The Influence of Femoral Lytic Tumors Segmentation on Autonomous Finite Element Analysis

Oren Rachmil[†], Kent Myers[⋆], Omri Merose[‡], Amir Sternheim[‡,♯], and Zohar Yosibash[†]

[†]Computational Mechanics & Experimental Biomechanics Lab, School of Mechanical Engineering, The Iby and Aladar Fleischman Faculty of Engineering, Tel-Aviv University, Ramat Aviv, 69978, Israel
[⋆]PerSimiO, Beer-Sheva, Israel
[♯]Faculty of Medicine, Tel Aviv University, Ramat Aviv, 69978, Israel
[‡]Dept. of Orthopedic Oncology, Tel-Aviv Sourasky Medical Center, Tel-Aviv, Israel

January 25, 2024

Abstract contains 260 words,
Manuscript excluding abstract, tables, figure, bibliography and appendix contains: 3992 words.

Corresponding Author: Prof. Zohar Yosibash, Head - Computational Mechanics and Experimental Biomechanics lab, School of Mechanical Eng, Faculty of Engineering, Tel Aviv University, Israel, yosibash@tauex.tau.ac.il

Abstract

Background: The validated CT-based autonomous finite element system *Simfini* [1] is used in clinical practice to assist orthopedic oncologists in determining the risk of pathological femoral fractures due to metastatic tumors. The finite element models are created automatically from CT-scans, assigning to lytic tumors a relatively low stiffness as if these were a low-density bone tissue because the tumors could not be automatically identified.

Methods: The newly developed automatic deep learning algorithm which segments lytic tumors in femurs, presented in [2], was integrated into *Simfini*. Finite element models of twenty femurs from ten CT-scans of patients with femoral lytic tumors were analyzed three times using: the original methodology without tumor segmentation, manual segmentation of the lytic tumors, and the new automatic segmentation deep learning algorithm to identify lytic tumors. The influence of explicitly incorporating tumors in the autonomous finite element analysis on computed principal strains is quantified.These serve as an indicator of femoral fracture and are therefore of clinical significance.

Findings: Autonomous finite element models with segmented lytic tumors had generally larger strains in regions affected by the tumor. The deep learning and manual segmentation of tumors resulted in similar average principal strains in 19 regions out of the 23 regions within 15 femurs with lytic tumors. A high dice similarity score of the automatic deep learning tumor segmentation did not necessarily correspond to minor differences compared to manual segmentation.

Interpretation: Automatic tumor segmentation by deep learning allows their incorporation into an autonomous finite element system, resulting generally in elevated averaged principal strains that may better predict pathological femoral fractures.

1 Introduction

A CT-based autonomous finite element (FE) system named Simfini was developed to compute the strains in femurs with tumors [1]. Simfini is used in clinical practice to assist orthopedic oncologists in determining the risk of pathological femoral fractures due to the presence of metastatic tumors [3, 4]. The FE model, created automatically by analyzing the CT-scan images, does not properly incorporate the material properties of lytic tumors because these could not be automatically identified. The proper treatment of these lytic tumors in the FE model and their influence on the FE results is addressed herein following the introduction of a deep learning (DL) algorithm which automatically segments lytic tumors in femurs from computed tomography (CT) scans, presented in the first part of our study. Specifically, the aim is to evaluate how different are the FE result in three different scenarios: the current implementation, incorporating lytic tumors in the FE model as identified by a manual segmentation, and incorporating the tumors automatically using the DL algorithm.

Several previous studies have employed FE models to assess the loads that may cause a pathological fracture in patients with metastatic bone disease. Keyak et. al [5, 6] established a complete workflow to construct FE models from quantitative computed tomography (QCT) images. Bone densities from QCT images were constructed having mechanical properties with no special treatment to tumors. Their FE models for the femoral shaft and upper femur matched the breaking point and peak force measured in human bone tests with a correlation coefficients $R^2 = 0.88 - 0.95$. These methods are lengthy, demand substantial FE expertise and cannot be applied in routine clinical practice. Also, the tumor locations are neglected in the models, introducing potential inaccuracies in the vicinity of the tumors.

Benca et. al [7] employed voxel-based meshing, optimizing automation and robustness against mesh distortion. Yet, their nonlinear, voxel-based model underestimated the fracture load by half.

Validating the efficacy of FE models, Goodheart et. al [8] analyzed CT scans of patients with femoral metastases. Their FE models, simulating walking conditions, showed comparable accuracy to Mirels' score in identifying fracture patients (sensitivity 80%), with superior specificity (86% vs. 43%). However, the specific tumor voxels were not segmented nor were treated as a different tissue than bone. Moreover, the process introduced is not automatic and cannot be performed by a non-FE expert thus cannot be considered as a clinical tool.

In [1] Yosibash et. al introduced a novel approach of autonomous finite element (AFE) analysis for femurs named Simfini. Simfini was tailored for use by clinicians as a decision-support system. It evaluates the fracture risk in patients with femoral tumors to assess the need for prophylactic surgery, and was validated by two retrospective clinical analyses [3, 4]. The automated workflow involves: segmentation of both femurs from CT scans using a U-Net network [9], a mesh generator, application of boundary conditions based on anatomical points, and an automated FE post-processor that creates a report for the clinician providing a clear assessment of bone fracture risk. The different components of the AFE are presented in Figure 1. However, Simfini follows the ideas presented by Keyak and collaborators and does not identify the lytic tumors, but assigns to the voxels with low Hounsfield Units (HU) a reduced Young's modulus as if these were soft bone tissues.

[Figure 1 about here.]

The objective is to quantify the influence of explicitly incorporating lytic tumors into the finite element (FE) model of femurs on principal strains. These serve as an indicator of femoral fracture and are therefore of clinical significance. The new approach is different compared to the prevailing methodology that derives the Young's modulus from CT's HU for tumor voxels, a strategy rooted in [10], suggesting that metastatic lesions minimally affect QCT's estimation of trabecular bone mechanical properties. In addition, the difference in computed principal strains when tumors are segmented manually or automatically by a DL algorithm (with adjusted Young's modulus for tumor-identified voxels) is quantified.

2 Methods

Ten lower abdominal CT scans of patients with femoral tumors are considered herein. These patients were randomly selected as the test case in [2] out of the fifty patients' CT scans randomly collected at Tel-Aviv Sourasky Medical Center (TASMC) after receiving approval from the institutional review boards (Helsinki committee approval number TLV-17-0532). Full details of these patients (Prosp7060, Prosp5060, Prosp1120, Prosp1190, Prosp7020, ProspD100, Prosp1140, Prosp5010, ProspB10, Prosp5050) are given in the Appendix of [2]. The 20 femurs in these scans served as the basis of our investigation. Tumors were present in 15 femurs at one or several locations.

