A probabilistic modelling scheme for analysis of long-term failure of cemented femoral joint replacements


Published in:

Document Version:
Publisher's PDF, also known as Version of record

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Download date: 20. Nov. 2018
A probabilistic modelling scheme for analysis of long-term failure of cemented femoral joint replacements

Pavel E Galibarov, Patrick J Prendergast and Alexander B Lennon


DOI: 10.1177/0954411912450959

The online version of this article can be found at:
http://pih.sagepub.com/content/early/2012/07/02/0954411912450959
A probabilistic modelling scheme for analysis of long-term failure of cemented femoral joint replacements

Pavel E Galibarov1, Patrick J Prendergast2 and Alexander B Lennon3

Abstract
Reliable prediction of long-term medical device performance using computer simulation requires consideration of variability in surgical procedure, as well as patient-specific factors. However, even deterministic simulation of long-term failure processes for such devices is time and resource consuming so that including variability can lead to excessive time to achieve useful predictions. This study investigates the use of an accelerated probabilistic framework for predicting the likely performance envelope of a device and applies it to femoral prosthesis loosening in cemented hip arthroplasty.

A creep and fatigue damage failure model for bone cement, in conjunction with an interfacial fatigue model for the implant–cement interface, was used to simulate loosening of a prosthesis within a cement mantle. A deterministic set of trial simulations was used to account for variability of a set of surgical and patient factors, and a response surface method was used to perform and accelerate a Monte Carlo simulation to achieve an estimate of the likely range of prosthesis loosening. The proposed framework was used to conceptually investigate the influence of prosthesis selection and surgical placement on prosthesis migration.

Results demonstrate that the response surface method is capable of dramatically reducing the time to achieve convergence in mean and variance of predicted response variables. A critical requirement for realistic predictions is the size and quality of the initial training dataset used to generate the response surface and further work is required to determine the recommendations for a minimum number of initial trials. Results of this conceptual application predicted that loosening was sensitive to the implant size and femoral width. Furthermore, different rankings of implant performance were predicted when only individual simulations (e.g. an average condition) were used to rank implants, compared with when stochastic simulations were used. In conclusion, the proposed framework provides a viable approach to predicting realistic ranges of loosening behaviour for orthopaedic implants in reduced timeframes compared with conventional Monte Carlo simulations.

Keywords
Hip replacement, probabilistic modelling, cemented hip replacement, stochastic modelling, stochastic versus deterministic, hip replacement failure

Introduction
Computer simulation of medical device behaviour is well established in pre-clinical development of new devices and is becoming increasingly applied to surgical decision making, e.g. pre-operative planning. However, achieving realistic predictions requires at least some incorporation of the variability that devices are subjected to during their service life. This can be particularly problematic for predicting the performance envelope of devices that fail over long time periods of several years, e.g. orthopaedic joint replacements, as the failure simulation must often be performed over numerous time points to capture the failure processes realistically. Stochastic approaches, such as Monte Carlo simulation,1 thus quickly become impractical for

1AnyBody Technology A/S, Denmark
2Trinity Centre for Bioengineering, Trinity College Dublin, Ireland
3School of Mechanical and Aerospace Engineering, Queen’s University Belfast, UK

Corresponding author:
Alexander B Lennon, School of Mechanical and Aerospace Engineering, Queen’s University Belfast, Belfast BT9 5AH, UK.
Email: a.lennon@qub.ac.uk
non-linear or time-adaptive analysis, and realistic prediction of the full range of device performance must frequently be sacrificed.

Several stochastic techniques, widely used by engineers in different fields, can be employed to reduce the computational burden of simulating the likely range of performance of devices and structures. First, it has been shown that sampling strategies, such as Latin Hypercube, can reduce the number of trials required to achieve a convergent prediction in Monte Carlo simulations. Further reductions can be achieved using either response surface or advanced mean value techniques. Another alternative is the probabilistic finite element method (PFEM), in combination with a suitable probability distribution function for the parameters of interest.

Several statistical and probabilistic methods have been applied to the assessment of variable response of joint replacements in a number of studies. Monte Carlo simulations were used by Viceconti et al., who investigated the influence of host bone size (using rigid scaling of the bone), tissue properties, patient body weight, and implant–bone interfacial condition on the primary stability of uncemented hip replacements. Dopico-Gonzalez et al. used Monte Carlo simulations with different sampling strategies to investigate: variability in mechanical moduli of the stem and bone, load magnitude and direction, anteversion of the prosthesis, and variable positions of two different implants in three different femora. Several studies have also used surrogate models based on response surfaces to investigate implant behaviour; for example, Bah and Browne investigated the effect of geometric uncertainty on cemented hip replacement behaviour. Advanced mean value has been used in several studies of knee replacement behaviour to investigate knee wear simulator mechanics. A different approach taken by Perez et al. used the PFEM to predict stochastic stress distributions arising from variable loading and damage to study failure of cemented joint replacements. One striking feature of many studies has been the reliance on an analysis of the initial state of the reconstruction, requiring only a single load cycle to be applied in the simulations. Relatively few studies, e.g. Perez et al. and Knight et al., have performed long-term non-linear simulations of failure approaching realistic proportions of service life.

