Model Validation and Uncertainty Quantification Applied to Cervical Spine Injury Assessment

Dr. Ben H. Thacker, Mr. W. Loren Francis and Dr. Daniel P. Nicolella
Southwest Research Institute®
Mechanical & Materials Engineering Division
6220 Culebra Road
San Antonio, Texas 78238, USA

bthacker@swri.org

ABSTRACT

Cervical spine injuries can occur as a result of large inertial forces such as those experienced by military aviators during ejection or evasive maneuvers. Mitigating the risk of injury during these types of extreme scenarios requires a fundamental understanding of the injury mechanisms associated with the transferral of loads through the human body and the ensuing kinematic response to these loads. This understanding can then be used to develop operational procedures, flight deck equipment and personnel protection gear to minimize the risk of injury. However, the wide range of loading scenarios, anatomical differences associated with aviator size and gender, and the obvious inability to perform full scale testing necessitate the use of numerical modelling and simulation to predict the uncertain behaviour and response of the human body to extreme loads. The credibility in these predictions is established through the application of formal model verification and validation (V&V) practices and procedures. A critical aspect to V&V applied to highly complex models where uncertainties are large and testing is especially difficult is the use of a validation hierarchy where the full system model is subdivided into its constituent subassemblies and components. Model V&V is then applied to each submodel in the hierarchy such that the error and confidence in the full system prediction can be quantified. This paper will briefly describe a model V&V process currently under development by the American Society of Mechanical Engineers and then provide an example of the hierarchical process applied to the development of a cervical spine model for predicting the risk of injury to military aviators.

1.0 INTRODUCTION

Numerical models are now routinely used to simulate complex behaviour in solid mechanics, dynamics, hydrodynamics, heat conduction, fluid flow, transport, chemistry, biology, and acoustics. These models can produce response measures such as deformation, stress, and velocity histories, and failure measures such as fatigue, fracture, and wear. The high-fidelity of these models, however, should not be confused with credibility, i.e., there is generally not a one to one relationship between the two. Model fidelity is the result of modelling tools, simulation software and computational speed; credibility, on the other hand, is the result of model verification and validation (V&V).

Model V&V are the primary processes for quantifying and building credibility in numerical models. Verification is the process of determining that a model implementation accurately represents the developer's

Thacker, B.H.; Francis, W.L.; Nicolella, D.P. (2007) Model Validation and Uncertainty Quantification Applied to Cervical Spine Injury Assessment. In *Computational Uncertainty in Military Vehicle Design* (pp. 26-1 – 26-30). Meeting Proceedings RTO-MP-AVT-147, Paper 26. Neuilly-sur-Seine, France: RTO. Available from: http://www.rto.nato.int.

conceptual description of the model and its solution. Validation is the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model [1]. Model V&V are processes that accumulate evidence of a model's correctness or accuracy for a specific scenario; thus, V&V cannot prove that a model is correct and accurate for all possible scenarios, but rather provide evidence that the model is sufficiently accurate for its intended use. Therefore, model V&V is a process that is never fully completed, but rather concluded when acceptable accuracy is achieved.

Most complex systems will comprise multiple (and complex) subsystems and components, each of which must be modelled and validated. It is not entirely uncommon for model developers to attempt to validate full system models directly from any available test data. This approach can be problematic if there are a large number of components or if subsystem models that contain complex connections or interfaces, energy dissipation mechanisms, and highly nonlinear behaviour. If there is poor agreement between the prediction and the experiment, it can be difficult, if not impossible, to isolate which subsystem is responsible for the discrepancy. Even if good agreement between prediction and experiment is observed, it is still possible that the model quality could be poor because of error cancellation among the subsystem models. Therefore, a hierarchical strategy must be used to conduct a careful sequence of submodel validations to build confidence in the ability of the full system model to provide accurate predictions.

Uncertainty and error quantification play a key role in model V&V. Nondeterminism refers to the existence of errors and uncertainties in the outputs of computational simulations due to uncertainties in the model. Likewise, the measurements that are made to validate these simulation outputs also contain errors and uncertainties. While the experimental data are used as the reference for comparison, the V&V process does not presume the experiment to be "right" or even more accurate than the simulation. Instead, the goal is to quantify the uncertainties in both experimental and simulation results such that the predictive accuracy of the model can be quantified.

Cervical spine injuries can occur as a result of impact or from large inertial forces such as those experienced by military pilots during ejections, carrier landings, and ditchings. Understanding the mechanisms and risk of cervical spine injury to military pilots during high-risk situations motivated the development of a numerical modelling capability that could be used to quantify the risk of injury and explore design modifications to reduce this risk as much as possible. In addition, the effect of gender on the risk of injury was also of interest due to the increasing number of female aviators.

In biological systems there is a great deal of uncertainty associated with the environment in which the structure is required to function as well as physical and mechanical properties and geometry of bone, ligaments, cartilage, joint and muscle. These uncertainties have a direct effect on the behavior and response of the system. The consequences associated with cervical spine injuries are also usually severe. Therefore, it is of great interest to have a validated predictive tool that can be used to design occupant safety systems to minimize the probability of injury. To do this, the designer must have quantified knowledge of the probability of injury due to different loading scenarios, and also understand which model parameters contribute the most to the injury probability.

2.0 MODEL VERIFICATION AND VALIDATION

Model verification and validation is undertaken to quantify confidence and build credibility in a numerical model for the purpose of making a prediction. Ref. [1] defines prediction as "use of a computational model to foretell the state of a physical system under conditions for which the computational model has not been validated." The predictive accuracy of the model must therefore reflect the strength of the inference being made from the validation database to the prediction. If needed, the predictive accuracy of the model can be improved as supported by additional experiments, information, or experience.

Verification is concerned with identifying and removing errors in the model by comparing numerical solutions to analytical or highly accurate benchmark solutions. Validation, on the other hand, is concerned with quantifying the accuracy of the model by comparing numerical solutions to experimental data. In short, verification deals with the mathematics associated with the model, whereas validation deals with the physics associated with the model. Obviously, verification must be performed before validation.

It is important to define and differentiate between the terms "code" and "model" in the context used herein. A code is the computer implementation of algorithms in specific computer software developed to facilitate the formulation and approximate solution of a class of models. A model includes the conceptual, mathematical, and numerical representation of a specific physical scenario and the equations comprising the model are solved using the algorithms in the code.

2.1 Validation Hierarchy

Figure 1 shows a schematic of generic model decomposition into a hierarchy of submodels. The top tier represents the complete system. Three more generic tiers are shown: assemblies, subassemblies and components. These tiers illustrate the decomposition of a complex system into a series of fundamental physical problems. The number of tiers needed to compose a complete problem may be more or less than that shown.

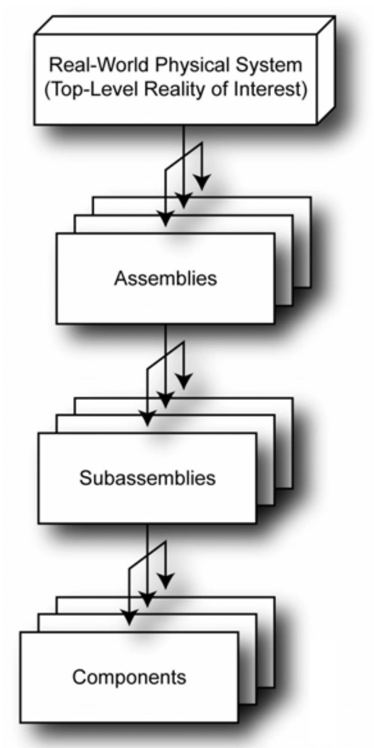


Figure 1: Validation hierarchy illustrating the decomposition of a full system model into submodels (components, subassemblies, and assemblies)

In Figure 1, component problems are typically physics-based and represent important problem characteristics that the model must be able to simulate accurately. Examples of unit problems include material coupon tests, interface or joint tests, and load environment tests. Component problems will typically involve simplifications involving idealized geometry, boundary conditions and applied loads. Also, some component problems are required to determine fundamental constitutive properties; i.e., they represent calibrations (or parameter estimation, but not validation) against the experimental data.

Careful construction of a validation hierarchy is of paramount importance because it defines the problem characteristics that the model must be able to simulate, the coupling and interactions between unit problems and the complete system, and arguably most importantly, the validation experiments that must be performed to validate the component, subsystem and assembly models. A procedure known as Phenomena Identification and Ranking Table (PIRT) has been successfully used to construct a validation hierarchy.

Each level of the hierarchy comprise problems that confirm the ability of the submodel to predict quantities and phenomenology that are important for the accurate simulation of the system performance. However, schedule or budget constraints may prohibit complete validation of every submodel in the validation hierarchy. If an upper level (and unvalidated) model is available, a sensitivity analysis can be performed to identify which submodels deserve the most attention. This focuses the validation effort on only those aspects that contribute significantly to the response of interest. The danger in this approach, of course, is in the use of an unvalidated model to determine, perhaps incorrectly, which submodels are important.

