Autonomous FEs (AFE) - A Stride Toward Personalized Medicine

Zohar Yosibash[a,b,1], Kent Myers[b], Nir Trabelsi[c,b], Amir Sternheim[d]

[a]*School of Mechanical Engineering, The Iby and Aladar Fleischman Faculty of Engineering, Tel-Aviv University, Ramat-Aviv 69978, Israel*
[b]*PerSimiO Ltd, Beer-Sheva, Israel*
[c]*Department of Mechanical Engineering, Shamoon College of Engineering, Beer-Sheva, Israel*
[d]*Sourasky Medical Center, Tel-Aviv, Isael*

Abstract

Finite element analysis (FEA), introduced more than half a century ago, requires a (qualified) analyst to generate the necessary input data, verify the output and post process the analysis results for a meaningful conclusion. The required expertise and labor efforts have precluded the use of FEA in daily medical practice by orthopedic surgeons for example.

Patient-specific analyses of the mechanical response of human bones may have a tremendous impact in clinical practice should they be easily accessible by orthopedic surgeons. Recent scientific advancements such as low dose CT scans, machine learning, and high order FEA which facilitates an inherent methodology for assessing numerical accuracy allow a fully autonomous process for assessing bone strength and fracture risk. This autonomous process, that we shall refer to here as "Autonomous Finite Element" (AFE) analysis, introduces a paradigm shift in the use of FEA.

We shall describe herein a novel process that utilizes AFE to produce a patient-specific assessment of bone strength. The process consists of an automatic segmentation of femurs from CT-scans by convolution neural networks, an automatic mesh generation and application of boundary conditions based on anatomical points, a high-order FE analysis with numerical error control, and finally an automatic report with a clear assessment of bone fracture risk. One specific application of AFE is the determination of the risk of fracture for patients with tumors of the femur and whether a prophylactic surgery is needed. To the best of our knowledge this is the first CE accredited AFE application being used by orthopedic surgeons in clinical practice.

Keywords: p-FEA, femur, CNN, V&V

[1]Corresponding author: yosibash@tauex.tau.ac.il

1. Introduction

Development of finite element methods (FEMs) to address structural mechanics problems started more than half a century ago. Real life applications emerged in the mid 1950's [1] and expanded enormously in the following 60 years. These days it is difficult to think of a mechanical or biomedical product that does not use FEMs during its development cycle. Although the expertise and effort required to perform a basic linear elastic finite element analysis (FEA) of a complex structure has decreased considerably during the last few years, the technology still requires an expert analyst at some point during the process.

Several technologies such as machine-learning algorithms, low dose computer tomography (CT) scans, automatic mesh generators and high order finite element methods [2] have emerged in the past two decades that allow a paradigm shift in the use of FEA. Similar to the well-known autonomous[2] vehicle which can drive itself without the assistance of a driver, we herein introduce the *Autonomous FE (AFE)* concept, a FEA that can be used without the assistance of an analyst. We describe the components of an AFE and demonstrate its use in clinical practice to determine the risk of fracture in a patient's femurs. AFE is possible because we are customizing the individual technologies to a fairly narrow specific problem and a specific goal: assessing bone strength in femurs. One may view AFE as an integration of technologies, including FE, that is used to analyze a specific problem without human intervention or expertise related to the component technologies. Just as an autonomous vehicle must be thoroughly validated before it can be safely used and trusted as a routine means of transportation, so must the components of AFE be rigorously verified to be accurate and reliable. After all components are verified, they must be validated. Validation in our case requires that predictions made by the AFE are supported by clinical outcomes on real patients.

The AFE described herein will satisfy an important unmet need by providing a quantitative assessment of fracture risk for patients who suffer from diseases that compromise bone strength. Although FEA for the simulation of the mechanical response of human femurs had started to appear in the scientific literature more than 40 years ago (2D FEA in [3, 4] and 3D FE models of the femur in [5, 6]), none have made their way into clinical practice. FE results from bone analyses were not sufficiently verified and validated, so it was impossible for medical professionals lacking engineering knowledge to successfully apply FE methods to individual patient cases.

In this manuscript we describe *Simfini*, an AFE developed to provide the orthopedic oncologist with a

[2]The word autonomous is borrowed from Greek as the combination of two words "autos + nomos" which means "self + govern".

tool for assessing fracture risk in patients who have metastatic tumors of the femur. In our view, providing a medical professional with an automatic, verified, and validated FE technology for patient specific bone strength assessment represents a paradigm shift.

In this application of AFE, the following components are integrated (see Figure 1):

a) Image analysis of CT scans to segment the femur automatically (identify the femur in the CT scan, determine the femur/pelvis interface which is unclear in most of the CT scans), i.e. determine which voxels in the CT scan belong to the femur,

b) Pointwise estimation of the inhomogeneous mechanical properties of the bone by analysis of image density in the CT scan,

c) Application of personalized physiologic boundary conditions (loads and constraints) that simulate daily activities or falls on the side,

d) Automatic generation of a finite element mesh (with curved faces) to represent the bone of interest,

e) An efficient high-order FE algorithm that solves the system of finite element equations and generates the data of interest,

f) A verification algorithm that quantifies and controls numerical errors within specified limits,

g) A post-processing algorithm that analyzes the finite element solution and automatically produces a report with the most relevant data for the medical doctor (MD),

h) A clinical study to validate the AFE recommendations by comparing them to measurable data.

Generation of the FE model and the application of the boundary conditions are the most difficult to automate in order to eliminate human interaction. Components a) and c) which have been only recently developed, will be described in Section 2. The other components of *Simfini* will also be briefly discussed in Section 2. Validation of *Simfini* in clinical practice is demonstrated in Section 3 by describing its use in the identification of patients with a high risk of pathological fracture due to femoral tumors.

2. Components of the AFE

An important feature required for an AFE is an automatic procedure that will accurately segment the domain of interest, i.e. the femur, from a CT scan. Recent developments in deep learning and particularly

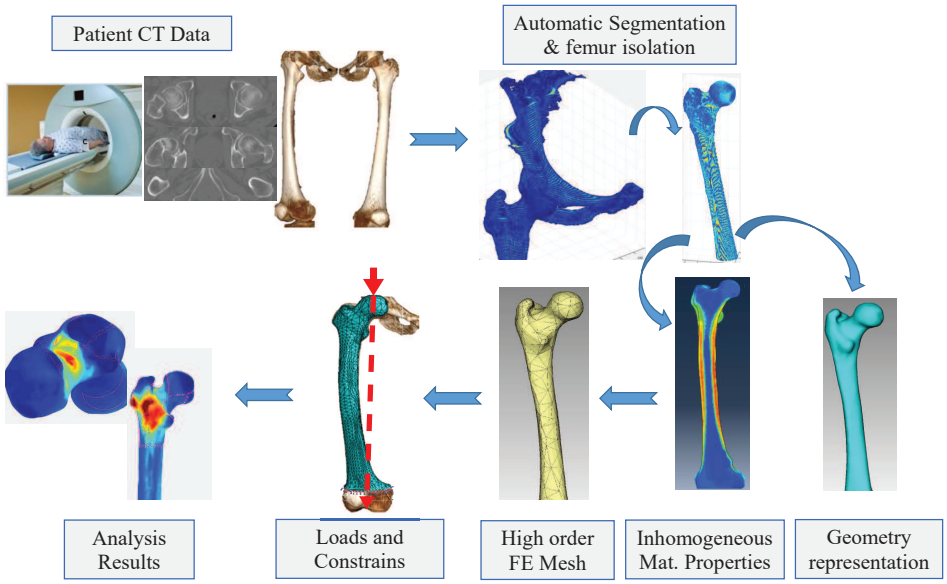


Figure 1: Schematic diagram of the components of the AFE *Simfini*.

convolution neural networks, provide the machinery to facilitate an accurate automatic segmentation of the femur's domain.

2.1. Automatic segmentation of patient's femurs from CT scans

Segmenting a femur from CT scans is the process by which the bone volume of interest is isolated from the rest of the image. This segmentation process poses difficulties for two main reasons. First, the cortical bone is not dense enough through the femur to be distinguishable from the tissue outside the bone. Second, the layer of cartilage that separates the femur from the pelvis is sometimes not visible, so the femur and pelvis appear as connected (Figure 2-right).

We consider clinical CT scans typically consisting of several hundreds of 2D slices taken at constant distances apart. These slices contain gray levels measured by the Hounsfield scale [7]. The Hounsfield units (HU) represent the density of the material, ranging from -1000 (air) to 3000 (metals). The algorithm segments one femur at a time, so a cropped volume consisting of a quarter of the full scan is considered on which initial segmentation by thresholding is performed. Pixels representing bone tissue (HU $>$ 150) are assigned a value of 1 while non-bone is assigned a value of 0, resulting in a binary mask.