2.1 Incorporating Tumors into the AFE Analysis

Femurs were segmented from the CT scans in Simfini by a U-Net architecture [9] with a DSC above 0.99 [11], resulting in two 3D masks that accurately represent both femurs. Subsequently,

the inhomogeneous mechanical properties of each voxel in this 3D mask were estimated. Although bone tissue is anisotropic, under a stance load condition the longitudinal Young's modulus is the dominant material property [12, 13, 14]. It was defined according to the HUs in the CT scans, validated in [15].

The coordinates and Young's modulus for each voxel were compiled into a text file in which each row corresponds to a point in the 3D mask of the femur, containing four columns, the coordinates (x, y, z), and the correspondingYoung's modulus. The file has approximately 500,000 rows, each representing a voxel in the femur's 3D mask. Importantly, the physical coordinates were listed in the single-femur-coordinate-system (not the same coordinate system as in the CT scan).

The text file was then converted into a NIFTI[1]-format. Tumors were segmented using the generated NIFTI file either manually or automatically by the DL algorithm described in the first part [2]. A binary mask was generated with all voxels with 1 assigned to tumor voxels and 0 otherwise. The corresponding tumor voxels in the material properties file "materialproperties.txt" were identified by comparing the coordinates from the binary mask with those in the material properties file. Then a reduced Young's modulus was assigned to these tumor voxels. A value of 3.66 ± 1.6 kPa was suggested for the femoral lytic tumors [16]. Given that tumors are substantially weaker than bone tissue, the strains within the tumor region are considerably larger. These strains, however, are irrelevant as our focus was on the strains within the bone tissue. Also, incorporating such low values may cause an ill-conditioned stiffness matrix. Hence, various different values for the tumor Young's modulus such as 1, 50, 100 kPa have been tested to evaluate their impact on the AFE. The goal was to accurately represent the contrast in tissue properties between tumors and bone while circumventing an ill-conditioned stiffness matrix.

2.2 Automated FE Analysis

A curved high-order finite element mesh was automatically generated [1]. This automated mesh generator also refines the mesh in regions of interest. Generally, a complete femur can be efficiently represented using roughly 4000-6000 high-order tetrahedral elements. The inclusion of tumor data presents a numerical challenge, as the tumor boundary may reside in one element, causing a sharp change in material properties within an element, leading to potential numerical errors. To mitigate these risks, mesh was refined close to the tumor . The entire process, ranging from the acquisition of the CT scan data to the creation of a mesh with refinements at tumor regions is illustrated in Figure 2.

[Figure 2 about here.]

A stance position load of magnitude equivalent to 2.5 times the body weight was applied, based on a statistical analysis of measured hip contact forces [17]. This data provides realistic peak hip contact forces for various daily activities, such as free walking and climbing stairs. The load was applied to the femoral head along a vector that connects the center of the femoral head to the estimated intercondylar notch, effectively emulating a stance position [1]. These anatomical points were determined for each individual femur, which influenced the location at which force was applied

[1]NIfTI (Neuroimaging Informatics Technology Initiative) is a file format commonly used in neuroimaging to store and exchange neuroimaging data. The format is supported by various software tools and libraries in the field.

on the surface of the femoral head. The distal part of the femur, defined by the proximal slice of the patella, was fully clamped.

Following the implementation of boundary conditions, the solution of the global system of equations was obtained using the Pardiso solver[2]. To verify the numerical error, an estimation of the error in energy norm was computed. The p-FE mesh remained unchanged, while the polynomial degree was increased in a hierarchical manner. A sequence of solutions, ranging from p = 6 to 8, was employed to determine the relative error in the energy norm. An estimated relative error less than 5% was considered sufficiently accurate for the subsequent post-processing phase. If a 5% relative error in energy norm at p = 8 was not reached, a more refined mesh was applied and the p-extension process was re-initiated.

In the post-processing phase, the principal strains were computed on the face of each element on the bone's surface, except within the tumor regions (Simfini has been updated to exclude tumor areas from strain calculations). The femur was partitioned into five anatomical regions: the neck, trochanter, proximal shaft, middle shaft, and distal shaft. Within each region, the maximum/minimum average principal strain within a circular region having a 5 mm radius, centered at this point, was computed. These values served as the criterion for estimating the fracture risk for each region [1].

The effect of the Young's modulus (E) assigned to tumor voxels was initially investigated. Patient Prosp1120 from the test cohort was selected as a representative example.Prosp1120 exhibited a significant deterioration in the cortical structure of the middle shaft of the right femur. The tumor's boundaries in this specific case permit precise segmentation. Furthermore, Prosp 1120's left femur, which is free of tumors, served as a control for this study. It allowed us to make a direct comparison of average mean principal strains to a tumor-free femur.

To assess the influence of incorporating the tumors into the FE model, a comparative analysis was performed on the CT scans of the ten patients (20 femurs). The mean principal strain in each region of both the right and left femurs was compared for the selected ten-patient dataset, considering scenarios with and without tumor presence. This comparison involved automatic tumor segmentations by DL and by manual segmentations. A comparison of the three strategies (no tumors, DL-identified, and manually identified tumors) was performed.

3 Results

3.1 Tumors Young's Modulus Influence on the FE Results

Figure 3 shows the right femur of Prosp 1120's alongside the corresponding automatically generated mesh, highlighting the region of the bone affected by the tumor

[Figure 3 about here.]

The Young's modulus E assigned to the tumor was adjusted to four values : 100 MPa, 50 MPa, 1 MPa, and finally 0.003 MPa, according to [16]. The tumor segmentation was performed manually and FE solutions from polynomial degree p=6 to p=8 were obtained to assess convergence. As

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

expected, varying tumors' E affected the strains in the middle shaft of the right femur, where the tumor was present. In the other regions of the femur the strains remained unchanged. Table 1 presents the average maximum principal strains in the middle shaft for differing E.

[Table 1 about here.]

Table 1 demonstrates a pattern where, as E decreases from 100 MPa to 0.003 MPa there is a minor change in the average maximum principal strains. The maximum change between $E = 100$ and $E = 0.003$ was less than 2% for ϵ_1 and less than 3% for ϵ_3. Furthermore, convergence is observed. The change between p-levels $p = 6$ to $p = 8$ for each E was within 1%. These results suggest that the influence of the mechanical properties of tumor voxels on the computed maximum principal strains is relatively modest. Thus choosing E of 100 MPa for tumors may well represent the mechanical properties of the tumor and avoid an ill-conditioned stiffness matrix.

