak et al.[7] In this material model, the post-failure material behaviour for each element is represented by an initial perfectly plasticity phase, followed by a strain softening phase and finally an indefinite perfectly plastic phase.[7] However, in case patients' limbs were supported by a cushion to diminish pain during CT scanning, an air gap between calibration phantom and patient was present, leading to an artefact in the calibration phantom at the diaphyseal level. In such cases, we used more proximal slices for calibration.

The FE simulations of the proximal femur were performed using MSC.MARC (2007r1; MSC Software Corporation). The FE models were loaded by displacing a cup on the head of the femur in the axial direction, while distally fixed at the knee joint centre by two bundles of high-stiffness (200 000 000 N/m) springs (Fig. 1), which roughly resembles single-legged stance. Force-displacement curves were made based on displacement and contact normal forces that were registered for each increment. The maximum total reaction force determined the failure load of the femur, which was normalized for BW. The failure location was defined by elements that had plastically deformed at the moment of structural failure, and was compared with the post-fracture radiograph.

Clinical assessment. To compare the FE predictions with clinical fracture risk assessments, we generated digitally reconstructed radiographs (DRRs) from the CT scans in this study.[20] We asked two radiation oncologists with broad expertise in palliative radiotherapy, who regularly discuss and refer patients to the orthopaedics department (C1 and C2), and one experienced orthopaedic oncology surgeon (C3) to individually assess the DRRs just as in