Following the initial segmentation, a search for the largest connected component in the volume is performed, by a Flood-Fill algorithm [8] with a 9×9 neighborhood matrix. All voxels not a part of the

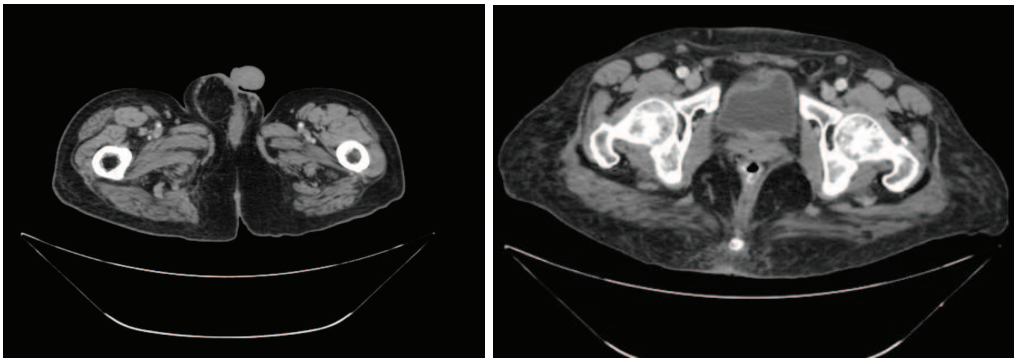


Figure 2: Left - A slice in a CT scan of a femur shaft. Right - CT scan of femoral head. The femoral head and the pelvis seem to be connected.

largest connected component are discarded by assigning a 0 to the associated voxels. The femur with the pelvis bone (still attached to it at this stage) as the largest hard tissue in the scan, is classified as "the object" (Figure 3).

Distal bone volume contains the shaft for which thresholding alone produces a satisfactory segmentation. At this stage, the proximal region of the femur seems to be connected to the pelvis. Thus the next step is the partition of voxels marked by 1 into a distal and proximal part (called the complex volume) in order to separate the proximal part from the pelvis. This partition is realized by identifying the CT slice above which the largest connected component is found. In the proximal part of the CT scan, convolutional neural networks (CNNs) are applied to segment the femur, i.e. tagging each voxel into one of three classes: femur, pelvis, or soft tissue. In such a CNN, convolution and max pooling actions reduce the dimensions of a CT image and classify all or parts of the image. The kernels of the convolution activities are determined by optimization. The CNN used is of typical architecture - a convolution layer, followed by a maxpooling layer followed by a layer with an activation function [9]. The convolution activities are repeated several times. During each repetition, the images become smaller, and the number of features increases (Figure 4). Finally, the features are mapped with a fully connected layer to an array of size 1024. From it, with another fully connected layer, the values are mapped to an array of size 3, the number of components we want to classify (femur, pelvis, other tissue). Before initiating the CNN algorithm, the number of voxels to be classified is increased by using a dilation morphological operator [10] on the binary mask. The structuring element (see [10]) used is a sphere with radius of 6 voxels. The dilation process extends the binary mask to include the femur boundary.

CNNs used for bone segmentation (especially femurs) were shown by the authors to perform very well by using three classifiers, two different networks of different volume sizes and applied only to the proximal

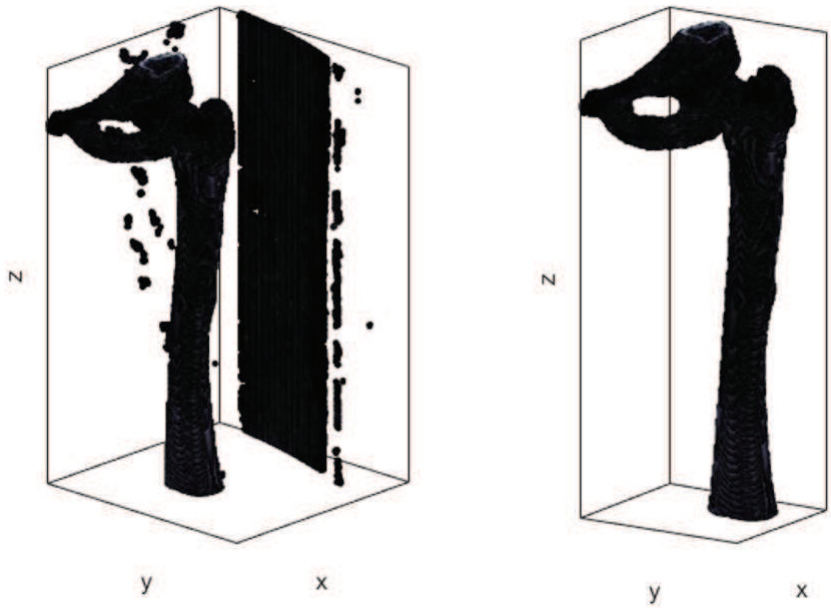


Figure 3: Initial thresholding and largest connected component. Left: Initial thresholding. The red parts are voxels with HU higher than 150. Right: The scan after thresholding and largest connected component finding. Small parts that are not connected to the largest connected component tagged as non-bone.

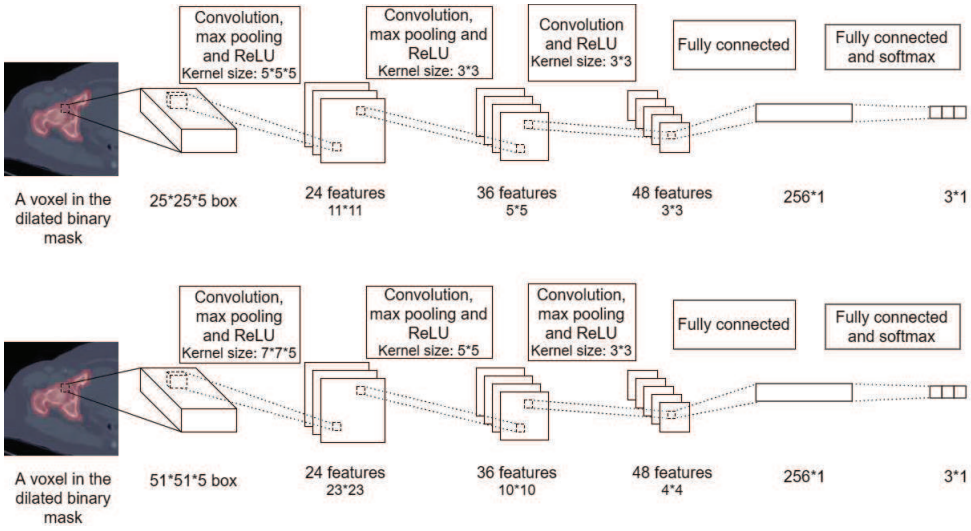


Figure 4: CNNs architecture.

part of the femur [11]. Later publications using different CNN architecture and algorithms show similar, although slightly less accurate results, see for example [12] in which CNNs are applied to segment the proximal femur from MRI scans.

2.1.1. Data used in the CNN algorithm

For each voxel in the dilated mask, the voxels surrounding it that are within box-shaped neighborhoods (centered at the current voxel) are the inputs to the CNNs. We use two boxes of different sizes $25 \times 25 \times 5$ and $51 \times 51 \times 5$ (voxels). The reason for creating two boxes is because each describes different properties of the voxel that should be classified. The smaller box classifies the voxels according to close neighborhood properties, while the large box considers global properties of the voxel neighborhood (location and its context within the whole bone). In both boxes, the size of each voxel is $1 \times 1 \times 2\ mm^3$, i.e., the volume of the smaller box is $25 \times 25 \times 10\ mm^3$ (for **Net25**), and the largest box volume is $51 \times 51 \times 10\ mm^3$ (for **Net51**). In this way, each box contains a known number of voxels representing a known physical volume.

Each network calculates the probability that a specific voxel belongs to either femur, pelvis or soft tissue. Each voxel is tagged to the class with the highest probability of the 6 probabilities (one for each tissue in each CNN) that were calculated.

Finally, additional minor smoothing operations are performed to obtain the final segmented femur (Figure 5). The smoothing is required because each voxel is classified by itself regardless of the neighboring voxels classification, and therefore a sharp curvature can be obtained in the resulting binary mask.

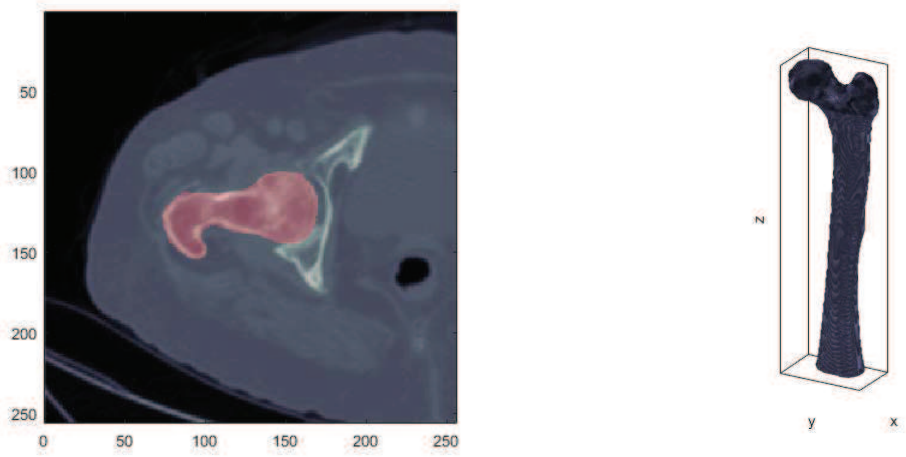


Figure 5: A typical slice of segmented bone (left): the red mask is the automatic segmentation, and the whole segmented bone (right).

