Prediction of the mechanical response of the femur with uncertain elastic properties

Hagen Wille[a], Ernst Rank[a], Zohar Yosibash[b,*]

[a] *Chair for Computation in Engineering, Technische Universität München, Munich, Germany*
[b] *Dept. of Mechanical Engineering, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel*

Abstract

A mandatory requirement for any reliable prediction of the mechanical response of bones,based on quantitative computer tomography, is an accurate relationship between material properties (usually Young's modulus E) and bone density ρ. Many such E-ρ relationships are available based on different experiments on femur specimens with a large spread due to uncertainties. The first goal of this study is to pool and analyze the relevant available experimental data and develop a *stochastic* E-ρ relationship. This analysis highlights that there is no experimental data available to cover the entire density range of the human femur and that some "popular" E-ρ relationships are based on data that contains extreme scatter, while others are based on a very limited amount of information.

The second goal is to use the newly developed stochastic E-ρ relationship in high-order finite element analyses (FEAs) for the computation of strains and displacements in two human proximal femurs, mimicking in vitro ex-

*Corresponding Author
Email address: zohary@bgu.ac.il (Zohar Yosibash)

periments. When compared with the experimental observations, the FEA predictions using the median of the stochastic E-ρ relationship follow the underlying distribution of the stochastic E-ρ relationship. Thus, most deviations of the FEA predictions from experimental observations can possibly be explained by uncertain elastic properties of the femur.
Keywords: femur, finite-elements, uncertainty, constitutive-relation, validation

1. Introduction

Patient-specific finite element (FE) analyses for predicting the mechanical response of human femurs are performed by many researchers nowadays (Bessho et al., 2007; Helgason et al., 2008b; Keyak et al., 1993; Schileo et al., 2008; Trabelsi et al., 2009; Yosibash et al., 2007b). The topology of the FE models is determined by manipulating quantitative computer tomography (qCT) scans, and isotropic heterogeneous material properties are usually assigned on the basis of a mathematical relationship between a density measure (either apparent density ρ_{app}, equivalent mineral density ρ_{eqm}, or ash density ρ_{ash}) and Young's modulus E. Most researchers contemplate a *deterministic* isotropic material with a heterogeneous Young's modulus related to the density from the point of view of a relationship based on experiments conducted on small pieces of bone tissue:

$$E = a\,\rho_{\text{x}}^{b}, \quad \rho_{\text{x}} \text{ in } [g/cm^3],\ E \text{ in } [MPa] \tag{1}$$

where ρ_{x} is the densitometric measure, and a and b are generic constants. Many such E-ρ relationships are documented in the literature, each determined by a single set of experiments, and adopted in the generation of

subject-specific FE models of the whole organ. None of the previous work pooled the various relevant experiments to derive a relationship based on all data.

All experimental data for E-ρ relationships show a large variation with scattering around the mean; see the review of Helgason et al. (2008a). These relationships have a strong influence on the mechanical response of the whole femur and several works investigated which of the deterministic relations "best represent the physical reality" - see (Pise et al., 2009; Taddei et al., 2007; Yosibash et al., 2007a), for example, and references therein.

Due to the uncertainty in the E-ρ relationships we employ *probabilistic* methods to investigate the uncertainty's influence on the simulated mechanical response. A review of probabilistic methods and their application to bonemechanics can be found in (Laz and Browne, 2010), and several relating particularly to the human-femur in (Dopico-Gonzalez et al., 2009; Laz et al., 2007; Nicolella et al., 2006; Taddei et al., 2006; Viceconti et al., 2006). Depending on the application, these studies also incorporate uncertainties related to geometry and loading conditions. However, some of these studies (Dopico-Gonzalez et al., 2009; Nicolella et al., 2006; Viceconti et al., 2006) apply a simplified homogeneous material which uses the same random variable everywhere to describe the Young's modulus and which is independent of ρ. In (Laz et al., 2007; Taddei et al., 2006) an E-ρ relationship in the form (1) is used, with a and b modeled as Gaussian random variables based on the experimental results reported by Keller (1994). In this way, the standard errors for the *estimation* of the regression parameters are used as standard deviations, while the estimates for a and b are used as mean values. Thus,

only the uncertainty related to the parameter estimation is investigated (resulting from a finite number of experimental observations), but not the uncertainty within the experimental data. From a purely statistical viewpoint, the appropriate mathematical procedure is to use a regression model with one random variable which describes the deviation of the data from the regression function, see (Abraham and Ledolter, 2006, Section 2.5 and Chapter 3). Apart from that, the cortical specimens tested by Keller (1994) are from both transversal and longitudinal directions so that the scatter in that data is very large compared to other experiments, which results in an overestimation of the uncertainty.

None of the reviewed studies investigated an E-ρ relationship based on *multiple* experimental studies or identified the influence of the variance within such a relationship on the mechanical response of a patient-specific FE-model.

The aim of this manuscript is to provide the answer to the following questions: a) Can a stochastic $E(\rho)$ relationship be derived for the human femur by pooling the data from past specimen tests, and how large is the scatter? b) How does the uncertainty manifested in the $E(\rho)$ relationship propagate through patient-specific finite element analyses aimed at predicting the mechanical response?

2. Methods

To answer the first question we scrutinized the experiments on which E-ρ relationships were determined for the *human femur*. We pooled E-ρ data from the different sources, analyzed it, and generated a new stochastic E-ρ relationship which incorporates the uncertainty related to the various exper-

iments. We then used the new relationship for patient-specific FE analyses of two femurs and investigated the propagation of the E-ρ relationship uncertainty. Finally, we compared the predicted mechanical response with in vitro experimental observations. The following section describes the various steps and methods in more detail.

2.1. Data acquisition

A literature survey on publications reporting on E-ρ relationships for human bone tissues (applicable to femurs) was performed and only data obtained from femurs was taken into account (Rice et al., 1988; Lotz et al., 1991; Keyak et al., 1994; Keller, 1994; Wirtz et al., 2000; Morgan et al., 2003; Kaneko et al., 2003, 2004; Helgason et al., 2008a).

These studies used no standard protocol and considered femurs from donors of both genders spanning a wide age group. In addition, data stemed from different specimen shapes and dimensions, different loading configurations, and E was employed as a function of different densitometric measures.

We choose ash density (ρ_{ash}) for pooling all relevant data, so other measures were converted to ρ_{ash} using the linear relations described by Keyak et al. (1994):

$$\rho_{\text{ash}}(\rho_{\text{dry}}) = 0.597\rho_{\text{dry}} - 0.00191 \tag{2}$$

$$\rho_{\text{ash}}(\rho_{\text{wet}}) = 0.551\rho_{\text{wet}} - 0.00478 \tag{3}$$

with ρ_{dry} denoting the dry density, defined as the ratio of defatted tissue mass to the overall volume of the specimen. On the other hand, ρ_{wet}, also known as ρ_{app}, stands for the apparent wet density, which is determined from the weight of the specimen after an additional rehydration procedure has been

applied. Table 4 summarizes the different experiments on specimens from femurs in a rigorous manner for a pooled analysis.