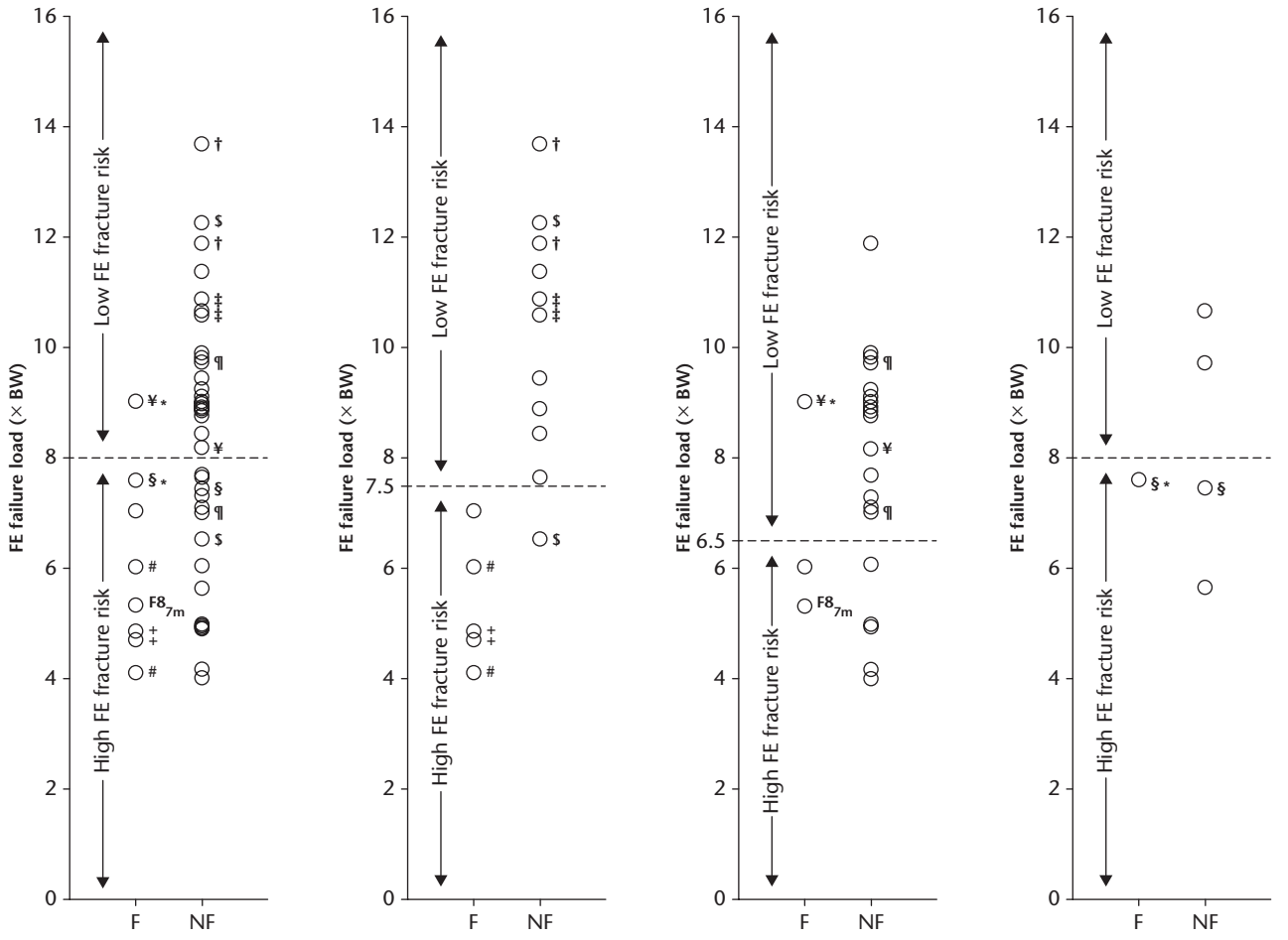

Fig. 2a Fig. 2b Fig. 2c Fig. 2d

Femoral failure load for patients who did (F) or did not (NF) sustain a femoral fracture during follow-up, corrected for body weight (BW), a) in all institutes without considering inter-scanner differences, and in b) institute 1, c) institute 2, and d) institute 3 separately. It should be noted that one femur (F8$_{7m}$) fractured one month after follow-up. The institutional thresholds were used to compare the predictive power of the finite element (FE) model *versus* experienced clinicians. *Femur fractured during unknown activity; all other symbols (#, +, †, $, ‡, ¶, ¥, §) indicate paired femurs.

daily practice, without providing any further information. First, they indicated whether the patient carried a high risk of fracture requiring elective surgery. Subsequently, we asked them to judge whether the cortical disruption caused by the metastasis was >30 mm.[14]

Statistical analysis. Baseline data were compared between fracture and non-fracture group on the femur level using chi-squared test, Fisher's exact test, or Mann–Whitney U test, where applicable. We compared the failure load corrected for BW between the fractured and the non-fractured femurs using Mann–Whitney U tests. For all tests, the level of significance was defined as $p<0.05$.

To compare the clinical assessments with the predictions by the FE model, a critical FE failure load was defined for the whole group, as well as for each institute separately, classifying a patient to a high or a low fracture risk. More specifically, diagnostic accuracy values (sensitivity and specificity, and positive and negative predictive values (PPV and NPV)) were calculated for different thresholds of a critical failure load using increments of 0.5 × BW. The threshold with the highest sum of specificity and sensitivity was chosen. For comparison with clinical assessments, we used the critical failure loads of the separate institutes.

Results

Patients. In all, 39 patients with predominant lytic painful bone metastases were included in this study. One of the patients sustained a fracture of the femur one month after follow-up, and was included in the F group. This F group consisted of seven patients sustaining nine fractures (Tables II and III). Two of these fractures occurred in the contralateral femur that was not irradiated (the irradiated femurs of these patients were included in the NF group). Additionally, two cases with an unknown cause of fracture were included. The NF group comprised a total of 34 patients with 38 treated non-fractured femurs (Table III). There were no significant differences in baseline characteristics between the F and NF group (Table IV).