2.1.2. Training

The CNNs were trained on 70 manually segmented femurs from 39 different patients. All CT scans were obtained from Philips scanners (model names: Brilliance 64, Ingenuity Core 128, iCT 256, Mx8000 IDT 16) performed at different hospitals. In each CT of the training set, the voxels belonging to femur, pelvis and other tissues were manually tagged. The training set contained about 15 million voxels with their tags and their neighborhoods. The networks were trained using a mini-batch [13] method with Adam optimizer [14] for 50 epochs. Each mini-batch contained 2000 boxes.

2.1.3. Verification of segmentation accuracy

The accuracy of the automatic segmentation was checked on 41 bones (which did not participate in the training process), of male and female, young and old patients, by comparison to manual segmentation. The quantitative measures for accuracy are the dice similarity coefficient (DSC) also known as the Sørensen-Dice Coefficient [15] and Average Surface Distance (ASD) [16]. DSC checks the overlap between the objects, using a range [0,1], where 1 is obtained for two completely overlapped volumes. Having two points sets P, R, their DSC index will be calculated as

$$DSC(P,R) = \frac{2|P \cap R|}{|P|+|R|}.$$

ASD is a two tuples metric, based on Hausdorff Distance (HD). HD metric measures "distance" between two subsets of metric space. Let $P, R \subseteq B$ where B is a metric space, for each $p \in P$ find the point $r \in R$ which is the closest point to p, the distance $d(p,r)$ is computed. HD is the maximal distance obtained in this way. The calculation is not symmetric, therefore, the opposite calculation should also be calculated:

$$ASD(P,R) = \frac{1}{2}\left[\frac{1}{|P|}\sum_{i=1}^{|P|}\min_{r\in R}\{d(p_i,r)\} + \frac{1}{|R|}\sum_{i=1}^{|R|}\min_{x\in P}\{d(p,r_i)\}\right].$$

The results of the index are in the range $[0,\infty)$. A small ASD index means closeness and high matching between the subsets.

ASD and DSC of the automatic segmentation algorithm were compared with corresponding measures reported in other publications that address the segmentation of the femur and the pelvis. Table 1 summarizes the various methods, the grades obtained in each one, whether the method is fully or semi-automatic, and the number of test bones as noted in the publications.

2.1.4. Perfomance

The computation time of segmentation was measured on a typical CT scan at an average resolution of $512 \times 512 \times 320$ voxels. The computations were performed on a PC with an Intel i7-6800K processor and 64

Method	DSC	ASD [mm]	Number of Test Bones
[17]	0.95±0.02	0.85±0.46	38
[18][1]	Unspecified	1.25±0.53	43
[19]	0.96	Unspecified	20
[20]	0.97±0.008	0.36±0.12	60
[21]	0.91±4.82	1.22±0.98	110
[22]	0.94	0.91	Unspecified
[23][2]	Unspecified	0.98	Unspecified
[24]	0.93±0.03	Unspecified	22
[25]	0.94±0.02	1.014±0.47	148
Current	**0.98±0.003**	**0.36±0.05**	**41**

[1] Semi-automatic method.
[2] Cited result was extracted from [21].

Table 1: Comparison of results between different automatic methods. The measures shown are average ± standard deviation.

GB RAM. The calculations for CNNs were done on the GPU using a Nvidia GeForce GTX 1080Ti graphic card. The average segmentation time was 12 minutes: 2.5 minutes to load files, cut them, align them and find FSCV, 2.5 minutes for running Net25, 5 minutes for running Net51, 2 minutes for postprocessing. The CNNs were performed using Python Tensorflow libraries.

2.2. Automatic assignment of inhomogeneous material properties to the femur

Bone tissue is inhomogeneous and anisotropic. However, in a stance position with load representing stair climbing, the longitudinal Young modulus is the most important material property that influences the mechanical response [26, 27, 28]. Thus, we use inhomogeneous relationships between longitudinal Young's modulus and ash density (determined by the HUs in the CT scan) for cortical [29] and trabecular bone tissue [30], validated in [31]:

$$\rho_{ash} = 0.877 \times 1.21 \times 10^{-3} \times a \times HU + 0.08, \quad [grm/cm^3], \tag{1}$$

$$E_{cort} = 10200 \times \rho_{ash}^{2.01} \quad [MPa], \rho_{ash} > 0.486[grm/cm^3] \tag{2}$$

$$E_{trab} = 2398[MPa], \quad 0.3 \leq \rho_{ash} < 0.4[grm/cm^3] \tag{3}$$

$$E_{trab} = 33900 \times \rho_{ash}^{2.2}[MPa], \quad \rho_{ash} < 0.3[grm/cm^3] \tag{4}$$

The Poisson ratio was set to the constant value of $\nu = 0.3$ [32, 31]. An estimation of Young modulus based on CT grey level (surrogate to bone density) is obtained from clinical CT scans performed without calibration phantoms. The linear relation between HU and bone density without phantoms is based on

conclusions drawn in previous experimental studies: a) Maximum bone density (along the entire femur) for all patient regardless of age and gender is almost constant. Therefore the maximum Young's modulus regardless of age and gender is almost constant. b) HU=0 is related to ash density=0. c) The maximum Young's modulus (E_{max}) is based on six relevant studies presented in Table 1 in [28], having a value of 20 GPa. Using $E_{max} = 20$ GPa in (2), $\rho_{ash_{max}}$ is computed.

Inserting $\rho_{ash_{max}}$ in (1), $a \times HU_{max}$ is computed. HU_{max} in patient's CT scan is known - it is evaluated by a histogram of the HU in CT's voxels: it is the value above which there are only 0.1% of the voxels in an entire femur. Thus, the linear slope parameter a is determined and finally one may use relationships (1)-(4) to determine the longitudinal Young's modulus all pixels. A file is generated with an E value for each voxel in the CT scan. This file is used to determine material properties at each integration point during the computation of element stiffness matrices. Due to "noisy data" and "artifacts" in CT scans at the bone's surface, boundary corrections are applied (details are provided in [33]).

The femur's response under physiological stance position loading is well described by the linear theory of elasticity and although the bone on the macroscopic level is orthotropic, excellent results have been obtained using isotropic inhomogeneous relations for stance position loadings, see [32, 31] and references therein. Therefore, a linear finite element analysis is performed for femurs.

2.3. Application of boundary conditions

Personalized boundary conditions were determined by a statistical analysis of measured hip contact force given in [34] following [35] and references therein. Instrumented hip implants in patients provided realistic data on the peak hip-contact force in various daily activities as free walking and going up the stairs. A simplified, yet realistic magnitude of the hip contact force, which is the most influential force during these activities, is approximately 2.5 times body weight. The mean direction of such a force is tilted by about 15 degrees to the shaft axis in the frontal plane and about 7 degrees in the sagital plane. To simplify the application of force on the bone's head, a vector is defined connecting two anatomic points, the center of the femur's head and the intercondylar notch (point P_2 in Figure 6, the deepest notch point between the rear surfaces of the medial and lateral epicondyles). These anatomic points are determined by analysis of the femur's geometry, when the entire femur is available in the CT-scan. This vector is usually at about 7 degrees in the frontal plane (see the schematic vector on femur's head in Figure 6). Otherwise, if the condyles are absent in the CT-scan (a short bone), the intercondylar point is estimated by another CNN algorithm (not addressed in herein). The hip contact force is projected to the proximal head surface

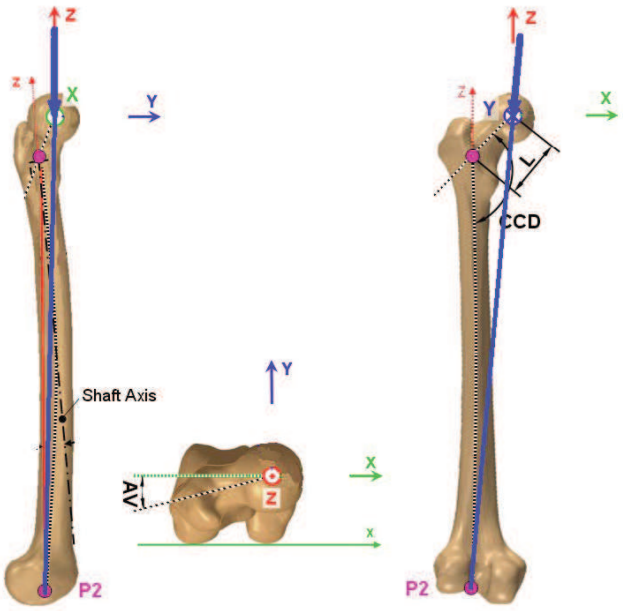


Figure 6: Anatomic points and contact hip force application on femur's head (augmented figure from [36]).

and applied as a distributed pressure on a circular area which is a part of the spherical head with a radius of 10 mm. The distal part of the femur is defined by the proximal slice of the patella and is fully clamped (zero displacements in all directions).