2.2. Determining a stochastic E-ρ relationship based on the pooled data

Experiments on small bone specimens with similar densities show a wide spread in the measured Young's modulus (cf. data in Figure 3). A regression analysis was conducted on the *pooled experimental data* in order to estimate the regression coefficients $\boldsymbol{\alpha}$ such that the regression function $f(\rho, \boldsymbol{\alpha})$ best represents the E-ρ relationship "on average". The corresponding regression model contains an additional error term X, which is a random variable representing the scatter for all experimental observations from the regression function. We assume X to be normally (Gaussian) distributed and to represent all relevant sources of deviations. After transferring the data into a double logarithmic scale[1], we used a linear regression function

$$f(\ln \rho, \boldsymbol{\alpha}) = \alpha_0 + \alpha_1 \ln \rho$$

suggesting a regression model of the type:

$$\ln E = \alpha_0 + \alpha_1 \ln \rho + X \tag{4}$$

with $X \sim \mathcal{N}(0, \sigma_X^2)$. This model assumes the error term to be independent of the regression function; both the mean and the variance of X are constants. Hence, we also require all experimental observations to be independent of each other.

The transformation into double logarithmic scale is required as the assumption of an error term with constant variance implies that (4) can only

[1]The terminology *logarithmic* refers to the natural logarithm.

be used on data which shows a homoscedastic[2] behavior. In the original scale the pooled data appears heteroscedastic (see Figure 3).

Following (Abraham and Ledolter, 2006), the method of maximum likelihood is used to obtain estimates for the unknown model parameters (regression coefficients and variance of X). For the model in question the maximum likelihood estimates for $\boldsymbol{\alpha}$ are equivalent to the least squares estimates $\hat{\boldsymbol{\alpha}}$, obtained by an ordinary least squares fit. Similarly, the estimate s^2 for the variance of the error term can be interpreted as the standard error of the residuals of the least squares fit. Further details are provided in Appendix A.

Given experimental data as a set of N independent measurements (E_i, ρ_i) with $i = 1, 2, \ldots, N$, the aforementioned estimates are used to derive the material relationship

$$\ln E = \hat{\alpha}_0 + \hat{\alpha}_1 \ln \rho + Y \tag{5}$$

with $Y \sim \mathcal{N}(0, s_Y^2)$. Equation (5) is reformulated, so the error term becomes a scaling factor of a power material relationship:

$$E^{\mathrm{sto}} = a\, \rho^{\, b} \cdot Z \tag{6}$$

with $Z = \exp(Y) \sim \ln \mathcal{N}(0, s_Y^2)$, $a = \exp(\hat{\alpha}_0)$ and $b = \hat{\alpha}_1$. In (6) the distribution of the random variable Z is *log-normal*, which is more realistic

[2] An assumption of the fitted model is that the standard deviations of the error terms are constant and do not depend on $\ln \rho$. Consequently, each probability distribution for $\ln E$ has the same standard deviation regardless of the $\ln \rho$ value. This assumption is called homoscedasticity.

than the normal distribution because Young's modulus cannot be negative. Note that $Z = \exp(0) = 1$ defines the *median* of the log-normal distributed error term in (6). Hence,

$$E^{\mathrm{det}} = a\,\rho^{b} \tag{7}$$

separates the higher (stiffer) half of all realizations from the lower (softer) half and we call E^{det} the *new deterministic* material relationship[3].

2.3. Statistical tests

The stochastic material relationship (6) was derived under the assumption that the error term Y is independent of ρ and normally distributed. Therefore, all residuals[4] have to show a homoscedastic behavior and must be normally distributed. Formulating these requirements as hypotheses makes it possible to use statistical tests to check whether the regression model (4) can be used on the data.

The Lilliefors test (Thode, 2002, Sec. 5.1.1) was used to test the hypothesis that the residual values come from a normally distributed population, and a studentized Breusch–Pagan test (Koenker, 1981) to check the hypothesis of homoscedastic residuals. Both tests were conducted using the statistical software environment R, version 2.13.1 (R Development Core Team, 2011).

Even though these tests are necessary, they are not sufficient. In other words, even if the hypothesis can not be rejected, this does not automatically prove the opposite. It is, however, possible to use characteristic plots of the

[3]Such deterministic relationships are used in almost all FE simulations.

[4]The residual r_i of a sample is defined as the deviation between the sample value and its fitted value, i.e. $r_i = \ln E_i - \ln E^{\mathrm{det}}(\rho_i)$.

data like a residual plot and a normal probability plot in order to support the hypotheses.

2.4. In vitro experiments and p-FE simulations of femurs

FE simulations of the mechanical response of human femurs are sensitive to the E-ρ relationship, see (Yosibash et al., 2007a; Helgason et al., 2008a). According to our past experience, the best deterministic relationship known so far is based on the combination of Keller's and Keyak's law and provides the closest FE results compared to in vitro experiments (Trabelsi et al., 2009, 2011). The "Keller–Keyak" relationship is:

$$E_{\text{Trab}} = 33900\,\rho_{\text{ash}}^{2.20} \qquad [MPa] \qquad \rho_{\text{ash}} \leq 0.27 \tag{8}$$

$$E_{\text{Trab}} = 469 + 5307\,\rho_{\text{ash}} \qquad [MPa] \qquad 0.27 < \rho_{\text{ash}} \leq 0.6 \tag{9}$$

$$E_{\text{Cort}} = 10200\,\rho_{\text{ash}}^{2.01} \qquad [MPa] \qquad 0.6 < \rho_{\text{ash}} \tag{10}$$

with $\nu = 0.3$. It is a heuristic model combining data for $\rho_{\text{ash}} \leq 0.27$ with data for $\rho_{\text{ash}} > 0.6$ by a linear relationship.

Both "Keller–Keyak" and the new deterministic relationship (7) were used in p-FE analyses of two human femurs to predict the mechanical response.

The p-FE analyses were validated by a series of in vitro experiments on fresh frozen femurs, as detailed in (Yosibash et al., 2007b; Trabelsi et al., 2009; Trabelsi and Yosibash, 2011). Here, we used the two most recent in vitro experimental results on the femurs called FF4 and FF5. Details of the CT scan resolution and load-rate in the experiments are summarized in Table 1. The femurs were qCT scanned and within one day of defrosting, experiments were conducted to mimic a simple stance position in which the femurs were

loaded by a 1000 N force through their head while inclined at 0 and 7 degrees (see Figure 1). We measured the vertical and horizontal displacements of the femur's head, and the strains at the inferior and superior parts of the neck, as well as on the medial and lateral femoral shaft. Data was successfully collected from about 12 strain gauges on each femur.

Figure 2 depicts the entire procedure from qCT scan to FEA. For more details on the used algorithms and the experimental protocol we refer the reader to (Yosibash et al., 2007b; Trabelsi et al., 2009; Trabelsi and Yosibash, 2011).

2.5. Propagation of material-relation uncertainty to the mechanical response

For a linear elastic analysis where only the Young's modulus is stochastic, $E^{\text{sto}} = Z \cdot E^{\text{det}}$, Hooke's law may be expressed by a material tensor $\mathbb{C}(\rho_{\text{ash}})$ that scales directly with the Young's modulus. This allows the factorization

$$\mathbb{C}^{\text{sto}} = Z \cdot \mathbb{C}^{\text{det}} \tag{11}$$

Therefore, the FE stiffness matrix is:

$$\mathbf{K}^{\text{sto}} = \int_{\Omega} \mathbf{B}^T \, \mathbb{C}^{\text{sto}} \, \mathbf{B} \, d\Omega = Z \cdot \int_{\Omega} \mathbf{B}^T \, \mathbb{C}^{\text{det}} \, \mathbf{B} \, d\Omega = Z \cdot \mathbf{K}^{\text{det}} \tag{12}$$

Table 1: Data of femurs and CT scan resolution.

