

When and where do patients with bone metastases actually break their femurs? A CT-based finite element analysis

Amir Sternheim^{1,2,*}, Frank Traub³, Nir Trabelsi^{4,5}, Solomon Dadia^{1,2}, Yair Gortzak^{1,2}, Nimrod Snir^{2,6}, Malka Gorfine⁷, Zohar Yosibash^{5,8}

¹ National Unit of Orthopaedic Oncology, Tel-Aviv Medical Center, Tel-Aviv, Israel

² Sackler Faculty of Medicine, Tel-Aviv University, Israel

³ Department of Orthopaedic Surgery, University of Tübingen, Germany

⁴ Dept. of Mechanical Engineering, Shamoon College of Engineering, Beer-Sheva, Israel

⁵ PerSimiO, Personalized Simulation in Orthopedics, Inc, Beer-Sheva, Israel

⁶ Orthopaedic Department, Tel-Aviv Medical Center, Tel-Aviv, Israel

⁷ Department of Statistics and Operations Research, Tel-Aviv University, Ramat-Aviv, Israel

⁸ School of Mechanical Engineering, Tel-Aviv University, Ramat-Aviv, Israel

Correspondence: amirst@tlvmc.gov.il

Take home message

- A retrospective in-vivo study on 41 patients with femoral MBD shows that CTFEA predicts the risk of an impending fracture by far better compared to Mirels' score (sensitivity of 100% and specificity of 67%).
- CTFEA predicts well the fracture location.
- The CTFEA is quick and automated and may be easily incorporated into CT scanners protocols.

Abstract (315 words)

Background: Accurate estimations of the risk of fractures due to metastatic bone disease in femurs are essential to avoid both under-treatment and over-treatment in patients with impending pathologic fractures. The purpose of the current retrospective in-vivo study was to utilize CT-based finite element analyses (CTFEA) to identify a clear quantitative differentiator between patients who are at imminent risk of fracturing their femur and those who are not, and to pinpoint the location of maximal weakness where the fracture is most likely to occur.

Patients and Methods: Data was collected on 82 patients with metastatic tumors in their femurs. Forty-one of them did not undergo a prophylactic surgery: 15 had a pathological fracture within 6 months following the CT scan, and 26 patients were fracture free during the five months following the CT scan. Mirels score and strain fold ratio (SFR) based on CTFEA was computed for all patients. A SFR value of 1.48 was used as the determinant of the threshold for a pathological femoral fracture.

Sensitivity, specificity, positive and negative predicted values for Mirels score and SFR predictions were computed on 9 patients who fractured against 24 who did not, as well as a comparison of areas under the receiver operating characteristic curves (AUC of the ROC curves).

Results: Sensitivity of SFR was 100% compared to 88% by Mirels score and specificity of SFR was 67% compared to 38% by Mirels score. AUC was 0.905 for compared to an AUC of 0.578 for the Mirels' score ($p = 0.0078$)

Conclusions: All the study patients who sustained a pathological fracture of the femur had an SFR >1.48 . The CTFEA was by far better compared to Mirels' score to accurately predict the risk of fracture in patients with metastatic tumors to the femur as well as the location of the

fracture. The CTFEA is quick and automated and can be incorporated into the protocol of CT scanners.

Level of Evidence: Level III

Keywords: Bone metastasis, biomechanical bone strength, finite element analysis, impending pathologic fractures

Manuscript 3855 Words.

Introduction

Long bone metastases, especially those involving the femur, are common events among patients with advanced cancer, occurring in up to 70% of them¹. Metastatic bone disease (MBD) is associated with substantial disability, pain, and expense. Currently, there are more than 250,000 new patients with MBD each year in the United States alone, representing an annual cost of 12 billion USD^{2,3}. MBD of the femur may result in pathologic fracture, and it represents a major contributor to the deterioration of the quality of life of patients with cancer. Impending and, even more so, actual pathologic fractures initiate the period of dependent care⁴. Surgery for patients with an impending femur fracture improves short-term survival, morbidity, functional outcome, and length of hospital stay⁵.

Accurate estimations of the risk of fracture and of the expected survival of a patient are essential to avoid both under-treatment and over-treatment in patients with impending pathologic fractures⁶. Clinical methods for identifying patients with femoral MBD who are at high risk of pathologic fractures are limited. Mechanical pain upon weight-bearing which increases over time may be the most important sign. Mirels' score⁷ is determined by location of the lesion, its size, the amount of cortical bone destruction, and mechanical pain. It is used to estimate the risk of a pathologic fracture and the need for prophylactic fixation. Mirels' score has been shown to be reproducible, but its low specificity (35%) may result in unnecessary surgery in two-thirds of the patients^{8,9,10}. In a cohort of 102 patients with painful femoral MBD, the Mirels' score correctly predicted fracture in all 14 patients who sustained a fracture, but it also erroneously predicted a fracture in 84 out of 88 patients⁹. Given its sensitivity, the Mirels' classification is, therefore, more valid as a screening tool⁸ by helping to identify those cases whose management needs to be decided upon. For example, based on Mirels' score, any patient with a proximal femur lesion

(a score of 3), which is not blastic (a score of 2), is somewhat painful (a score of 2), and causes minimal cortical destruction (a score of 1) reaches a final score of 8 and is considered to need prophylactic fixation. A recent study evaluating the validity of Mirels' score system¹¹ in 62 patients with metastatic tumors in their lower limb reported on the following three take home messages: "The Mirels score is neither reproducible nor repeatable. Clinicians should exercise caution when using the Mirels score to inform patient management plans. A new objective, reproducible and repeatable clinical tool is required to predict impending pathological fractures in these patients."

Computerized tomographic (CT)-based finite element analysis (CTFEA) can assess the risk of fracture by creating a patient-specific finite element model of the affected bone, load the bone model with physiological loads typical of daily activity, and compute the femur's mechanical response. It is adjusted to patient's weight, the accurate geometry and nonhomogeneous material properties of the femur, and the geometry and location of the tumor. Recent studies confirm the benefits of these CTFEA to accurately predict the risk of a pathological fracture¹²: *"Our findings indicate that the FE method is useful for the prediction of the pathological fracture. This method shows a versatile potential for the prediction of pathological fracture and might aid in judging the optimal treatment to prevent fracture"*.

Importantly, CTFEAs have been validated by our group ex-vivo^{13, 14} and in a retrospective clinical study in-vivo¹⁵. The whole process is completely automatic and takes one hour to complete. The outcome of the analysis is the tensile and compressive strains (in microstrain) at any location along the femur, which are used to determine the risk of fracture when compared to the typical values in a healthy femur. These results are also compared to those for the patient's healthy contralateral bone that is routinely scanned as well. The algorithm used

to generate the FE model from the CT scan is depicted schematically in Figure 1. This method has been validated against the experimental mechanical loading of twelve femurs in a double-blind study¹⁶ and on fourteen fresh-frozen femurs with real metastatic tumors¹⁴, with a high correlation between the predicted and the observed yield load. A retrospective clinical study by our group on 45 patients with metastatic bone lesions of the femur demonstrated that about 39% of patients who underwent operative fixation based on clinical evaluation and application of the Mirels' criteria may not have required surgery¹⁵.