In real life, there are activities that result in a torsion moment applied to the femur. Unfortunately, to the best of our knowledge, there are no measurements or estimates of such torsion loads available. To consider torsion influence on a diseased femur, a comparison analysis is performed between the response of the diseased and healthy femora under a torsional load alone. A load of 1000 N is applied on the femur head on the same surface on which the stance position is applied but in the y direction. The results of this loading conditions are used in the post processing phase to comparatively assess the mechanical response of the diseased versus the healthy contralateral femur because of a torsional load.

2.4. Automatic tetrahedra p-mesh generator

Having voxels that define the femur's geometry, the SimModeler[3] libraries [37] are used first to generate a geometric surface that represents the surface of the femur, and then to automatically generate a curved p-FE mesh. In p-FE technology the mapping of the elements to the standard element is performed by

[3]SimModeler is provided by Simmetrix, 10 Executive Park Drive, Clifton Park, NY 12065

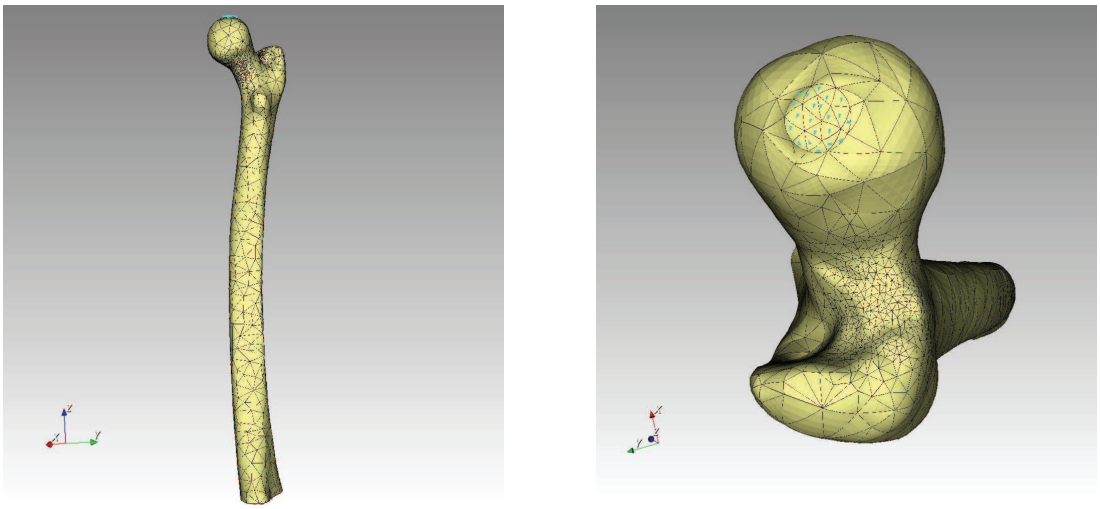


Figure 7: Typical p-mesh for an entire femur with boundary conditions.

blending-function mappings (see [38, 2] for 2D FE elements) then an accurate representation of element's surfaces is obtained independently of approximation space. Blending mapping of a curved face requires the coordinates of many points on faces of an element (known also as Chen-Babuska points [39]) to generate a polynomial approximation of element surfaces.

The automatic mesh generator also generates a graded mesh in regions of high curvatures such as the femur neck, for example. In the interior of the domain, flat faced tetrahedral elements are generated to improve computational efficiency. Typically, a whole femur is meshed by 4000 to 6000 tetrahedral elements (see a typical mesh for a femur in Figure 7).

2.5. p-FE implementation

The *Simfini* framework code was written in the Java programming language because: a) It is platform independent. b) It has a mature class library. c) It supports automatic garbage collection which simplifies programming and debugging. d) It is relatively easy to integrate native libraries written in C++, FORTRAN, etc. for multiple computational platforms (Windows, Linux) using Java Native Interface (JNI) wrappers. e) Development and prototyping is more rapid than using explicitly compiled and linked languages like C++. One can easily change code on the fly and continue development without having to stop the execution to recompile and link. f) Developing multi-threaded applications is relatively simple. g) Many libraries and frameworks are available for deploying applications for the web.

p-FEMs were chosen because they provide several advantages relative to conventional h-FEMs. These advantages are of major importance for the generation of AFE analysis of bones: a) The mesh is kept

unchanged and only the polynomial degree p of the shape functions is increased systematically to allow an automatic verification of the numerical error. b) p-Elements are much larger, may be far more distorted, and may have larger aspect ratios (much more robust in terms of deterioration of the numerical accuracy) and yet produce considerable faster convergence rates compared to their h-FEM counterparts [2]. Hierarchical shape functions for each p level (up to $p = 8$) for tetrahedral elements were computed and the number of shape functions with their association to vertex, edge, face or internal modes are given in Table A.2 in Appendix A. The p-element stiffness matrix computation is a relatively lengthy procedure for an element which has varying material properties. This is because there are hundreds of degrees of freedom at high p for each element, and the Young modulus at each integration point must be retrieved from a file generated during segmentation, which provides the property evaluated at the center of each CT-voxel. To reduce computational time, a parallel implementation uses multiple threads to distribute the computation of element stiffness and load to different CPU cores. For the computation of the element stiffness matrix, the majority of computation time is spent on numerical integration. We used the non-classical cubature scheme by Zhang et al. [40] for the numerical integration of the tetrahedral stiffness matrix and load vector for which the integration points are distributed evenly within the standard tetrahedron. It requires fewer cubature points compared to classical Gauss cubature. For example, for a monomial of order eleven (in one variable) in Gauss cubature, 216 integration points are required for an exact integration, while a Zhang et. al. scheme requires only 96 integration points. However, the Zhang et al. scheme suffers from a non-monotonic convergence. Finally, the solution of the global system of equations is obtained using the Pardiso[4] solver which supports parallel computing on multiple threads. It is worth mentioning that the computation of element stiffness matrices, assembly and solution of the global system of equations consume most of the CPU time during the finite element analysis phase of the process.

Estimation of the error in energy norm used for the automatic verification of the numerical error: The p-FE mesh is unchanged while increasing the polynomial order of the trial and test functions, and mapping is independent of the p-level thus a hierarchical FE space is obtained, i.e. the FE space of the solution at p is a sub-space of the FE solution at $p+1$. This implies that the error in energy norm of the FE solution at $p+1$ is equal or smaller compared to the error in energy norm of the FE solution at p, i.e. a monotonic convergence of the error in energy norm is obtained when using p-FEA. This in turn allows the use of three

[4]https://pardiso-project.org

consecutive FE solutions with increasing polynomial orders, call them FE_1, FE_2, FE_3, with increasing number of degrees of freedom $N_1 < N_2 < N_3$, and three potential energies $\Pi_{FE_1} \geq \Pi_{FE_2} \geq \Pi_{FE_3}$. Having these, the exact potential energy Π_{EX} can be estimated by solving the simple non-linear equation [2]:

$$\frac{\Pi_{EX} - \Pi_{FE_3}}{\Pi_{EX} - \Pi_{FE_2}} \approx \left(\frac{\Pi_{EX} - \Pi_{FE_2}}{\Pi_{EX} - \Pi_{FE_1}} \right)^{\frac{\log N_2 - \log N_3}{\log N_1 - \log N_2}} \tag{5}$$

Once Π_{EX} is determined by (5), the relative error in energy norm[5] for each FE solution in percentage can be estimated by:

$$\frac{\| \mathbf{u}_{FE_i} - \mathbf{u}_{EX} \|_{\mathcal{E}}}{\| \mathbf{u}_{FE_i} \|_{\mathcal{E}}} = 100 \sqrt{\frac{\Pi_{FE_i} - \Pi_{EX}}{\frac{1}{2}\mathcal{B}(\mathbf{u}_{FE_i}, \mathbf{u}_{FE_i})}} \tag{6}$$

These error estimates are progressively better as N increases, and are used to automatically determine whether the quality of the solution is sufficient to draw conclusions from the latest FE solution, or a more refined FE solution is required.