This study addresses the problem of applying a probabilistic framework to model failure owing to aseptic loosening in cemented total hip replacement (THR). Particular attention is given to incorporating the influence of long-term failure processes, rather than inferring response from the initial post-operative state. In a conceptual demonstration of the potential of the framework, an estimate of the influence of variability in patient-specific geometry, prosthesis design, and prosthesis position on post-operative prosthesis migration is undertaken. Implications of using deterministic simulations to assess risk are investigated by comparing predictions from the stochastic framework with predictions from a set of deterministic simulations.

**Methods**

**Failure model for aseptic loosening**

As a basis for a single deterministic simulation a failure modelling scheme, developed by Lennon et al. to assess revision risk based on prosthesis migration was adopted. Multiple deterministic simulations performed over a range of input parameters were used to build a training set for fitting a response surface. A combination of creep, fatigue damage, and interfacial debonding was used to simulate prosthesis migration. Cement creep was calculated using a Maxwell creep model. Fatigue damage accumulation was based on an anisotropic continuum damage mechanics formulation. Additionally, interfacial fatigue was modelled using a shear fatigue model proposed by Perez et al. which was further adapted to include tensile fatigue failure. Repeated loading over time, and high stresses in the cement and at the prosthesis interface, induce microcracks, creep deformation, and interfacial debonding in the model, causing degradation of the cement material properties; this, in turn, results in prosthesis migration over time. Each failure mode was included within a finite element simulation using customizable subroutines available within the MSC.Marc (MSC Software, Inc., USA) finite element code.

Prostheses were assigned properties of stainless steel ($E = 210 \text{ GPa}, \nu = 0.33$) and undamaged bone cement was also assumed to be linear isotropic ($E = 2.28 \text{ GPa}, \nu = 0.3$). Cortical bone was modelled as transversely isotropic with properties taken from Reilly and Burstein: longitudinal elastic modulus of 17 GPa, transverse modulus of 11.5 GPa, longitudinal shear modulus of 3.3 GPa, and transverse Poisson’s ratio of 0.46 and transverse component of 0.58. Cancellous bone was assumed to have an apparent isotropic modulus of 1.5 GPa and Poisson’s ratio of 0.33. Cement material properties were assigned to integration points based on results of a ray-tracing test to determine which integration points were inside a superimposed surface mesh of the cement mantle. Muscle loads were applied to simulate both walking and stair-climbing activities using the proximal femur loading dataset proposed by Heller et al. For a particular muscle (gluteus maximus, vastus medialis, vastus lateralis) attachment locations were identified by performing a landmark transform. The muscle standardized femur was used as a source of muscle attachment surface patches. Each muscle was characterized as a centroid of the muscle attachment patch. These points were located on the different femur surfaces included in the study using the `vtkLandmarkTransformFilter` class available within the Visualization Toolkit (VTK) (Kitware Inc., NY, USA). The muscle loads were distributed to

![Image](https://via.placeholder.com/150)
several closest nodes on the surface of the target mesh. Load magnitudes of both joint and muscle loads were proportional to body weight,22 which was assumed to be constant for all models and was equal to 70 kg. A weekly load profile was utilized consisting of alternating blocks of walking followed by stair-climbing. Duration of each activity, measured in cycles, was calculated based on activity data from Morlock et al.24 This weekly profile was repeated 52 times to simulate 1 year of activity.

Case-specific finite element meshes (see ‘Response surface construction (sampling)’) were created using an automated mesh generation technique developed in-house. Python scripts were utilized to set up finite element (FE) models in a semi-automated manner within the MSC.Mentat pre-processor (MSC Software, Inc., USA).

Response surface methodology

Response surface methods (RSM) use design space sampling to construct an estimate of the response surface of a process.25 Frequently, minimum, maximum, and ridge analyses of such a surface are used to optimize an experiment design. Another application is to use the estimated function as an inexpensive surrogate model to perform sensitivity studies or even Monte Carlo simulations in a fraction of the time that would be required by the more expensive initial sampling procedure;3 this was the approach taken in this study.

A second degree polynomial was used for definition of the response surface owing to ease of use and application,26 and because it is unusual to require a third-order model within many real-world design spaces.27 A response surface was defined as follows

\[ y_r = b_0 + x^T b + x^T B x \]  

where

\[ x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \]

is a vector of explanatory variables

\[ b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \]

are linear regression coefficients, and

\[ B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} \]

are quadratic coefficients for a particular response surface.

Choice of explanatory and response variables

In order to perform stochastic simulations explanatory variables, or factors, and response variables need to be defined. For this study, resultant prosthesis migration at 1 year was chosen as the response variable, with higher migration assumed to indicate higher risk of failure. Explanatory variables were defined by combinations of hypothetical femur geometry, prosthesis choice, and prosthesis position.