2.2 Validation Process

The validation process as defined in Ref. [1] is shown in Figure 2. The “Reality of Interest” at the top represents the particular problem being validated, whether a component, subassembly, assembly, or the complete system. V&V of a complete system will then require the process shown in Figure 2 to be repeated as the model development progresses upward from the bottom of the hierarchy shown in Figure 2.

In Figure 2, the left branch illustrates the process of developing and exercising the model, and the right branch illustrates the process of obtaining relevant and high-quality experimental data via physical testing. The closed boxes denote objects or data, connectors in black solid lines denote modelling or experimental activities, and the connectors in red dashed lines denote assessment activities. The validation process is fully described in Ref. [1].

2.3 Uncertainty Quantification

Nondeterminism is generally modeled through the theory of probability. Much of the nondeterministic information that exists in a numerical model can be identified and treated in order to quantify its effects and in some cases even reduce these effects. Uncertainties associated with the model input parameters include material behavior, geometry, loads, initial conditions, and boundary conditions. The variability (irreducible uncertainty) of these parameters can be estimated using repeated experiments to establish a statistically significant sample. In lieu of sufficient data, expert knowledge may be used to estimate distribution parameters or bounds; however, the imprecise nature of expert opinion must be reflected as additional uncertainty in the simulation outcomes.

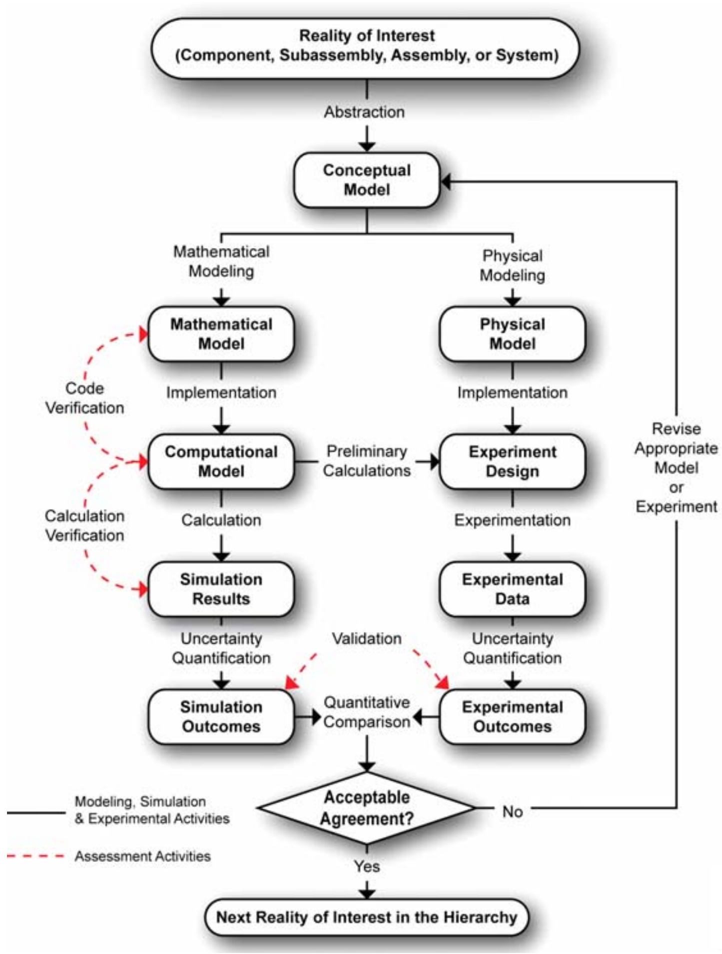


Figure 2: Detailed model development, verification and validation process

When the variability in the model input parameters has been established, this variability can be propagated through the simulation to establish an expected variability on the simulation output quantities. Sampling-based propagation methods (Monte Carlo, Latin Hypercube, etc.) are relatively straightforward techniques for propagating variabilities. These methods draw samples from the input parameter populations, evaluate the deterministic model using these samples and then build a distribution of the appropriate response quantities. Sampling methods can be made more efficient via the use of local response surface approximations (e.g., metamodel, reduced-order model, etc.) of the model being studied. However, the error involved in the use of a response surface must also be estimated. Sensitivity-based methods (First Order Reliability Method, Advanced Mean Value, etc.) may also be used to propagate input uncertainties to uncertainties on the response quantities, but again, the errors introduced by the approximations must be recognized and quantified.

Uncertainty quantification plays a key role in model V&V. It provides the basis for quantifying and understanding the effect of uncertainties in experimental data and numerical predictions, as well as for quantifying the predictive accuracy of the model. As compared to traditional deterministic analysis, quantifying the effect of uncertainties requires additional effort to collect and characterize input data, to perform the uncertainty analysis, and to interpret the results.

2.4 Validation Metrics

Comparing simulation and experimental outcomes requires the selection of one or more features and a metric for comparison. A feature could be a direct response measure such as stress, strain or frequency, or a derived quantity such as an average response, principal component, or integrated acceleration. Desirable characteristics of features are that they should possess physical interpretation, be relevant to the intended use of the model, and be reasonably stable such that they are not significantly affected by numerical or experimental noise.

Metrics are the basis for comparing features from experimental data with model predictions. A metric could be the comparison of the average feature values, the variance of the feature values, or the distribution of the feature uncertainty. Graphical overlays of experimental and model features do not quantify the measure of agreement and therefore are insufficient for validation. Desirable characteristics of metrics are that they should utilize both experimental and simulation uncertainty, quantify the agreement between the experiment and simulation, and reflect the level of confidence in the comparison as a function of the number of experimental measurements.

A simple uncertainty-based metric is the t-test. The t-test provides a measure of the statistical difference (p-value) between the means of two distributions. Typically a p-value of less than 5% is considered significantly different. The t-test assumes that the distributions are normal and have equal variances. In model V&V, this assumption will not be valid if the t-test is applied to feature values. In spite of this fact, however, the t-test is still a useful metric because it is a function of the number of experimental measurements, and provides a quantitative measure of agreement between the experiment and simulation. The p-value is computed as follows:

$$t_c = \frac{\bar{x}_{\exp} - \bar{x}_{\text{mod}}}{s_n / \sqrt{n}}, \qquad \text{p-value} = 2\Pr\left(t \geq |t_c|\right)$$

where s_n is the experimental standard deviation and n is the sample size. For the t-test we calculate a p-value from a two tailed t-distribution.

Statistical tests that compare two distributions should generally be more useful for model V&V. Standard statistical tests, such as the Kolmogorov-Smirnov test, can be used; however, most of these tests were designed to test for normality and in some cases do not perform well for small sample sizes.

A simple metric proposed here is based on the root-mean-square (RMS) of the difference between the experimental distribution and the model distribution. At each point in the experimental CDF the response of the model is subtracted from the experimental response. The RMS is calculated for all available CDF points in the experimental distribution:

$$\text{RMS} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_{i-\exp} - x_{i-\text{model}}\right)^2}$$

Another measure is the bias, which is the measure of the difference between the means:

$$\text{BIAS} = \left| \bar{x}_{\text{exp}} - \bar{x}_{\text{mod}} \right|$$

where $\bar{x}_{\text{exp}}$ and $\bar{x}_{\text{mod}}$ are the sample mean values from the experiment and model, respectively. When BIAS and RMS are equal and non-zero, the model is predicting the shape of the distribution but the predictions are offset by an amount equal to BIAS. If the model and experimental distributions are the same, then BIAS and RMS are zero.

3.0 CERVICAL SPINE VALIDATION

The purpose of this study is to develop, verify and validate a probabilistic parametric cervical spine finite element (FE) model for use in predicting the risk of cervical spine injury. Four FE models of the cervical spine representing two weight groups of male and female volunteers were developed using a parametric modelling procedure.

The V&V procedure consisted of validating the model at increasing levels of complexity, each step building upon the previous one. Four hierarchical levels, as shown in Figure 3, were used in this study: the component level (individual tissue properties), the meso-component level (the intervertebral disc construct), a complete motion segment, and the final full cervical spine column. At each hierarchical level above the component level, the model predictions were validated against experimental data without adjusting the model parameters to fit the experiment results.