3.2 Verification of FE Results

To ensure the convergence of the FE results, manual segmentation for annotating the tumors was used, the tumors were "inserted" in the FE models of the twenty femurs, and $E = 100MPa$ was assigned to the tumor's voxels. A p-extension was performed increasing p from 6 to 8 and the error in energy norm was monitored so as to decrease below 5% at $p = 8$. The values of ϵ_1 and ϵ_3 and the number of degrees of freedom (DOF) are presented in Appendix B in Tables 4 and 5 respectively as the p level is increased.

In several femurs, an especially dense mesh resulted in DOFs surpassing 5 million. This was notably observed in the case of Prosp7020, which had a substantial tumor volume as depicted in Appendix B in Figure 8. This led to a dense mesh encompassing a significant portion of the femur, thereby increasing computational time.

In most of the regions of interest, the percent difference between the results at $p = 6$ and $p = 7$ is less than 5% suggesting that $p = 6$ provides results of sufficient accuracy. There are few exceptions such as the Middle Shaft of the right femur of Prosp7020 which has a difference of -16.9% in ϵ_1 between $p = 6$ to $p = 7$, however the value does not change when increasing the p-level to $p = 8$ (0% difference). Similarly, at the distal shaft a significant difference is between $p = 7$ to $p = 8$ (35.3%) indicating a lack of convergence at this region. Also the Distal Shaft in the left femur of ProspD100 shows a difference of 18.8% from $p = 6$ to $p = 7$ but the difference decreases to -0.6% for ϵ_1 and 0.9% for ϵ_3.

3.3 Assessing the Impact of Tumor Segmentation on AFE Analysis

Each of the ten patients' CT scans underwent two AFE pipelines: one using the traditional method [1] while the other integrated the manual tumor segmentation. The tumors were assigned a Young's modulus of 100 MPa for this analysis. The tumor manual segmentation was used for an accurate comparison between the tumor-inclusive and tumor-exclusive analysis (these serve also as a benchmark for DL segmentation as will be shown).

A comparison of the average maximum/minimum principal strains reported by the AFE post-processing stage was conducted. These serve as Simfini's fracture risk assessments [1]. Table 2

presents the average maximum principal strains within each tumor-affected femur region and the total volume of tumors within these regions for the patient cohort. In Table 2, only regions affected by tumors are reported, as no variations in strains were observed in tumor-free regions.

[Table 2 about here.]

Incorporating the lytic tumors in the FE models resulted in larger absolute values for both ϵ_1 and ϵ_3, though the extent of this increase varied among patients and femoral regions. In particular, a notable increase in strain values in the distal shaft region of patient Prosp7020 R was observed, with ϵ_1 and ϵ_3 rising by 74% and 52% respectively. Conversely, in some regions like the proximal and middle shaft for patient ProspB10 L, and the trochanter region for patient Prosp7020 L, no change in strain values was observed. There does not seem to be a connection between the volume of the tumor and its impact on the calculated principal strains. For example, in the Prosp7020 Right femur a relatively small tumor in the trochanter region of 12.4 cm^3 resulted in an increase in ϵ_1 by 59%. On the other hand, in the Prosp7060 Left femur a relatively large tumor in the proximal shaft of 71.8 cm^3 resulted in a minor increase in ϵ_1 by 4%.

In Fig. 4, histograms for the maximum and minimum average principal strains are presented emphasizing the effect of manual tumor segmentation. Each bar quantifies the number of cases within a specific range of difference compared to traditional FE models.

[Figure 4 about here.]

Explicit incorporation of the tumors in the FE models had diverse effects on the strain characteristics of the affected regions. A minor change of less than 6% was observed in 8 out of 23 tumor-affected regions. In 9 (out of 23) regions, the difference in ϵ_1 exceeded 18%, reaching up to 74% in one femur. Similarly, in 11 regions ϵ_3 exceeded 21% difference. These results emphasize the important role of lytic tumors on the biomechanical response.

3.4 Influence of DL Segmentation on the AFE Results

The integration of lytic tumor segmentation using DL into the Autonomous Finite Element (AFE) pipeline was executed seamlessly. The AFE analysis was conducted twice for each patient in our test cohort: once utilizing manual segmentation and once employing automated DL segmentation. The computed average principal strains were then compared. Lytic tumors were assigned a modulus of $E = 100$ MPa, and the results were computed at a polynomial order of $p = 6$. The focus was on variations in tumor-affected regions, as non-affected regions exhibited almost no differences in strain values. The differences between manual and DL segmentation are summarized in Table 3. Additionally, the reported Dice Similarity Coefficient (DSC) scores are presented based on voxels segmented throughout the entire femur. Furthermore, the differences in tumor volume between manual and DL segmentations are presented, specifically accounting for the tumor volume within each region.

The results show an overall agreement between manual and automatic segmentations, with the majority of cases exhibiting less than 20% difference. However, this agreement is not consistently observed, as there are instances where a high Dice Similarity Coefficient (DSC) does not result in

a small difference in average principal strains between manual and DL segmentations. Also, no direct correlation was found between variations in tumor volume (as segmented manually and by DL) and differences in principal strains.

[Table 3 about here.]

The maximum DSC was 0.91 for Prosp7060 L and the minimum was 0 for Prosp7020 R, ProspB10 L and both the left and right femurs of Prosp5050. A large difference in strains was obtained in the right femur of Prosp 1120 between the DL and manual segmentations, despite a high DSC value of 0.85 for the DL segmentation.The mesh for this femur, along with refinements at the tumor and ϵ_1 are illustrated in Figure 3. A significant difference in ϵ_1 is evident; Figure 5 presents the principal strains on the bone's surface near the tumor (indicated by the blue region) in the middle shaft of the Prosp1120 right femur. The DL algorithm did not segment the entire tumor, leading to formation of "isolated islands" of assumed healthy bone tissue within the tumor region. Consequently, as these "islands" exhibited low Hounsfield Unit (HU) values, indicating "soft" tissue, the strains were large, resulting in a considerably larger averaged principal strain (highlighted by red circles in Fig. 5). This artifact contributed to erroneously larger average maximum principal strains compared to the manual segmentation. Conversely, a low DSC of 0.46 for Prosp7060 R at the trochanter region, resulted in only an 11% difference in ϵ_1 and virtually no difference in ϵ_3.

[Figure 5 about here.]

The histograms in Fig. 6 provide a visual representation of the differences in mean principal strains between manual and DL segmentations. Each bar in the histogram was annotated with the number of cases within that specific range of strain difference. For ϵ_1, 18 out of the total 23 cases had differences of less than $\pm 23\%$. For ϵ_3, 19 out of the 23 had differences of less than $\pm 21\%$. Figure 7 displays a scatter plot of the average maximum principal strain in the tumor-affected regions for the three different tumors representation methodologies. In 18 of the 23 locations, the no-segmentation method yielded the smallest strains. The discrepancy between the segmentation methods and the traditional no-segmentation approach (which disregards tumors) revealed differences between the two methods.