Femur geometry. Only the proximal part of the femur was used for this study and femur geometry was described by three values: medial–lateral width of the femur at the isthmus, \( W_i \), medial–lateral width at the lesser trochanter, \( W_{lt} \), and the distance between the lesser trochanter and the distal border of the isthmus, \( L_i \) (Figure 1(a)).

Prosthesis type. Prostheses were characterized quantitatively using the implant head offset, \( L_{offset} \), stem cross section diameter, \( D_{cs} \), femoral neck angle, \( \alpha_{neck} \), and length of the stem, \( L_{stem} \) (Figure 1(b)).
Medial–lateral (ML), anterior–posterior (AP), and rotations (g, b, c) were restrained at one node, simulating the use of a central canal. For the deviations, the tip of the implant was defined parametrically within a simulation.

Prosthesis positioning. Prosthesis position was limited to rotational deviation from a reference position achieved by means of standard templating in the medullary canal. For the deviations, the tip of the implant was restrained at one node, simulating the use of a centralizer, and rotations (α, β, γ) were applied about the medial–lateral (ML), anterior–posterior (AP), and proximal–distal (PD) axes, respectively (Figure 2).

Response surface construction (sampling)

A combination of these three categories of explanatory variable results in ten factors that need to be included in the response surface. Furthermore, the use of a quadratic response surface requires three levels for each factor in order to perform the necessary regression to fit a response surface to the sample points. Frequently, a ‘design of experiments’ approach is used to specify the required combinations of factors to achieve a realistic approximation of the response surface; e.g. full or fractional factorial design, central composite design, or Box–Behnken design. However, to achieve this type of sampling in this study would require parameterized models of both the femur and prosthesis that enabled independent variation of each of the explanatory variables. As a conceptual investigation of the utility of the response surface method was of prime consideration, a considerably reduced design was considered, restricted primarily by the use of a limited number of femur and prosthesis models.

First, three different prostheses were used to include prosthesis variability: flanged charnley with 40 mm offset (F40), flanged extra-heavy charnley (F40EH), and long-neck long-stem charnley (LNLS). These implants belong to the same family, and, therefore, dimensions can be defined parametrically within a simulation. Furthermore, surgeons frequently tend to use prostheses from a particular range, selecting a specific size depending on patient hip geometry. Second, three publicly available femur geometries were included in the study to represent femur variability: the third generation composite femur model, the Standardized Femur and the Visible Human femur. Measurements from each of the models were used to generate the set of sampling factor combinations related to the prosthesis design (i.e. L_stem, D_cs, α_neck) and femur geometry (i.e. L_i, W_i, W_b). For prosthesis positioning only independent rotation about the ML, AP, and PD axes was considered to simulate surgeon-dependent variability (i.e. α = ±0.5°, β = ±1°, γ = ±1° and the centre point α = 0°, β = 0°, γ = 0°). This resulted in 63 simulations in total (three femora, three prostheses, and seven orientations).

Predicted end-point migration data were collected from the 63 deterministic failure simulations and used to construct response surface coefficients in MATLAB (The Mathworks Inc., MA, USA). These coefficients were used to implement response surfaces as quadratic functions of the explanatory variables with coefficients. Six response surfaces were constructed to estimate the individual components of the migration vectors for the two activities of walking and stair-climbing simulated in the deterministic sample cases. Numerical evaluation of the deterministic response surfaces for a particular combination of explanatory variables could be executed in milliseconds, compared with between 24 and 48 hours for each finite element-based failure simulation.

Monte Carlo simulations

A Monte Carlo simulation was performed by evaluating the response surface function for a randomly generated set of explanatory parameters. Implant- and position-related parameters were assumed to be independent and uniformly distributed in a given range. A uniform distribution was chosen for these parameters because the total range of positions and orientations used were relatively small; thus, a reasonable first approximation was to assume that the probability of every implantation scenario was equal. Femur-specific parameters typically have a normal distribution (see Noble et al.); however, for this study, a small number of femur geometries representing a relatively small range were used, so the distribution was considered to be uniform between minimum and maximum values of femur dimensions obtained from the models. Table I describes the ranges of explanatory variables used in the study.

Normally, to ensure reliability of the results, Monte Carlo trials are performed repetitively until a stable solution is observed; mean resultant migrations owing to walking and stair-climbing of the implant tip were used as indicators of the converged solution in the present study. Sensitivity analysis based on Pearson correlation coefficients and generalized linear regression analysis was performed on the resulting dataset. Finally, based on the results of stochastic and
deterministic datasets, the implants were ranked by migrations, which indicate the risk of THR failure.34

Test of response surface prediction

Nine additional deterministic simulations were carried out on the third-generation composite femur model in order to assess the ability of the response surface approach to predict parameter combinations that have not been used for the construction of the response surface. Good candidates for such a check are combinations of parametric limits (i.e. ‘corner’ points of the design space) and intermediate values not included in the original deterministic simulations used to construct the response surface. Therefore, three combinations of the placement parameters \((\alpha, \beta, \gamma) = (1,1,0), (0,1,-1), \) and \((0,1,1), \) and six intermediate deviations \((\alpha = \pm 0.5, \beta = \pm 0.5, \gamma = \pm 0.5)\) were used to perform this test.