Donor label	Side	Age (years)	Gender	Slice thickness (mm)	Pixel size (mm)	Load rate (mm/sec)
FF4	R	63	male	1.25	0.195	1/60, 1/6, 1
FF5	R	56	male	1.25	0.26	1/2

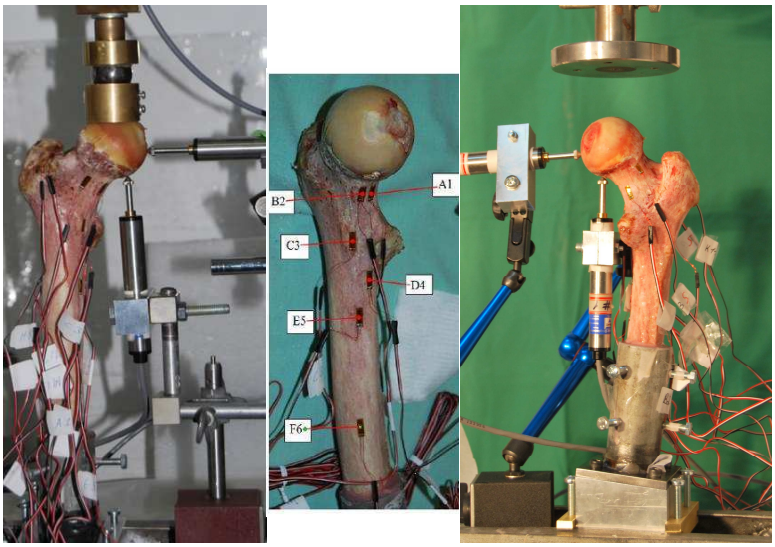


Figure 1: Typical experiments on FF4 and FF5 at 0^o and 7^o inclination angles.

where $\mathbf{B}$ is the matrix containing the shape function derivatives (Zienkiewicz et al., 2005, p. 204). The stochastic p-FE solution (the displacements) $\mathbf{u}^{\text{sto}}$ is accordingly computed by

$$\mathbf{u}^{\text{sto}} = \left(\mathbf{K}^{\text{sto}}\right)^{-1} \mathbf{f} = 1/Z \cdot \left(\mathbf{K}^{\text{det}}\right)^{-1} \mathbf{f} = 1/Z \cdot \mathbf{u}^{\text{det}} \tag{13}$$

where $\mathbf{f}$ denotes the load vector (which is deterministic in our case) and $\mathbf{u}^{\text{det}}$ the p-FE displacement vector for a deterministic simulation. Since Z is a random variable with log-normal distribution and mean zero, its reciprocal has the same identical distribution. In other words, if $Z \sim \ln \mathcal{N}(0, s^2)$, then $1/Z \sim \ln \mathcal{N}(0, s^2)$. Therefore, the stochastic response is the product of the

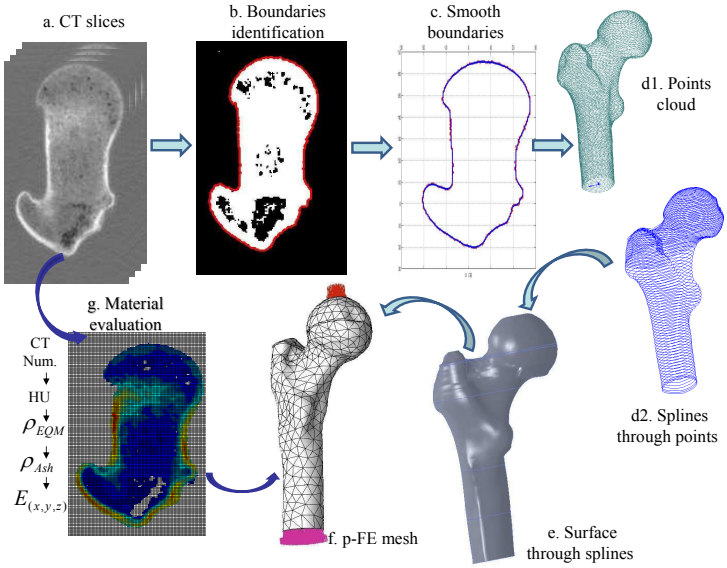


Figure 2: Schematic flowchart describing the generation of the p-FE model from qCT scans. a - Typical CT-slice, b. - Contour identification, c. - Smoothing boundary points, d1. - Points cloud representing the bone surface. d2. - Closed splines for all slices, e. - Bone surface, f. - p-FE mesh and g. - Material evaluation from CT data.

deterministic result and the random variable Z:

$$\mathbf{u}^{\text{sto}} = Z \cdot \mathbf{u}^{\text{det}}. \tag{14}$$

One may observe that stochastic strains are computed by applying the gradient operator to the stochastic displacement field, resulting in the same probability distribution. Accordingly, just a single simulation based on the new deterministic material relationship is needed to quantify the uncertainty in the femur's mechanical response.

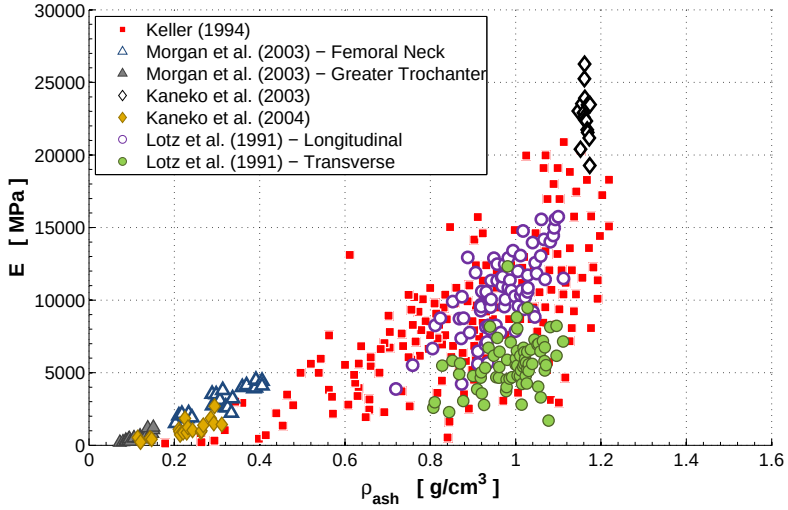


Figure 3: All pooled experimental data, partially recovered from the literature and converted into the same density measure (ρ_{ash}) using Eq. (2)-(3).

3. Results

3.1. A stochastic E-ρ relationship

Figure 3 shows the pooled data from publications listed in Table 4. Since the transversal Young's modulus in the cortical bone is significantly lower than the longitudinal, the recorded data by Keller (1994) has to be discarded as his study makes no distinction between the different directions. This also explains the large scatter in that study. The analysis of the data from Lotz et al. (1991), where a clear distinction is made between longitudinal and transverse directions, indeed shows that it covers the entire span of Keller's results. The final data chosen for inclusion in the analysis contains $N = 161$ measurements and appears homoscedastic in a double-logarithmic scale (see Figure 4).