[Figure 6 about here.]

[Figure 7 about here.]

4 Discussion

An automatic lytic tumor segmentation algorithm was integrated into the Autonomous Finite Element (AFE) pipeline, and the resulting impact on average strains, serving as a surrogate for fracture risk, was assessed across a cohort of 10 patients. The conventional AFE pipeline, which categorizes tumors as 'soft bone tissues' based on their low Hounsfield Units (HUs), was compared to the AFE pipeline following manual tumor segmentation by an analyst. Additionally, the fully automatic AFE pipeline, incorporating tumor segmentation by DL, was included in the comparative analysis.

Because the average principal strains in the bone tissue are the ones that determine the fracture risk, strains in the tumors are of course of no interest and should be discarded. It was first demonstrated that the very low Young's modulus suggested for the lytic tumors in [16] (0.003 MPa) which can result in ill-conditioned stiffness matrices, can be replaced by a Young modulus Young's modulus of 100 MPa, with an influence on the average strains which is less than 3%. A Young's modulus of 100 MPa for the lytic tumor is two orders of magnitude smaller than the healthy cortical bone tissue. As expected, the low Young's modulus only influences the strains in the close vicinity of the tumor, but not in the tumor-free locations.

The convergence analysis performed on the average maximum principal strains, ϵ_1 and ϵ_3, in tumor-affected femoral regions, offers key insights into the numerical accuracy of the results. A relative difference between solutions at $p = 6$ and $p = 8$ smaller than 5% was considered as a converged verified solution. This was the case for many regions, indicating successful convergence already at $p = 6$. In few regions a larger difference was noticed, suggesting the need to use $p = 8$ solutions. Nevertheless, a recommendation is to use $p = 6$ as it offers satisfactory accuracy within an acceptable computational time of a couple of hours on a desktop.

The influence of integrating manual tumor segmentation into the AFE pipeline was then evaluated. Comparing the average maximum principal strains between the traditional method (tumors are not specifically segmented and a trabecular-like Young's modulus is assigned as if it was a bone tissue with a relatively low HU) and the new approach, showed generally an increase in the absolute values of both ϵ_1 and ϵ_3. The degree of this increase varied among patients and across different femoral regions and tumor sizes. The analysis indicated that segmenting the tumor resulted in a broad spectrum of effects on the strain characteristics within tumor-affected femoral regions. These findings highlighted the influence of lytic tumors on the biomechanical response in the affected femoral regions and advocated for the explicit representation of the tumors in the FE models.

The histogram in Figure 6 represents the average max and min principal strain obtained by the manual and DL segmentation methodologies. The differences in most of the cases were within 22% difference. Out of the five outliers, three were the result of a DSC of 0. For example, in the Proximal Shaft and Middle Shaft of ProspB10 left femur, a significant difference was noticed between the manual and DL segmentation methodologies. The DSC for this femur was 0, i.e. the

manual segmentation did not identify any tumor whereas the DL algorithm erroneously segmented a tumor (False-Positive occurrences). This erroneous identification of a tumor had a large impact with strain difference of 49% for the middle shaft region. This phenomenon, however, was not detected in other femoral regions.

4.1 Conclusions

In most cases the results obtained using manual segmentation are close to these obtained by the automatic DL segmentation, except for several outliers, implying the possible use of the later in an automatic finite element algorithm. Nonetheless, high DSC scores for the DL algorithm or a small difference in tumor size do not guarantee small differences in strains. This implies that minor alterations at the tumor boundaries may lead to shifts in fracture risk assessment.

4.2 Limitations and Future Required Research

Several limitations in our study may be addressed by future required research activity:

- A limited cohort size for the test dataset was considered, which might not capture the full range of variability present in a larger population.
- Our analysis was confined to a single loading condition, potentially limiting the comprehensiveness of the biomechanical insights derived.
- The main focus in this work is on lytic tumors. The findings might not be directly applicable to other tumor types or pathologies. Mixed and blastic tumors impact on the AFE results are planned to be investigated.
- The implications of our results will be validated in a follow-up publication by clinical observations. Specifically, the results will be cross-referenced with actual patients who have experienced femur fractures due to tumors. A similar study as in [3] is warranted to determine whether explicit segmentation of tumors may better predict patients at high risk of a pathological fracture.
- Mesh refinement in the tumor region may be optimized by employing two distinct meshes for the healthy bone tissue and tumor and low polynomial degrees could be used for the shape functions in the elements in the tumor region.
- The medullary cavity is represented in Simfini as a soft bone tissue known to have a minor effect on the biomechnical response in healthy femurs. The impact of such a representation on the biomechanical response should be re-evaluated for femurs with lytic tumors.

Acknowledgements

OR, ZY and AS acknowledge the support of this research by the Israel Ministry of Science and Technology under the Tenth Call of Israel-Italy Scientific collaboration.

References

[1] Z. Yosibash, K. Myers, N. Trabelsi, and A. Sternheim. Autonomous FEs (AFE) - A stride toward personalized medicine. *Comp. Math. App.*, 80:2417–2432, 2020.

[2] O. Rachmil, M. Artzi, M. Iluz, I. Druckmann, O. Merose, Z. Yosibash, and A. Sternheim. Autonomous finite element analyses of femurs with metastatic tumors. Part I - Segmentation of femoral tumors by nnU-Net. Submitted for Publication, 2023.

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

[4] A. Sternheim, F. Traub, N. Trabelsi, S. Dadia, Y. Gortzak, N. Snir, M. Gorfine, and Z. Yosibash. When and where do patients with bone metastases actually break their femurs? A CT-based finite element analysis. *Bone & Joint Jour*, 102B(5):638–645, 2020.

[5] JH Keyak, TS Kaneko, J Tehranzadeh, and HB Skinner. Predicting proximal femoral strength using structural engineering models. *Clinical Orthopaedics and Related Research*, 437:219–228, August 2005.

[6] J.H. Keyak, T.S. Kaneko, S.A. Rossi, M.R. Pejcic, J. Tehranzadeh, and H.B. Skinner. Predicting the strength of femoral shafts with and without metastatic lesions. *Clinical Orthopaedics and Related Research*, pages 161–170, 2005. 2005b.

[7] E. Benca, A. Synek, M. Amini, F. Kainberger, L. Hirtler, R. Windhager, W. Mayr, and D.H. Pahr. Qct-based finite element prediction of pathologic fractures in proximal femora with metastatic lesions. *Sci. Rep.*, 9:1–9, 2019.

[8] J.R. Goodheart, R.J. Cleary, T.A. Damron, and K.A. Mann. Simulating activities of daily living with finite element analysis improves fracture prediction for patients with metastatic femoral lesions. *J. Orthop. Res.*, 33:1226–1234, 2015.