The purpose of the current retrospective in-vivo study was to identify a clear quantitative differentiator between patients who are at imminent risk of fracturing their femur and those who are not, and to pinpoint the location of maximal weakness where the fracture is most likely to occur. This analysis was conducted on the largest cohort to date, and its findings may serve to improve our ability to identify the patients who may benefit from a preventive fixation.

Methods

Study population

Permission to review the patients' files was given by the local IRB, which waived informed consent for this retrospective study. Data were collected on a total of 82 patients with metastatic tumors in their femurs who were referred to orthopedic consultation during 2014-2018. Patients were included in this study if they had metastatic tumors in their femurs that warranted consideration for prophylactic fixation. Other criteria for study entry were a CT scan of at least one-half of the proximal femur, age twenty years or older (Note: CTFEA has not been validated in the pediatric population before achieving skeletal maturity), and no earlier prophylactic surgery. Patients with an implant in one of their femurs were excluded because such implants

may create artifacts in the CT scan that affects the Hounsfield Units (HU) and thus corrupt the identification of the material properties in the CTFEA. Due to the retrospective nature of this study, the CTFEA results did not influence the tumor board's clinical decision with regard to surgery.

CTFEA methodology

A femur's mechanical response to physiological loads can be predicted by solving the equations of the linear theory of elasticity in the nonhomogeneous femur domain. Using an individual CT scan, a CTFEA is performed following the procedure documented in detail elsewhere.^{9,10} Briefly, the analysis starts by processing the individual Digital Imaging and Communication in Medicine (DICoM) format of a CT scan in which the two femurs are segmented and aligned with the z axis, retaining only those pixels that belong to the femur and discarding the surrounding soft tissue pixels. The HUs of each pixel include information on the local bone density that is correlated to the Young modulus using relationships between the Young modulus and the ash density for cortical¹¹ and trabecular bone¹², validated in⁷. The metastatic tumor is assigned the same material properties, with lytic tumors having a low density that is assigned a low Young modulus. The following relations are used to determine the Young modulus in the femur:

$$\rho_{K_2HPO_4} = 10^{-3} (a \times HU + b) \quad [gm/cm^3] \quad (1)$$

$$\rho_{ash} = 0.877 \times 1.21 \times \rho_{K_2HPO_4} + 0.08 \quad [gm/cm^3] \quad (2)$$

$$E_{cort} = 10200 \times \rho_{ash}^{2.01} [MPa], \quad \rho_{ash} \geq 0.486 [gm/cm^3] \quad (3)$$

$$E_{trab} = 2398 [MPa], \quad 0.3 < \rho_{ash} < 0.486 [gm/cm^3] \quad (4)$$

$$E_{\text{trab}} = 33900 \times \rho_{\text{ash}}^{2.2} [\text{MPa}], \quad \rho_{\text{ash}} \leq 0.3 [\text{g}/\text{cm}^3] \quad (5)$$

The Poisson ratio is set to the constant value of $\nu = 0.3$ ^{7,9}.

A tetrahedral FE mesh is created automatically for each femur that divides it into approximately 4000-5000 finite elements. A physiological load of a magnitude of 2.5 body weight that represents the contact hip force is applied to the femoral head to mimic a stance position loading¹³. The hip contact force is applied on the femoral head and directed along the line connecting its center to the intercondylar region of the distal femur. The load is applied at an angle of 7° to the shaft axis for CT scans that included only the proximal part of the femur.

A linear elastic analysis is then performed by solving the system of generated equations, obtaining the principal strains at any region of interest in the femur, particularly in the regions adjacent to the tumor lesions. These quantitative strain measures are compared to typical values in a healthy bone and to the values at the same locations in the contralateral healthy bone, hence determining the risk of fracture. The flowchart of the CTFEA algorithm is illustrated in Figure 1.

Figure 1 – about here.

Estimation of femur strength and fracture risk

The median tensile or compressive principal strains of 10 regions of the femur (neck, trochanters, proximal shaft, middle shaft, and distal shaft) of disease-free femurs were computed based on 12 femurs as detailed elsewhere¹⁵. The maximum principal strain was computed at all points on the femoral surface: superior neck 2850 μstrain , inferior neck -2750 μstrain , lateral (medial) trochanters 1375 (-2100) μstrain , lateral (medial) proximal shaft 1375 (-2100) μstrain , lateral (medial) middle shaft 1325 (-1850) μstrain , lateral (medial) distal shaft 625 (-1100) μstrain .

The ratio between the absolute maximum principal strain in the diseased femur and the median strain in the same anatomical region of the disease-free femurs documented in Sternheim et al¹⁵

was calculated and labeled the “strain fold ratio” (*SFR*). A *SFR* value of 1.48 was used as the determinant of the threshold for a pathological femoral fracture¹⁵. This threshold was set based on the minimum *SFR* among five patients who experienced a pathological fracture and had a CT scan available prior to the fracture in Sternheim et al¹⁵. The *SFR* was computed for both femurs of each patient. An additional analysis of torsional load was performed for patients who demonstrated a low fracture risk, with fracture risk defined according to a difference larger than 50% between the diseased and healthy femur of same patient. The torsion load is a horizontal load applied on femur’s head perpendicular to the frontal plane. The fully automatic CTFEA generates a report as presented in Figure 2.

Figure 2 – about here.

The Mirels’ score was determined for all patients by experienced orthopedic oncologists. The *SFR* was compared to the Mirels score, and the level of prediction of the fracture location by the CTFEA was investigated.

Estimation of location of pathological fracture

The location at which the highest *SFR* (larger than 1.48) was obtained in the CTFEA was estimated to be the location of the expected pathological fracture (in some cases there were more than one region at which $SFR > 1.48$). Actual fractures detected by X-ray radiographs were compared to the estimated CTFEA to compare the accuracy of the predicted location.

Statistical analysis

Various popular diagnostic measures along with their respective exact binomial 95% confidence intervals were calculated for the *SFR* predictions. Sensitivity (“true positive rate”) was defined as the proportion of femurs with a fracture that were correctly predicted as having

one. Specificity (“true negative rate”) was defined as the proportion of femurs with no fracture that were correctly predicted as not having one. The positive (negative) predicted value was the proportion of true positive (negative) rates out of the total number predicted as positive (negative) rates. Diagnostic accuracy was defined as the proportion of correct results. Youden’s index was the difference between the true positive rate (sensitivity) and the false positive rate (1 minus specificity), and it ranged from -1 to 1 with values closer to 1 if both sensitivity and specificity were high. We also compared the areas under the receiver operating characteristic curves (AUC of the ROC curves) of the Mirels’ score versus the *SFR* based on non-parametric bootstrap and the percentile method. The higher the AUC, the better the model will be at predicting the true fracture status. A value of 0.5 implies that the model has no better prediction ability than a flip of a coin.

Since the study design divides the patients into those who had sustained a fracture that did not warrant prophylactic surgery and those who eventually were fracture free, the sensitivity, specificity, positive predicted value, negative predicted value, and diagnostic accuracy of prediction of an impending fracture based on these data are all 0. The statistical analysis was by the *epiR*, *DTCComPair*, and *pROC* packages of R.