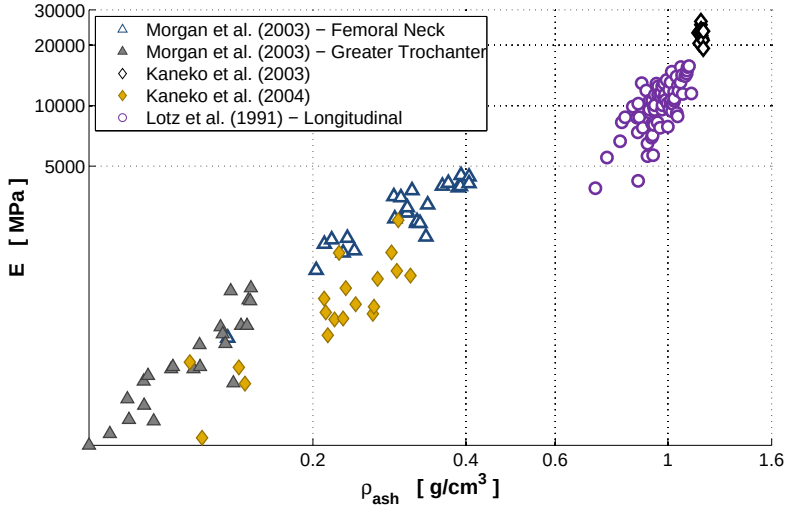


Figure 4: Final included data presented in double logarithmic scale, number of experimental results $N = 161$.

A regression analysis is performed on this data using the model described in (4), which yields the stochastic relationship

$$E^{\text{sto}} = 12\,000\,\rho_{\text{ash}}^{1.45} \cdot Z, \quad Z \sim \ln \mathcal{N}(0, s^2 = 0.1). \tag{15}$$

All regression coefficients were found to be statistically significantly different from zero ($P < 0.001$), and the coefficient of determination $R^2 = 0.941$ indicated a strong relation.

Figure 5 presents the new deterministic material relationship corresponding to (15) together with the data in Table 4 (excluding Keller's data and Lotz' data pertaining to the transverse direction). Figure 6 shows the distribution of E^{sto} for three different values of ρ_{ash}.

The highest value of ρ_{ash} obtained from HU with the K_2HPO_4 phantoms in qCT scans is about $\rho_{\text{ash}} \sim 1.6$ g/cm^3 corresponding to $E^{\text{det}} \sim 23\,700$

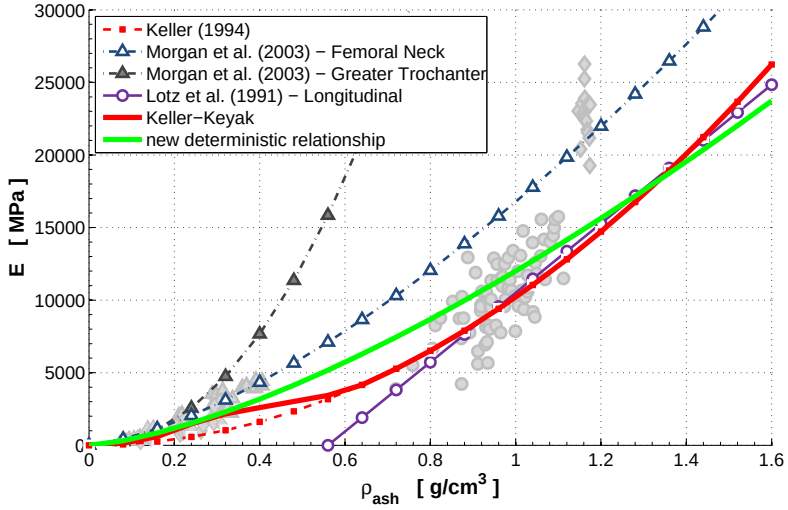


Figure 5: The new deterministic material relationship (green line) derived from the relevant experimental data (grey marks in the background) is compared to the relationships ascertained from the literature research (Morgan et al., 2003; Kaneko et al., 2003, 2004; Lotz et al., 1991) and the "Keller–Keyak" relationship (red line). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

MPa according to (15). It is, however, apparent that, for $\rho_{\text{ash}} = 1.6$, the 95% confidence interval of the stochastic material relationship is bounded by $E^{\text{low}} = 12\,840$ MPa and $E^{\text{up}} = 44\,290$ MPa, respectively[5].

[5]For this density value, 2.5% of all possible Young's moduli are smaller than E^{low} and 2.5% are higher than E^{up}, so $\text{P}(E^{\text{low}} \leq E^{\text{sto}} \leq E^{\text{up}}) = 95\%$. One may observe that, for $\rho_{\text{ash}} = 1.6$, the upper bound of the 95% confidence interval becomes unrealistically "stiff". Due to the considerable uncertainty in the cortical range, a value of $E^{\text{up}} = 44\,290$ MPa by far exceeds the maximum Young's modulus of about 30 000 MPa obtained by nanoindentation experiments at the osteon level (Franzoso and Zysset, 2009).

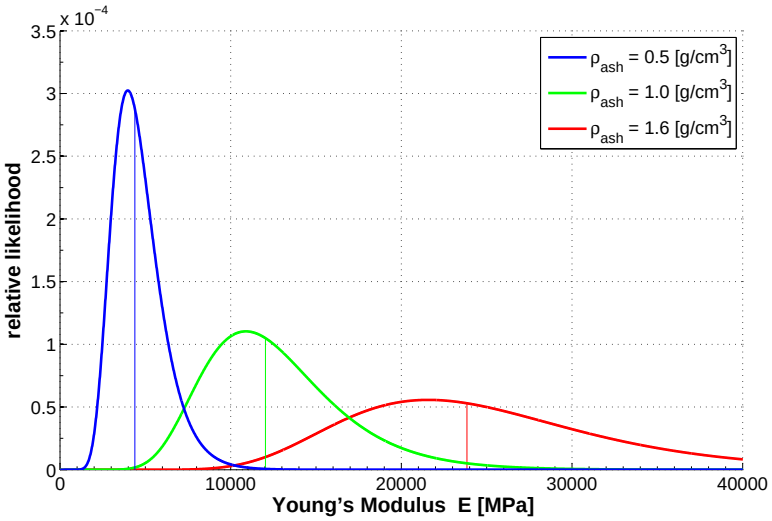


Figure 6: Schematic representation of the probability density function of E^{sto} for three different values of ρ_{ash}. In every distribution, the vertical line marks the median, which is equivalent to the value E^{det} obtained with the new deterministic material relationship.

We should bear in mind that the stochastic material relationship was derived under the assumption of homoscedastic and normally distributed residual values. To ensure that the relevant pooled data does not violate these assumptions, we carried out a studentized Breusch–Pagan test (Koenker, 1981) and a Lilliefors test (Thode, 2002, Sec. 5.1.1). None of the test statistics makes it possible to reject the respective hypothesis (Breusch-Pagan: $P = 0.1585 > 0.05$; Lilliefors: $P = 0.4578 > 0.05$).

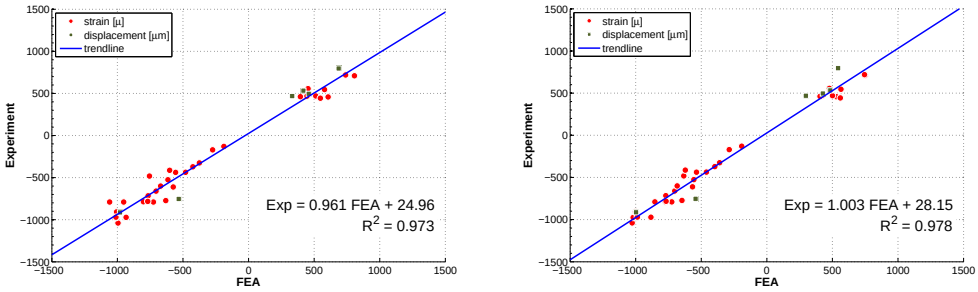
In addition, the assumption of homoscedasticity was checked within a residual plot. When plotting the residual values against the fitted values of the used data, no increase in variance was observed. Moreover, the stan-

dardized residual values were plotted against the theoretical quantiles of a standard normal distribution. None of the data points showed any significant deviations from the diagonal line $y = x$, which supports the assumption of normality.