[9] O Ronneberger, P Fischer, and T Brox. U-Net: Convolutional networks for biomedical image segmentation. In N Navab, J Hornegger, WM Wells, and AF Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241. Springer International Publishing, 2015.

[10] TS Kaneko, JS Bell, MR Pejcic, J Tehranzadeh, and JH Keyak. Mechanical properties, density and quantitative ct scan data of trabecular bone with and without metastases. *J Biomech*, 37(4):523–530, Apr 2004.

[11] Z. Yosibash, Y. Katz, N. Trabelsi, and A. Sternheim. Femurs segmentation by machine learning from CT scans combined with autonomous finite elements in orthopedic and endocrinology applications. *Comp. Math. Appl.*, 2023. In Print.

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

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

[14] N. Trabelsi and Z. Yosibash. Patient-specific fe analyses of the proximal femur with orthotropic material properties validated by experiments. *ASME J. Biomech. Eng.*, 155:061001–1–061001–11, 2011.

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

[16] C.M. Whyne, S.S. Hu, K.L. Workman, and J.C. Lotz. Biphasic material properties of lytic bone metastases. *Annals of Biomedical Engineering*, 28:1154–1158, 2000.

[17] G. Bergmann. OrthoLoad Database. http://www.OrthoLoad.com, 2008.

[Table 5 about here.]

[Figure 13 about here.]

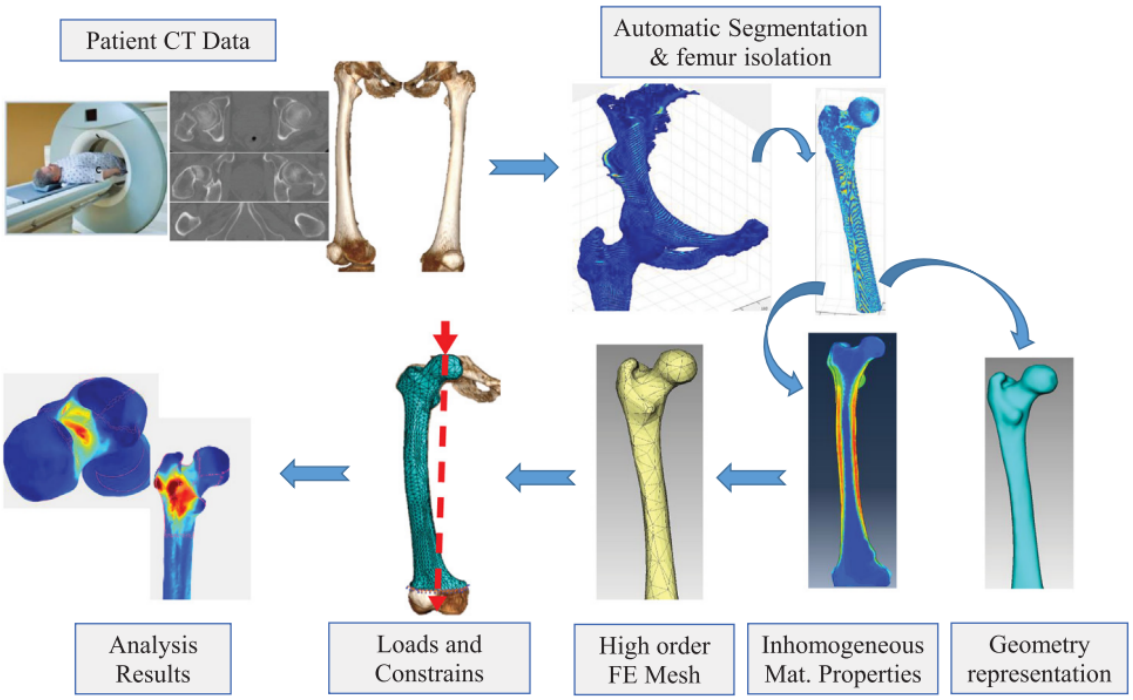


Figure 1 Illustration depicting the distinct components of Simfini (from [1].)

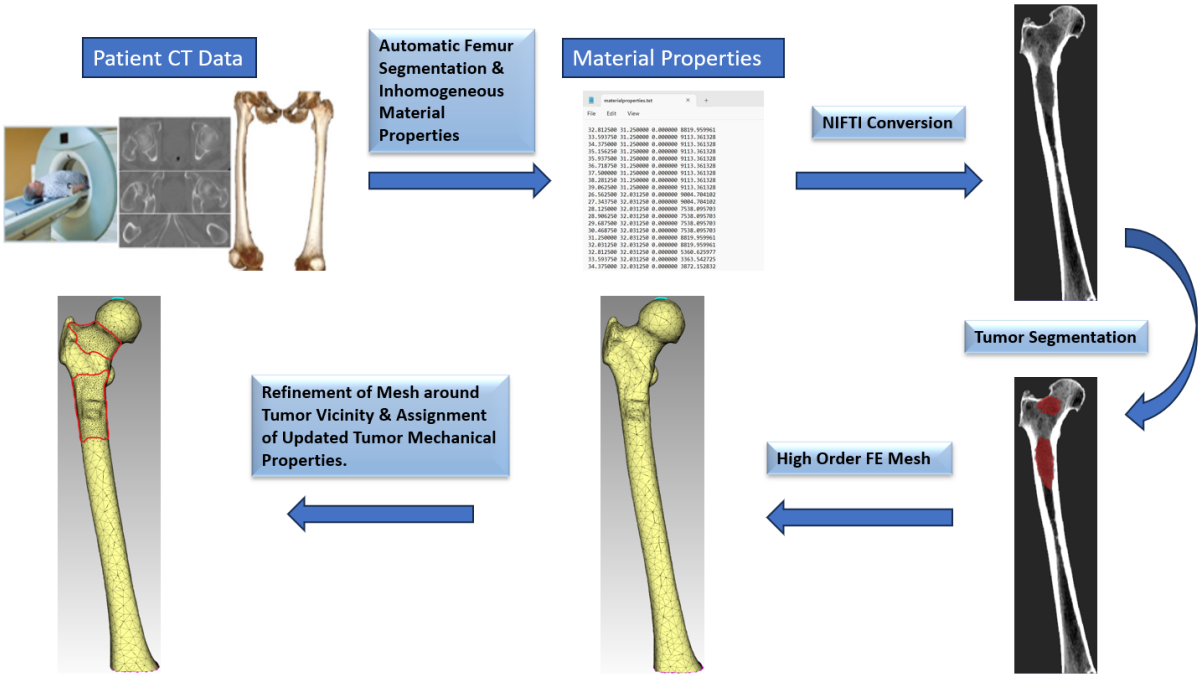


Figure 2 Mesh generation and refinement workflow.