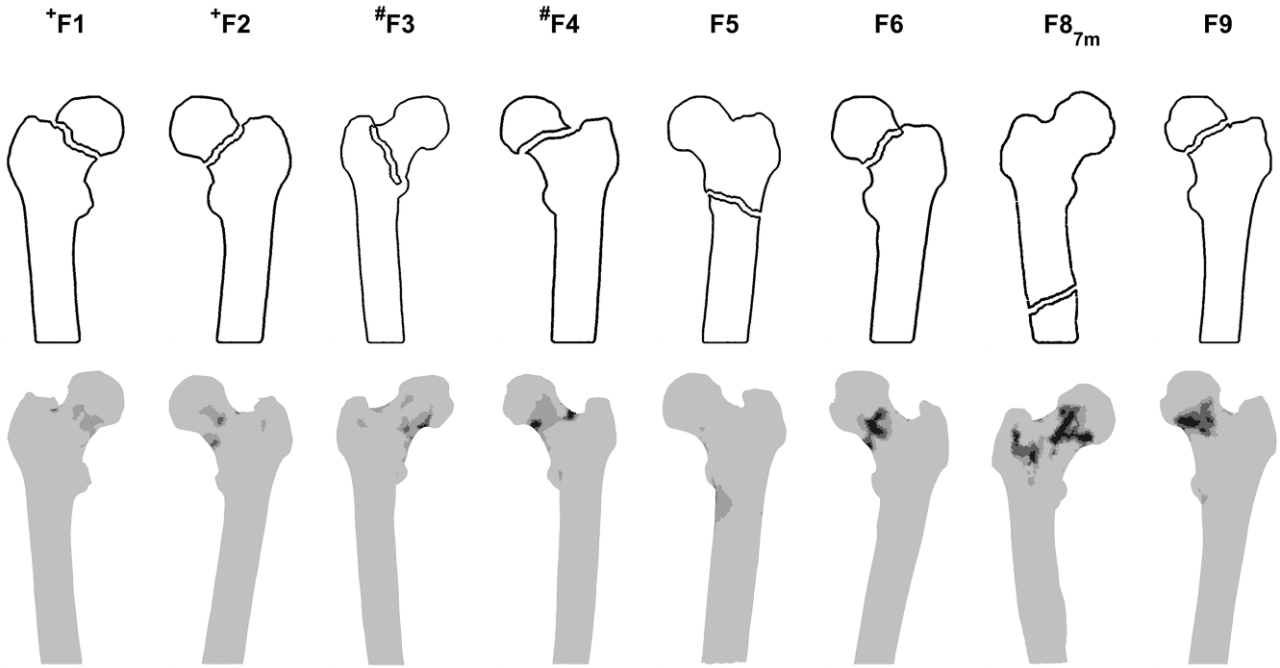

Fig. 3

Schematic overview of clinical fracture locations (upper panel), indicated by an experienced clinician who was blind to the predicted fracture locations, and the fracture locations at failure (mid-coronal plane) predicted by the finite element models (lower panel). Femurs indicated with + and # are paired femurs. F8 fractured one month after follow-up (7m). There was no clinical information about fracture location available for F7.

FE models. Figure 2 shows the BW-corrected failure loads for all femurs in this study together (Fig. 2a), as well as split to the three institutes (Figs 2b-d). The median failure load of all fractured femurs together was significantly lower compared with failure load of all non-fractured femurs (6.03 (interquartile range (IQR) 4.80 to 7.33) *vs* 8.93 (IQR 7.10 to 9.85), p = 0.002). After splitting to individual institute to avoid inter-scanner differences, the median failure load of fractured femurs was significantly lower in institute 1 compared with failure load of the non-fractured femurs (4.89 (IQR 4.41 to 6.55) *vs* 10.60 (IQR 8.46 to 11.90), p = 0.001). This was not the case for the femurs from institutes 2 and 3 (6.03 *vs* 8.78 (IQR 7.04 to 9.12), p = 0.5 and 7.61 *vs* 8.61 (IQR 6.12 to 10.44), p = 1).

We compared the actual fracture location with those predicted by the FE models for eight out of nine femurs. In one case, no radiological information on clinical fracture location was available. Six out of eight FE fracture locations resembled the actual fractures on post-fracture radiographs (Fig. 3). In the two other cases, the FE models predicted femoral neck fractures, whereas these patients clinically presented with a pertrochanteric and diaphyseal fracture, respectively.

Clinical assessment. A critical failure load was defined for the whole group, as well as for each institute. The critical failure loads were 8.0 × BW for the whole group and 7.5 × BW, 6.5 × BW, and 8.0 × BW for institutes 1, 2, and 3, respectively (Fig. 2). When each institute was analyzed individually, the sensitivity remained the same, while the specificity increased from 0.63 to 0.79.