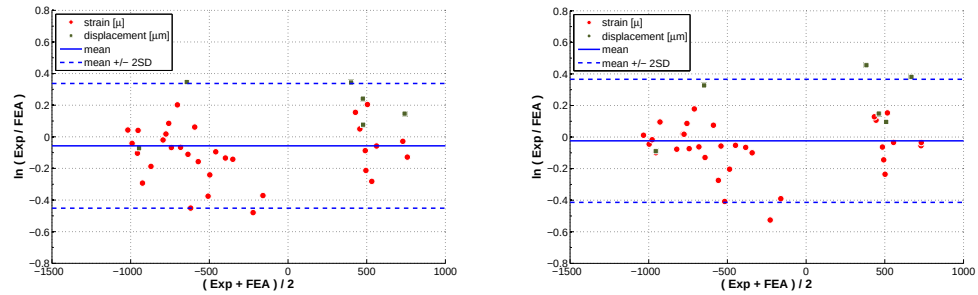
3.2. Experimental observations versus deterministic simulation results

In all experiments, a linear response was observed between force and displacements and strains beyond the 200N preload. The experimental error was within the $\pm 5\%$ range. p-FE simulations that mimic the in vitro experiments for FF4 and FF5, presented in (Yosibash et al., 2010; Katz, 2011) using the "Keller–Keyak" material relationship, were performed again with the new deterministic material relationship. By increasing the polynomial order from $p = 1$ to $p = 5$, we verified that the numerical errors were lower than 5% in the energy norm and that the strains and displacements at the points of interest had converged. Details on the process are given in the study of Trabelsi et al. (2009). Altogether, there are 39 data points for comparison, 11 strains and 2 displacements for FF4 at 0° inclination experiment, and 22 strains and 4 displacements for FF5 at 0° and 7° inclination experiments. The p-FE results versus the experimental observations are shown as a regression line in Figure 7a and in a modified Bland–Altman plot in Figure 7b.

The slope of the regression line in Figure 7a increased from 0.961 (for the hitherto best "Keller–Keyak" relationship) to 1.003 (for the new deterministic relationship) but the coefficient of determination remained nearly the same. For both material relationships we obtained a very strong linear relation between the experimental observation and the deterministic prediction



(a) regression lines



(b) Bland-Altman plots

Figure 7: Both strains and displacements measured in the FF4 at 0° and FF5 at a 0° and 7° inclination angle vs. p-FE predictions. Left: FEA with Keyak–Keller relationship. Right: FEA with new deterministic relationship.

($R^2 = 0.973$ and 0.978, respectively).

In the Bland–Altman plot we chose the "y-axis" to represent $\ln|\text{EXP}| - \ln|\text{FEA}| = \ln(\text{EXP}/\text{FEA})$, i.e., we compared the absolute errors of the logarithms. This was motivated by the log-normal distribution of the material uncertainty. If every p-FE result matched its corresponding experimental value, then the Bland–Altman plot would present all data points on the horizontal line $y = 0$. Deviations from this line represent an error showing that the physical reality is under- or over- estimated. For an unbiased

Table 2: Summary of characteristic values from Figure 7a and 7b.

	regression lines		Bland-Altman plots	
material relationship	slope	R^2	mean	standard deviation (SD)
Keller–Keyak	0.961	0.973	-0.057	0.197
new deterministic	1.003	0.978	-0.024	0.195

prediction, the mean of all ordinates in the Bland–Altman plot should be close to zero, while the spread of the mismatch can be quantified with the standard deviation of the ordinates. The Bland–Altman plots show that the mean of the logarithmic ratios is closer to zero for the new deterministic relations (-0.024) than for the Keller–Keyak relationship (-0.057), whereas the spread measured by the standard deviation remained nearly the same (0.195 and 0.197). Note that the mean and standard deviation shown in the Bland–Altman plots are based on a comparison between in vitro experiments and the deterministic prediction of the strains and displacements and should not be confused with the uncertainty in the stochastic material relationship.

All the characteristic values are summarized in Table 2 again for the sake of comparison.

3.3. Experimental observation versus stochastic response

The previous section compared all experimental observations to a characteristic value (median) of the stochastic response. Although the match is impressive *on average*, individual predictions do deviate from experimental observations. For the simulation results based on the new deterministic relationship we found relative errors up to $\sim 70\%$. This motivates further examinations of the predicted strains and displacements in comparison to

the experimental observations.

We should bear in mind that $\mathbf{u}^{\mathrm{sto}}$ is the product of the deterministic simulation result and a random variable Z describing the "input" uncertainty. Therefore, at each and every point within the bone the stochastic response is the result of the underlying distribution of Z scaled by the deterministic prediction. Treating all experimental observations as realizations of this (theoretical) stochastic response allows us to quantify the actual "output" uncertainty. Accordingly, every experimental observation is divided by its corresponding deterministic prediction and compared to the expected distribution of the stochastic response. Figure 8 depicts the normalized histogram of all $N = 39$ ratios together with the probability density function of Z.

The distribution of all ratios seems to follow the distribution of the random variable Z. Since Z is log-normally distributed, estimates for the mean and standard deviation of the actual "output" uncertainty are computed from the logarithmic ratios. Performing a Lilliefors test on the logarithmic ratios did not lead to a rejection of the hypothesis of log-normal distributed ratios on a high level of significance ($P = 0.0421 > 0.01$). All estimates have already been depicted in Table 2 as they are also used for characterizing the Bland-Altman plots. Please note that for the new material relationship the mean is very close to zero (-0.024), while the variance of all logarithmic ratios ($\mathrm{SD}^2 = 0.038$) is smaller than the variance of the material uncertainty ($s^2 = 0.1$).

This corroborates the hypothesis that most of the uncertainty is related to the material relationship, so Z may be used to quantify the uncertainty related to a deterministic prediction of the mechanical response. For instance,

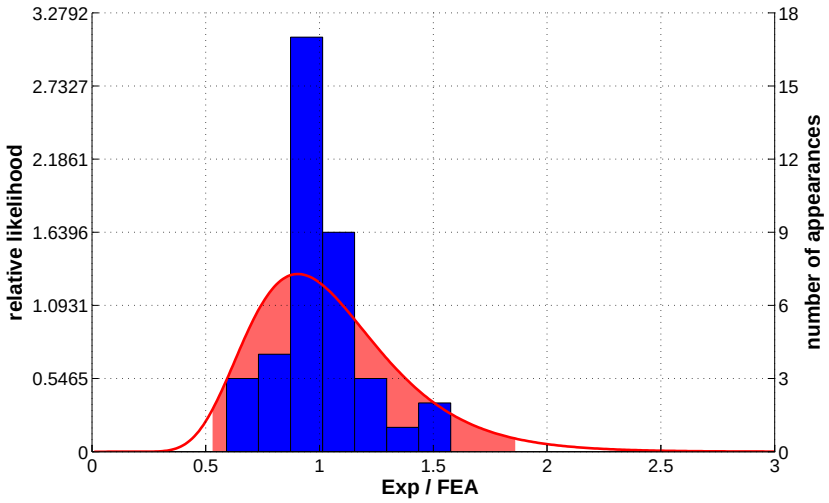


Figure 8: Histogram of all ratios between an experimental observation and its corresponding prediction (with the new deterministic material relationship) compared to the uncertainty related to the stochastic material law, which is the probability density function (PDF) of Z. The red-shaded area under the PDF marks its 95% confidence interval. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

as $Z \sim \ln \mathcal{N}(0, 0.1)$ in the present study, the 95% confidence interval of Z is bounded by $[0.5385; 1.8570]$, i.e., 95% of the experimental observations will be within 0.538 and 1.857 of the deterministic predictions. The anticipated spread in the results is quite high because the experimental value has a 5% probability of being less than 53.85% and greater than 185.70% of the predicted result.