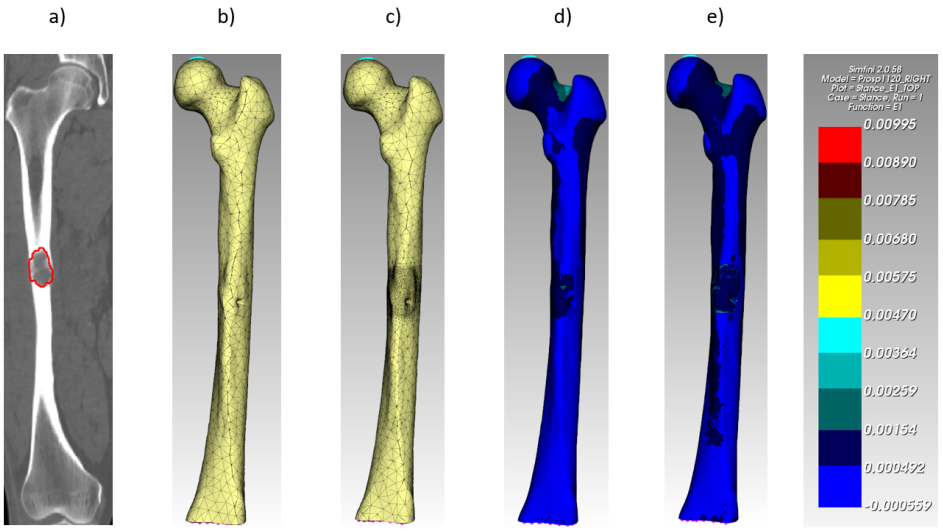


Figure 3 Prosp1120's right femur. (a) Displays the bone in the CT scan with delineated tumor around the middle shaft as observed by the analyst (marked by a red polygonal boundary). (b) + (c) Presents the p-mesh, with and without local refinement around the tumor in the middle shaft region. (d) shows the map of ϵ_1 principal strain on the original mesh while (e) shows it on the refined mesh. $E = 100MPa$ was used.

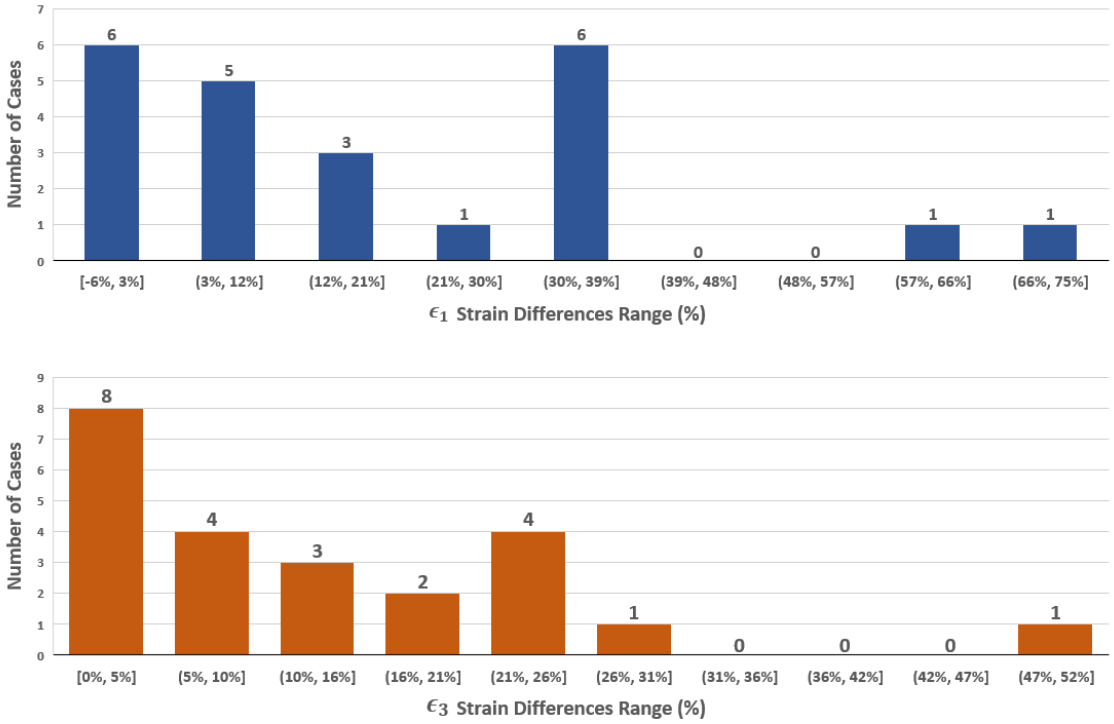


Figure 4 Histogram of average principal strain differences between traditional and manual tumor segmented FE models. Top: ϵ_1, Bottom: ϵ_3.

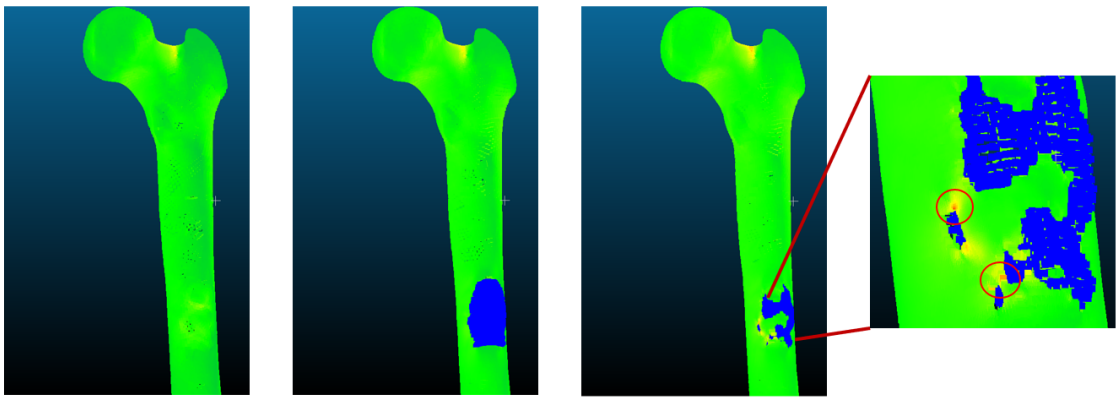


Figure 5 Principal strains ϵ_1 for Prosp1120. Left: No segmentation of tumor, Middle: Manually segmented tumor, Right: DL segmented tumor with zoom-in on the infected tumor area.

