Clinical Biomechanics 29 (2014) 373–380

Contents lists available at ScienceDirect

Clinical Biomechanics

journal homepage: www.elsevier.com/locate/clinbiomech

Review

How accurately can we predict the fracture load of the proximal femur using finite element models?

Sven van den Munckhof, Amir Abbas Zadpoor *

Department of Biomechanical Engineering, Faculty of Mechanical, Maritime, and Materials Engineering, Delft University of Technology (TU Delft), Mekelweg 2, Delft 2628 CD, The Netherlands

ARTICLE INFO

Article history:
Received 10 July 2013
Accepted 31 December 2013

Keywords:
Femoral fracture
Osteoporosis
Finite element method
Patient-specific model
Accuracy assessment
Review

ABSTRACT

Background: Current clinical methods for fracture prediction rely on two-dimensional imaging methods such as dual-energy X-ray absorptiometry and have limited predictive value. Several researchers have tried to integrate three-dimensional imaging techniques with the finite element (FE) method to improve the accuracy of fracture predictions. Before FE models could be used in clinical settings, a thorough validation of their accuracy is required. In this paper, we try to evaluate the current state of accuracy of subject-specific FE models that are used for prediction of the fracture load of proximal femora.

Methods: All the studies that have used FE for prediction of fracture load and have compared the predicted fracture load with experimentally measured fracture loads in vitro are identified through a systematic search of the literature. A quantitative analysis of the results of those studies has been carried out to determine the absolute prediction error, percentage error, and linear correlations between predicted and measured fracture loads.

Findings: The reported coefficients of determination (R^2) vary between 0.773 and 0.96 while the percentage error in prediction of fracture load varies between 5 and 46% with most studies reporting percentage errors between 10 and 20%.

Interpretation: We conclude that FE models, which are currently used only experimentally, are in general more accurate than clinically used fracture risk assessment techniques. However, the accuracy of FE models depends on the details of their modeling methodologies. Therefore, modeling procedures need to be optimized and standardized before FE could be used in clinical settings.

1. Introduction

In Europe, around 22.1% of the female population and 6.6% of the male population aged 50 and more are diagnosed with osteoporosis (Hernlund et al., 2013). The prevalence of the disease increases with age. For example, in Sweden, 47.2% of the female population and 16.6% of the male population aged between 80 and 84 years are diagnosed with osteoporosis (Kanis et al., 2000). Osteoporosis often results in bone fracture. Hip fractures are among the most detrimental for both patient and society. Patients with hip fracture suffer from pain and loss of mobility. Almost all patients must be hospitalized and undergo surgical intervention (Woolf and Pfleger, 2003). A systematic review of the literature found that persons who have experienced hip fractures exhibit an excess mortality rate of 8.4–36% during the first year following the fracture (Abrahamsen et al., 2009). Moreover, a recent study in Ireland showed that only 55% of individuals who were independent before fracture maintained their independence 120 days after fracture (Brewer et al., 2011). In 2000, the total number of osteoporotic fractures in Europe was estimated at 3.7 million of which 24% (890,000) were hip fractures (Kanis and Johnell, 2005). The total healthcare cost in Europe as a consequence of osteoporotic fracture was estimated at €36.2 billion from which €24 billion can be attributed to hip fractures (Kanis and Johnell, 2005). European health care costs for osteoporotic fractures are estimated to increase to €76 billion in 2050 (Kanis and Johnell, 2005).

An accurate estimation of fracture risk is required for proper treatment of osteoporotic patients. In clinical practice, osteoporosis is often diagnosed by dual-energy X-ray absorptiometry (DXA). It has been shown that decreased Bone Mineral Density (BMD) determined by DXA, independent of ethnic background, sex, or age is related to an increased risk of hip fracture (Barrett-Connor et al., 2005; Hillier et al., 2007; Johnell et al., 2005; Trémollieres et al., 2010). A decrease of one unit in the T-score scale roughly doubles the risk of hip fracture (Hillier et al., 2011; Johnell et al., 2005). However, assessment of absolute fracture risk by BMD is limited due to poor specificity of BMD in predicting the actual fracture event (Hillier et al., 2011; Trémollieres et al., 2010). Indeed, many patients with normal BMD values develop osteoporotic fractures (Wainwright et al., 2005). Studies on the relationship between densitometric measures and in vitro femoral fracture load show that DXA is a limited predictor for femoral fracture load (Cody et al., 1999; Lochmüller et al., 2000). Lochmüller et al. report a correlation of $R^2 = 0.449$ ($P < 0.01$) between Bone Mineral Content (BMC) and mechanical fracture load (Lochmüller et al., 2000).

* Corresponding author.
E-mail address: a.a.zadpoor@tudelft.nl (A.A. Zadpoor).

0268-0033/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.clinbiomech.2013.12.018

Cody et al. report a cross validation correlation of $R^2 = 0.57$ between fracture load and BMD value (Cody et al., 1999).

In 2008, Kanis et al. developed, on behalf of the World Health Organization WHO, a clinical fracture risk assessment tool called FRAX (Kanis et al., 2008). Hillier et al. conducted a study on the predictive value of FRAX on a population of women aged >65 (Hillier et al., 2011). They present AUC (area under the curve) as a measure of the probability of a person who has sustained a femoral fracture to be in the group of people who were predicted to experience a fracture. The AUC value of FRAX combined with BMD was found to be 0.62 for subjects classified as osteoporotic (Hillier et al., 2011). A study on women aged 54 ± 4 (mean ± SD) found AUC values between 0.56 and 0.69 ($P < 0.05$) (Trémollieres et al., 2010).