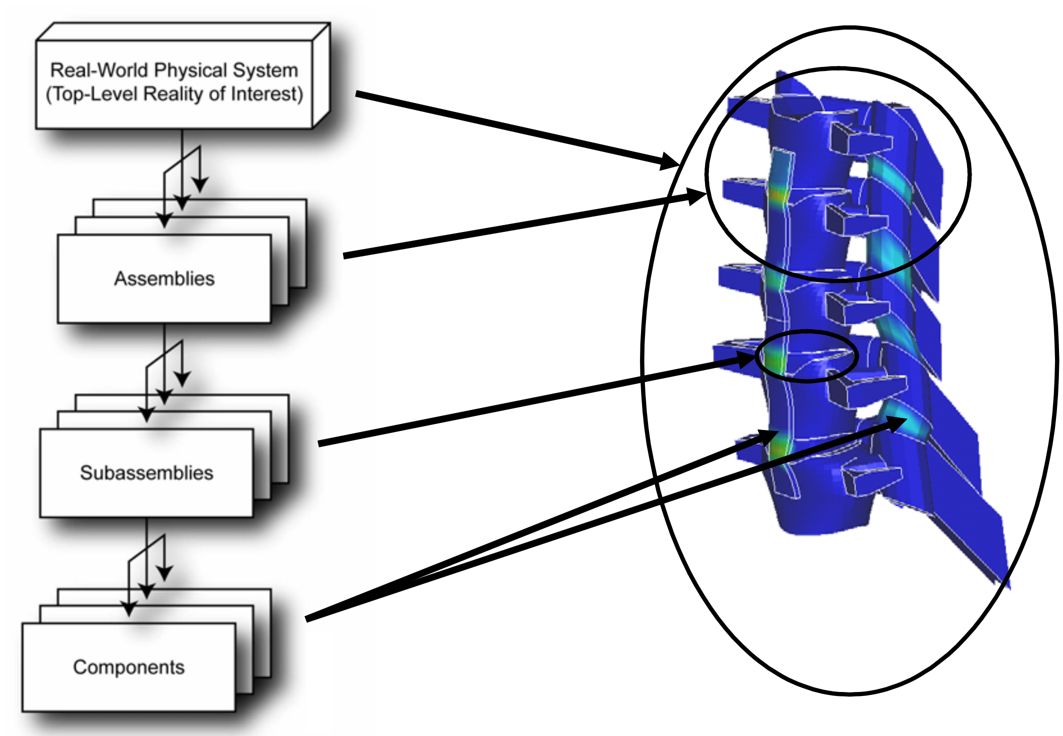


Figure 3: Validation hierarchy for the cervical spine

The first level of the hierarchy consisted of the individual ligaments. At this level, the objective was to calibrate the stress-strain response of each ligament using measured ligament stress-strain behavior. This level also confirmed the choice of element formulation and material model used to predict the behavior of the ligaments.

The second level consisted of an isolated C5 vertebral body, the intervertebral disc (IVD) between C5 and C6, and the C6 vertebral body. The objective of this level was to calibrate the behavior of the IVD material model. The component level model parameters were fixed from this point forward.

The third level comprised cervical motion segments for C3-C4, C4-C5 and C5-C6. Each motion segment included the associated vertebral bodies, interconnecting ligaments and IVD. Each motion segment FE model was validated against experimental measurements under flexion, extension, lateral bending, and axial rotation.

The fourth (top) level consisted of the complete cervical spine column. At this level the FE model was validated using both quasi-static *in vitro* experiment data as well as *in vivo* sled test experimental data.

A total of 100 (50 female and 50 male) healthy volunteers ranging in age from 18-40 years were scanned using computed tomography (CT). Selected key dimensions from the CT scans were then measured from three-dimensional reconstructions of the CT scans [2]. The parameter measurements were then averaged into four different groups: small female (106 lb – 120 lb), large female (136 lb – 150 lb), small male (166 lb – 180 lb), and large male (226 lb – 240 lb). Models representing each group were then constructed using an all-hexahedral finite element mesh. The cancellous material of the vertebral bodies and articular processes were modeled using hex elements; the cortical shell of those components was modeled with shell elements.

The mechanical simulations were performed using the LS-DYNA® explicit dynamic software (LSTC®, Livermore, CA). The probabilistic response was simulated using the NESSUS® (Version 8.3, Southwest Research Institute®, San Antonio, TX) probabilistic analysis software by incorporating measured variabilities at the component level into the model input parameters and propagating those uncertainties through the model. The probabilistic analysis approach yields a probability distribution for the model predicted responses rather than a single deterministic quantity (e.g., ligament strain), which then allows the response to be statistically compared to the experimental data for validation.

3.1 Ligament Calibration

Five ligaments that govern the overall kinematic response of the cervical spine (shown in Figure 4) were modelled using either three dimensional continuum elements or one dimensional springs. The anterior longitudinal ligament (ALL) and posterior longitudinal ligament (PLL) were created with eight-node hexagonal elements using previously determined values for the cross sectional area and the length.

The LS-DYNA soft tissue material model was used for the ligaments. The model is transversely isotropic hyperelastic with a quasi-linear viscosity (QLV) constitutive model [3]. One dimensional nonlinear spring elements were used to model the joint capsular ligaments (JC), the interspinous ligament (ISL) and the ligament flavum (LF). The inelastic spring model in LS-DYNA was used along with experimental force-displacement data from the literature [4]. Published quasi-static and dynamic experimental data [5, 6] was used to calibrate the soft tissue material model. The mean, standard deviation and distribution type for the material parameters (bulk modulus, K, and viscoelastic coefficients, G_1, G_2, and G_{inf}) and ligament cross section area were selected based on statistical analysis of experimental data.

Quasi-static tension simulations and experimental for the ALL, PLL and ISL ligaments are compared in Figure 5. The results show good agreement throughout the range of normalized stretch.

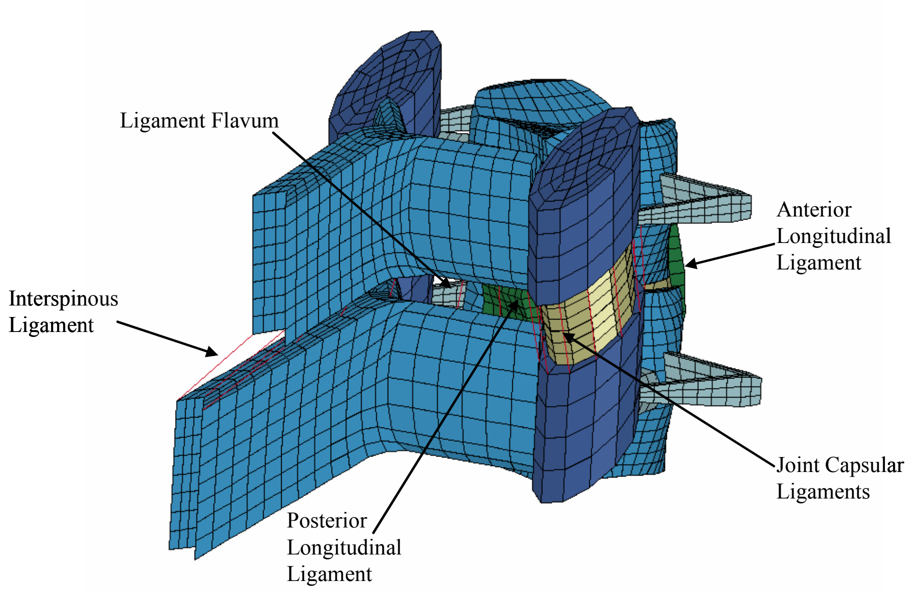


Figure 4: Motion segment model with the associated ligaments.

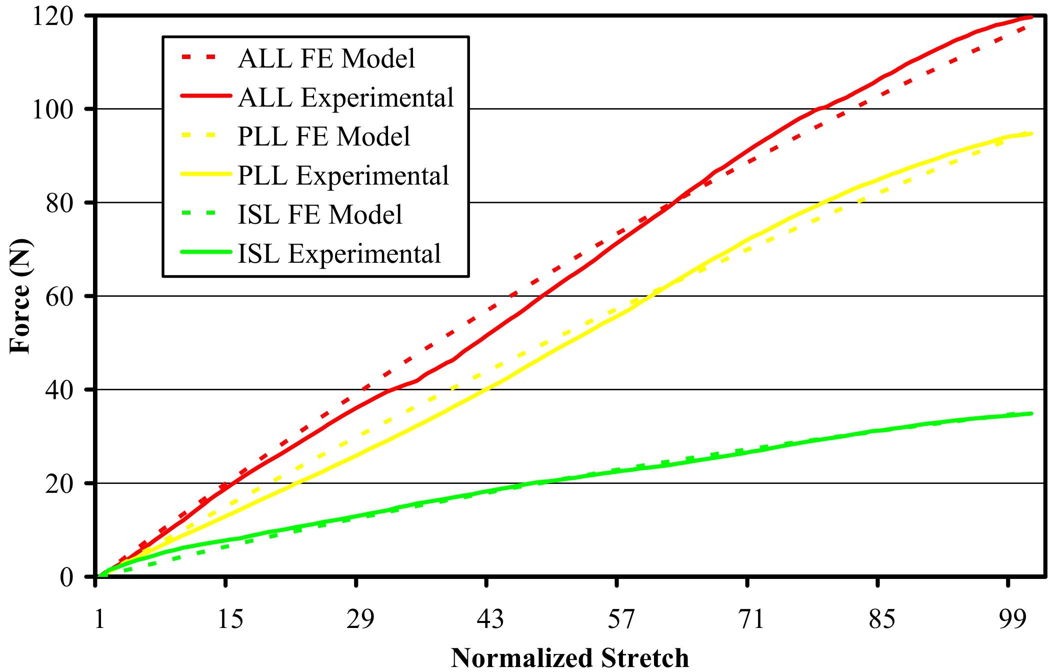


Figure 5: Ligament finite element models compared to the quasi-static experimental data for the ALL, PLL and ISL

The viscoelastic response of the ligaments was determined using dynamic relaxation data. After fitting the relaxation time constants to the mean results of the relaxation data, a Monte Carlo uncertainty analysis with 10,000 samples was performed to confirm that the model represented the mean and the variation observed in the experimental data. The QLV material model was fit to each relaxation experiment and the coefficient of variation (COV) for the set of experiments was calculated for each of the four time constants. Each relaxation time parameter in the QLV model was considered random and assigned the experimental COV.