4. Discussion

By investigating the various deterministic relations that predict the longitudinal Young's modulus as a function of a density measure in a femur, we conclude that the two popular relationships frequently used in FE-models, those proposed by Morgan et al. (2003) and by Keller (1994) are inappropriate when used with ρ_{ash}. In the former case, it is due to the fact that it is based on bone specimens with low bone densities alone and predict unrealistically high Young's modulus for high bone densities. And in the latter case it is because it includes specimens from both the longitudinal and the transversal directions (with no distinction between them), which may explain the very large scatter.

We were not able to trace any experimental data on the E-ρ relationship beyond $\rho_{\text{ash}} = 1.22$ g/cm^3 (see Table 4 and Figure 3), although this range of densities accounts for a significant part of femur's volume. This is demonstrated in Table 3 where the percentage of bone volume with densities beyond $\rho_{\text{ash}} = 1.22$ g/cm^3 is summarized for seventeen femurs—see also Trabelsi et al. (2011).

The discretization error is guaranteed to be small (under 5% error in the energy norm and in averaged strains and displacements at locations of interest) compared to the uncertainties involved in our analysis. Thus, for reliable simulations using FE methods, new well thought out experiments on femur specimens where $\rho_{\text{ash}} > 1.2$ g/cm^3 are required because such densities determine a large portion of the femur's volume. As the various specimen tests were performed in the past by different researchers , moreover, each using a different experimental protocol, different specimen sizes, different

methods for determining Young's modulus, and different density measures, the spread in the pooled data is fairly large. *This calls for new research efforts based on a well-defined protocol that will establish the E to Hounsfield-Unit relationship in the longitudinal and transverse directions of the femur tissue and cover the entire density range.*

Based on the pooled data consisting of $N = 161$ samples, we determined a stochastic E-ρ_{ash} relation (in the longitudinal direction). Compared with individual studies, pooling the data increased the coefficient of determination for the regression function significantly ($R^2 = 0.94$ compared to 0.67 in (Keller, 1994), 0.67 in (Lotz et al., 1991), and 0.85 in (Morgan et al., 2003) - see Table 4).

The median of the stochastic relationship, the new deterministic relationship, was shown to be better than the best material relationship known so far in respect of the prediction of p-FE strains and displacements when a simple stance loading condition is applied to the femur and results are compared to in vitro experiments. Nevertheless, the improvement provided by the new deterministic relationship is only marginal compared to the "Keller–Keyak" relationship. The slope of the experimental results versus p-FE predictions is 1.003 with $R^2 = 0.978$, and the mean error in the Bland–Altman plot is very close to zero, demonstrating an unbiased error. This excellent match is not surprising because, in a simple stance position, it is reasonable to consider the isotropic relationship with E as being the one in the "longitudinal direction", see (Trabelsi and Yosibash, 2011).

Most recent FEA studies, see (Yosibash et al., 2007a; Bessho et al., 2007; Helgason et al., 2008b; Schileo et al., 2008) for example, used one of the

E-ρ relationships given in Table 4 and record an accurate estimation of experimental results. It is interesting to note that the overall quality of the predictions, as demonstrated by correlation analysis, has always been good and largely independent of the underlying FE models. At this point we would like to stress that the usual quality measures - slope and coefficient of determination - are insufficient criteria for judging the reliability and robustness of a method since they only characterize the average behavior and since individual predictions can still deviate considerably from experimental observations even though slope and coefficient of determination indicate an excellent match. In the present study based on $N = 39$ experimental results, the maximum absolute relative error in predicting strains and displacements was found to be $\sim 70\%$. This remaining simulation uncertainty correlates well with the uncertainty related to the material relationship, which was shown to propagate directly to the mechanical response under the assumption of linear, isotropic elasticity.

Even though only one source of uncertainty was taken into account, the probabilistic approach presented in this paper indicates that most of the observed uncertainty is explained by the material relationship. Other sources of uncertainty should be included in future studies, but these are expected to have a smaller influence in the case of well-controlled in vitro experiments. For example, a femur's FE domain is determined with relative errors of less than 3–4% when compared with measurements of the actual geometry (Trabelsi et al., 2009). The same holds for modeling the load that is applied during an in vitro experiment. Of course, if more information is available on the gender or age of the donor associated with each of the tested specimens,

for instance, then a multivariate regression can be performed. This would make it possible to perform simulations with stochastic material relationships which are dependent on a patient's gender or age, for instance.

This study has a number of limitations. One important limitation is the use of the connections (2) and (3) to obtain E associated with one density measure for all experimental data. This limitation may also result in the excessively stiff bone response in Morgan's relationship for $\rho_{\text{ash}} > 0.4$ g/cm^3. Another limitation is that the comparison between FE predictions and in vitro experiments considers two fresh-frozen femurs only. This comparison should be extended to cover a much larger experimental cohort. In addition, the femurs were assumed to be inhomogeneous and isotropic, with E being dependent on the ash density and a single random variable. Due to the lack of experimental information on $\nu(\rho)$, it was kept constant at the value of 0.3 (in (Yosibash et al., 2007a) we demonstrated that the Poisson ratio has a minor influence on the longitudinal strains measured in the experiment). In future studies, based on an anisotropic material law, or based on ν as an additional stochastic variable, more intensive numerical computation will be required, employing Monte-Carlo techniques or a variant of the polynomial-chaos approach as in Foo et al. (2007).

Finally, we strongly advocate that future FE studies on the mechanical response of bones describe not only a regression line representing the average behavior of all FE predictions, but also quantify the uncertainty associated with the FE simulation by using either Bland–Altman plots or plots showing the distribution of all ratios between an experimental observation and its corresponding deterministic prediction.

Table 3: Summary of percentage of volume in femurs with densities in excess of 1.22 g/cm^3 and maximum estimated ρ_{ash} from qCT scans.