Forty-one patients underwent prophylactic surgery so were excluded from the study, leaving a group of 41 patients. Fifteen patients had a pathological fracture within 6 months following the CT scan, and 26 patients were fracture free during the five months following the CT scan. Data on the demographics of these 41 patients (gender, age, weight, and height), type of lesion, CT findings, and CTFEA results are presented in Appendix Tables A.1 and B.2.

Of the 15 patients who sustained a pathological fracture within 6 months following the CT scan, six were males and nine were females whose average age/weight were 66 years/78.8 kg with a standard deviation of 13.3 years/28.3 kg. The types of metastases were multiple myeloma (n = 4), lung (n = 3), breast (n = 4), prostate (n = 1), transitional cell carcinoma (n = 1), Angiomatoid fibrous histiocytoma (n=1) and sarcoma (n = 1). Pathological fractures occurred at mid-shaft and distal shaft (n = 7), head and neck (n = 5), and close to the lesser trochanter (n = 3). Since the first five patients were used for calibration of the SFR threshold for fracture ¹⁵, they were not included in the statistical analysis to avoid over-optimistic results when evaluating the SFR outcome. Patient #7 (Appendix Table 1) who experienced a pathological fracture had a Mirels' score of 12 and an *SFR* of 9.6 (no residual healthy bone tissue). He sustained a fracture while waiting for surgery and was excluded from the statistical analysis, *leaving a total of nine study patients that sustained fractures and whose data were statistically analyzed.*

The 26 patients who did not undergo prophylactic surgery and did not sustain any pathological fractures within 5 months following the CT scan included eleven males and fifteen females. Their average age/weight were 61.9 years/74.6 kg with a standard deviation of 14.5 years/24.3 kg. The metastases types in this group were breast (n = 9), multiple myeloma (n = 7), renal cell carcinoma (n = 2), lung (n = 1), prostate (n = 1), colon (n = 1), uterus (n = 1), lymphoma (n = 1), nasopharynx (n = 1), and unknown (n = 2). Two patients (#2 and # 13 Appendix Table 1) who did not sustain a fracture in spite of having an SFR of 4 were excluded from the analysis because one was in a wheelchair during the entire period (until his demise) after being diagnosed with a high risk of fracture, and the other had been instructed not to bear weight on the injured limb. *The data of the twenty-four study patients who were fracture-free*

were statistically analyzed. A flowchart that summarizes patients excluded and included in the statistical analysis is presented in Figure 3.

Table 1 presents the demographic characteristics and statistical comparison of the Fracture and Non-Fracture groups. Table 2 presents the true fractures versus the *SFR* and the Mirels' score predictions. Table 3 summarizes the diagnostic measures along with their respective exact binomial 95% confidence intervals. Figure 4 presents the ROC curves and AUC for the *SFR* and Mirels' score respectively. Figures 5 and 6 are descriptive analyses of the *SFR* and the Mirels' score, respectively, for all 38 patients (including the first 5 who fractured and used for calibration of the *SFR* threshold for fracture¹⁵). They show that a perfect prediction would place all the triangles (patients who sustained fractures) above the threshold, and all the circles (patients who did not sustain any fractures) below the threshold.

Figures 3-6 – about here.

The receiver-operating characteristic (ROC) curves for the CTFEA and the Mirels' score where the X axis is 0-100% specificity and the y axis is 0-100% sensitivity showed an area under curve (AUC) of 0.905 for *SFR* with a 95% Confidence Interval (CI): (0.785, 1.0), and an AUC of 0.578 for the Mirels' score with a 95% CI: (0.364, 0.781). The p-value for comparing the two AUCs was 0.0078, where under the alternative hypothesis the true difference in AUC is not equal to 0.

As a tool for predicting fracture status, *SFR* performed far better than the Mirels' score, especially in terms of sensitivity and negative predicted value.

The location of the pathological fracture was correctly predicted for all fifteen patients who fractured their femurs. A typical example of the predicted location and the actual fracture is shown in Figure 7.

Figures 7 – about here.

Discussion

A CTFEA of femurs of patients with a MBD was shown on a cohort of 41 patients that it could clearly delineate between patients who went on to sustain a fracture and those who did not. The *SFR* threshold criterion of 1.48 was reconfirmed on a much larger cohort, and CTFEA was significantly more sensitive than a Mirels' score 8 and above.

Mechanical pain upon ambulation that increases over time may be the best predictor of an impending pathologic fracture related to MBD in the femur. A large lytic lesion in the proximal femur as depicted on plain radiographs is the best radiographic predictor. The Mirels' score is a numerical classification for predicting future fractures, so that tumors with a score of 8 and higher are suggested to undergo a prophylactic internal fixation⁷. Mirels estimated in the same paper that a tumor with a score of 9 had a greater than 33% chance of fracture.

A painful lesion of the proximal femur usually indicates a recommendation for surgery. It is important to note that clinical decisions on the need of a prophylactic fixation takes into consideration the Mirels score, but also consider patient's history as well as physical examination. Thus, there are in our study patients with a Mirels score of 8 and above that did not undergo a prophylactic surgery and experienced a pathological fracture.

The sensitivity of the Mirels' score has several practical limitations, such as the difficulty in detecting small lesions on radiographs even though their locations significantly weaken the bone. The Mirels' score also cannot account for small lytic lesions of the lesser trochanter, small lesions of the femoral neck, and other lesions which are not directly anteroposterior to the lateral plane.

CT provides higher resolution multi-planer imaging than plain radiography. Surgeons and radiologists who assess CTs for mechanical failure should always take into account the weight of the patient since there is a linear correlation to mechanical failure. Pointwise mechanical properties and a strength analysis used for accounting for the risk of fracture is an ability beyond that of the human eye. It is a basic capability needed to quantify bone strength. In our model for this study, we assumed that patients break their femur under normal hip contact daily loads, and that the process begins with mechanical failure followed by breaking of the bone, and only then will the patient fall. Therefore, the finite element models of the femurs were loaded by a hip load of a magnitude equal to 2.5 patients' body weight as experienced by the bone during normal walking. We tested for tension and compression failure and for torsion if they showed normal strains.

CTFEA is a decision-aiding tool which has been validated ex-vivo and in one retrospective clinical study by our group ^{14,15}. The limitations of that study was that most of the patients underwent prophylactic fixation and we could not clinically validate fracture outcomes. Currently, our database of prospectively collected cases includes over 81 analyses. We retrospectively identified our current group of 15 patients who fractured their femurs and had a current pre-fracture CT. Our control groups included 26 patients who had CT-based analyses and did not undergo any surgical intervention and remained mobile and free of fracture over the next 5 months. Since the current analytic process is completely automatic and takes only 1 hour to complete, it provides real-time quantitative insight for clinical determinations of bone strength, the need for protected weight bearing, and the need for prophylactic surgery, making it clinically applicable for patient management.