The results, shown in Figures 6 and 7, indicate that the probabilistic ligament model simulates both the mean and variation of the ALL and PLL relaxation response well (p-values well above 0.05). The cumulative distribution function (CDF) of the peak force response of the experimental data and the uncertainty analysis were also computed and compared, shown in Figure 8 and 9. The CDF of the peak force is another feature for judging the accuracy of the FE model. The results show that the ligament model predicts the mean peak force and variation well.

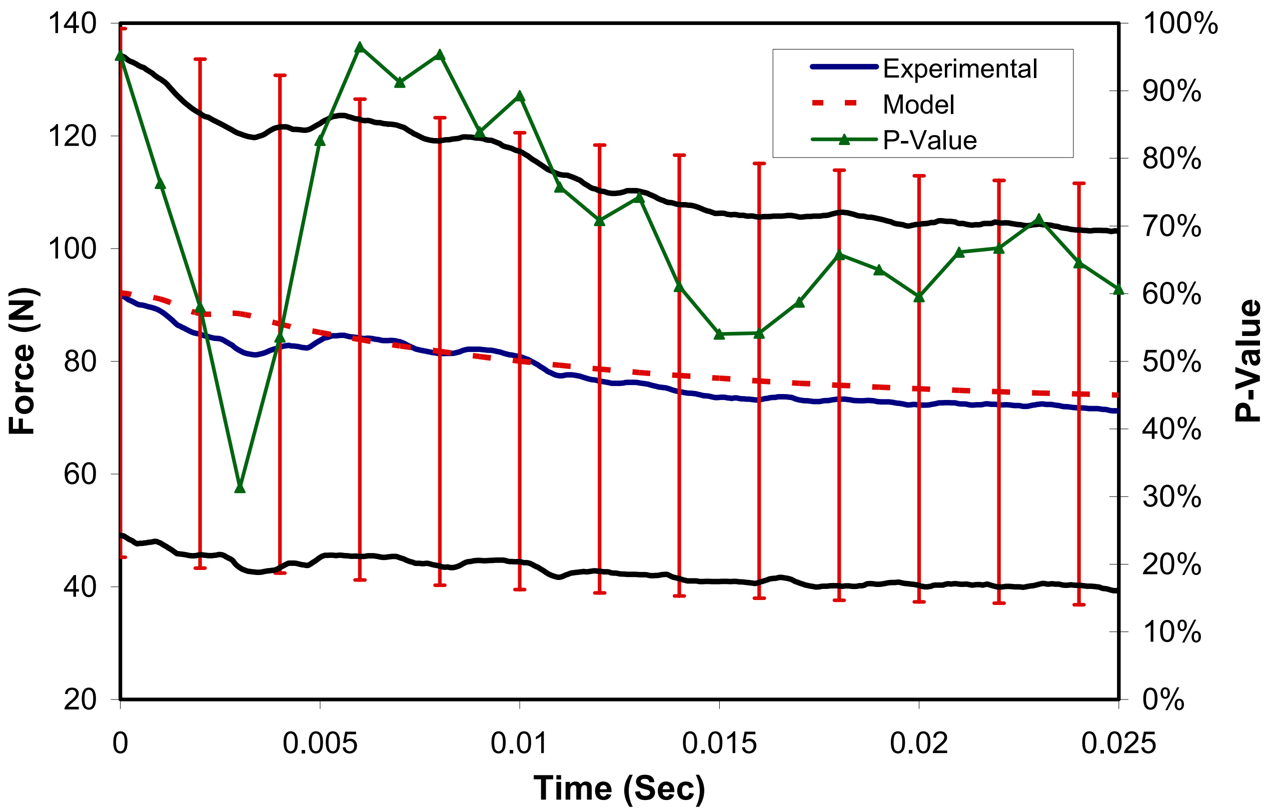


Figure 6: The experimental mean and one standard deviation corridors for the relaxation response of the ALL is shown with the mean and one standard deviation corridors for the probabilistic model response of the ALL

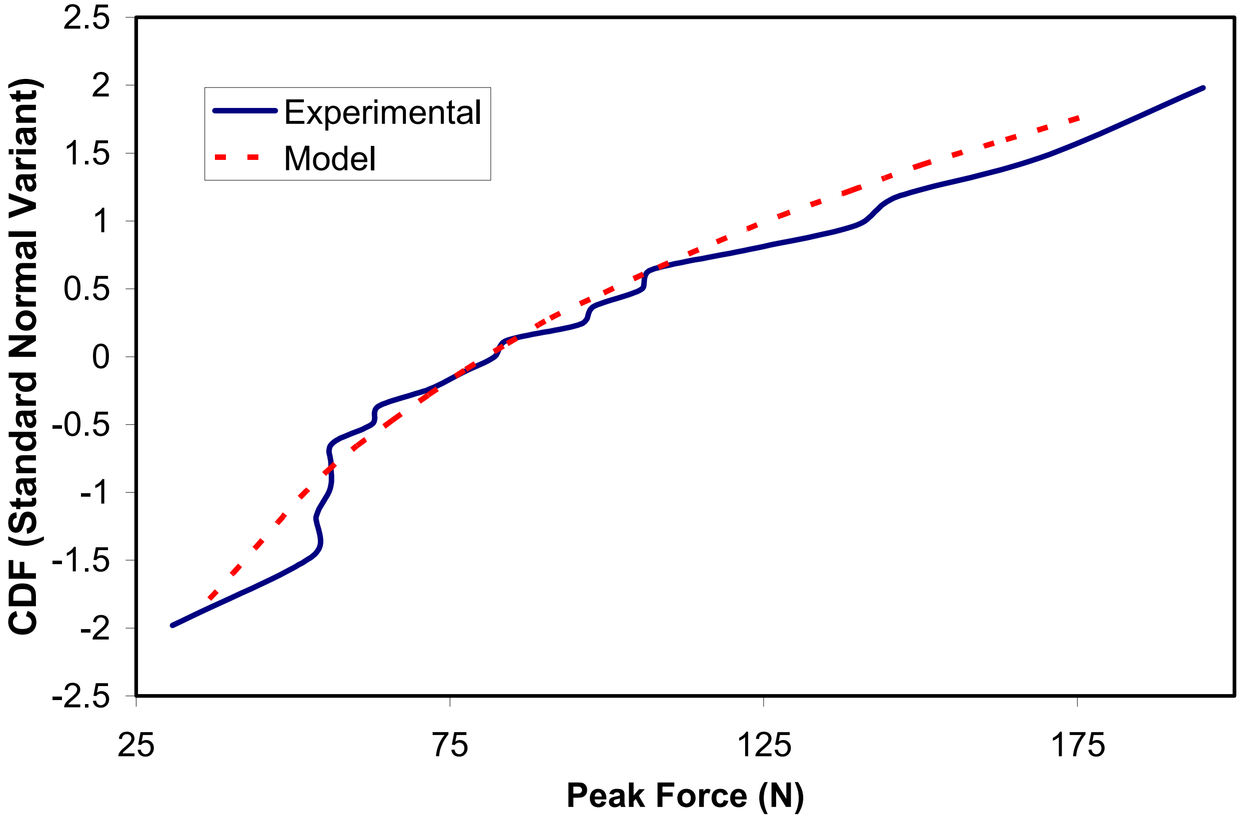


Figure 7: The experimental and model cumulative distribution functions of the maximum force during the relaxation test and simulation of the ALL