Verification of the AFE implementation and its efficiency were evaluated by considering a curved cube as shown in Figure 8 subjected to traction $\vec{T} = (1 \quad 2 \quad 3)^T [\frac{N}{mm^2}]$ on face $z = 2$ and homogenous BC on face $z = 0$. The cube is made of an isotropic homogenous material ($\nu = 0.3, E = 1000[\frac{N}{mm^2}]$). Although

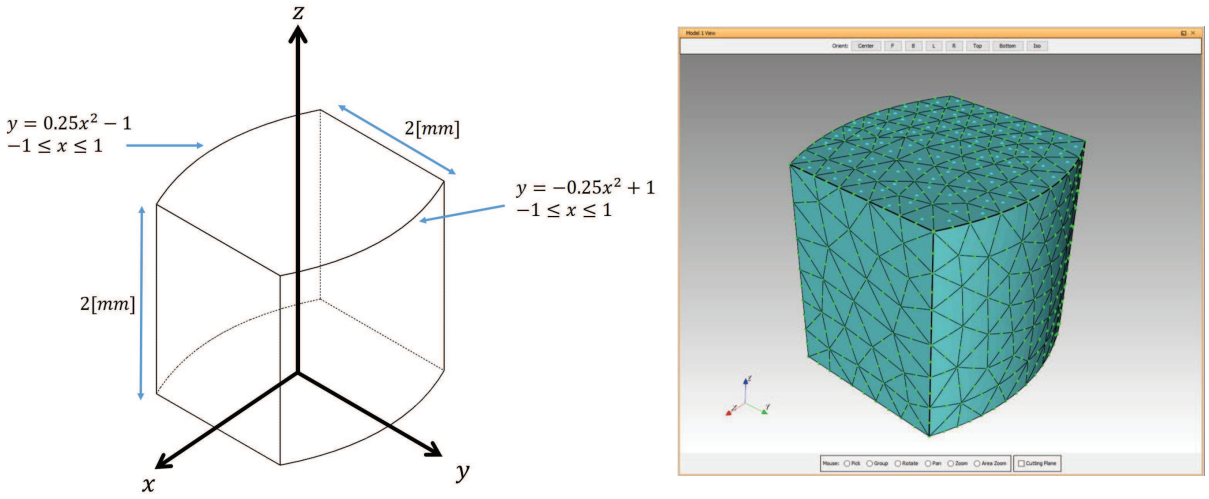


Figure 8: Curved cube dimensions and mesh.

the best method for verification is using a manufactured solution, we chose a shortcut, and instead we

[5]The energy norm of the vector function $\mathbf{w}$ is defined as $\| \mathbf{w} \|_{\mathcal{E}} \stackrel{\text{def}}{=} \sqrt{\frac{1}{2}\mathcal{B}(\mathbf{w}, \mathbf{w})}$, where $\mathcal{B}(\bullet, \bullet)$ is the bilinear form of the linear elasticity operator.

computed the potential energy and the displacement at an arbitrary point in the domain $(0, 0, 1)$ in an "overkilled mesh" with 1,212,336 degrees reported of freedom (DOF) solved by StressCheck V9.2[6]. These were considered as the "exact" values used for comparison.

The advantage of parallelization has been checked on two systems. System1: Dell Precision T7500 with two Intel Xeon x5647 @ 2.93GHz processors with 8 threads and 12 GB RAM. System2: Dell Precision Tower 5810 with a single processor Intel Xeon E5-1620 @ 3.5GHz with 8 threads and 16 GB RAM. For the curved cube with 1632 elements the convergence in energy norm as a function of DOFs obtained by *Simfini*, StressCheck and the popular h-FE code Abaqus[7] as well as the computational time are presented in Figure 9. The parallel implementation of *Simfini* out-performed StressCheck v9.2, especially at higher p-levels.

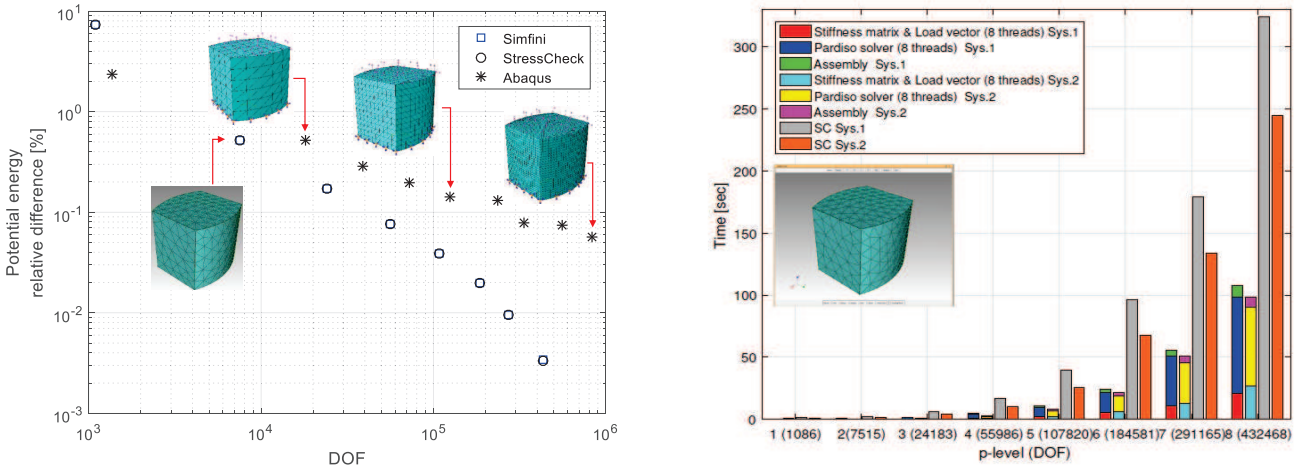


Figure 9: Curved cube with 1632 FEs. Right: Convergence of the error in energy norm as a function of number of DOF. Left: Computational time for the various tasks as p increases: *Simfini* vs StressCheck.

Also in terms of CPU *Simfini* results converged faster compared to StressCheck and Abaqus when both the error in energy norm or the pointwise error in the displacements magnitude at an arbitrary point in the domain are of interest, as shown in Figure 10.

2.6. Automatic post-processing and verification of numerical errors

An identical AFE analysis is performed for each of the two femora in the CT scan of a patient, and for each of the femora both a stance position loading and a torsional loading is applied separately. A series

[6]StressCheck is trademark of Engineering Software Research and Development, Inc, St. Louis, MO, USA.
[7]Abaqus is a trademark of Dassault Systemes Simulia Corp. , Johnstron, RI 02919, USA.

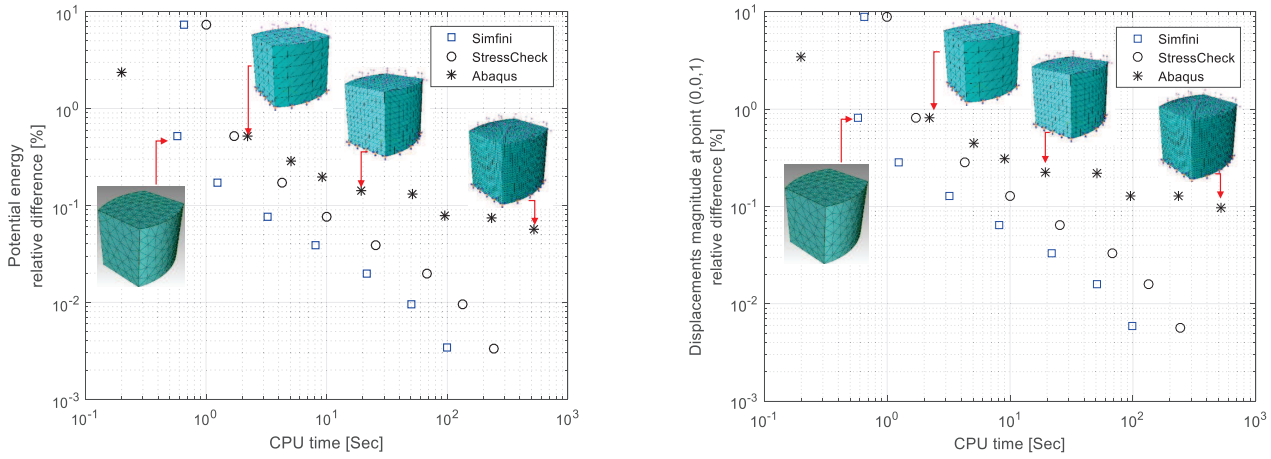


Figure 10: Curved cube with 1632 FEs. Left: Convergence of the error in energy norm as a function of CPU. Right: Convergence of the magnitude of the displacement at point (0,0,1).

of $p = 1$ to 6 solutions are used to estimate the relative error of each solution in energy norm. If the FE solution at $p = 6$ has an estimated relative error of less than 5%, it is considered to be accurate enough for post-processing, otherwise the p-level is increased. If at $p = 8$ the solution is still not converged within 5% relative error in energy norm, then a more refined mesh is applied and a p-extension is re-performed.

Once the solution is converged in energy norm, strains are computed on the faces of the elements that coincide with the bone surfaces. The strain tensor and principal strains are computed at 15 points on the face of each element that lies on the bone surface, for each p-level.

Since a single pointwise strain value may be largely influenced by numerical errors and since the bone failure is assumed to depend on the averaged strain in a given area, the median strain over a specific area needs to be computed. Principal strains are calculated over femur's surface with the maximum determined on a surface of an area of 5 mm^2 around a local extremum. Any strains close to the edge surrounding the applied load or constrains known to be in the vicinity of singularities are filtered out

In the segmentation stage, each femur is divided into 5 regions - neck, trochanters, proximal shaft, middle shaft, and distal shaft (see Figure 11). The maximum median tensile and compressive principal strains in each region is computed and used as the indicators in each of these 5 regions. For verification purpose these maximum median strains are computed for the last three p levels to assure these had converged.