Our results showed that a *SFR* of 1.48 could clearly delineate between patients who went on to sustain a fracture and those who did not. CTFEA was significantly more sensitive than a Mirels' score 8 and above ($p = 0.025$). The AUC ROC was also significantly different. The *SFR* AUC was 0.905, which is considered a "good" level of accuracy, compared to a Mirels' score AUC which was 0.578, a result that is not much better than chance. The comparison of the AUC for the two methods yielded a significant difference ($p = 0.008$).

These results are comparable to alternative techniques, such as the CT-based rigidity analysis (CTRA)¹⁷. The difference is that the CTFEA is completely automatic, significantly faster, and has fewer technical limitations. Moreover, numerical errors in the CTFEA are controlled whereas the CTRA may include such errors which may not be quantified. In one study that assessed CTRA compared to Mirels' score on 78 patients, the CTRA provided higher sensitivity (100% versus 66.7%), specificity (60.6% versus 47.9%), positive predictive value (17.6% versus 9.8%), and negative predictive value (100% versus 94.4%) compared with the classic Mirels' score (impending fracture risk for score ≥ 9), although there was considerable overlap in the confidence intervals. The ROC curve analysis found the CTRA to be better than the Mirels' score regardless of the impending fracture risk score used in the latter¹⁸. The CTRA, however, has some limitations. Specifically, the analysis can be performed only if one femur is tumor-free, it is a manual procedure that must be performed by an experienced biomedical engineer, it needs calibration phantoms in the clinical CT scan which are not commonly available and, finally, it does not consider the weight of the patient and physiological loads. There are only two recent studies by others that investigated the use of manual CTFEA in a retrospective clinical study. Goodheart et al.¹⁹ reported 38 patients (44 femurs with MBD), among whom there were only five pathological fractures and 28 non-fractures (11 femurs

underwent prophylactic surgery) with a 4-month follow-up. Eggermont et al.²⁰ investigated only patients with lytic tumors (39 patients) that included nine fractures (seven patients) and 32 non-fractures. Their patients were one or two years post-CT, and mostly had a low Mirels' score so that they were not referred to a fixation surgery.

Limitations and future research.

The limitations of this study are that it is partially retrospective in nature. To achieve maximal statistical strength, we used very stringent inclusion and exclusion criteria. The difficulty in finding a group of patients who sustained a fractured femur and also had a recent CT scan of the femur resulted in a small study group. A multicenter randomized prospective study with the CTFEA as a decision-aiding tool is currently underway (35 patients already enrolled) and its findings will contribute to the next level of validation.

The median tensile or compressive principal strains for healthy bone used for the computation of the *SFR* were determined on only 12 femurs. These strains should be tested on a much larger cohort of healthy individuals, and checked to see whether there are any gender-related differences.

In conclusion, all the study patients who sustained a pathological fracture of the femur had an *SFR* >1.48. The CTFEA was able to accurately predict the risk of fracture in patients with metastatic tumors to the femur as well as the location of the fracture. The CTFEA is quick and automated and can be incorporated into the protocol of CT scanners.

FIGURE LEGENDS

Figure 1 – Illustrative flowchart of the consecutive steps of the CTFEA from the retrieval of the appropriate CT scan from the hospital's PACS system to the segmentation step, to the FE mesh generation, to the FE analysis, and to the presentation of principal strains.

Figure 2 – An example of the generated CTFEA report for the surgeon for a patient that has an SFR >1.48 in the right femur at the intertrochanteric region.

Figure 3 – Flowchart that summarizes patients excluded and included in the statistical analysis.

Figure 4 – ROC curves comparing the two fracture prediction tools showing the greatest AUC for the SFR.

Figure 5 – SFRs for patients who eventually sustained a pathological fracture (orange triangles) compared to patients who did not within the 6 months following the CT scan (blue circles).

Figure 6 – The Mirels' score for patients who eventually sustained a pathological fracture (orange triangles) compared to patients who did not within the 6 months following the CT scan (blue circles).

Figure 7 – Typical CT scan of femurs (a), CTFEA with the predicted location at which fracture is expected (b) and X-ray after fracture with the actual fracture (c).

Table 1. Demographic characteristics and statistical comparison of the two groups.

	Fracture (n=15)	Non-Fracture (n=26)	p-Value
Mean age \pm STD(y)	66 \pm 13.3	61.9 \pm 14.5	0.37*
Mean weight \pm STD (kg)	77.8 \pm 28.3	74.6 \pm 24.3	0.70*
Sex (male/female)	6/9	11/15	NA

* Since p-Value is considerably larger than 0.05, the means are not significantly different

Table 2. True fractures versus the strain fold ratio (SFR) and the Mirels' score predictions

		SFR		Mirels' score	
		Yes (>1.48)	No	Yes (8&above)	No
True fracture status	Yes	9	0	7	1*
	No	8	16	15	9
	Total	17	16	22	10*

*A Mirels' score was not available for one patient who sustained a fracture.

Table 3. Statistical measures of the strain fold ratio (SFR) and the Mirels' score predictions

Measure	SFR (CI95%)	Mirels' score (8&above) (CI95%)
Sensitivity	1.00 (0.66, 1.00)	0.88 (0.47, 0.99)
Specificity	0.67 (0.45, 0.84)	0.38 (0.19, 0.59)
Positive predicted value	0.53 (0.28, 0.77)	0.32 (0.14, 0.55)
Negative predicted value	1.00 (0.79, 1.00)	0.90 (0.55, 1.00)
Diagnostic accuracy	0.76 (0.58, 0.89)	0.50 (0.32, 0.68)
Youden's index	0.67 (0.11, 0.84)	0.25 (-0.34, 0.59)

CI, confidence interval

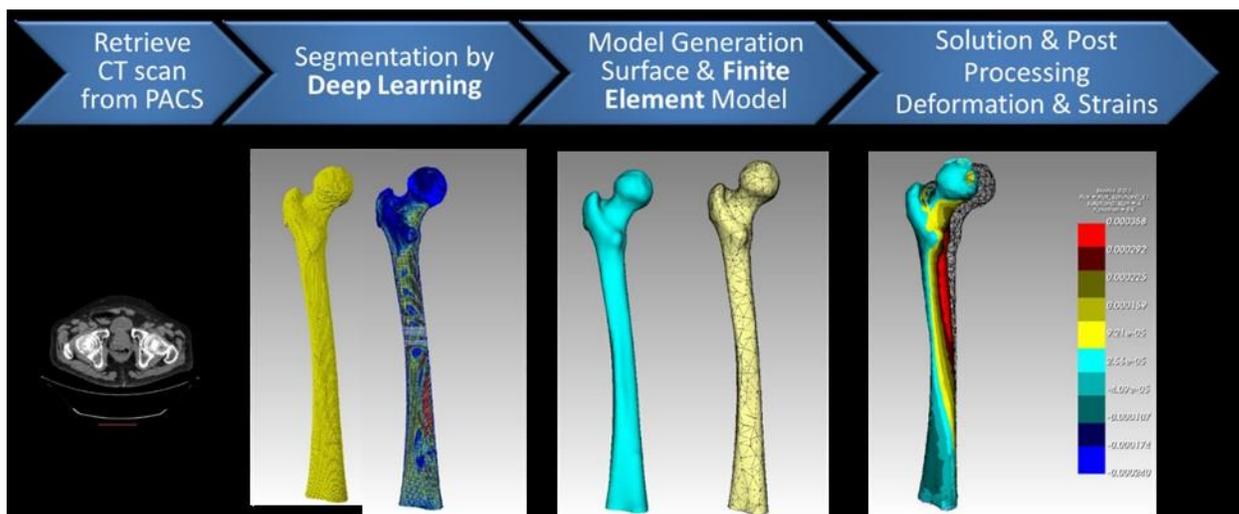


Figure 1 – Illustrative flowchart of the consecutive steps of the CTFEA from the retrieval of the appropriate CT scan from hospital’s PACS system to the segmentation step, to the FE mesh generation to the FE analysis and to the presentation of principal strains.