Results

Computation time

Each deterministic failure simulation used to generate the response surface functions was solved in approximately 24–48 h. A total of 20,000 Monte Carlo trials, using a desktop computer (Intel Pentium(R) 4 CPU, 3.00 GHz, 1.00 GB RAM), were performed in 3 s, giving a mean time to execute one trial of the response surface functions of 0.15 ms, illustrating the dramatic speed-up achievable with surrogate model approaches. Mean migration values for each activity converged within 10,000 trials without employing optimization strategies (Figure 3(a)).

Prosthesis performance

Ranking of prostheses based on mean migrations predicted the same ranking for walking and stair-climbing

Table 1. Explanatory variable ranges used in Monte Carlo trials.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Femur</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W_{ist}</td>
<td>Femur width at isthmus (mm)</td>
<td>26</td>
<td>32.3</td>
</tr>
<tr>
<td>W_{lt}</td>
<td>Femur width at lesser trochanter (mm)</td>
<td>43</td>
<td>47</td>
</tr>
<tr>
<td>L_{ist}</td>
<td>Distal isthmus border (mm)</td>
<td>127.3</td>
<td>143.4</td>
</tr>
<tr>
<td>Prosthesis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L_{head}</td>
<td>Implant head offset (mm)</td>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td>\alpha_{neck}</td>
<td>Implant neck angle (degrees)</td>
<td>138</td>
<td>155</td>
</tr>
<tr>
<td>D_{cs}</td>
<td>Stem cross-section</td>
<td>13.75</td>
<td>17.02</td>
</tr>
<tr>
<td>L_{stem}</td>
<td>Stem length (collar tip) (mm)</td>
<td>122.5</td>
<td>126</td>
</tr>
<tr>
<td>Placement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>\alpha</td>
<td>Rotation about AP axis (degrees)</td>
<td>-0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>\beta</td>
<td>Rotation about ML axis (degrees)</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>\gamma</td>
<td>Rotation about PD axis (degrees)</td>
<td>-1</td>
<td>1</td>
</tr>
</tbody>
</table>


Figure 3. Migration for walking and stair-climbing activities, predicted by Monte Carlo simulations, utilizing response surface surrogate models: (a) mean and standard deviations for each activity, illustrating convergence by 10,000 trials; (b) histograms illustrating the distribution of predicted migrations for each activity.
activities using both deterministic and stochastic simulations: lowest migration was for the F40 stem, while the LNLS was predicted to migrate most on average (Table 2). Activity was also predicted to affect prosthesis migration, with stair-climbing predicted to cause greater migration than walking (Figure 3(b)).

Results for an individual femur were compared with results of all femurs to assess whether different prosthesis rankings would arise for patient-specific simulations versus simulations over a range of patients. When the deterministic FE simulations were carried out on a single femur (the third-generation composite femur), ranking of the implants by migration (small to large) were F40EH, F40, and LNLS. However, when more femur models were used (to simulate inter-patient variability) the predicted ranks were F40, F40EH, and LNLS (Table 2).

Analysis of variance demonstrated that the deterministic training set could only partially discriminate implant performance – mean walking migration for the F40 implant was significantly different from mean walking migration for the LNLS prosthesis ($p < 0.05$), but no significant difference was predicted between the F40 and F40EH or the F40EH and LNLS. Similar results were predicted for stair-climbing – migration for the F40 was significantly smaller than for both F40EH and LNLS ($p < 0.05$); however, the F40EH implant did not migrate significantly less than the LNLS. Statistical discrimination between the F40 and other prostheses was strengthened considerably using the stochastic approach ($p < 0.0005$). However, the F40EH and LNLS stems were still not statistically different when considering only walking. Stair-climbing was a more discriminatory activity, in that all prostheses for the stochastic simulations were significantly different from each other ($p < 0.0005$).

### Sensitivity of parameters

For the deterministic training set, the most sensitive parameter based on correlation coefficients was the femoral width variable, $W_{lt}$ ($-0.2030$ for walking and $-0.3012$ for stair-climbing: Table 3). For the stochastic simulations, two variables were predicted to be considerably more sensitive than other parameters: the length variable, $L_i$ (0.4488 for walking and 0.4672 for stair-climbing) and the femoral width variable (0.4783 for stair-climbing: Table 3). Examination of the sign of the correlation coefficients and the regression coefficients (Table 3) indicates that predicted migration tended to increase with reduced femoral width and increased femoral length (i.e. was inversely proportional to femoral width and proportional to length).