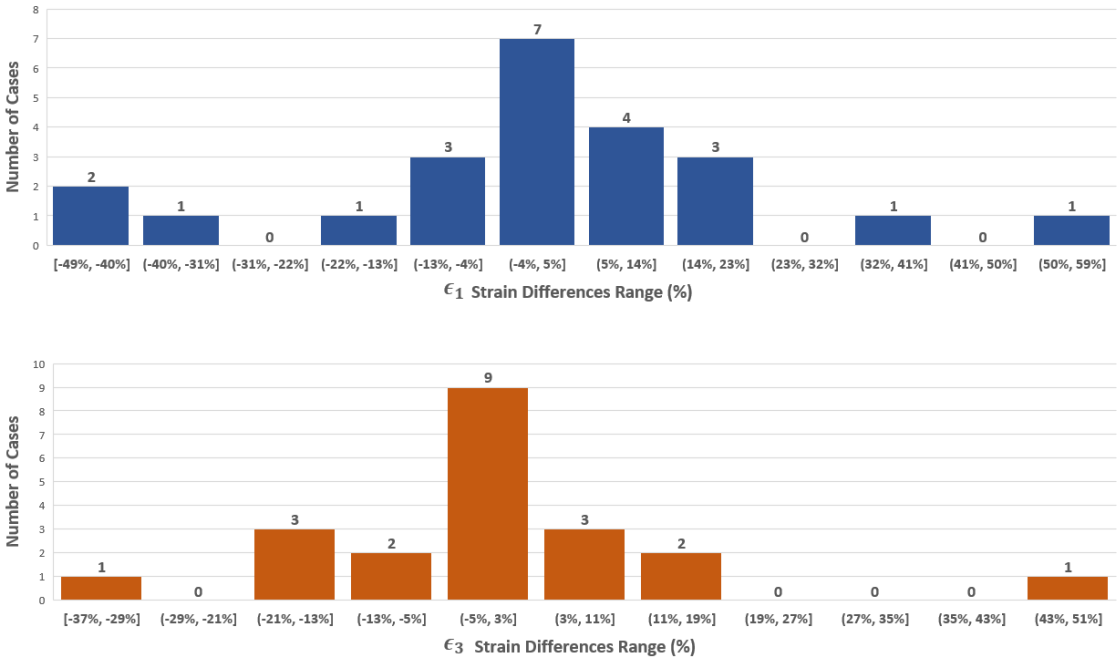
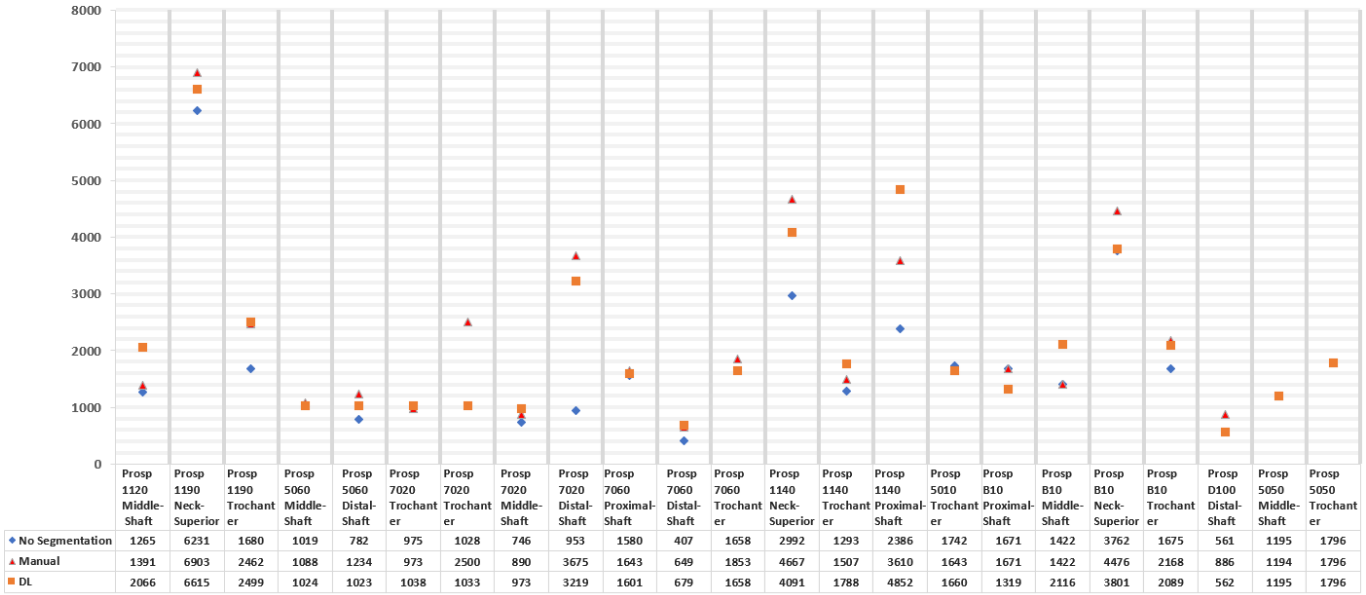


Figure 6 Histogram illustrating differences in computed average maximum principal strains (ϵ_1 and ϵ_3) between manual and DL automatic tumor segmentation. Top: ϵ_1, Bottom: ϵ_3.



	Prosp 1120 Middle-Shaft	Prosp 1190 Neck-Superior	Prosp 1190 Trochanter	Prosp 5060 Middle-Shaft	Prosp 5060 Distal-Shaft	Prosp 7020 Trochanter	Prosp 7020 Trochanter	Prosp 7020 Middle-Shaft	Prosp 7020 Distal-Shaft	Prosp 7060 Proximal-Shaft	Prosp 7060 Distal-Shaft	Prosp 7060 Trochanter	Prosp 1140 Neck-Superior	Prosp 1140 Trochanter	Prosp 1140 Proximal-Shaft	Prosp 5010 Trochanter	Prosp B10 Proximal-Shaft	Prosp B10 Middle-Shaft	Prosp B10 Neck-Superior	Prosp B10 Trochanter	Prosp D100 Distal-Shaft	Prosp 5050 Middle-Shaft	Prosp 5050 Trochanter
◆ No Segmentation	1265	6231	1680	1019	782	975	1028	746	953	1580	407	1658	2992	1293	2386	1742	1671	1422	3762	1675	561	1195	1796
▲ Manual	1391	6903	2462	1088	1234	973	2500	890	3675	1643	649	1853	4667	1507	3610	1643	1671	1422	4476	2168	886	1194	1796
■ DL	2066	6615	2499	1024	1023	1038	1033	973	3219	1601	679	1658	4091	1788	4852	1660	1319	2116	3801	2089	562	1195	1796

Figure 7 Scatter plot of the maximum average ϵ_1 for three tumor segmentation strategies (no-segmentation, manual, and DL) in all femoral regions affected by tumors. Each region is delineated by a thick horizontal line, dividing the plot into distinct columns. The table below provides the calculated strains for each method per region, alongside a legend for a clearer interpretation of the graph.