1 General Details

Analysis Id	190327.101339	Weight[kg]	90
Analysis Date	27/03/2019	Gender	
Ref. Clinician	oded	Birth Date	10/01/1950
Patient ID	ICP10B0007	Age	69
Patient Name	ICP10B014^ICP10B		

2 Location of risk of fracture

Area of interest	Right Femur	Left Femur
Neck	Low	Low
Greater / Lesser Trochanter	High	Low
Proximal Shaft	Low	Low
Middle Shaft	Low	Low
Distal Shaft	Low	Low

The risk of fracture in the right femur is **high**.

The risk of fracture in the left femur is low.

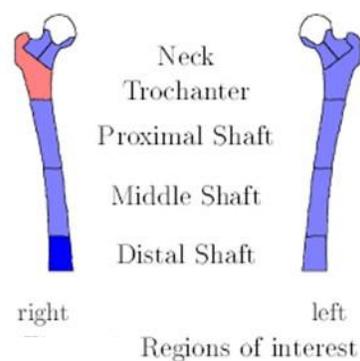


Figure 2 – An example of the generated CTFEA report for the surgeon for a patient that has an $SFR > 1.48$ in the right femur at the intertrochanteric region.

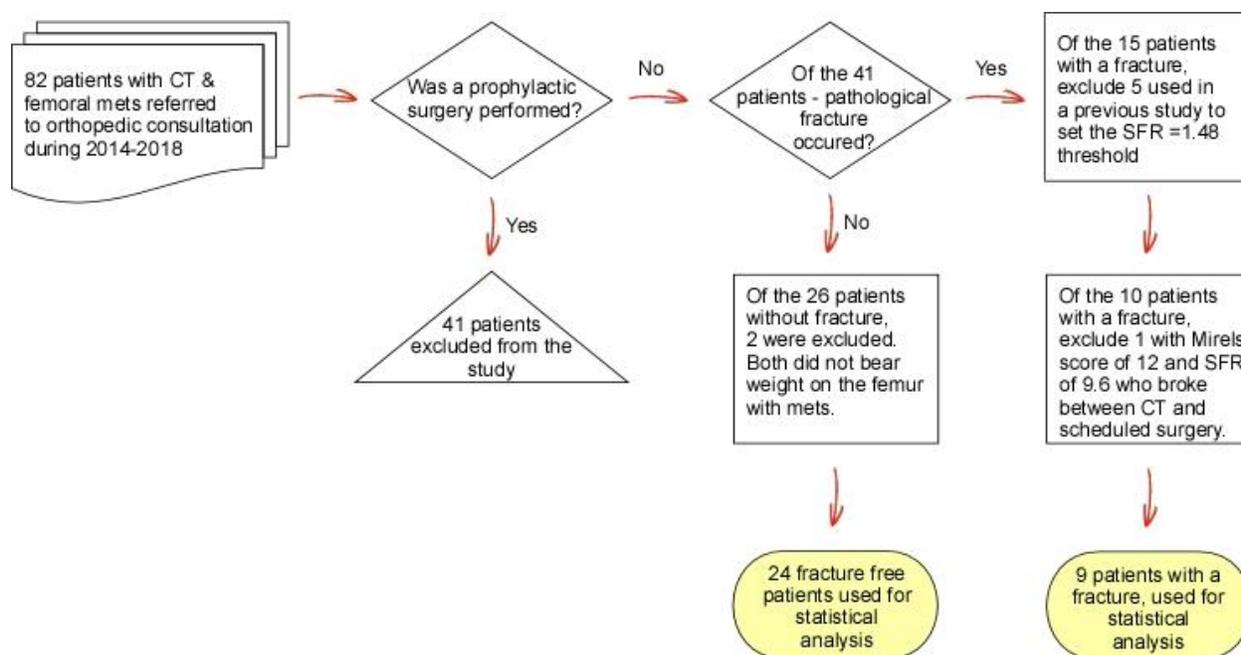


Figure 3 – Flowchart that summarizes patients excluded and included in the statistical analysis.

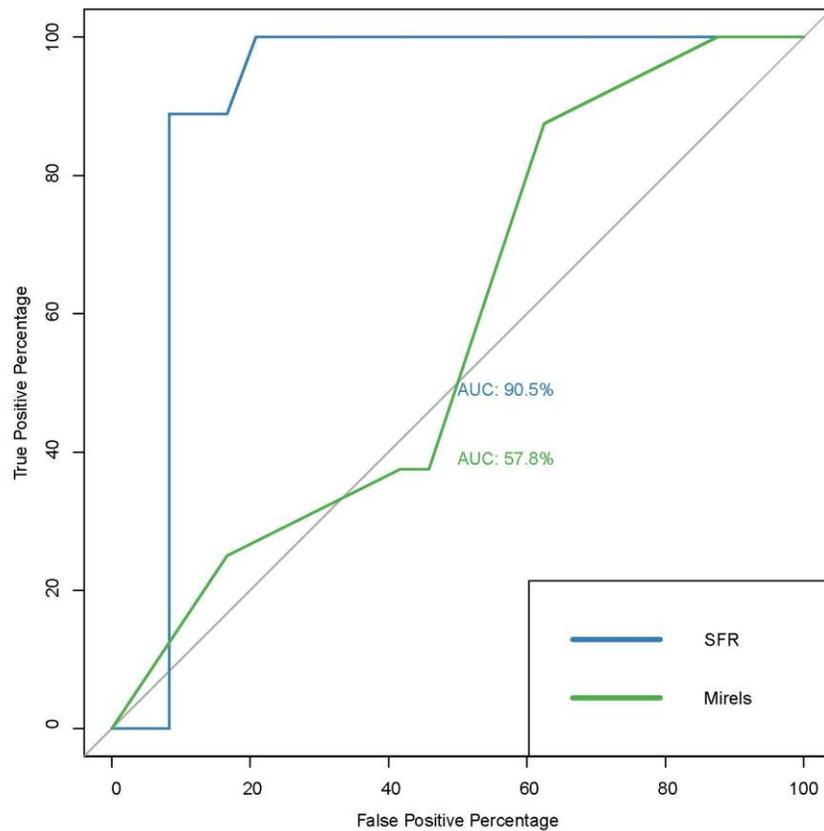


Figure 4 – ROC curves comparing the two fracture prediction tools show the greatest AUC for the SFR.