2.7. Failure criterion to be used in clinical practice to determine risk of a pathological fracture

In realistic daily activity, there are several loads applied by muscles on a femur, and the distal part is not really clamped as in the FEA. Nonetheless, the hip contact load is the most important one so it is used as the load to be considered in the failure criterion. It has been demonstrated that one of the best criteria of femoral fracture is the maximum elastic principal strain averaged over a surface area (see [41, 42]), therefore the same median principal strain criterion is considered.

The "typical median principal strains" for the 5 regions in a "healthy femur" were computed based on 12 femurs as detailed elsewhere [43]. 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 femur is 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 fracture [43]. The SFR was computed for both femurs of each patient. An additional analysis of torsional load is performed for patients who demonstrated a low fracture risk under stance position loading. A high fracture risk in torsion is defined if the difference between the diseased and healthy femur is larger than 50% . The AFE generates a report for the physician as presented in Figure 11. 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).

3. Validation

The verified results must be shown to well represent the physical phenomenon (validation), i.e. that the strains and displacements computed by FEA under a hip contact force are as observed in in-vitro experiments under the same conditions. To this end ex-vivo experiments on a large cohort of healthy and diseased human fresh-frozen femurs were performed. Seventeen fresh frozen femurs were CT-scanned, instrumented by strain-gauges and loaded, and p-FE analyses were performed to mimic the experiments. Twelve of these were used in a double-blind comparison (one institution performed the experiments and another institution performed the analysis, both being blind to the data of the other until after both FEA and experiments were complete). Figure 12 presents a typical femur during the mechanical experiment and the FE results.

Uniaxial strains were FE computed along the direction of the bonded strain gauges and total displacements $u_{tot} = \sqrt{u_x^2 + u_y^2 + u_z^2}$ at the optical markers on the bone surface were also calculated (details on

1 General Details

Analysis Id	190921.192934	**Weight[kg]**	46
Analysis Date	21/09/2019	**Gender**	M
Ref. Clinician	Nir	**Birth Date**	25/01/1956
Patient ID	ICpro36EH00036	**Age**	63
Patient Name	5		

2 Location of risk of fracture

Area of interest	Right Femur	Left Femur
Neck	Low	Low
Greater / Lesser Trochanter	**High**	Low
Proximal Shaft	**High**	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.

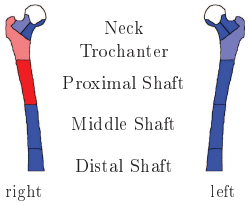


Figure 1: Regions of interest

3 CT Parameters

CT Date	16/09/2019	**Study ID**	...084607680429	**Convolution Kernel**	B
Institute Name	modiCAS GmbH	**Study Description**	Not Specified	**KVP**	120
Manufacturer	Philips	**CT Series number / Description**	Num:3/Desc:Not Specified	**Slice Thickness [mm]**	2.00
Model	Brilliance 64	**Filter Type**	B	**Pixel Size [mm/pixel]**	0.79

Figure 11: An example of the AFE report for the surgeon for a patient that has an SFR¿1.48 in the right femur at two location - in the shaft the risk of fracture is higher.

these experiments are provided in [44]). A total of 102 displacements and 161 strains on the 17 femurs were used to assess the validity of the FE simulations. In Figure 13 the pooled FE strains and displacements are compared to experimental observations by a linear regression plot, inspecting the slope, intersection and R^2 of the linear regression between the experimental observations and FE predictions, and by a Bland-Altman plot [45] emphasising the errors between the two, and the 95% confidence limits. One may notice the excellent match between the predicted and measured data for femurs: the slope and R^2 of the linear regression are very close to 1, and the average error in the Bland-Altman is zero.

In view of the next section, it is important to emphasize that *Simfini* was also validated by ex-vivo experiments on fourteen femurs with metastatic tumors [31].

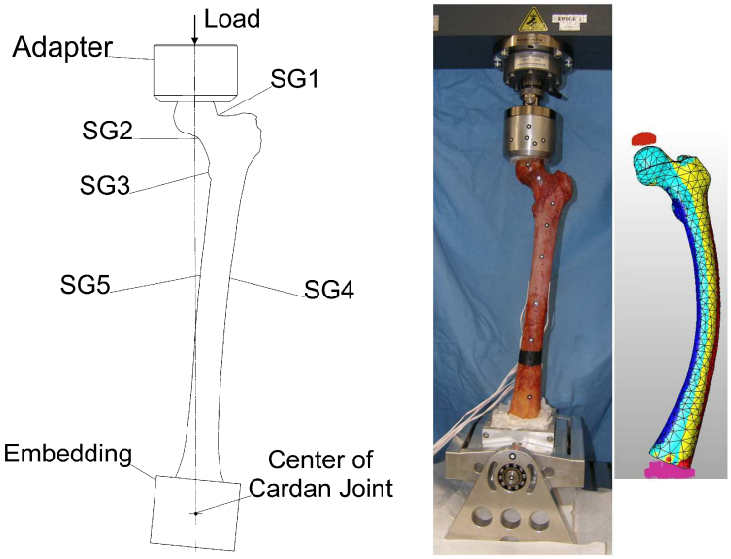


Figure 12: (a-left) Sketch of the frontal plane of an embedded and instrumented left femur. (b-right) Experimental setup with the optical markers on an instrumented left femur and its corresponding deformed (magnified) FE model (Figure from [44])

4. AFE in clinical practice: Predicting risk of a pathological fracture in femurs with metastatic tumors

Our ultimate goal is to make *Simfini* available for use in clinical practice. The clinical application presently addressed by *Simfini* is the determination of femoral risk of fracture for patients with metastatic lesions in their femur. Metastasis of the femur may weaken the bone to the point at which a pathologic fracture may occur under daily activity, 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 care. A preventive (prophylactic) surgery for patients with an impending femur fracture improves short-term survival, morbidity, functional outcome, and reduces length of hospital stay. An accurate estimation of the risk of fracture is essential to avoid both under-treatment and over-treatment in patients with impending pathologic fractures.

CT-scans and weights were collected on 38 patients with metastatic tumors in their femurs that did not undergo a prophylactic surgery: 14 had a pathological fracture within 6 months following the CT scan, and 24 patients were fracture free during the five months following the CT scan. The strain fold ratio (SFR) was computed by the AFE for all patients, with the value of 1.48 used as the determinant of the threshold for a pathological femoral fracture. Sensitivity, specificity, positive and negative predicted values were computed.

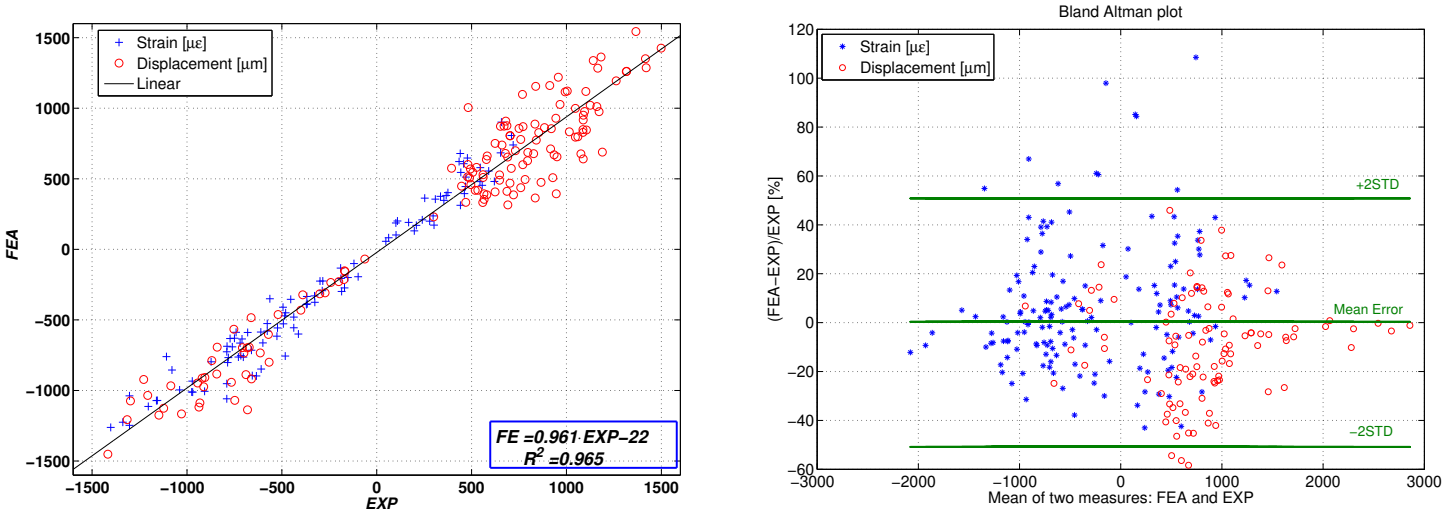


Figure 13: Comparison of the computed strains $+, *$ and displacements $\circ$ to the experimental observations normalized to 1000 N load. (Left) Linear regression, (Right) Bland-Altman plot.