Donor ID	Age (years)	Height (cm)	Weight (Kg)	Gender	% Vol with $\rho_{\text{ash}} > 1.22$ g/cm^3 Left/Right	Max(ρ_{ash}) Left/Right
FF1	30	n.r.	n.r.	male	13 / -	1.53/ -
FF2	20	n.r.	n.r.	female	- / 17	- /1.54
FF3	54	n.r.	n.r.	female	20 / -	1.6/ -
FF4	63	n.r.	n.r.	male	- / 32	- /1.59
FF5	56	n.r.	n.r.	male	- / 13	- /1.6
Donor 1	54	178	161	male	32 / 31	1.58/1.57
Donor 2	58	185	86	male	28 / 26	1.6/1.6
Donor 3	64	168	136	female	25 / 22	1.6/1.6
Donor 4	48	170	55	male	30 / 21	1.6/1.6
Donor 5	53	193	98	male	19 / 16	1.6/1.58
Donor 6	59	180	96	female	1 / 0	1.45/1.39

Acknowledgements We would like to acknowledge the generous support of the Technische Universität München–Institute for Advanced Study, funded by the German Excellence Initiative. Special thanks are extended to Mr. Nir Trabelsi and Mr. Alon Katz, two graduate students supervised by the third author, for their help with FE simulations. We would also like to thank three colleagues from the Technische Universität München, Mr. Iason Papaioannou and Dr. Stephan Haug, for their support on the statistical part, and Dr. Stefan Kollmannsberger for the helpful discussions.

Conflict of interest None of the authors have any conflict of interest to declare that could prejudice this work.

References

Abraham, B. and Ledolter, J. (2006). *Introduction to Regression Modeling*. Thomson Brooks/Cole.

Bessho, M., Ohnishi, I., Matsuyama, J., Matsumoto, T., Imai, K., and Nakamura, K. (2007). Prediction of strength and strain of the proximal femur by a CT-based finite element method. *J. Biomech.*, 40(8):1745–1753.

Dopico-Gonzalez, C., New, A. M., and Browne, M. (2009). Probabilistic analysis of an uncemented total hip replacement. *Med. Eng. Phys.*, 31(4):470 – 476. Finite Element Modelling of Medical Devices.

Foo, J., Yosibash, Z., and Karniadakis, G. E. (2007). Stochastic simulation of riser-sections with uncertain measured pressure loads and/or uncertain material properties. *Computer Methods in Applied Mechanics and Engineering*, 196(41-44):4250 – 4271.

Franzoso, G. and Zysset, P. (2009). Elastic anisotropy of human cortical bone secondary osteons measured by nanoindentation. *ASME J. Biomech. Eng.*, pages 021001/1 – 021001/11.

Helgason, B., Perilli, E., Schileo, E., Taddei, F., Brynjolfsson, S., and Viceconti, M. (2008a). Mathematical relationships between bone density and mechanical properties: a literature review. *Clinical Biomechanics*, 23:135 – 146.

Helgason, B., Taddei, F., Palsson, F., Schileo, E., Cristofolini, L., Viceconti, M., and Brynjolfsson, S. (2008b). A modified method for assigning material properties to FE models of bones. *Med. Eng. Phys.*, 30:444–453.

Kaneko, T. S., Bella, J. S., Pejcica, M. R., Tehranzadeha, J., and Keyak, J. H. (2004). Mechanical properties, density and quantitative CT scan data of trabecular bone with and without metastases. *J. Biomech.*, 37:523–530.

Kaneko, T. S., Pejcic, M. R., Tehranzadeha, J., and Keyak, J. H. (2003). Relationships between material properties and CT scan data of cortical bone with and without metastatic lesions. *Med. Eng. Phys.*, 25:445–454.

Katz, A. (2011). The Mechanical Response of Femurs Fixed by Metal Devices. MSc thesis, Dept. of Mechanical Engineering, Ben-Gurion University of the Negev, Beer-Sheva, Israel.

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

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

Keyak, J. H., Lee, I. Y., and Skinner, H. B. (1994). Correlation between orthogonal mechanical properties and density of trabecular bone: use of different densiometric measures. *J. Biomedical Mater. Res.*, 28:1329–1336.

Koenker, R. (1981). A note on studentizing a test for heteroscedasticity. *Journal of Econometrics*, 17(1):107 – 112.

Laz, P. J. and Browne, M. (2010). A review of probabilistic analysis in orthopaedic biomechanics. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 224(8):927–943.

Laz, P. J., Stowe, J. Q., Baldwin, M. A., Petrella, A. J., and Rullkoetter, P. J. (2007). Incorporating uncertainty in mechanical properties for finite element-based evaluation of bone mechanics. *J. Biomech.*, 40(13):2831 – 2836.

Lotz, J. C., Gerhart, T. N., and Hayes, W. C. (1991). Mechanical properties of metaphyseal bone in the proximal femur. *J. Biomech.*, 24:317–329.

Morgan, E. F., Bayraktar, H. H., and Keaveny, T. M. (2003). Trabecular bone modulus-density relationships depend on anatomic site. *J. Biomech.*, 36:897–904.

Nicolella, D. P., Thacker, B. H., Katoozian, H., and Davy, D. T. (2006). The effect of three-dimensional shape optimization on the probabilistic

response of a cemented femoral hip prosthesis. *J. Biomech.*, 39(7):1265 – 1278.

Pise, U., Bhatt, A., Srivastava, R., and Warkedkar, R. (2009). A B-spline based heterogeneous modeling and analysis of proximal femur with graded element. *J. Biomech.*, 42:1981 – 1988.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Rice, J. C., Cowin, S. C., and Bowman, J. A. (1988). On the dependence of the elasticity and strength of cancellous bone on apparent density. *J. Biomech.*, 21:155–168.

Schileo, E., DallAra, E., Taddei, F., Malandrino, A., Schotkamp, T., Baleani, M., and Viceconti, M. (2008). An accurate estimation of bone density improves the accuracy of subject-specific finite element models. *J. Biomech.*, 41:2483–2491.

Taddei, F., Martelli, S., Reggiani, B., Cristofolini, L., and Viceconti, M. (2006). Finite-element modeling of bones from CT data: Sensitivity to geometry and material uncertainties. *IEEE Trans. Biomed. Eng.*, 53:2194–2200.

Taddei, F., Schileo, E., Helgason, B., Cristofolini, L., and Viceconti, M. (2007). The material mapping strategy influences the accuracy of CT-based finite element models of bones: An evaluation against experimental measurements. *Med. Eng. Phys.*, 29(9):973–979.

Thode, H. C. (2002). *Testing for Normality*. Marcel Dekker, New York.

Trabelsi, N. and Yosibash, Z. (2011). Patient-Specific Finite-Element Analyses of the Proximal Femur with Orthotropic Material Properties Validated by Experiments. *ASME J. Biomech. Eng.*, 133(6):061001-1–11.

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

Trabelsi, N., Yosibash, Z., Wutte, C., Augat, R., and Eberle, S. (2011). Patient-specific finite element analysis of the human femur—A double-blinded biomechanical validation. *J. Biomech.*, 44(9):1666–1672.

Viceconti, M., Brusi, G., Pancanti, A., and Cristofolini, L. (2006). Primary stability of an anatomical cementless hip stem: A statistical analysis. *J. Biomech.*, 39(7):1169–1179.

Wirtz, D., Schiffers, N., Pandorf, T., Radermacher, K., Weichert, D., and Forst, R. (2000). Critical evaluation of known bone material properties to realize anisotropic FE-simulation of the proximal femur. *J. Biomech.*, 33:1325–1330.

Yosibash, Z., Padan, R., Joscowicz, L., and Milgrom, C. (2007a). A CT-based high-order finite element analysis of the human proximal femur compared to in-vitro experiments. *ASME J. Biomech. Eng.*, 129(3):297–309.

Yosibash, Z., Tal, D., and Trabelsi, N. (2010). Predicting the yield of the proximal femur using high order finite element analysis with inhomoge-

neous orthotropic material properties. *Philosophycal Transaction of the Royal Society: A*, 368:2707–2723.