For comparison with clinical assessments, we used the critical failure loads of the separate institutes. More patients were correctly identified with a high fracture risk by the FE model than by clinicians who relied on their clinical experience (Fig. 4), resulting in higher sensitivity of the FE model (0.89) compared with the clinicians (ranging from 0.00 to 0.33; Table V). The FE model identified 16 femurs with a high fracture risk, eight of which actually fractured during follow-up (PPV = 0.50). The PPV for clinicians ranged between 0.00 and 0.50. Of the 38 non-fractured femurs, the FE model correctly identified 30 femurs as having a low fracture risk (specificity = 0.79). The specificity values for clinicians were slightly higher and ranged between 0.84 and 0.95, although the 95% confidence intervals overlapped. The FE model identified 31 femurs with a low fracture risk, of which 30 indeed did not fracture (NPV = 0.97). NPV for clinicians were lower and ranged between 0.78 and 0.85.

When the experienced clinicians were asked to base their decision on 30 mm axial cortical disruption (Table V), their diagnostic accuracy values were comparable with the predictions based on clinical experience.

Discussion

Previously, we have shown that FE models calculated the femoral load to failure comparably with those measured in mechanical experiments.[11] In the current study, we

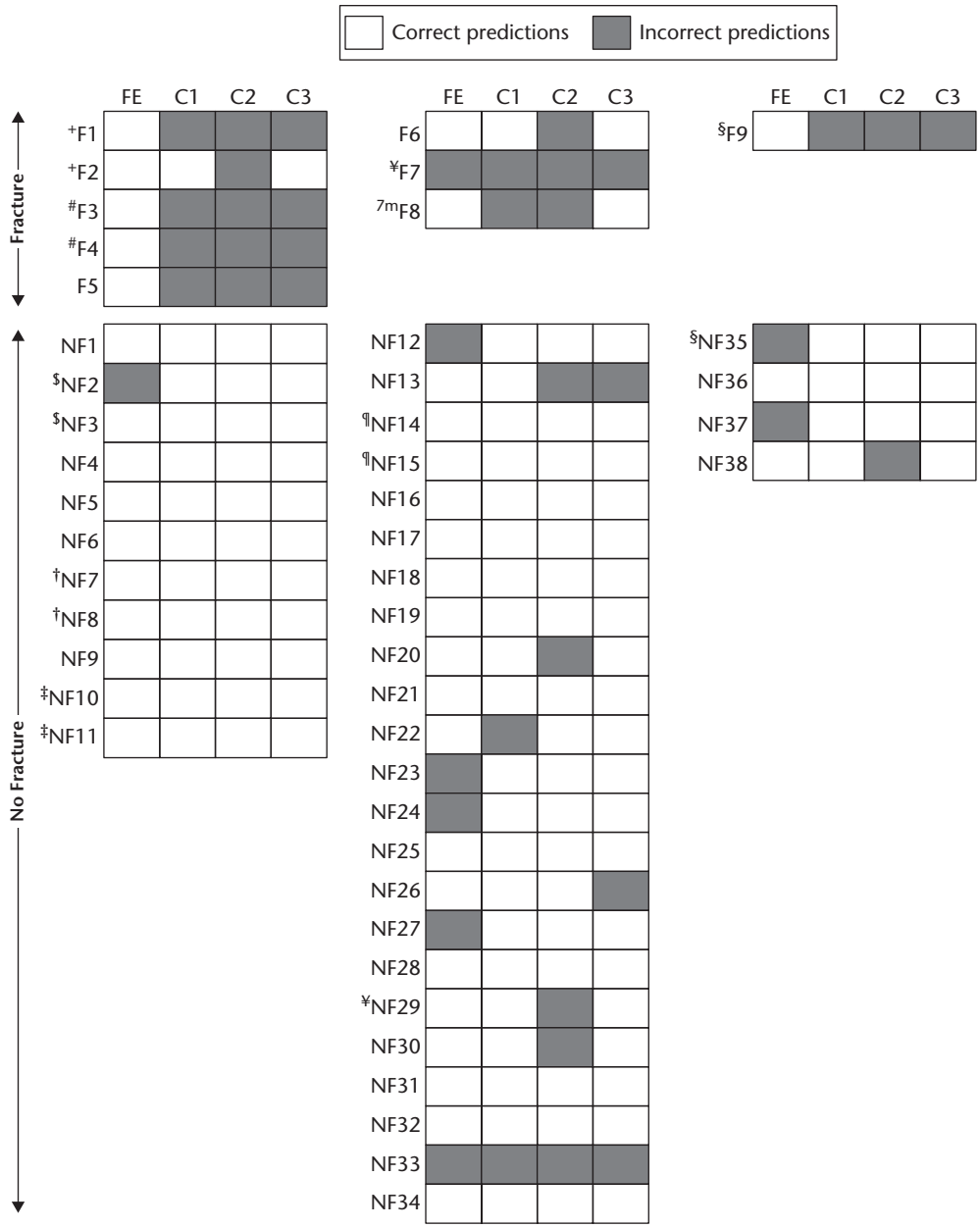

Fig. 4a Fig. 4b Fig. 4c

Correct and incorrect fracture predictions by the finite element (FE) model and the experienced clinicians (C1, C2, C3) for a) institute 1, b) institute 2, and c) institute 3. Clinicians judged the reconstructed radiographs of the patients based on their experience, without any further guidelines prescribed. For the FE predictions thresholds of 7.5 × BW, 6.5 × BW, and 8.0 × BW for institute 1, 2, and 3, respectively, were used to indicate fracture (F) or non-fracture (NF). Results are shown per group (F and NF). Symbols (+, #, $, †, ‡, ¥, ¶, §) indicate paired femurs. F8 fractured one month after follow-up (7m).