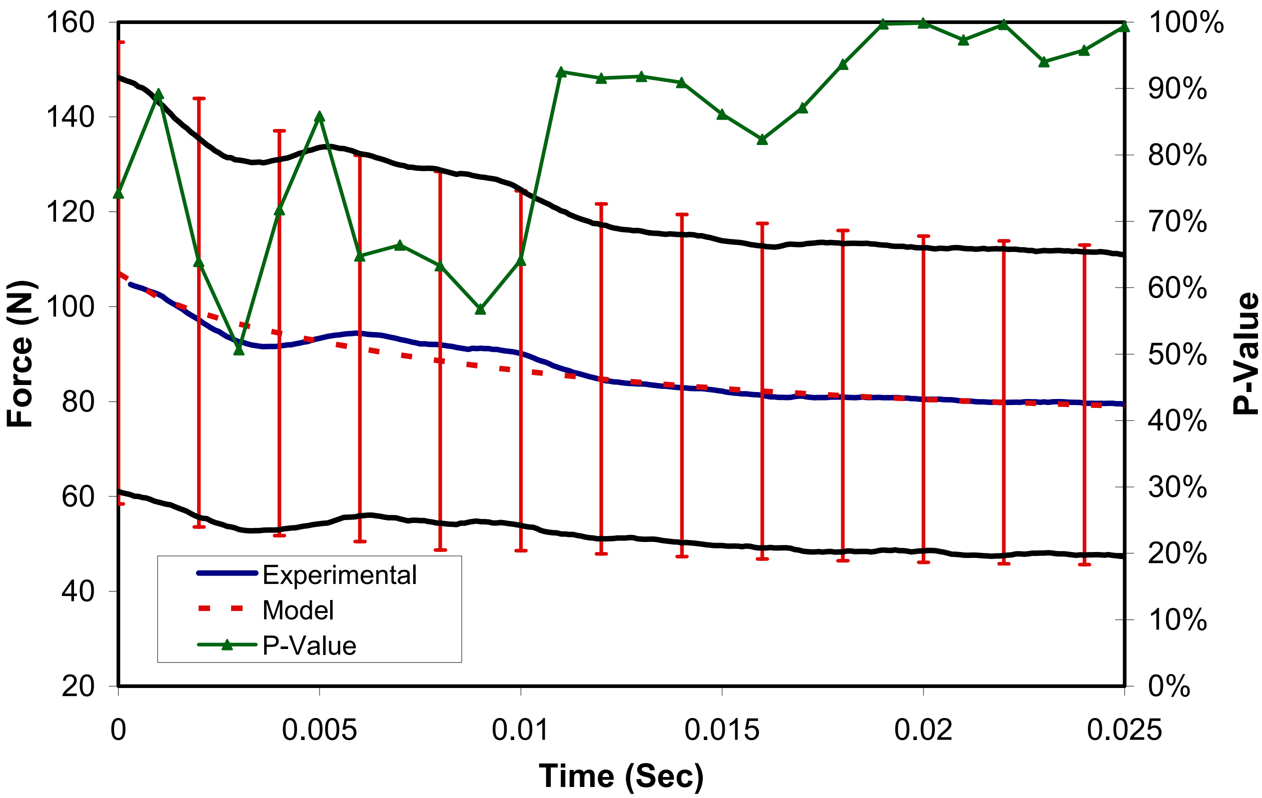


Figure 8: The experimental mean and one standard deviation corridors for the relaxation response of the PLL is shown with the mean and one standard deviation corridors for the probabilistic model response of the PLL

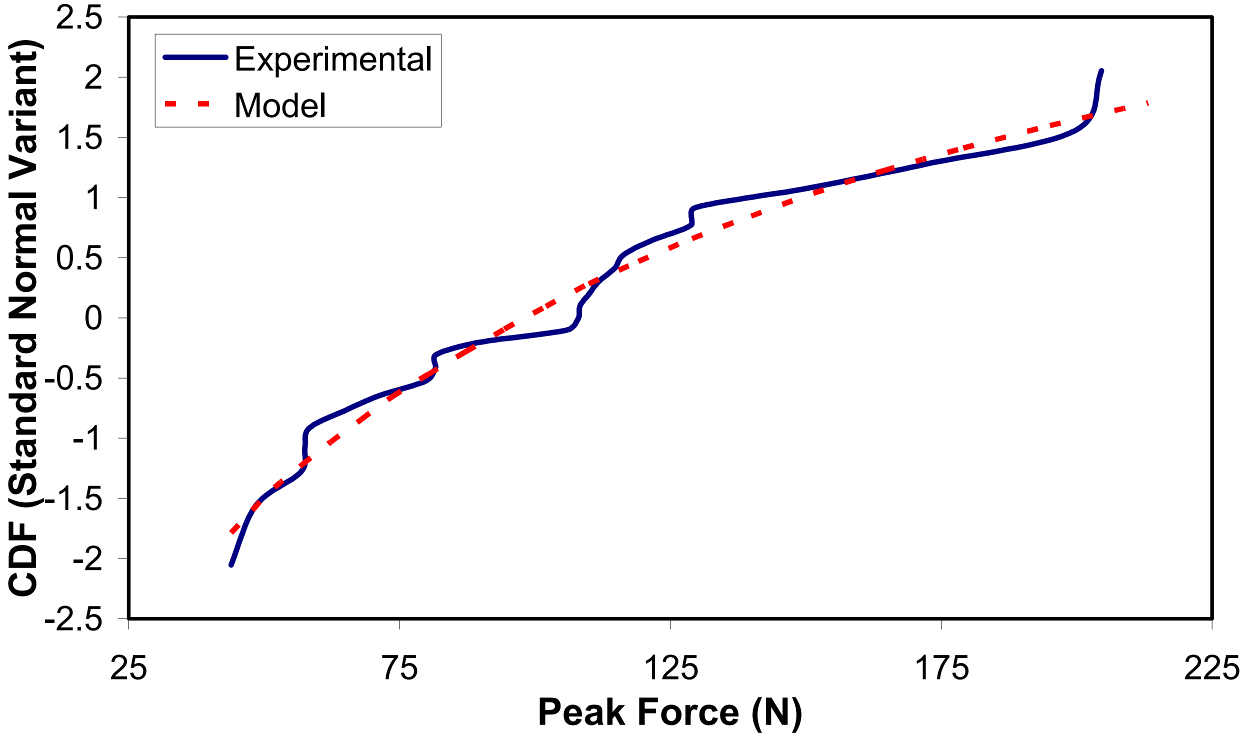
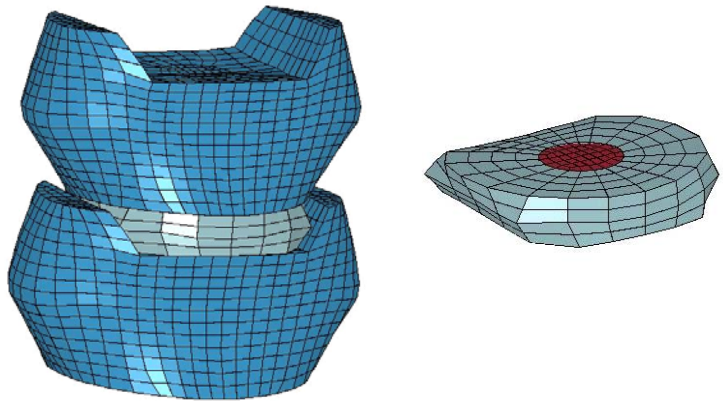


Figure 9: The experimental and model cumulative distribution functions of the maximum force during the relaxation test and simulation of the PLL

3.2 Intervertebral Disc Calibration

The geometry of the intervertebral disk (IVD) is based on a simplified representation of an anatomical disk. The model is created entirely of brick elements and is separated into two parts; the annulus and the nucleus, shown on the right in Figure 10. An incompressible fluid material model is used for the nucleus; the annulus is modeled using the same soft tissue model as the ligaments.

Published values [7] for the nucleus and annulus were used as a starting point for the material properties of the disk. A suite of experimental data generated by the Medical College of Wisconsin (MCW) was used to calibrate the model. For the IVD tension and compression experiments, a cadaveric disc was isolated with only the upper and lower vertebral bodies attached. With the vertebral bodies potted and constrained, quasi-static displacement controlled tests were performed in tension and compression. The experiment was simulated using the finite element model shown in Figure 10 in LS-DYNA such that the disk material properties could be calibrated to the test data. The posterior surface of the C6 vertebral body was fully constrained while the anterior surface of the C5 vertebral body had a displacement applied.



(a) (b)

Figure 10: (a) Intervetrebral disk FE model with attached upper and lower C5 and C6 vertebral bodies. (b) Intervertebral disk FE model

Starting with the component level shown in the schematic of the hierarchical method in Figure 3, a viscoelastic material model was fit to both the static (Figure 5) and dynamic relaxation (Figures 6 and 8) experiments. As a check, we predicted the variance on the response as well (Figures 6 and 8). The CDF for the ALL and PLL probabilistic model predicted that the maximum force during the relaxation tests are statistically equivalent to the experimental CDF (Figures 7 and 9). Using the same material modeling methodology, the disc model accurately predicted the mean loading and unloading force response in both compression and tension.