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

Zienkiewicz, O. C., Taylor, R. L., and Zhu, J. Z. (2005). *Finite Element Method–Its Basis and Fundamentals*. Elsevier, 6th edition.

Appendix A. Determination of the E-ρ relation using experimental data and the principle of maximum likelihood

Given experimental data as a set of N measurements (E_i, ρ_i) with $i = 1, 2, \ldots, N$ and a simple linear regression model of type

$$E = f(\rho, \boldsymbol{\alpha}) + X = \alpha_0 + \alpha_1 \rho + X \tag{A.1}$$

with $X \sim \mathcal{N}(0, \sigma^2)$, the principle of maximum likelihood can be used to obtain estimates for the regression coefficients $\boldsymbol{\alpha}$ and the variance of the error term σ^2.

Using (A.1), the probability of observing a single experiment (E_i, ρ_i) can be computed directly from the normal distribution for given $\boldsymbol{\alpha}$ and σ^2:

$$\Pr(E = E_i \,|\, \boldsymbol{\alpha}, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{E_i - f(\rho_i, \boldsymbol{\alpha})}{\sigma}\right)^2\right] \tag{A.2}$$

Assuming all experimental observations to be statistically independent of one another, the total likelihood $\mathrm{L}(\boldsymbol{\alpha}, \sigma)$ of observing all N measurements is equal to the product of all individual probabilities:

$$\mathrm{L}(\boldsymbol{\alpha}, \sigma) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{E_i - f(\rho_i, \boldsymbol{\alpha})}{\sigma}\right)^2\right] \tag{A.3}$$

$$= \frac{1}{\sqrt[N]{2\pi\sigma^{2N}}} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{N} \left(E_i - f(\rho_i, \boldsymbol{\alpha})\right)^2\right] \tag{A.4}$$

The maximum likelihood estimates are those values for the parameters $\boldsymbol{\alpha}$ and σ that maximize the likelihood function. The $\boldsymbol{\alpha}$ which maximizes L is the same as the least squares estimates $\hat{\boldsymbol{\alpha}}$, since maximizing the above likelihood for $\boldsymbol{\alpha}$ is equivalent to minimizing the sum of squares in the exponent of (A.4).

In the case of a linear regression function $f(\rho, \boldsymbol{\alpha}) = \alpha_0 + \alpha_1 \rho$ the estimates for α_0 and α_1 accordingly are

$$\hat{\alpha}_0 = \bar{E} - \hat{\alpha}_1 \bar{\rho} \tag{A.5}$$

$$\hat{\alpha}_1 = \frac{\sum_{i=1}^{N} (\rho_i - \bar{\rho}) \left(E_i - \bar{E}\right)}{\sum_{i=1}^{N} (\rho_i - \bar{\rho})^2} \tag{A.6}$$

with $\bar{E}$ and $\bar{\rho}$ being the sample mean. Those estimates are unbiased, i.e., $\mathbb{E}[\hat{\alpha}_0] = \alpha_0$ and $\mathbb{E}[\hat{\alpha}_1] = \alpha_1$.

Substituting the maximum likelihood estimate $\hat{\boldsymbol{\alpha}}$ into (A.4) and taking the logarithm of the result makes it possible to find an estimate for the variance σ^2 by setting the partial derivative of the logarithmic likelihood function equal to zero, i.e.:

$$\frac{\partial \ln \mathrm{L}}{\partial \sigma^2} = 0 = \frac{-N}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{N} \left(E_i - \hat{E}_i\right)^2 \tag{A.7}$$

with $\hat{E}_i = f(\rho_i, \hat{\boldsymbol{\alpha}})$ representing the fitted Young's modulus for a certain density value of the data. Solving for σ^2 yields the maximum likelihood estimate s^2_{mle} for the variance of the error

$$s^2_{mle} = \frac{1}{N} \sum_{i=1}^{N} \left(E_i - \hat{E}_i\right)^2 \tag{A.8}$$

Since this estimate is known to be biased, we use the unbiased estimate s^2 instead

$$s^2 = \frac{1}{N-2} \sum_{i=1}^{N} \left(E_i - \hat{E}_i\right)^2 \tag{A.9}$$

In the end, all three maximum likelihood estimates can be used to infer from the experimental data a linear model:

$$E \approx \hat{\alpha}_0 + \hat{\alpha}_1 \rho + Y \tag{A.10}$$

with $Y \sim \mathcal{N}(0, s^2)$.

Table 4: Summary of the test details for E-ρ relationship on human femurs.

Ref	bone	type	anatomical site	info. on donors			# specimens	orientation	density	reported $E[MPa]$	R^2	ρ_{ash}-range	Test	Geom. LxWxH	strain rate
				#	gender	age	(recovered)			(recovered $E[MPa]$)	(recov.)	[g/cm^3]	type	or DxL [mm]	[1/sec.]
Lotz et al.	left	cortical	metaphyseal cortical	5	n.r.	28 - 90	79	longitudinal	ρ_{dry}	$14261\rho_{\text{dry}} - 13430$	0.67	0.72 - 1.11	3PB	7 x 5 x 0.18-0.4	0.05
(Lotz et al., 1991)	femur		shell (neck and				(76)			$(14131\rho_{\text{dry}} - 12521)$	(0.50)				
			intertrochanteric)												
			and diaphysis				80	transverse		$4979\rho_{\text{dry}} - 3122$	-	0.81 - 1.11			
							(74)			$(5190\rho_{\text{dry}} - 3040)$	(0.12)				
Keller	femur	cortical &	lesser trochanter & part of	2	$2 \times M$	46, 67	259	n.r.	ρ_{ash}	$10200\rho_{\text{ash}}^{2.01}$	0.67	0.25 - 1.22	Platen	8 x 8 x 8	0.01
(Keller, 1994)		trabecular	lateral supra-condylar ridge				(184)			$(9882\rho_{\text{ash}}^{1.98})$	(0.62)				
Morgan et al.	femur	trabecular	neck	23	$15 \times M$	57 - 101	27	on-axis	ρ_{wet}	$6850\rho_{\text{wet}}^{1.49}$	0.85	0.14 - 0.41	End-caps	8 x 25	0.005
(Morgan et al., 2003)					$8 \times F$		(27)			$(6843\rho_{\text{wet}}^{1.49})$	(0.85)				
			greater trochanter	21	$16 \times M$	49 - 101	23	on-axis		$15010\rho_{\text{wet}}^{2.18}$	0.82	0.07 - 0.15			
					$5 \times F$		(23)			$(15199\rho_{\text{wet}}^{2.19})$	(0.82)				
Kaneko et al.	femur	cortical	diaphysis	2	$1M$	n.r.	16	longitudinal of	ρ_{ash}	-	-	1.14 - 1.17	Dumbbell	6 x 2 x 5	0.001
(Kaneko et al., 2003)					1 n.r.		from author	diaphyseal axis		-	-				
Kaneko et al.	femur	trabecular	distal	4	$2 \times M$	78	19	superior- inferior	ρ_{ash}	-	-	0.11 - 0.31	Platen	15 x 15 x 15	0.01
(Kaneko et al., 2004)					$2 \times F$	67, 88	from author			-	-				
Present study		trab & cort					161		ρ_{ash}	$(12000\rho_{\text{ash}}^{1.45})$	0.94	0.07 - 1.22			