Sensitivity of SFR was found to be 100% (compared to 88% by the standard Mirels score used in practice today) and sensitivity was 67% (compared to 38% by Mirels score). All study patients who sustained a pathological fracture of the femur had an SFR >1.48. Figure 14 is a descriptive analysis of the SFR for all 38 patients. It shows 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. The presented AFE is by far better compared to Mirels' score for the accurate prediction of the risk of fracture in patients with metastatic tumors to the femur. Actual fractures detected by X-ray radiographs post fracture were compared to the estimated AFE locations. The location of the pathological fracture was correctly predicted for all 14 patients who fractured their femurs. A typical example of the predicted location and the actual fracture is shown in Figure 15. A recent retrospective study using AFE analysis on 50 patients who were referred to undergo prophylactic stabilization predicted that 39% of these patients may not have needed surgery [43]. A prospective randomized clinical trial evaluating AFE as a criterion for determining the need for surgical stabilization in patients with metastatic tumors in the femur is underway on 162 patients at the Sourasky Medical Center in Tel-Aviv, Isael.

5. Summary and conclusions

We presented an autonomous finite element (AFE) analysis that introduces a paradigm shift in the use of FEA. Such an AFE is now possible thanks to recent scientific advancements such as CT scans, machine

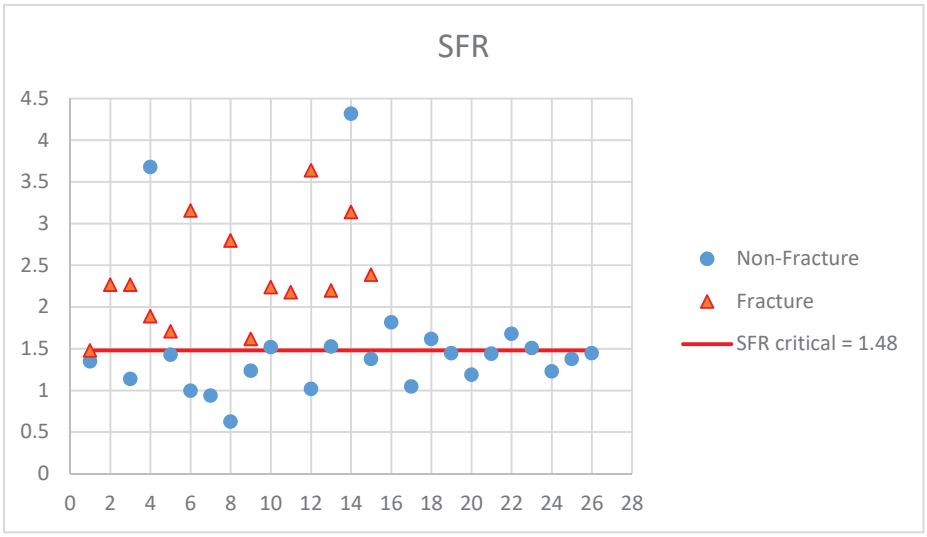
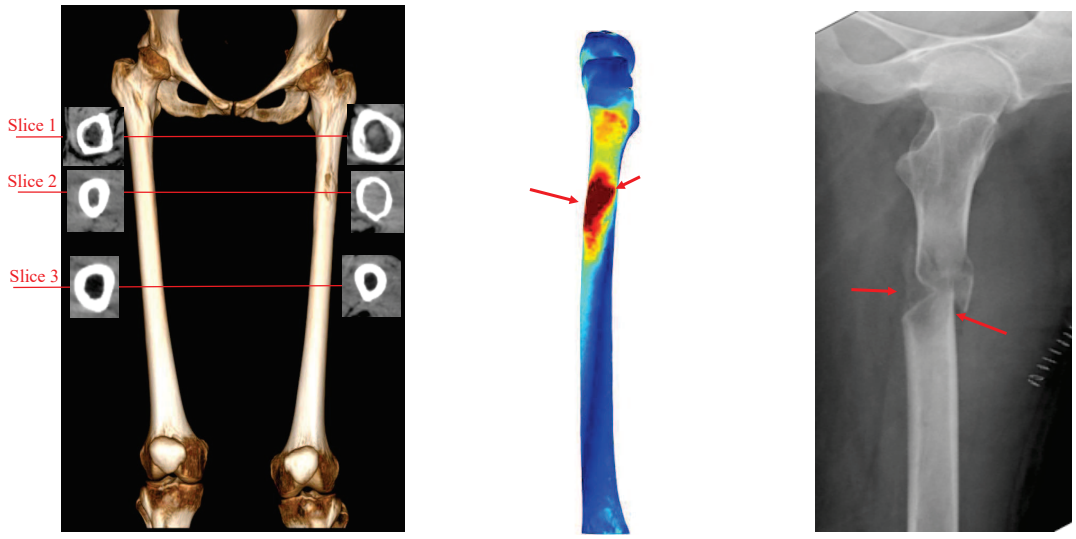


Figure 14: SFRs for patients who eventually sustained a pathological fracture (orange triangles) compared to patients who did not within the 6 months following the CT scan (blue circles)..



a. CT and slices in the ROI

b. Prediction of fracture location byAFE

c. X-ray after fracture

Figure 15: Typical CT scan of the femurs of a patient with metastatic tumors (a). AFE results with the predicted location at which fracture is expected (b). X-ray after fracture (c).

learning, and high order FEA which allows an inherent verification methodology of the numerical accuracy. A novel AFE for patient-specific analysis of human femurs (named *Simfini*) was described that utilizes an automatic segmentation of femurs from CT-scans by convolution neural networks, an automatic mesh generation and application of boundary conditions based on anatomical points, a high-order FE analysis with numerical error control, and finally an automatic report with a clear assessment of bone fracture risk. One specific application of AFE is the determination of the risk of fracture for patients with tumors of the femur to determine whether a prophylactic surgery is needed. It was shown that AFE can serve as a decision support system in orthopedics, providing the clinical community advanced tools that may improve clinical treatment. To the best of our knowledge this is the first CE accredited AFE application being used by orthopedic surgeons in clinical practice.

Simfini is intended to be used as a decision support system that enables MDs to provide more effective treatment in daily clinical practice. It is being enhanced to assist MDs in determining the risk of femoral fracture as a result of fall on the side in the elderly population with osteoporosis, and determine the outcome of vertebroplasty and kyphoplasty procedures. It is hoped that with further exposure to the clinical community, and by demonstrating accurate and reliable patient-specific assessments of fracture risk, this AFE will find its way into medical practice as a means of improving treatment for patients who suffer from bone strength related diseases.

Acknowledgments

The authors thank G. Book, O. Levi and E. Cohen, former employees of PerSimiO for their assistance in the segmentation algorithm, and Itay Yacobi, a former graduate student of the first author for his investigation on p-tetrahedral elements. The authors thank Dr. Xian Guo for his valuable assistance developing efficient high order finite element mapping, constraints, matrix assembly and solution procedures.

Appendix A. Association of Tetrahedral shape functions to vertex, edge, face, and interior

Table A.2: Tetrahedral shape functions for each p-level. The association of the shape function to vertex, edge, face and internal is given by the color.

$p=1$	N_1	N_2	N_3	N_4																											
$p=2$	N_5	N_6	N_7	N_8	N_9	N_{10}																									
$p=3$	N_{11}	N_{12}	N_{13}	N_{14}	N_{15}	N_{16}	N_{17}	N_{18}	N_{19}	N_{20}																					
$p=4$	N_{21}	N_{22}	N_{23}	N_{24}	N_{25}	N_{26}	N_{27}	N_{28}	N_{29}	N_{30}	N_{31}	N_{32}	N_{33}	N_{34}	N_{35}																
$p=5$	N_{36}	N_{37}	N_{38}	N_{39}	N_{40}	N_{41}	N_{42}	N_{43}	N_{44}	N_{45}	N_{46}	N_{47}	N_{48}	N_{49}	N_{50}	N_{51}	N_{52}	N_{53}	N_{54}	N_{55}	N_{56}										
$p=6$	N_{57}	N_{58}	N_{59}	N_{60}	N_{61}	N_{62}	N_{63}	N_{64}	N_{65}	N_{66}	N_{67}	N_{68}	N_{69}	N_{70}	N_{71}	N_{72}	N_{73}	N_{74}	N_{75}	N_{76}	N_{77}	N_{78}	N_{79}	N_{80}	N_{81}	N_{82}	N_{83}	N_{84}			
$p=7$	N_{85}	N_{86}	N_{87}	N_{88}	N_{89}	N_{90}	N_{91}	-	-	-	-	-	-	-	20	-	-	-	-	-	-	N_{110}	N_{111}	-	-	10	-	N_{120}			
$p=8$	N_{121}	N_{122}	N_{123}	N_{124}	N_{125}	N_{126}	N_{127}	-	-	-	-	-	-	-	24	-	-	-	-	-	-	N_{150}	N_{151}	-	-	15	-	N_{165}			
$p=9$	N_{166}	N_{167}	N_{168}	N_{169}	N_{170}	N_{171}	N_{172}	-	-	-	-	-	-	-	28	-	-	-	-	-	-	N_{199}	N_{200}	-	-	21	-	N_{220}			

References

[1] M. J. Turner, R. W. Clough, M. H. C., L. J. Topp, Stiffness and deflection analysis of complex structures, J. Aero. Sci. 23 (9) (1956) 805–823.