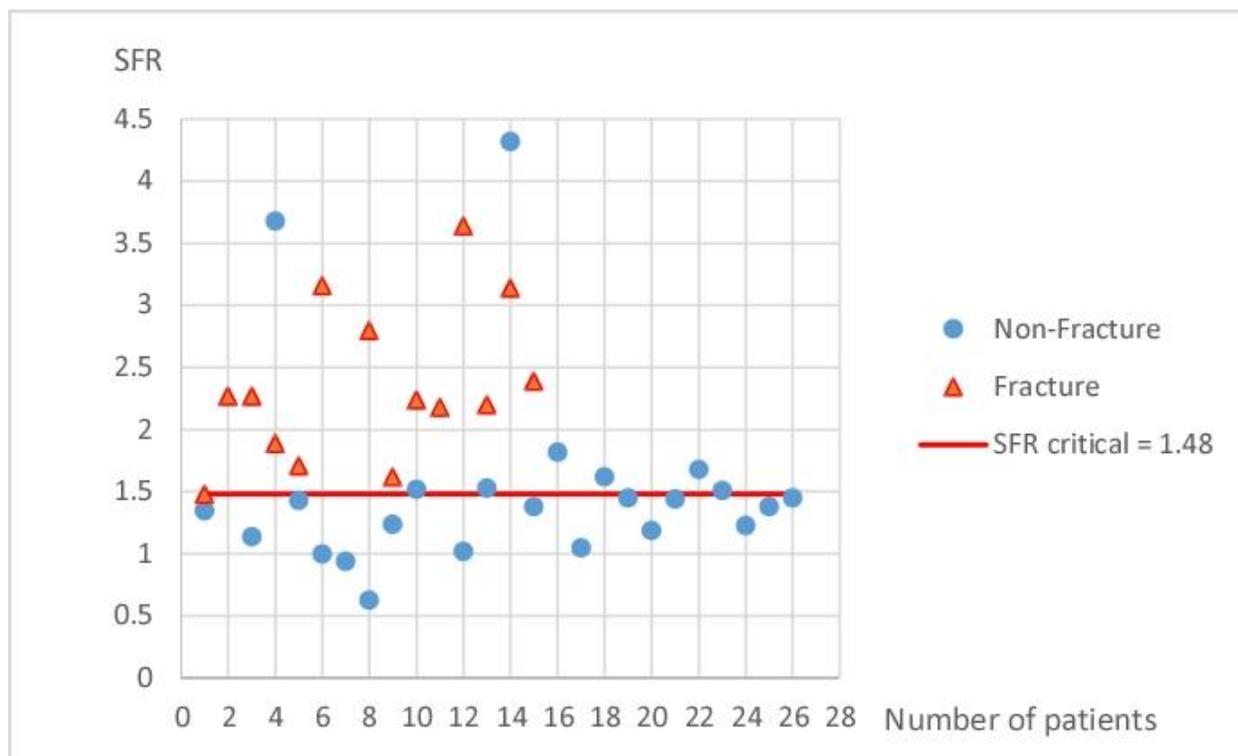


Figure 5 – SFRs for patients who eventually sustained a pathological fracture (orange triangles) compared to patients who did not within the 6 months following the CT scan (blue circles).

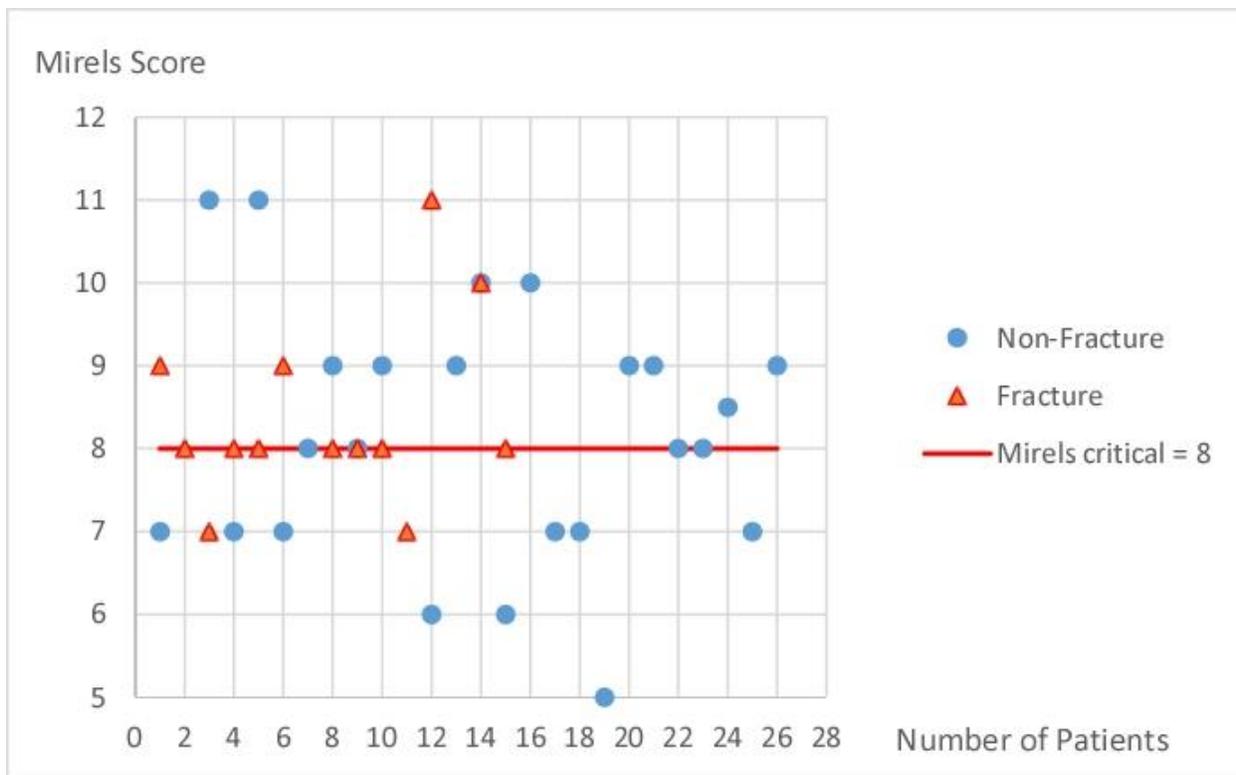


Figure 6 – The Mirels' score for patients who eventually sustained a pathological fracture (orange triangles) compared to patients who did not within the 6 months following the CT scan (blue circles).