### Deterministic versus stochastic

In order to investigate if stochastic simulation could yield different conclusions compared with conventional deterministic simulation, predicted rankings of the implants in one of the femurs (third-generation composite femur) for the stochastic simulations were compared with individual deterministic predictions for the template positions of the implants in the same femur. When ranked according to migration during walking, the ranking, based on the simulations for the template positions, were different compared with the ranking

### Table 3. Sensitivity analysis for patient-specific geometry and surgical technique variables = correlation and linear regression coefficients.

<table>
<thead>
<tr>
<th>Sensitivity measure</th>
<th>Variable/case</th>
<th>D/W</th>
<th>D/S</th>
<th>S/W</th>
<th>S/S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation coefficient</td>
<td>$\alpha$</td>
<td>-0.0474</td>
<td>-0.0036</td>
<td>0.1127</td>
<td>0.2075</td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>-0.1112</td>
<td>0.0524</td>
<td>-0.1455</td>
<td>0.1980</td>
</tr>
<tr>
<td></td>
<td>$\gamma$</td>
<td>-0.0263</td>
<td>0.0030</td>
<td>0.0256</td>
<td>0.0076</td>
</tr>
<tr>
<td></td>
<td>$W_{lt}$</td>
<td>0.0950</td>
<td>-0.0377</td>
<td>-0.0160</td>
<td>0.0060</td>
</tr>
<tr>
<td></td>
<td>$W_{lt}$</td>
<td>-0.2030</td>
<td>-0.3012</td>
<td>-0.2643</td>
<td>-0.4783</td>
</tr>
<tr>
<td></td>
<td>$L_i$</td>
<td>0.1752</td>
<td>0.0382</td>
<td>0.4488</td>
<td>0.4672</td>
</tr>
<tr>
<td>Linear regression coefficient</td>
<td>$\alpha$</td>
<td>-0.0027</td>
<td>0.0029</td>
<td>0.0074</td>
<td>0.0179</td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>-0.0096</td>
<td>0.0052</td>
<td>-0.0094</td>
<td>0.0170</td>
</tr>
<tr>
<td></td>
<td>$\gamma$</td>
<td>-0.0023</td>
<td>0.0003</td>
<td>0.0021</td>
<td>0.0013</td>
</tr>
<tr>
<td></td>
<td>$W_{lt}$</td>
<td>-0.0163</td>
<td>-0.0238</td>
<td>-0.0002</td>
<td>0.0016</td>
</tr>
<tr>
<td></td>
<td>$W_{lt}$</td>
<td>-0.0218</td>
<td>-0.0220</td>
<td>-0.0082</td>
<td>-0.0152</td>
</tr>
<tr>
<td></td>
<td>$L_i$</td>
<td>0.0120</td>
<td>0.0139</td>
<td>0.0037</td>
<td>0.0060</td>
</tr>
</tbody>
</table>

D/W: deterministic approach, walking migration; D/S: deterministic approach, stairclimbing migration; S/W: stochastic approach, walking migration; S/S: stochastic approach, stairclimbing migration.
predicted by the stochastic simulations (Table 4). Rankings based on stair-climbing were equivalent.

Expanding the deterministic case to include all the training set simulations also indicated differences between stochastic and deterministic approaches. First, a difference between mean values predicted by the deterministic and stochastic simulations was computed. Typically, the deterministic approach tended to predict higher values with the only exception, in the case of stair-climbing, being the LNLS implant (Table 2).

Second, correlation coefficients for the stochastic simulations were larger, i.e. sensitivity of the parameters was stronger. The strongest explanatory variables did not change between deterministic and stochastic approaches.

Both sets of deterministic simulations, walking and stair climbing, typically were not sufficient for statistically significant ranking of all implants and during pair-wise comparison (Figure 4). In contrast, the stochastic approach was capable of predicting stronger differences between implants.

**Response surface accuracy**

Comparing the results for the verification set of parameters, using the deterministic approach (FE simulation) with RSM predictions (Table 5), found that errors were lower for the case of intermediate deviations of single parameters (mean prediction errors were 0.019 mm for walking activity and 0.007 mm for stair-climbing), compared with cases in which combinations of parameters were used (mean prediction errors were 0.029 mm for walking and 0.035 mm for stair-climbing activity).

**Discussion**

The aim of this work was to develop a stochastic failure model of femoral component loosening incorporating such uncertainties as patient geometry, implant design and surgical technique. This was achieved by utilizing an automated mesh generation technique and simulating 63 case-specific scenarios (three femurs, three implants, seven positions per implant–femur combination) to obtain a training set for the construction of a stochastic predictive model. An investigation of implant performance in a number of femurs (emulating inter-patient variability) was undertaken to investigate the framework’s capabilities for assessing revision risk in a population sample.

For this study several limitations related to the FE and aseptic loosening models were inherited from Lennon et al.16

1. Material properties of each tissue type (cortical and cancellous bone, and marrow) were assumed to be homogeneous and constant. No microstructure was taken into account. No bone remodelling processes were included to allow adaptation to the altered tissue loading arising post-implantation.