While the IVD tension and compression tests allowed determination of the IVD material model values for the bulk modulus and the viscoelastic parameters for the long time constants (quasi-static behavior), the model also needed material properties for dynamic loading. Therefore, high speed indentation tests were conducted on isolated discs and QLV material model parameters were determined resulting in experimentally derived values for the G_2 and G_4 viscoelastic relaxation parameters [5] The results from the quasi-static tension compression and the high speed indentation tests were combined to derive the material parameters for the IVD.

For the quasi-static IVD response, the results of the tension and compression experiments for both a female and male C5-C6 vertebral body/IVD complex was compared to the model response. The bulk modulus and viscoelastic parameters of the disc annulus were calibrated to give the best fit to the male and female tension and compression tests. The results show that the FE model captures the hysteresis response of the experimental IVD in both compression and tension (Figures 11 and 12).

3.3 Motion Segment Validation

After calibrating the ligaments and IVD to experimental data, the models were integrated into several motion segment models (e.g., Figure 13). Sliding contact surfaces were used for the facet joints and between the spinous processes. Validation of the model was performed in quasi-static flexion, extension, lateral bending, and axial rotation; the results were then compared to experimental data [8].

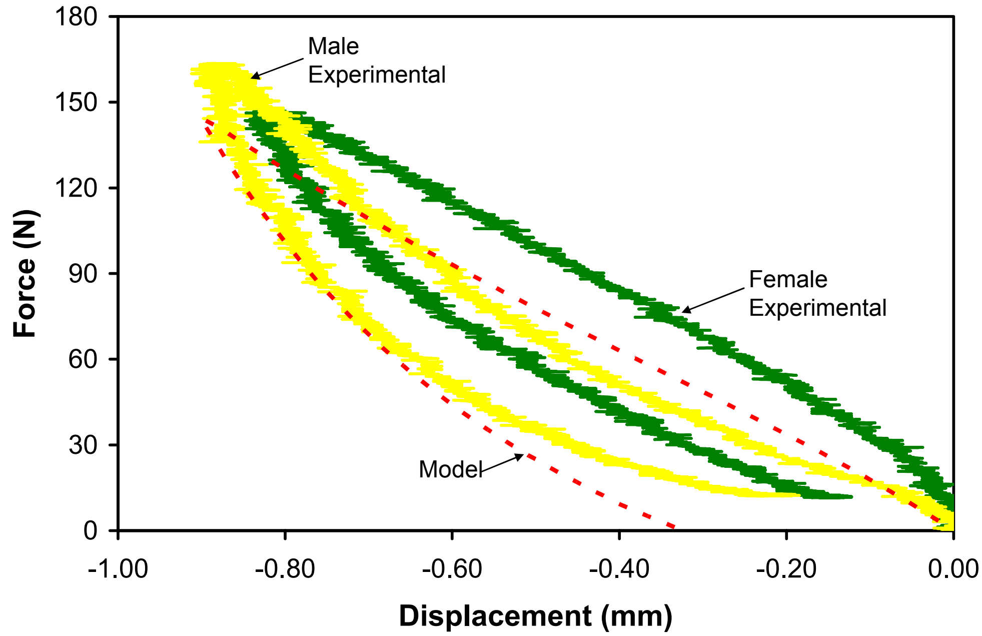


Figure 11: Disc finite element model compared to the experimental data for the disc in compression between two vertebral bodies

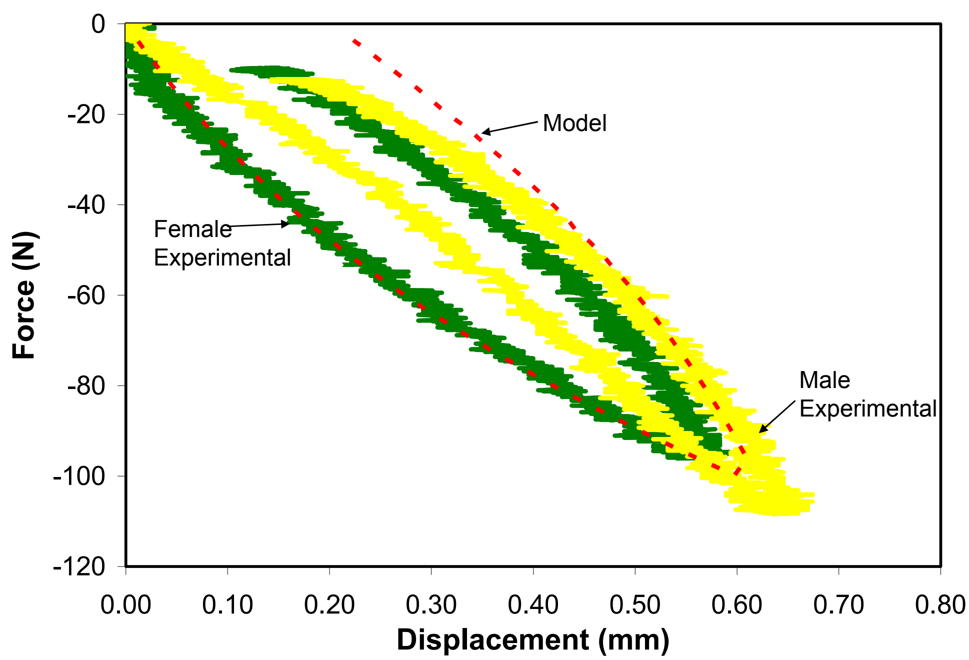


Figure 12: Disc finite element model compared to the experimental data for the disc in tension between two vertebral bodies

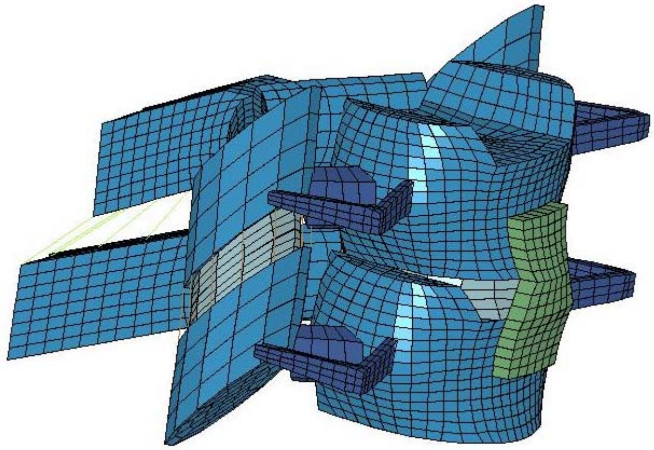


Figure 13: Small female C5-C6 motion segment finite element model

For the motion segment validations, the material parameters values determined from the component and meso-component level analysis were not changed. The point must be emphasized in order to draw a distinction between the hierarchical method of model development and validation and those traditionally used in this field. This method does not allow for "tuning" of the models once they have be incorporated into the motion segment level and complete FE model. If one changes model values in order to achieve a particular model response then confidence in the ability of that model to predict a wide range of response is lost.

The model boundary conditions were defined to match the experimental constraints and boundary conditions as closely as possible. To exercise the model in flexion, extension, lateral bending and rotation a set of massless rigid beams were connected to the superior surfaces of the upper vertebral body and spinous process. A 2 N-m moment was applied by imposing a 10 N force to each end of the beam at a distance of 0.1 m from the center of the posterior edge of the upper vertebral body. The sign of the force and beam orientation was changed in order to achieve the desired applied bending moment. For all modes of loading, the inferior surface of the lower vertebral body was fully constrained. Throughout the quasi-static application of the moment to the motion segment, the rotation of the upper vertebra was monitored resulting in a complete response curve that was then compared to the experimental results.

A probabilistic analysis of the C3-C4, C4-C5, and C5-C6 motion segments was performed using the parameter statistics that were generated for the ligaments material properties, ligament areas, and disc material properties. For each motion segment, gender and weight group, 100 Latin Hypercube Samples (LHS) were used to determine the mean response and variation in quasi-static axial rotation, lateral bending, flexion and extension. LHS was used instead of Monte Carlo because it is generally more efficient than Monte Carlo for small sample sizes. Lognormal distributions were used for all variables and a 10% COV was assumed for ISL area and the joint capsule ligament area.

The results for flexion and extension model simulations of the C5-C6 motion segment are shown in Figure 14. Good correlation between model and experimental is seen between ±2 N-m of loading. The experimental one standard deviation corridors are shown. The predicted COV, illustrated by the one standard deviation bars for the small female model, ranged from 40% to 50% and is consistent with all four models.