applied these FE models *in vivo* by comparing the model predictions with clinical follow-up data in a prospective cohort of patients with cancer and painful femoral metastases who were referred for palliative radiotherapy. We verified whether the model could have predicted the pathological fractures that some of the patients with painful bone metastases unexpectedly sustained during follow-up.

We showed a difference in median failure load between patients who sustained a pathological fracture and those who did not when we analyzed the whole group together as well as in institute 1. However, this difference was not present in the two other institutes, probably because of the low number of fractures. Additionally, for two femurs in the latter institutes, the activity during which the fracture occurred was not recorded. We could therefore not confirm whether these fractures were pathological. If these fractures were traumatic, the high predicted failure loads as calculated by the FE model would have been expected, with improved

Table V. Summary statistics for the prediction accuracy of the finite element (FE) model and the experienced clinicians when relying on their experience and when judging whether cortical involvement was larger than 30 mm.[14] 95% confidence intervals are given between brackets

		F, n	NF, n	Sensitivity	Specificity	PPV	NPV
FE model whole group*	F predicted	8	14	0.89 (0.52 to 1.00)	0.63 (0.46 to 0.78)	0.36 (0.26 to 0.48)	0.96 (0.79 to 0.99)
	NF predicted	1	24				
FE model split to institute†	F predicted	8	8	0.89 (0.52 to 1.00)	0.79 (0.63 to 0.90)	0.50 (0.34 to 0.66)	0.97 (0.82 to 0.99)
	NF predicted	1	30				
Experience							
Clinician 1	F predicted	2	2	0.22 (0.03 to 0.60)	0.95 (0.82 to 0.99)	0.50 (0.14 to 0.86)	0.84 (0.78 to 0.88)
	NF predicted	7	36				
Clinician 2	F predicted	0	6	0.00 (0.00 to 0.34)	0.84 (0.69 to 0.94)	0.00	0.78 (0.76 to 0.80)
	NF predicted	9	32				
Clinician 3	F predicted	3	3	0.33 (0.07 to 0.70)	0.92 (0.79 to 0.98)	0.50 (0.19 to 0.81)	0.85 (0.78 to 0.90)
	NF predicted	6	35				
30 mm cortical involvement							
Clinician 1	F predicted	2	1	0.22 (0.03 to 0.60)	0.97 (0.86 to 1.00)	0.67 (0.17 to 0.95)	0.84 (0.79 to 0.88)
	NF predicted	7	37				
Clinician 2	F predicted	1	6	0.11 (0.00 to 0.48)	0.84 (0.69 to 0.94)	0.14 (0.02 to 0.55)	0.80 (0.75 to 0.84)
	NF predicted	8	32				
Clinician 3	F predicted	0	1	0.00 (0.00 to 0.34)	0.97 (0.86 to 1.00)	0.00	0.80 (0.80 to 0.81)
	NF predicted	9	37				