---

**Table 4.** Predicted prosthesis ranking in the third-generation composite femur model for deterministic simulations of template positions, compared with stochastic simulations including variable positioning. Template implant positions were the neutral positions (i.e. $\alpha$, $\beta$, $\gamma$ = 0).

<table>
<thead>
<tr>
<th></th>
<th>$W_{\text{template}}$</th>
<th>$W_{\text{stochastic}}$</th>
<th>$S_{\text{template}}$</th>
<th>$S_{\text{stochastic}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highest to</td>
<td>F40</td>
<td>LNLS</td>
<td>LNLS</td>
<td>LNLS</td>
</tr>
<tr>
<td>Lowest</td>
<td>LNLS</td>
<td>F40</td>
<td>F40</td>
<td>F40 EH</td>
</tr>
</tbody>
</table>

$W$: walking-based ranking; $S$: stair-climbing based ranking; template: template simulations; stochastic: Monte Carlo simulations based on a response surface model generated specifically for the third-generation composite femur model.

**Table 5.** Comparison of predicted migration by a single deterministic simulation and RSM-based response, mm: Italic fonts: two largest prediction errors for walking and stair-climbing.

<table>
<thead>
<tr>
<th>Model</th>
<th>Det.Walk</th>
<th>RS.Walk</th>
<th>$\delta$</th>
<th>Det.StCl</th>
<th>RS.StCl</th>
<th>$\delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$+a/2$</td>
<td>0.1766</td>
<td>0.1674</td>
<td>0.0091</td>
<td>0.2203</td>
<td>0.2076</td>
<td>0.0127</td>
</tr>
<tr>
<td>$-a/2$</td>
<td>0.1820</td>
<td>0.1610</td>
<td>0.0210</td>
<td>0.2034</td>
<td>0.2015</td>
<td>0.0020</td>
</tr>
<tr>
<td>$+b/2$</td>
<td>0.1858</td>
<td>0.1573</td>
<td>0.0285</td>
<td>0.2028</td>
<td>0.2032</td>
<td>-0.0005</td>
</tr>
<tr>
<td>$-b/2$</td>
<td>0.1859</td>
<td>0.1712</td>
<td>0.0147</td>
<td>0.2128</td>
<td>0.2008</td>
<td>0.0120</td>
</tr>
<tr>
<td>$+y/2$</td>
<td>0.1857</td>
<td>0.1593</td>
<td>0.0264</td>
<td>0.2145</td>
<td>0.2042</td>
<td>0.0103</td>
</tr>
<tr>
<td>$-y/2$</td>
<td>0.1837</td>
<td>0.1683</td>
<td>0.0154</td>
<td>0.2102</td>
<td>0.2040</td>
<td>0.0063</td>
</tr>
<tr>
<td>$+b + a$</td>
<td>0.1927</td>
<td>0.1612</td>
<td>0.0315</td>
<td>0.2195</td>
<td>0.2010</td>
<td>0.0184</td>
</tr>
<tr>
<td>$+b - a$</td>
<td>0.1320</td>
<td>0.1530</td>
<td>-0.0210</td>
<td>0.2285</td>
<td>0.1894</td>
<td>0.0392</td>
</tr>
<tr>
<td>$+b + \gamma$</td>
<td>0.1902</td>
<td>0.1627</td>
<td>0.0355</td>
<td>0.2396</td>
<td>0.1923</td>
<td>0.0473</td>
</tr>
</tbody>
</table>

Det: predicted using deterministic approach; RS: predicted by evaluating a response function; $\delta$: difference between predicted values; Walk: walking activity; StCl: stair climbing.
The same material properties were used across different patients. This limitation could be overcome in future by using patient-specific material properties obtained from medical images. Furthermore, bone remodelling algorithms could also be incorporated to simulate adaptation of the bone to the altered loading experienced by the femur following introduction of a prosthesis (see Scannell and Prendergast); however, their inclusion will require enlargement of the training set owing to the increase in parameters.

2. Only aseptic loosening owing to a cement damage accumulation failure scenario was investigated. Other failure scenarios, as described by Huiskes, could be included, e.g. a particulate reaction failure scenario, by including accelerated bone loss owing to excessive wear of the articulating surfaces of the joint reconstruction.

3. Physical activity was simulated assuming the same number of walking and stair-climbing loading cycles for each femur (or patient). However, patient activity is also subject to variability. Questionnaires, such as Western Ontario and McMaster Universities Arthritis Index (WOMAC) or University of California, Los Angeles (UCLA) activity scores, could be utilized to improve activity modelling for patient-specific studies. Datasets on daily activity (e.g. Morlock et al.) could also be used to develop a stochastic activity model.