List of Tables

Table 1 Comparison of average maximum principal strains (Compression and Tension) for varying E at different polynomial orders (p) in the middle shaft of the right femur of patient Prosp1120.

E **(MPa)**	**Tension** (ϵ_1) $[\mu Strain]$			**Compression** (ϵ_3) $[\mu Strain]$		
	p=6	**p=7**	**p=8**	**p=6**	**p=7**	**p=8**
100	1391	1397	1400	−3705	−3716	−3729
50	1400	1406	1409	−3753	−3764	−3779
1	1410	1417	1420	−3805	−3821	−3838
0.003	1411	1418	1422	−3808	−3828	−3859

Table 2 Difference in the average principal strain at tumor-affected femoral regions between Traditional (T), i.e. no segmentation of tumors, and manual tumor segmentation (MS) FE models. FE solutions at $p = 6$ were used.

Femur #	**Region**	**Tumor Vol. (cm^3)**	ϵ_1		**Diff ϵ_1 (%)**	ϵ_3		**Diff ϵ_3 (%)**
			T	**MS**		**T**	**MS**	
Prosp5010 R	Trochanter	5.9	1742	1643	−6.0	−2836	−2924	3.0
Prosp7060 L	Distal Shaft	49.3	407	649	37.0	−1032	−1210	14.7
Prosp7060 L	Proximal Shaft	71.8	1580	1643	4.0	−2681	−2902	7.6
Prosp5060 L	Middle Shaft	6.3	1019	1088	6.0	−1769	−2053	13.8
Prosp5060 L	Distal Shaft	25.5	782	1234	37.0	−1761	−2387	26.2
Prosp1120 R	Middle Shaft	21.5	1265	1391	9.0	−3485	−3705	5.9
Prosp1140 R	Trochanter	10.0	1293	1507	14.0	−3935	−4851	18.8
Prosp1140 R	Proximal Shaft	29.8	2386	3610	34.0	−6932	−9078	23.6
Prosp1140 R	Neck Superior	15.2	2992	4667	36.0	−2687	−3615	25.6
Prosp1190 L	Neck Superior	11.3	6231	6903	10.0	−4889	−6931	29.4
Prosp5050 R	Trochanter	0.3	1796	1796	0.0	−2941	−2941	0.0
Prosp7020 R	Middle Shaft	102.2	746	890	16.0	−1600	−2030	21.1
Prosp7020 R	Distal Shaft	30.9	953	3675	74.0	−2228	−4668	52.2
Prosp7020 R	Trochanter	12.4	1028	2500	59.0	−1699	−1784	4.7
Prosp5050 L	Middle Shaft	1.7	1195	1194	0.0	−1616	−1616	0.0
Prosp1190 R	Trochanter	14.9	1680	2462	32.0	−2586	−2969	12.9
Prosp7020 L	Trochanter	18.8	975	973	0.0	−1582	−1582	0.0
ProspB10 R	Neck Superior	8.1	3762	4476	16.0	−4270	−4690	8.9
ProspB10 R	Trochanter	35.8	1675	2168	23.0	−3296	−3661	9.9
ProspD100 L	Distal Shaft	15.3	561	886	37.0	−1074	−1292	16.8
Prosp7060 R	Trochanter	2.7	1658	1853	11.0	−2768	−2779	0.3
ProspB10 L	Proximal Shaft	0	1671	1671	0.0	−2652	−2652	0.0
ProspB10 L	Middle Shaft	0	1422	1422	0.0	−1893	−1893	0.0

Table 3 Comparison of Average Principal Strains, Tumor Volumes, and Dice Similarity Coefficients (DSC) in Tumor-Affected Femoral Regions between Manual Tumor Segmentation (MS) and Automatic Tumor Segmentation by DL. Finite Element (FE) solutions at $p = 6$ were utilized for the analysis.

Femur#	**Region**	**DSC**	**Tumor Vol. Diff (cm^3)**	ϵ_1		**Diff ϵ_1 (%)**	ϵ_3		**Diff ϵ_3 (%)**
				DL	**MS**		**DL**	**MS**	
Prosp5010 R	Trochanter	0.88	-0.1	1660	1643	−1	−2956	−2924	−1
Prosp7060 L	Distal Shaft	0.87	11.1	679	649	−5	−2814	−2902	3
Prosp7060 L	Proximal Shaft	0.87	13.7	1601	1643	3	−1460	−1210	−21
Prosp5060 L	Middle Shaft	0.85	-0.9	1024	1088	6	−1993	−2053	3
Prosp5060 L	Distal Shaft	0.85	-3.7	1023	1234	17	−2132	−2387	11
Prosp1120 R	Middle Shaft	0.85	3.3	2066	1391	−49	−5065	−3705	−37
Prosp1140 R	Trochanter	0.84	0.7	1788	1507	−19	−3279	−3615	9
Prosp1140 R	Proximal Shaft	0.84	1.6	4852	3610	−34	−5462	−4851	−13
Prosp1140 R	Neck Superior	0.84	0.6	4091	4667	12	−8807	−9078	3
Prosp1190 L	Neck Superior	0.84	2.31	6615	6903	4	−6756	−6931	3
Prosp5050 R	Trochanter	0.83	0.1	1796	1796	0	−2941	−2941	0
Prosp7020 R	Middle Shaft	0.75	1.9	973	890	−9	−1717	−1784	4
Prosp7020 R	Distal Shaft	0.75	3.2	3219	3675	12	−2131	−2030	−5
Prosp7020 R	Trochanter	0.75	20.5	1033	2500	59	−4071	−4668	13
Prosp5050 L	Middle Shaft	0.71	0.4	1195	1194	0	−1616	−1616	0
Prosp1190 R	Trochanter	0.67	-2.3	2499	2462	−2	−3034	−2969	−2
Prosp7020 L	Trochanter	0.65	-15.0	1038	973	−7	−1607	−1582	−2
ProspB10 R	Neck Superior	0.62	-14.2	3801	4476	15	−5624	−4690	−20
ProspB10 R	Trochanter	0.62	-20.8	2089	2168	4	−3557	−3661	3
ProspD100 L	Distal Shaft	0.52	7.3	562	886	37	−1090	−1292	16
Prosp7060 R	Trochanter	0.46	-0.8	1658	1853	11	−2768	−2779	0
ProspB10 L	Proximal Shaft	0	-16.8	1319	1671	21	−1474	−2652	44
ProspB10 L	Middle Shaft	0	-39.7	2116	1422	−49	−2177	−1893	−15