*Without considering inter-scanner differences
†Prediction of the FE models based on different thresholds for each institute (institute 1: 7.5 × body weight (BW), institute 2: 6.5 × BW, institute 3: 8.0 × BW)
F, fracture group; NF, non-fracture group; PPV, positive predictive value; NPV, negative predictive value

predictions as a result. Nevertheless, the results from institute 1 show that FE models were able to comprehend many factors that contribute to the *in vivo* load capacity of metastatic femurs, such as the bone quality and the bone geometry, or compromise it, such as the location and the size of the lesion. Goodheart et al[10] recently found that FE models can be used to distinguish between metastatic femurs that would and would not fracture. Positive findings were also shown in the field of osteoporosis (e.g. Keyak et al,[21] Kopperdahl et al[22]), where FE strength was found to highly correlate with fracture[21] and FE bone strength remained predictive for fracture after correction for total hip areal bone mineral density (aBMD) in men and women.[22]

In the present study, the FE predictions demonstrated higher sensitivity compared with clinical assessments. This suggests a better identification of patients who will sustain a fracture by the FE model, resulting in prevention of more pathological fractures. Specificity of the FE model was relatively high but slightly lower compared with the clinicians. However, NPV were very high (97%), indicating that if the FE model predicts a low fracture risk, a fracture almost never occurs. As a result, the FE model could be clinically used to prevent unnecessary surgery. Diagnostic values of the 30 mm cortical involvement from a previous study (sensitivity 86%, specificity 58%, PPV 23%, NPV 97%)[5] were quite different from the clinical assessment of the current study (Table V), showing that these values may be dependent on the clinicians and/or the studied patient group.

In six out of eight cases, the predicted fracture location resembled the actual clinical fracture location. In cases F3 and F8, the FE fracture locations did not resemble the clinical fracture lines. F8 suffered from mixed metastases and clinically fractured through a lesion with higher CT density, which could explain why the FE model did not predict the correct fracture location. Since F3 fractured spontaneously, the axial load applied in this study might be inappropriate to simulate the correct fracture line in this femur. Modelling more and realistic loading conditions may further improve the predicted fracture location.

Although the results in this study are promising, some limitations should be mentioned here. First of all, we realize that the sample size in this study is limited, especially after splitting to institute. As a result, in institutes 2 and 3, the critical failure loads were based on only a few fractures (n = 3 and n = 1), indicating the need for larger data sets in the near future. For that purpose, a solution to overcome inter-scanner differences should be developed, something we are currently working on.[23]

A second limitation in this study relates to the modelling of metastatic tissue. We excluded patients with predominant blastic femoral lesions from our analyses, as blastic lesions generally show very high CT intensities. In the current FE model, these CT intensities would have been converted to material behaviour using relationships that are defined based on experiments with human tissue affected by metastases as well as healthy bone.[7] Therefore, the empirical relationships have to be adapted for blastic metastatic tissue. So far, differences in microarchitecture have been described for metastases (e.g. Sone et al[24]), but the mechanical behaviour has not yet been established unequivocally.[25,26] Moreover, adapted material models did not yet improve the predictive power of FE models with metastatic lesions.[27] Hence, further research is required to determine the mechanical behaviour of different types of metastatic tissue.

Third, the use of strain softening as a material property can cause mesh sensitivity and its use to capture localization can be questioned. In the past, our group performed a sensitivity analysis by varying mesh density with or without applying a correction for element size. Based on a fit with experimental results, we decided to use the current mesh density (average element volume of 1 mm^3) without applying the correction. Subsequently, we have been using the same protocol[11] to minimize the differences between bones.