4. An ideal shape of cement mantle was assumed. No defects were taken into account. Homogeneous material properties were assumed. Size and position was assumed to be slightly larger than the implant in the reference position. A stochastic approach, for instance, similar to Lennon and Prendergast, to simulate variability of bone cement damage owing to random porosity, could be employed to improve this modelling aspect. Moreover, variability of cement mantle sizes, shapes, and positions could also be investigated.

In addition to assumptions related to the mechanical simulations, there were also several assumptions for the probabilistic aspects of the framework.

1. Approximation of the migration prediction was based on quadratic functions and a small number of deterministic simulations, i.e. three femoral shapes, three implants, and seven positions for each femur–implant combination (reference and six deviations). A full, fractional, or two-level fractional (with some centre-point runs) factorial design may be used to construct more reliable response surfaces. Efficacy of the approximation was assessed by simulating additional possible scenarios and computing a corresponding prediction using the response surfaces. The predictions were more reliable for the individual parameter deviations than for the combined deviation of pairs of parameters. This was most likely owing to the fact that combinations of parameters were not used for the construction of the response surfaces. Thus, the next step is to increase the number of deterministic simulations and construct more reliable response surfaces, ensuring simultaneous variation of multiple parameters are used in the training set.

2. Selection of explanatory variables should be more systematic and based on a statistical analysis performed on the modelled entity. For instance, an implant design can be described by a larger number of variables than presented in this study; however, only a subset of them may describe the unique nature of a given design. Such relationships can be revealed by performing principal component analysis (PCA) of the implant dimensions. Also, implant positions can be described by a larger number of variables than the three rotations used in this study, i.e. more complex three-dimensional transforms using both rotational and translational transformations could be used as a variable in this case.

3. Ranges of explanatory variables were obtained by measuring the computer-aided design (CAD) models used in the study. It is possible to define ranges of variables from the literature, e.g. femur dimensions could be taken from Noble et al., if a custom implant is being designed for a patient there are patient-specific limitations on its dimensions. Recently, statistical shape modelling approaches have emerged as a useful tool to incorporate anatomical variability in computational studies.

4. Moreover, with inclusion of realistic ranges for explanatory variables, realistic probability distribution functions need to be utilized as well, i.e. not all the variables are distributed uniformly.

5. Finally, the framework cannot be considered fully tested unless a Monte Carlo simulation, using full-scale FE trials rather than response surface functions, is undertaken.

These assumptions primarily limit this FE-based framework to a proof-of-concept demonstration. This study demonstrates a possible application of the stochastic framework on a small group of hypothetical patients for a particular failure mode (aseptic loosening) rather than a demonstration of comprehensive risk assessment in a realistic population sample. In order to enhance predictive capability of the framework it requires increasing the number of training simulations, as well as resolving the aforementioned limitations. However, many of the limitations of this study can be readily overcome in developing a more general tool; e.g. it is possible to include previously developed bone adaptation algorithms (including patient-specific variability), add more parameters taken from other stochastic investigations (such as muscle attachments and load directions), apply patient-specific material properties from CT/DEXA scans, and so on.
One advantage of performing parametric investigation, even with a relatively small set of deterministic simulations, is the ability to identify unforeseen results and predict trends. This makes computational assessment more relevant to clinical interpretation. One such finding from this study was the prediction from the deterministic trials that a position of the implant achieved by the standard templating procedure does not necessarily result in the lowest implant migration; this finding was reinforced by the Monte Carlo simulation. Another interesting prediction was that the extra-heavy Charnley (F40EH) had the lowest migrations in the case of the third-generation composite femur, whereas the standard Charnley was the least migrating prosthesis for all femurs studied. This highlights the often competing requirements of an individual, compared with those of a population, and represents a significant challenge to implant developers. One more interesting observation can be made based on Figure 4. The mean migration computed for the

Figure 4. Mean predicted migrations and confidence intervals (95%) for each activity for both deterministic- and stochastic-response surface simulations. Stochastic simulations predicted stronger differences between prosthesis types, as illustrated by the narrower confidence intervals when compared with the deterministic results. Implant colour-coding: (green/left) flanged Charnley 40 mm; (red/middle) extra-heavy Charnley 40 mm; and (blue/right) long-neck long-stem Charnley.
extra-heavy Charnley, in the case of stochastic simulation for stair-climbing activity, does not belong to the confidence interval of the deterministic trial. This suggests that two absolutely different results can be achieved by using stochastic and deterministic approaches. Appendix 1 supports this point by providing an analytic proof.

All these examples can be considered as surgical effects in that they result from selection and positioning of an implant. Patient-specific effects were also predicted using the probabilistic framework. For example, sensitivity analysis predicted that size of the host bone had a larger effect on implant migrations than surgical procedure. Moreover, negative linear regression coefficients suggest that patients with larger bones are less susceptible to high implant migration; a similar result has also been predicted for primary stability of cementless femoral prostheses by Viceconti et al. Implant migrations owing to stair-climbing activity were more indicative of a difference in implant behaviours – therefore, stair-climbing could be a better predictor of implant failure and is thus a critical activity to include in the failure modelling scheme.