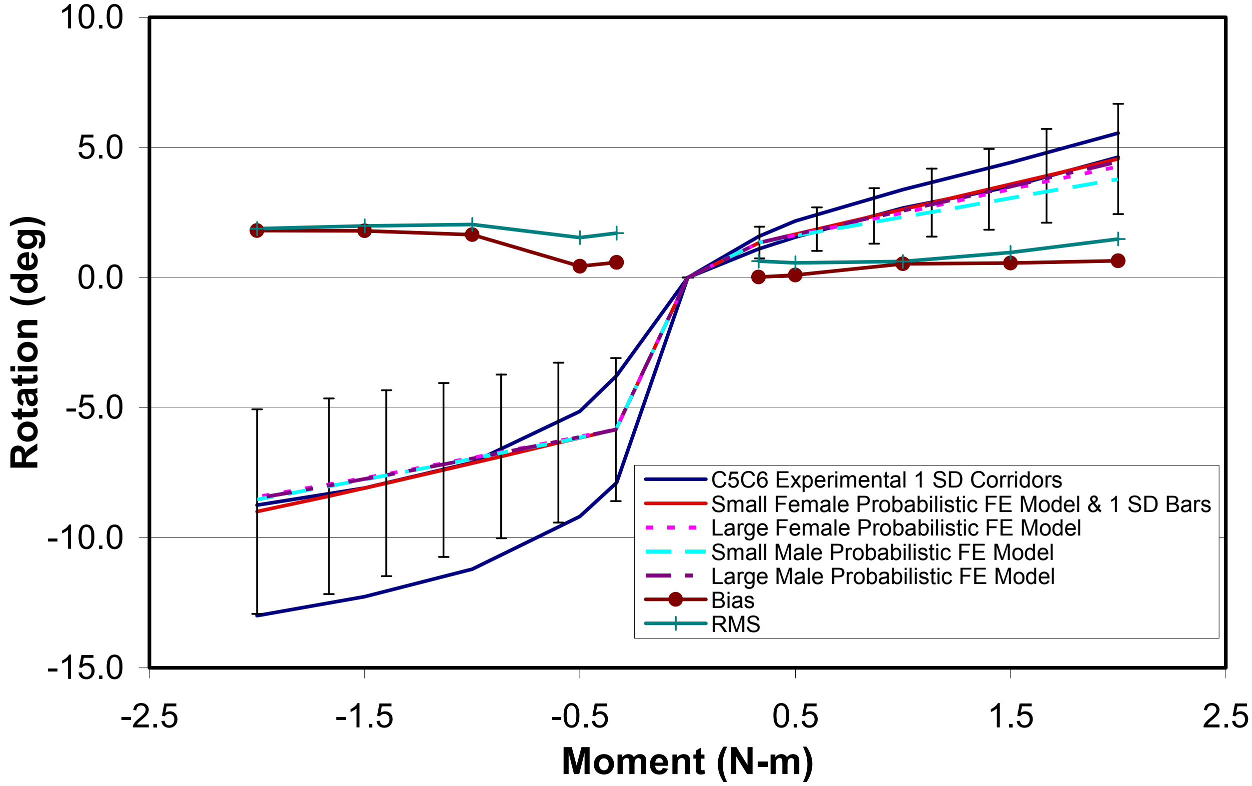


Figure 14: Probabilistic response of a C5-C6 motion segment in flexion and extension. Four models are shown representing large and small males and females. The one standard deviation corridors are shown for the experimental results as well as for the small female probabilistic FE model

The RMS and BIAS validation metrics are also shown in Figure 14. Between -2 N-m and -1 N-m the BIAS and RMS are approximately the same and equal to about 2º. This indicates that the distributions have the same shape, but are shifted from each other. Between -1 and -0.25 N-m, BIAS decreases to less than 1º but the RMS stays near 2º. This indicates that the two distributions have nearly the same mean, but slightly different shape. The same reasoning applies to the 0.25 to 2.0 N-m loading range. Overall, these metrics correctly indicate excellent agreement between the experimental and simulation uncertainty distributions.

Under axial rotation loading the computed stiffness is very similar to the experimental results and the COV is large (40% to 50%) throughout the range of motion (Figure 15). In this comparison the t-test metric is used for demonstrative purposes. The high p-values correctly reflect the excellent agreement between experiment and simulation.

Model predicted results for lateral bending, shown in Figure 16, also show good correlation to experimental data in stiffness in the -2 to 2 N-m loading range as well as 40% to 50% COV for all four models (only small female standard deviation is shown). In this case, however, the p-values are much lower for the higher loading magnitudes. The decrease in level of agreement can be visually seen by comparing the experimental corridors to the model one standard deviation (vertical) bars.

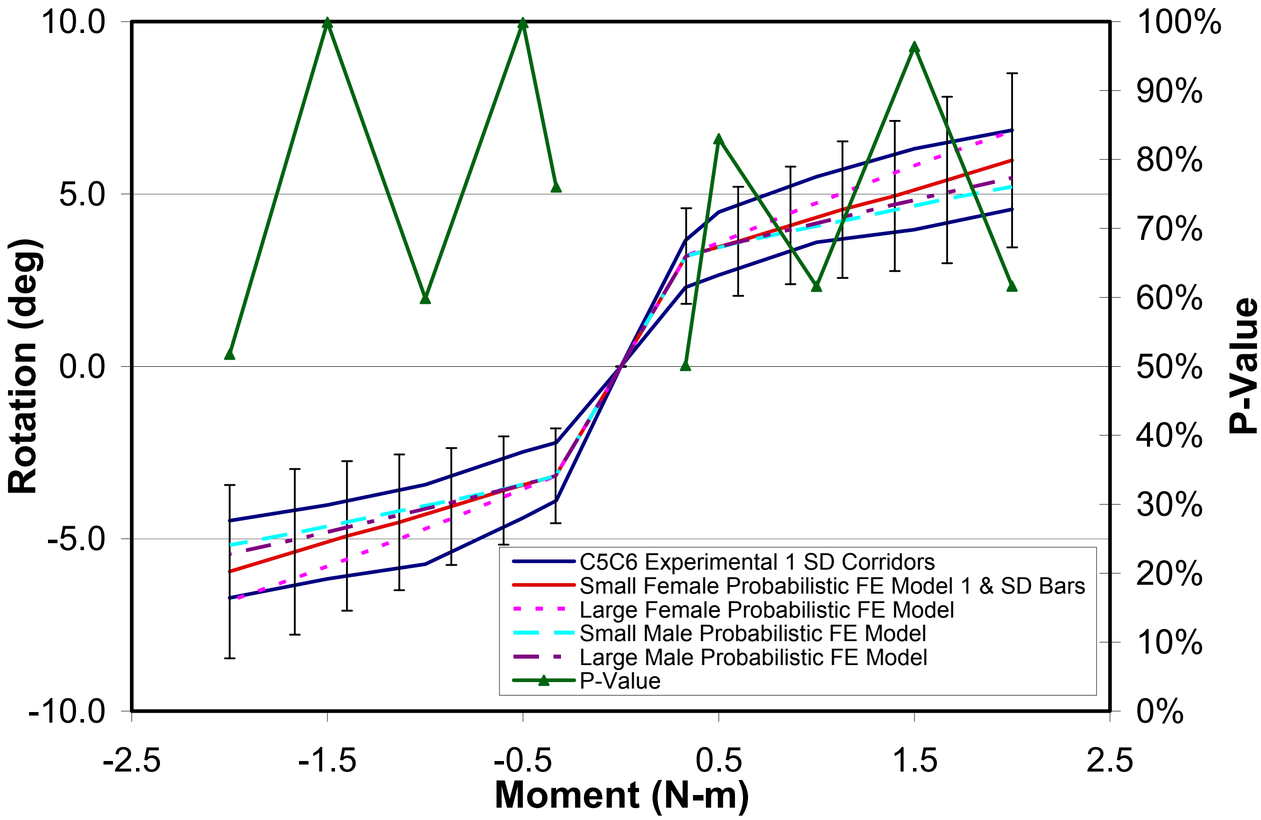


Figure 15: Probabilistic response of a C5-C6 motion segment in axial rotation. Four models are shown representing large and small males and females. The one standard deviation corridors are shown for the experimental results as well as for the small female probabilistic FE model