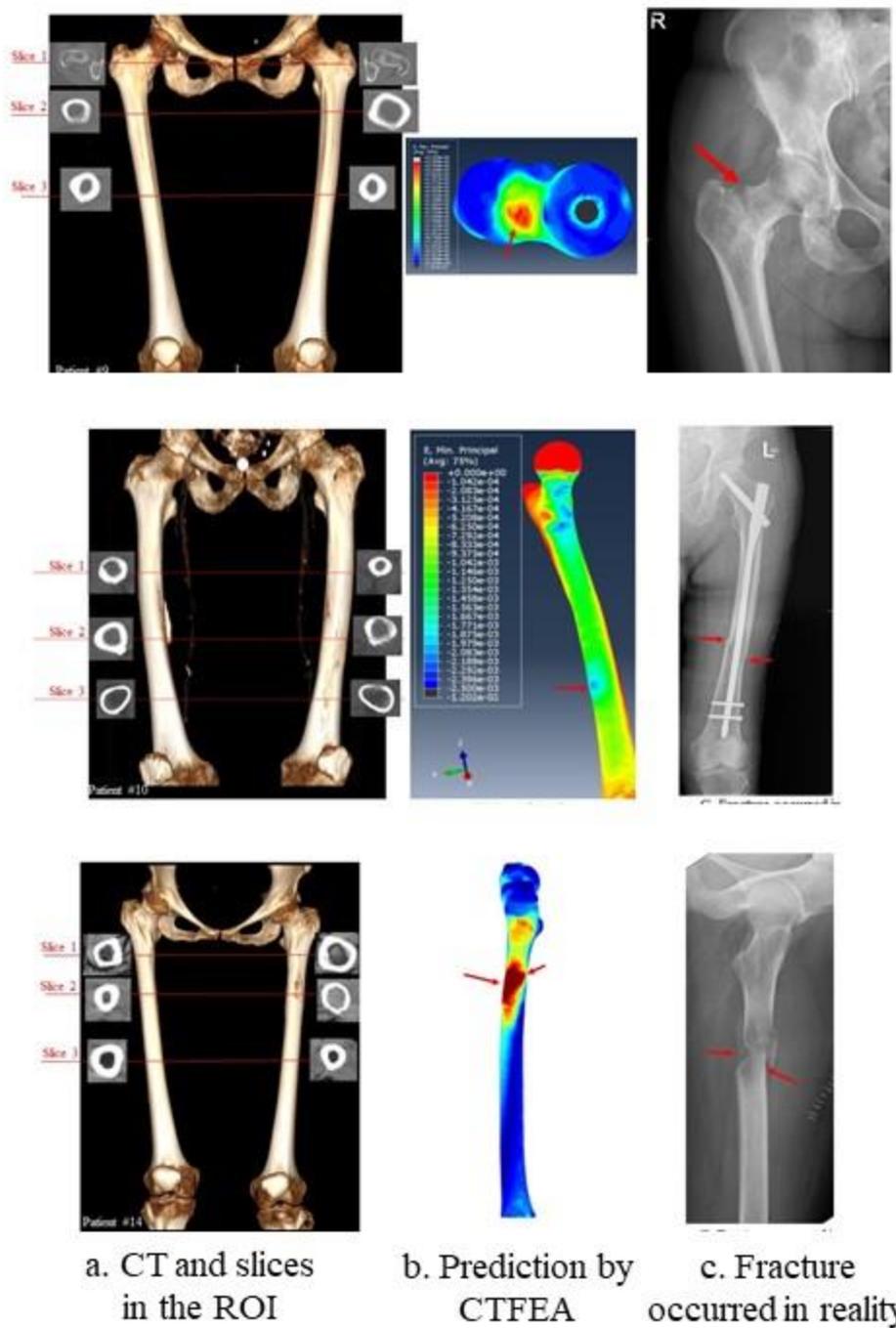


Figure 7 – Typical CT scan of femurs (a.), CTFEA with the predicted location at which fracture is expected (b.) and X-ray after fracture with the actual fracture (c.).

Bibliography

1. Galasko CSB. Monitoring of Bone Metastases. *Schweiz Med Wschr.* 1981; 111:1873-5.
2. Guy GP, Ekwueme DU, Yabroff KR, Dowling EC, Li CY, Rodriguez JL, et al. Economic Burden of Cancer Survivorship Among Adults in the United States. *J Clin Oncol.* 2013; 31:3749-57.
3. Li S, Peng Y, Weinhandl ED, Blaes AH, Cetin K, Chia VM, et al. Estimated number of prevalent cases of metastatic bone disease in the US adult population. *Clin Epidemiol.* 2012; 4:87-93.
4. Bickels J, Dadia S, Lidar Z. Surgical Management of Metastatic Bone Disease. *J Bone Joint Surg Am.* 2009; 91a:1503-16.
5. McLynn RP, Ondeck NT, Grauer JN, Lindskog DM. What Is the Adverse Event Profile After Prophylactic Treatment of Femoral Shaft or Distal Femur Metastases? *Clin Orthop Relat R.* 2018; 476:2381-8.
6. Willeumier JJ, van de Sande MAJ, van der Wal RJP, Dijkstra PDS. Trends in the surgical treatment of pathological fractures of the long bones based on a questionnaire among members of the Dutch orthopaedic society and the European musculo-skeletal oncology society (EMSOS). *Bone Joint J.* 2018; 100b:1392-8.
7. Mirels H. Metastatic Disease in Long Bones - a Proposed Scoring System for Diagnosing Impending Pathologic Fractures. *Clin Orthop Relat R.* 1989; 249:256-64.
8. Jawad MU, Scully SP. In brief: classifications in brief: Mirels' classification: metastatic disease in long bones and impending pathologic fracture. *Clin Orthop Relat Res.* 2010; 468:2825-7.
9. Damron TA, Morgan H, Prakash D, Grant W, Aronowitz J, Heiner J. Critical evaluation of Mirels' rating system for impending pathologic fractures. *Clin Orthop Relat R.* 2003:S201-S7.
10. Benca EP, Janina; Mayr, M. Winfried ; Pahr, Dieter H. ; Windhager, Reinhard The insufficiencies of risk analysis of impending pathological fractures in patients with femoral metastases: a literature review. *Bone Reports.* 2016; 5:51-6.
11. Howard EL, Shepherd KL, Cribb G, Cool P. The validity of the Mirels score for predicting impending pathological fractures of the lower limb. *The Bone & Joint Journal.* 2018; 100-B:1100-5.
12. Kawabata Y, Matsuo K, Nezu Y, Kamiishi T, Inaba Y, Saito T. The risk assessment of pathological fracture in the proximal femur using a CT-based finite element method. *J Orthop Sci.* 2017; 22:931-7.
13. Yosibash Z, Trabelsi N, Milgrom C. Reliable simulations of the human proximal femur by high-order finite element analysis validated by experimental observations. *J Biomech.* 2007; 40:3688-99.
14. Yosibash Z, Mayo RP, Dahan G, Trabelsi N, Amir G, Milgrom C. Predicting the stiffness and strength of human femurs with real metastatic tumors. *Bone.* 2014; 69:180-90.
15. Sternheim A, Giladi O, Gortzak Y, Drexler M, Salai M, Trabelsi N, et al. Pathological fracture risk assessment in patients with femoral metastases using CT-based finite element methods. A retrospective clinical study. *Bone.* 2018; 110:215-20.