Due to the limitations of DXA, application of other measurement techniques such as quantitative ultrasound (QUS) (Hartl et al., 2002; Krieg et al., 2008), quantitative computed tomography (qCT) (Lotz and Hayes, 1990), and magnetic resonance imaging (MRI) (Majumdar and Genant, 1995) for predicting the chance of osteoporotic fractures is being studied. Every one of the above-mentioned techniques has certain limitations. However, the main limitation of the above-mentioned techniques is that they, as imaging techniques, are not capable of combining the structural properties, anatomical shape, and applied forces in a mechanically consistent way. The fracture risk determined using these methods is therefore bound to be limited in accuracy.

Finite element (FE) modeling goes beyond imaging by incorporating mechanics, geometry, and shape into one single model. FE is shown to be a promising method for predicting the risk of osteoporotic fractures (Cody et al., 2000a,b; Cong et al., 2011; Grassi et al., 2011; Keyak and Rossi, 2000; Lotz et al., 1991; Luo et al., 2011; MacNeil et al., 2012; Ota et al., 1999; Schileo et al., 2007, 2008b; Testi et al., 1999; Yosibash et al., 2007, 2010; Zdero et al., 2010). Patient-specific finite element models are created based on the specific conditions of individual patients. They are therefore considered to be potentially the most accurate models for estimating the risk of osteoporotic fracture. However, there are many methodological choices, modeling assumptions, and parameter values that in practice determine the accuracy of FE models in predicting the risk of osteoporotic bone fracture. The methodological aspects of FE models are not reviewed in this paper, as several excellent review papers (Bouxsein, 2008; Christen et al., 2010; Engelke, 2012; Geusens et al., 2010; Keaveny, 2010) have been recently published on the methodological aspects of FE modeling and other fracture assessment techniques.

Probably the most important question regarding FE models is: 'how accurate are patient-specific FE models in predicting the risk of femoral fractures?' This review tries to answer this question through a systematic review of the literature. The studies that use finite element modeling for predicting the fracture load of proximal femora and compare the predicted values with the fracture loads measured through in vitro mechanical testing are found through a systematic search of the literature. A quantitative analysis of the accuracy of finite element models is carried out to clarify what kind of accuracies is achievable with the currently available technology for patient-specific FE modeling.

2. Methods

2.1. Literature search method

A literature search was carried out on PubMed in October 2012 (latest search: 25 October). The following keywords were used for the literature search: "finite element" AND ("hip" OR "femur" OR "femoral") AND "fracture".

The literature search resulted in 327 results that were manually scanned to determine whether they satisfied the inclusion criteria. ISI Web of Knowledge and Google Scholar were used as additional databases to ensure that all studies are included. The reference lists of most relevant studies were scanned and crosschecked to identify all the studies that satisfy the inclusion criteria.

The following inclusion criteria were used to select the papers that were included in the study:

1. Both FE-predicted and experimentally measured femoral fracture loads are presented in the paper.
2. The experimental fracture loads are determined using in vitro experiments.
3. The size of the study is large enough ($n > 3$).
4. The FE model contains at least the most proximal intact part of the human femur.
5. The modeling strategy has the potential to be clinically feasible, excluding the studies that rely on μ-CT images.
6. The FE models are three-dimensional and are based on CT images of the same bones that are experimentally tested.

Twelve studies (Table 1) satisfied the above-mentioned inclusion criteria.

2.2. Quantitative analysis of fracture data

The fracture loads presented in the included studies were quantitatively analyzed to determine certain error quantifiers that are not reported in all studies including average and standard deviation of experimental and predicted fracture loads, average and standard deviation of absolute prediction error, and the percentage error.

The data needed for calculating the above-mentioned error quantifiers were extracted from the studies. Wherever the data was not presented in tabular form, the figures of the paper were digitized as high quality images (600 dpi) and the reported values were back calculated from the figures. In order to check the accuracy of the extracted data, the coefficient of determination (R^2) was calculated using the extracted data and was compared with the reported value (Table 2). The reported and calculated values were in all cases either very close or identical (Table 2).

Using the extracted data, several additional parameters were calculated including the average and standard deviation of experimental and predicted fracture loads. The average and standard deviation of the absolute error value was calculated by calculating the absolute values of the difference between the predicted and measured fracture loads. The percentage error, E_{rel}, was calculated as:

$$E_{rel} = \frac{1}{n}\sum_{i=1}^{n}\frac{F_{exp} - F_{FE}}{F_{exp}} \times 100 \quad (1)$$

where F_{exp}, F_{FE}, and n are measured fracture load, predicted fracture load, and the number of bones modeled in the study, respectively.

Other relevant parameters that were reported in the studies including the parameters of the linear equations that were fitted to the data were also collected. In the studies where both stance and sideways fall loading conditions were reported, the results of both loading conditions were analyzed separately.

The studies were scanned to see which ones report on the predictive capability of other techniques and compare them with the predictive capability of FE models. Three of the included studies (Cody et al., 1999; Dall'Ara et al., 2013; Dragomir-Daescu et al., 2011) reported also on the predictive capability of other techniques such as densitometry-based techniques.

3. Results

Among the twelve included studies, nine studies considered the stance loading and five studies considered the sideways fall loading condition with two studies considering both loading conditions. All studies combined, the coefficient of determination varies between 0.73 and 0.96 (Table 2). The reported coefficient of determination (R^2)