Response surface methods assume that a sample of the input parameter space can be used to construct a reasonable approximation of system behaviour. Provided this is the case, the response surface is suitable as a surrogate model to quickly compute predictions for combinations of input parameters not included in the initial approximation of the response, and thus increase statistical power. For example, in the current study, using the response surface to perform a Monte Carlo simulation enabled statistically significant differences to be predicted that were not achievable from the deterministic trial (Figure 4). Furthermore, different risk rankings were predicted when only individual template positions for the implants were considered, compared with when variable positions (using the training set) and random positions (using the Monte Carlo simulations) were considered. This illustrates that the proposed stochastic approach is more likely to capture the performance of an implant under varying conditions and that misleading conclusions can result from relying on individual simulations of average or idealized cases.

Realistic assessment of revision risk of implant performance in a population is a daunting, yet important, goal for computational biomechanics. This study demonstrates a way in which some of the complexity of long-term performance (i.e. high cycle fatigue damage accumulation, creep of the cement layer, and debonding of the prosthesis) can be incorporated in a risk assessment scheme that includes relatively large numbers of random variables in simulations that can be executed in achievable computational times. It is not difficult to envisage a framework based around the developed toolset that would allow investigation of several failure modes at the same time, e.g. damage accumulation, particulate reactions, and bone resorption. Provided that clinical corroboration of the predictions can be achieved, several applications can be envisaged, for example, pre-operative decision support, post-operative monitoring to highlight patients most at risk of early revision, design of new implants to perform well in large populations, and even custom implants for individual patients.

Conclusions

A stochastic framework capable of incorporating patient femur geometry, implant design, and surgical technique was developed. This study suggested that femoral dimensions were important factors for failure risk of cemented THR: with increase of the femur width at the lesser trochanter level, predicted migration levels decreased. Stair-climbing was predicted to be more discriminatory in ranking prostheses than walking alone and, thus, inclusion of stair-climbing is recommended for simulations attempting to assess aseptic loosening risk of hip prostheses. Benefits of the stochastic framework were an ability to increase sensitivity and statistical power when compared with predictions from a deterministic parameter study. Furthermore, it was demonstrated that reliance on individual deterministic simulations can lead to different conclusions regarding prosthesis performance compared with stochastic simulations, and thus approaches accounting for at least some variability are to be recommended over individual simulations of idealized or average cases. Although subject to a number of limitations and in need of further development, an accelerated Monte Carlo simulation framework based on response surface methods, or other surrogate models, shows considerable promise as a tool for computational risk assessment of orthopaedic implants, such as cemented hip replacements.

Funding

This research was funded by Science Foundation Ireland under its Research Frontiers Programme (05/ RFP/ENG0097).

References


**Appendix I**

Figure 4 shows mean migration values and 95% confidence intervals for different physical activities. In the case of the F40EH stochastic study for stair-climbing, the mean migration value lies outside of the confidence interval for the deterministic simulation. This can be explained by looking at the form of the response surface functions used in the study.

For example, let us consider a two-dimensional case: a polynomial of the second order was fitted to points \((x_0, y_0), (x_0 + d, y_1), \) and \((x_0 + 2d, y_2)\). Let us write this polynomial as

\[ f(x) = ax^2 + bx + c \]

where \( f(x_0) = y_0, \) \( f(x_0 + d) = y_1, \) \( f(x_0 + 2d) = y_2.\)

Then, the computed mean over three function values can be written as

\[
\begin{align*}
 f_{\text{mean}} &= \left[ ax_0^2 + bx_0 + c + a(x_0 + d)^2 + b(x_0 + d) + c \right] / 3 \\
 &= a(x_0 + d)^2 + b(x_0 + 2d) + c \\
 &= c + a(x_0 + 2d)^2 + b(x_0 + 2d) + c
\end{align*}
\]

This is equivalent to the deterministic mean in the case of the F40EH study. On the other hand, the mean value for a stochastic simulation, with uniform distribution of the explanatory variable, approaches the following value, when the number of stochastic trials approaches infinity

\[
 f_{\text{stoch.mean}} = \frac{1}{2d} \int_{x_0}^{x_0 + 2d} f(x) dx = \left[ \frac{x^3}{3} + \frac{b x^2}{2} + c x \right]_{x_0}^{x_0 + 2d}
\]

Therefore, a situation when \( f_{\text{stoch.mean}} < f_{\text{mean}} - CI, \) can be simplified by computing the integral and cancelling corresponding terms to a quadratic inequality

\[
4d^2 - 5d - 3\alpha/a < 0
\]

here \( a \) is the coefficient of the highest degree in \( f, \) and \( \alpha \) is a positive length of the confidence interval. In case of a positive \( a, \) this inequality can be satisfied in the range \( d = [0,1], \) which corresponds to the positive ranges used in this study. This means that both approaches, stochastic and deterministic, can predict different mean migrations and do not necessarily belong to the confidence intervals of each other.