16. Trabelsi N, Yosibash Z, Wutte C, Augat P, Eberle S. Patient-specific finite element analysis of the human femur--a double-blinded biomechanical validation. *J Biomech.* 2011; 44:1666-72.
17. Nazarian A, Entezari V, Zurakowski D, Calderon N, Hipp JA, Villa-Camacho JC, et al. Treatment Planning and Fracture Prediction in Patients with Skeletal Metastasis with CT-Based Rigidity Analysis. *Clin Cancer Res.* 2015; 21:2514-9.
18. Damron TA, Nazarian A, Entezari V, Brown C, Grant W, Calderon N, et al. CT-based Structural Rigidity Analysis Is More Accurate Than Mirels Scoring for Fracture Prediction in Metastatic Femoral Lesions. *Clin Orthop Relat R.* 2016; 474:643-51.
19. 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; 33:1226-34.
20. Eggermont F, Derikx LC, Verdonschot N, van der Geest ICM, de Jong MAA, Snyers A, et al. Can patient-specific finite element models better predict fractures in metastatic bone disease than experienced clinicians?: Towards computational modelling in daily clinical practice. *Bone Joint Res.* 2018; 7:430-9.

Appendix A

Table A.1 - Summary of patients' data who sustained a fracture

#	CT date	M / F	Age	Weight [kg]	Cancer type	R/L	Part of Femur in CT	Mir els	SFR	Fx	Scanner	Kvp	Filter
1	30.9.13	F	59	68	MM	R	166 mm Proximal	9	1.48	A month after CT	Philips-59640ed	140	YD-Bone
2	6.10.13	M	75	75	MM	L	Entire	8	2.27	On 16.10.13 less than 2 weeks after CT	Philips Briliance 64	120	Bone-D
3	21.6.16	F	68	62	MM	L	Entire	7	2.27	2 days after CT. Did not predict by MDs	Philips Briliance 64	120	B
4	10.10.16	F	54	67	Adenocarcinoma intrauterine	L	Entire	8	1.89	??	Philips Ingenuity Core 128		A
5	22.2.17	M	62	90	Lung	R	10-15 Prox	8	1.71	Within 4 weeks after CT	Philips Core Ingenuity	120	A
6	18.12.17	F	73	93	MM	R&L	Entire	9	3.16	6 months after CT	Philips IDT16	140	A
7	28.12.17	M	43	110	TCC	L	152mm	12	9.6	Day after CT	Philips iCT256	120	B
8	3.10.18	F	65	77	Breast	L	Entire	8	2.8	A week after CT	GE Discover RT	120	Body

9	29.8.18	F	57	60	Breast	L	Entire	8	1.62	About 6 weeks after CT	Philips Ingenuity Core 128	120	A
10	23.05.18	M	94	62	Prostate	L	Entire	8	2.24	2 days after CT	Siemens Somaton	100	FLAT / Bf40d/2
11	26.6.18	M	58	65	Lung	L	Prox up to 2 cm below lesser	?	1.75	6-7 months after CT	Siemens	120	I30f\1
	29.9.18						Prox up to 2 cm below lesser	7	2.18	4 months after CT	Siemens	100	I30f\1
12	31.12.18	M	79	72	Lung	R	Prox up to 3-4 cm below lesser	11	3.64	Couple of days after CT	Siemens	100	B30/f
13	30.12.18	F	83	61	Breast	R	Entire	?	2.2	Couple of days after CT	Siemens Somaton	100	FLAT / Bf32d
14	23.11.16	F	69	75	Breast	R	Entire	10	3.14	2 months after CT	Philips iCT256	120	B
15	11.8.19	F	46	54	Angiomatoid Fibrous Histiocytoma	L	Entire	8	2.39	Couple of days after CT	Philips	120	B

First 5 gray patients were used for calibration. Patient 7 was predicted by the orthopedic surgeon to have a pathological fracture.

Appendix B

Table B.2 - Summary of patient data who did not fracture

#	CT date	M/ F	Age	Wei ght [kg]	Cancer type	R/L	Part of Femur in CT	Mire ls score	SFR	Scanner	Kvp	Filter
1	?	M	89	80	Prostate	L	187 mm SHAFT only	7	1.35	Philips iCT256	120	B
2	15.1.2015	M	80	75	MM	R	Entire	11	4	Philips iCT256	120	B
3	6.1.2016	F	74	73	Breast	L	Entire	11	1.14	Philips Brilliance 64	120	A
4	6.11.2015	M	71	170	Renal Cell Carcinoma (RCC)	L	Entire	7	3.68	Philips Brilliance 64	120	B
5	15.1.2015	F	42	63	Breast	L	Entire	11	1.43	Philips Brilliance 64	120	C
6	22.7.2016	F	39	70	Breast	L	Entire	7	1	Philips Brilliance 64	120	B
7	20.1.2017	M	64	65	Multiple Myeloma	R	Entire	8	0.94	Philips Brilliance 64	120	A
8	8.3.2017	F	64	78	Multiple Myeloma	R	Entire	9	0.63	Philips iCT256	120	A
9	11.9.2017	M	59	103	TCC urine?	L&R	Entire	8	1.24	Philips Brilliance 64	120	B
10	27.11.2017	F	57	40	Breast	R&L	Entire	9	1.52	GE Discovery CT750 HD	120	Body
11	10.1.2018	M	64	68	Lung	L	Entire	9	4	Philips IDT16	120	B
12	14.1.2018	F	53	70	Breast	L?	Entire	6		Philips iCT256	120	B
	2.5.2018									GE Discovery 690	120	
	8.8.2018			72					1.02	Philips Brilliance 64	120	A

13	14.12.2016	F	58	72	?	Both	Entire	?	1.53	Siemens Gemini TF TOF 16	120	B
14	29.4.2018	F	??	90	MM	R	Entire	?	4.32	Philips Brilliance 64	120	B
15	6.6.2018	F	54	78	Breast	R	Entire	6	1.38	Philips Brilliance 64	120	A
16	19.6.2018	F	87	65	RCC	L	Entire	10	1.82	Philips Ingenuity	120	B
17	17.7.2016	M	72	45	Gastric	R	Entire	7,8	1.05	Philips Ingenuity Core	120	C
18	26.3.2018	F	54	68	Lymphoma	L	Entire	7	1.62	GE Optima CT660	120	BODY FILTER/ STANDARD
19	30.4.2018	M	32	68	Nasopharynx Carcinoma	L	Entire	5	1.45	GE Discovery RT	120	BODY FILTER/ STANDARD
20	25.4.2018	F	73	46	MM	R/L	Entire	8	1.19	Philips Ingenuity	120	B
21	14.10.2018	M	68	92	Prostate	R	Entire	9	1.44	Philips IDT1 6	140	A
22	28.10.2018	M	71	76	MM		Entire	8	1.68	Philips Brilliance64	120	A
23	3.12.18	F	61	70	Breast	L	Entire	8	1.51	Philips Brilliance64	120	A
24	31.12.2018	F	42	76	Breast	R	Entire	9,8	1.23	Philips Iqon Spectral	120	B
25	12.5.2015	M	72	90	MM	R	Entire	7	1.38	Philips iCT256	120	B
26	28.4.2017	F	47	49	Breast	R	Entire	9	1.45	Philips iCT256	120	A