As a fourth limitation, it should be mentioned that the clinicians pointed out that the quality of the DRRs was suboptimal compared with the conventional radiographs they normally use, which may have affected their assessments.

In conclusion, we showed that patient-specific FE models are a potential tool to improve clinical fracture risk predictions in patients with metastatic bone disease. The FE models provided an accurate identification of patients with high fracture risk in one of the three institutes. Future work in a larger patient population should confirm the higher predictive power of the FE models compared with current clinical guidelines. However, a robust solution to overcome inter-scanner differences should be developed before the FE models can be extensively used for clinical fracture risk assessments in a multicentre setting. In the future, the individual FE outcome may help patients and their clinicians to weigh the chance of fracturing against choosing the most appropriate treatment, which is either non-invasive radiotherapy to treat pain, or surgery to restore stability.

Supplementary material

A figure showing the substantially higher failure loads of femurs with blastic lesions in comparison with femurs with lytic or mixed lesions.

References

1. **Coleman RE.** Skeletal complications of malignancy. *Cancer* 1997;80(Suppl):1588-1594.
2. **Gralow JR, Biermann JS, Farooki A, et al.** NCCN Task Force Report: Bone Health in Cancer Care. *J Natl Compr Canc Netw* 2009;7(Suppl 3):S1-S32.
3. **Johnson SK, Knobf MT.** Surgical interventions for cancer patients with impending or actual pathologic fractures. *Orthop Nurs* 2008;27:160-171.
4. **Coleman RE.** Clinical features of metastatic bone disease and risk of skeletal morbidity. *Clin Cancer Res* 2006;12:6243s-6249s.
5. **Van der Linden YM, Dijkstra PD, Kroon HM, et al.** Comparative analysis of risk factors for pathological fracture with femoral metastases. *J Bone Joint Surg [Br]* 2004;86-B:566-573.
6. **Bessho M, Ohnishi I, Matsuyama J, et al.** Prediction of strength and strain of the proximal femur by a CT-based finite element method. *J Biomech* 2007;40:1745-1753.
7. **Keyak JH, Kaneko TS, Tehranzadeh J, Skinner HB.** Predicting proximal femoral strength using structural engineering models. *Clin Orthop Relat Res* 2005;437:219-228.
8. **Lenaerts L, van Lenthe GH.** Multi-level patient-specific modelling of the proximal femur. A promising tool to quantify the effect of osteoporosis treatment. *Philos Trans A Math Phys Eng Sci* 2009;367:2079-2093.
9. **Schileo E, Taddei F, Cristofolini L, Viceconti M.** Subject-specific finite element models implementing a maximum principal strain criterion are able to estimate failure risk and fracture location on human femurs tested in vitro. *J Biomech* 2008;41:356-367.
10. **Goodheart JR, Cleary RJ, Damron TA, Mann KA.** Simulating activities of daily living with finite element analysis improves fracture prediction for patients with metastatic femoral lesions. *J Orthop Res* 2015;33:1226-1234.
11. **Derikx LC, van Aken JB, Janssen D, et al.** The assessment of the risk of fracture in femora with metastatic lesions: comparing case-specific finite element analyses with predictions by clinical experts. *J Bone Joint Surg [Br]* 2012;94-B:1135-1142.
12. **Derikx LC, Vis R, Meinders T, Verdonschot N, Tanck E.** Implementation of asymmetric yielding in case-specific finite element models improves the prediction of femoral fractures. *Comput Methods Biomech Biomed Engin* 2011;14:183-193.
13. **Tanck E, van Aken JB, van der Linden YM, et al.** Pathological fracture prediction in patients with metastatic lesions can be improved with quantitative computed tomography based computer models. *Bone* 2009;45:777-783.
14. **Van der Linden YM, Kroon HM, Dijkstra SP, et al.** Simple radiographic parameter predicts fracturing in metastatic femoral bone lesions: results from a randomised trial. *Radiotherapy & Oncology* 2003;69:21-31.
15. **Koswig S, Budach V.** Remineralization and pain relief in bone metastases after after different radiotherapy fractions (10 times 3 Gy vs. 1 time 8 Gy). A prospective study. *Strahlenther Onkol* 1999;175:500-508.
16. **Karnofsky DA, Burchenal JH.** The clinical evaluation of chemotherapeutic agents in cancer. In: *Evaluation of Chemotherapeutic Agents*. MacLeod CM, ed. New York, New York: Columbia University Press, 1949:191-205.
17. **Healey JH, Brown HK.** Complications of bone metastases: surgical management. *Cancer* 2000;88(Suppl):2940-2951.
18. **Eggermont F, Derikx LC, Verdonschot N, et al.** Limited short-term effect of palliative radiation therapy on quantitative computed tomography-derived bone mineral density in femora with metastases. *Adv Radiat Oncol* 2017;2:53-61.
19. **Carpenter RD, Saeed I, Bonaretti S, et al.** Inter-scanner differences in in vivo QCT measurements of the density and strength of the proximal femur remain after correction with anthropomorphic standardization phantoms. *Med Eng Phys* 2014;36:1225-1232.
20. **Jacobs F, Sundermann E, De Sutter B, Christiaens M, Lemahieu I.** A fast algorithm to calculate the exact radiological path through a pixel or voxel space. *J Comput Inf Technol* 1998;6:89-94.
21. **Keyak JH, Sigurdsson S, Karlsdottir GS, et al.** Effect of finite element model loading condition on fracture risk assessment in men and women: the AGES-Reykjavik study. *Bone* 2013;57:18-29.
22. **Kopperdahl DL, Aspelund T, Hoffmann PF, et al.** Assessment of incident spine and hip fractures in women and men using finite element analysis of CT scans. *J Bone Miner Res* 2014;29:570-580.
23. **Eggermont F, Derikx LC, Free J, et al.** Effect of different CT scanners and settings on bone CT values for finite element predicted failure load [abstract]. *European Society of Biomechanics*, 2016.
24. **Sone T, Tamada T, Jo Y, Miyoshi H, Fukunaga M.** Analysis of three-dimensional microarchitecture and degree of mineralization in bone metastases from prostate cancer using synchrotron microcomputed tomography. *Bone* 2004;35:432-438.
25. **Kaneko TS, Bell JS, Pejcic MR, Tehranzadeh J, Keyak JH.** Mechanical properties, density and quantitative CT scan data of trabecular bone with and without metastases. *J Biomech* 2004;37:523-530.
26. **Kaneko TS, Pejcic MR, Tehranzadeh J, Keyak JH.** Relationships between material properties and CT scan data of cortical bone with and without metastatic lesions. *Med Eng Phys* 2003;25:445-454.
27. **Keyak JH, Kaneko TS, Rossi SA, et al.** Predicting the strength of femoral shafts with and without metastatic lesions. *Clin Orthop Relat Res* 2005;439:161-170.