[2] B. A. Szabó, I. Babuška, Finite Element Analysis, John Wiley & Sons, New York, 1991.

[3] W. A. M. Brekehnans, H. W. Poort, T. J. J. H. Slooff, A new method to analyse the mechanical behaviour of skeletal parts, Acta Ortho. Scand. 43 (1972) 301–317.

[4] E. F. Rybicki, F. A. Simonen, W. E. B., On the mathematical analysis of stress in the human femur, J. Biomech. 5 (1972) 203–215.

[5] R. Scholten, Ueber die berechnung der mechanischen beanspruchung in knochenstrukturen mittels fuer den flugzeugbau entwickelter rechenverfahren, Med. Ortho. Technik 6 (1975) 130–138.

[6] H. Olofsson, Three dimensional FEM calculation of elastic stress field in human femur, Ph.D. thesis, Institute of Technology, University of Uppsala, Uppsala, Sweden (1976).

[7] M. G. Kivelson, C. T. Russell, Introduction to space physics, Cambridge University Press, 1995.

[8] S. Torbert, Appl. Comp. Sci., Springer-Verlag, 2016.

[9] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proceedings of the IEEE 86 (11) (1998) 2278–2324.

[10] R. M. Haralick, S. R. Sternberg, X. Zhuang, Image analysis using mathematical morphology, IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-9 (4) (1987) 532–550.

[11] G. Book, Z. Yosibash, N. Trabelsi, Automated bone segmentation in images, US Application Patent 62/641,483 (March 2018).

[12] C. Deniz, S. Xiang, R. Hallyburton, A. Welbeck, J. Babb, S. Honig, K. Cho, G. Chang, Segmentation of the proximal femur from MR images using deep convolutional neural networks, Sci Rep 8 (2018) 16485.

[13] S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, arXiv preprint arXiv:1502.03167.

[14] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980.

[15] L. R. Dice, Measures of the amount of ecologic association between species, Ecology 26 (3) (1945) 297–302.

[16] G. Gerig, M. Jomier, M. Chakos, Valmet: A new validation tool for assessing and improving 3D object segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2001, pp. 516–523.

[17] Y. Xia, J. Fripp, S. S. Chandra, R. Schwarz, C. Engstrom, S. Crozier, Automated bone segmentation from large field of view 3D MR images of the hip joint, Phys. Med. Biology 58 (20) (2013) 7375.

[18] J. Schmid, J. Kim, N. Magnenat-Thalmann, Robust statistical shape models for MRI bone segmentation in presence of small field of view, Medical Image Analysis 15 (1) (2011) 155–168.

[19] J. Fripp, S. Crozier, S. K. Warfield, S. Ourselin, Automatic segmentation of the bone and extraction of the bone–cartilage interface from magnetic resonance images of the knee, Phys. Med. Biology 52 (6) (2007) 1617.

[20] C. Chu, C. Chen, L. Liu, G. Zheng, Facts: fully automatic CT segmentation of a hip joint, Annals Biomed. Eng. 43 (5) (2015) 1247–1259.

[21] Y. Cheng, S. Zhou, Y. Wang, C. Guo, J. Bai, S. Tamura, Automatic segmentation technique for acetabulum and femoral head in CT images, Pattern Recognition 46 (11) (2013) 2969–2984.

[22] R. A. Zoroofi, Y. Sato, T. Sasama, T. Nishii, N. Sugano, K. Yonenobu, H. Yoshikawa, T. Ochi, S. Tamura, Automated segmentation of acetabulum and femoral head from 3-D CT images, IEEE Transactions on Information Technology in Biomedicine 7 (4) (2003) 329–343.

[23] F. Yokota, T. Okada, M. Takao, N. Sugano, Y. Tada, N. Tomiyama, Y. Sato, Automated CT segmentation of diseased hip using hierarchical and conditional statistical shape models, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2013, pp. 190–197.

[24] T. Gangwar, J. Calder, Tak, T. Takahashi, J. E. Bechtold, D. Schillinger, Robust variational segmentation of 3D bone CT data with thin cartilage interfaces, Med./ Image Anal. 47 (2018) 95–110.

[25] D. F. Almeida, R. B. Ruben, J. Folgado, P. R. Fernandes, E. Audenaert, B. Verhegghe, M. De Beule, Fully automatic segmentation of femurs with medullary canal definition in high and in low resolution CT scans, Med. Eng. Phys. 38 (12) (2016) 1474–1480.

[26] L. Peng, J. Bai, Z. Zeng, Y. Zhou, Comparison of isotropic and orthotropic material property assignments on femoral finite element models under two loading conditions, Med. Eng. Phys. 28 (2006) 227–233.

[27] H. Yang, X. Ma, T. Guo, Some factors that affect the comparison between isotropic and orthotropic inhomogeneous finite element material models of femur, Med. Eng. Phys. 32 (2010) 553–560.

[28] N. Trabelsi, Z. Yosibash, Patient-specific FE analyses of the proximal femur with orthotropic material properties validated by experiments, ASME J. Biomech. Eng. 155 (2011) 061001–1 – 061001–11.

[29] T. S. Keller, Predicting the compressive mechanical behavior of bone, J. Biomech. 27 (1994) 1159–1168.

[30] J. . Keyak, M. G. Fourkas, J. M. Meagher, H. B. Skinner, Validation of automated method of three-dimensional finite element modelling of bone, ASME J. Biomech. Eng. 15 (1993) 505–509.

[31] Z. Yosibash, R. Plitman Mayo, G. Dahan, N. Trabelsi, G. Amir, C. Milgrom, Predicting the stiffness and strength of human femurs with realistic metastatic tumors, Bone 69 (2014) 180 – 190.

[32] Z. Yosibash, N. Trabelsi, C. Milgrom, Reliable simulations of the human proximal femur by high-order finite element analysis validated by experimental observations, J. Biomech. 40 (2007) 3688–3699.

[33] N. Trabelsi, Z. Yosibash, C. Milgrom, Validation of subject-specific automated p-FE analysis of the proximal femur, J. Biomech. 42 (2009) 234–241.

[34] G. Bergmann (Ed.), OrthoLoad database, Charité - Universitätsmedizin Berlin, 2008, www.OrthoLoad.com, accessed on June 10, 2012.

[35] Z. Yosibash, H. Wille, R. E., Stochastic description of the peak hip contact force during walking free and going upstairs, J. Biomech. 48 (2015) 1015 – 1022.

[36] G. Bergmann, A. Bender, J. Dymke, G. Duda, P. Damm, Standardized loads ccting in hip implants, PLoS ONE 11 (5) (2016) e0155612.

[37] O. Klaas, M. Beall, M. Shephard, Construction of Models and Meshes of Heterogeneous Material Microstructure from Image Data, in: J. Zhang (Ed.), Image Based Geometric Modelinga and Mesh Generation: Lecure Notes in Computational Vision and Biomechanics, Vol. 3, Springer, 2013, pp. 171–194.

[38] W. J. Gordon, C. A. Hall, Transfinite element methods: Blending functions interpolation over arbitrary curved element domains, Numer. Math. 21 (1973) 109–129.

[39] Q. Chen, I. Babuska, Approximate optimal points for polynomial interpolation of real functions in an interval and in a triangle, Computer Meth. Appl. Mech. Engrg. 128 (1995) 405–417.

[40] L. Zhang, T. Cui, H. Liu, A set of symmetric quadrature rules on triangles and tetrahedra, Jour. Computational Math. (2009) 89–96.

[41] E. Schileo, F. Taddei, L. Cristofolini, M. Viceconti, 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. 41 (2) (2008) 356–367.

[42] Z. Yosibash, D. Tal, N. Trabelsi, Predicting the yield of the proximal femur using high order finite element analysis with inhomogeneous orthotropic material properties, Phylosophical Transaction of the Royal Society: A 368 (2010) 2707–2723.

[43] A. Sternheim, O. Giladi, Y. Gortzak, M. Drexler, M. Salai, N. Trabelsi, C. Milgrom, Z. Yosibash, Pathological fracture risk assessment in patients with femoral metastases using CT-based finite element methods. A retrospective clinical study, Bone 110 (2018) 215–220.

[44] N. Trabelsi, Z. Yosibash, C. Wutte, R. Augat, S. Eberle, Patient-specific finite element analysis of the human femur - a double-blinded biomechanical validation, J. Biomech. 44 (2011) 1666 – 1672.

[45] J. Bland, D. Altman, Statistical methods for assessing agreement between two methods of clinical measurement, Lancet 1 (8476) (1986) 307 – 310.