Funding Statement

- This work was supported by the Dutch Cancer Society (KUN 2012-5591), the Dutch Science Foundation NWO-STW (NPG.06778), and Fonds NutsOhra (1102-071).

Acknowledgements

- The authors would like to thank Wouter Gevers for his help in generating the digitally reconstructed radiographs, and Femke Peters, MD, PhD, and Tom Rozema, MD for their input to this study.

Author Contributions

- F. Eggermont: Contributed to acquisition of data, Analyzing and interpreting the data, Drafting and critically revising the manuscript.
- L. C. Derikx: Contributed to acquisition of data, Analyzing and interpreting the data, Drafting and critically revising the manuscript.
- N. Verdonschot: Contributed to study conception and design, Interpreting the data, Critically revising the manuscript.
- I. C. M. van der Geest: Contributed to study conception, Acquisition of data, Critically revising the manuscript.
- M. A. A. de Jong: Contributed to acquisition of data, Critically revising the manuscript.

- A. Snyers: Contributed to study conception and design, Acquisition of data, Critical revision of the manuscript.
- Y. M. van der Linden: Contributed to study conception and design, Acquisition of data, Interpreting the data, Critically revising the manuscript.
- E. Tanck: Contributed to study conception and design, Interpreting the data, Critically revising the manuscript.
- F. Eggermont and L. C. Derikx contributed equally to this work.

Conflict of Interest Statement
- None declared