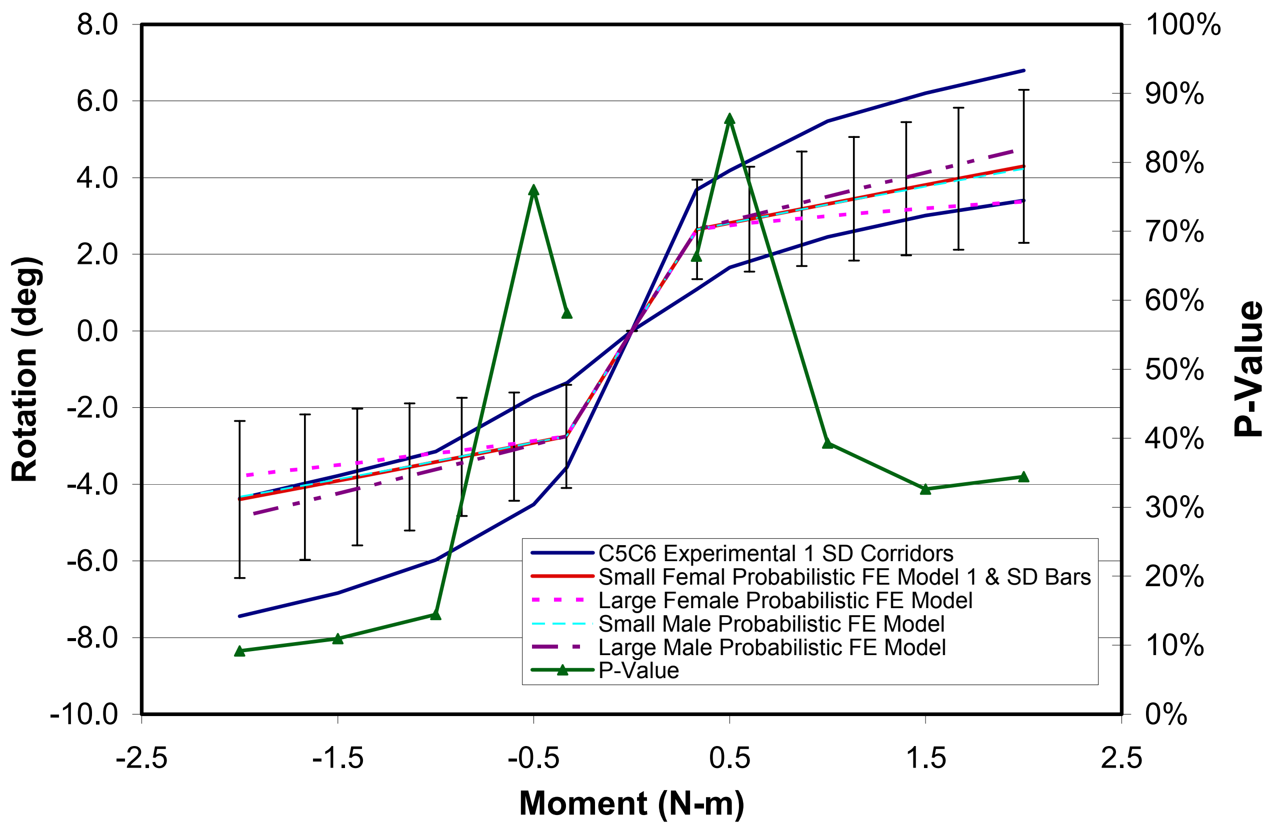


Figure 16: Probabilistic response of a C5-C6 motion segment in lateral bending. Four models are shown representing large and small males and females. The one standard deviation corridors are shown for the experimental results as well as for the small female probabilistic FE model

The results shown in Figures 14 thru 16 represent four separate FE models. Differences in the responses are due to the variations in the geometry of the four gender and weight groups that the models were based on. Figure 14 shows the flexion and extension of a C5-C6 motion segment under a range of moment loading. The degree of rotation for all four models is on the low side of the experimental corridors, however, and more importantly, the stiffness of the response is quite close to experimental values. Perhaps of most importance is that the 100 sample Latin hypercube analysis yielded a predicted variance that closely matches the experimental variance.

Figures 15 and 16 are the results of the axial and lateral bending respectively. These results show a larger variation between the four models. This is due to the fact that the joint facet angles play a large role in the degree of lateral bending and axial rotation. Therefore, the differences in responses can be correlated to the different joint facet angles in the models. Again, the results for the axial (Figure 15) and lateral bending (Figure 16) illustrate the ability of the model to predict the variance of the response.

3.4 Cervical Column (C3-T1) Validation

The three motion segments were combined to create a C3-C7 model representing the small female group. The T1 vertebra was added along with the C7-T1 disc and ligaments resulting in the model shown in Figure 17. To validate the C3-T1 cervical column, the inferior surface of the T1 vertebra was fully constrained and a 2 N-m moment was applied to superior surface of C3 using the same technique used for the motion segments described earlier. The model was exercised in flexion, extension, axial rotation and lateral bending. The rotation of the each vertebra was monitored throughout the application of the 2 N-m moments.

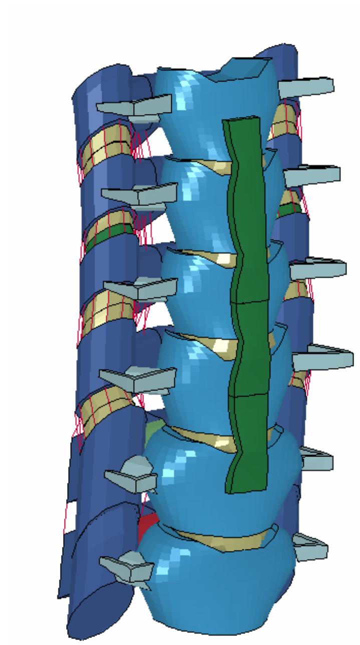


Figure 17: C3-T1 finite element model

Experimental data from Ref. [8] was used to validate the results of the model. The results for the 2 N-m moment quasi-static flexion, extension, lateral bending and axial rotation model simulations for the C3 rotation with respect to T1 are shown in Figures 18-21. Figure 18 shows the rotation under extension loading including one standard deviation corridors. Also shown is the mean response for the experimental data along with the mean response of the FE model. Figure 19 shows the flexion results while Figures 20 and 21 show the results for lateral bending and axial rotation respectively. These results are for the small female model only; other groups were similar. A probabilistic analysis was not performed for this set of simulations like those performed on the motion segments.

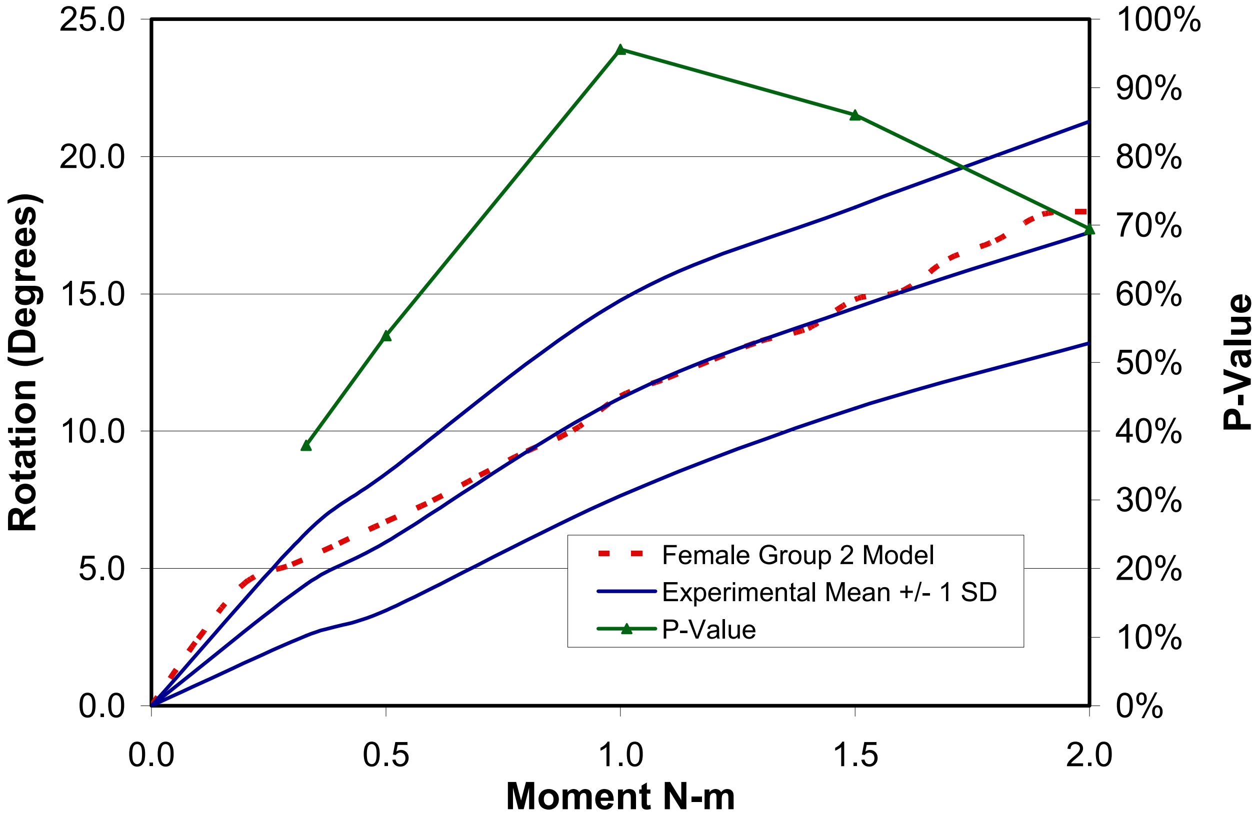


Figure 18: C3 extension rotation with respect to T1 mean value model response with one standard deviation corridors and mean experimental results

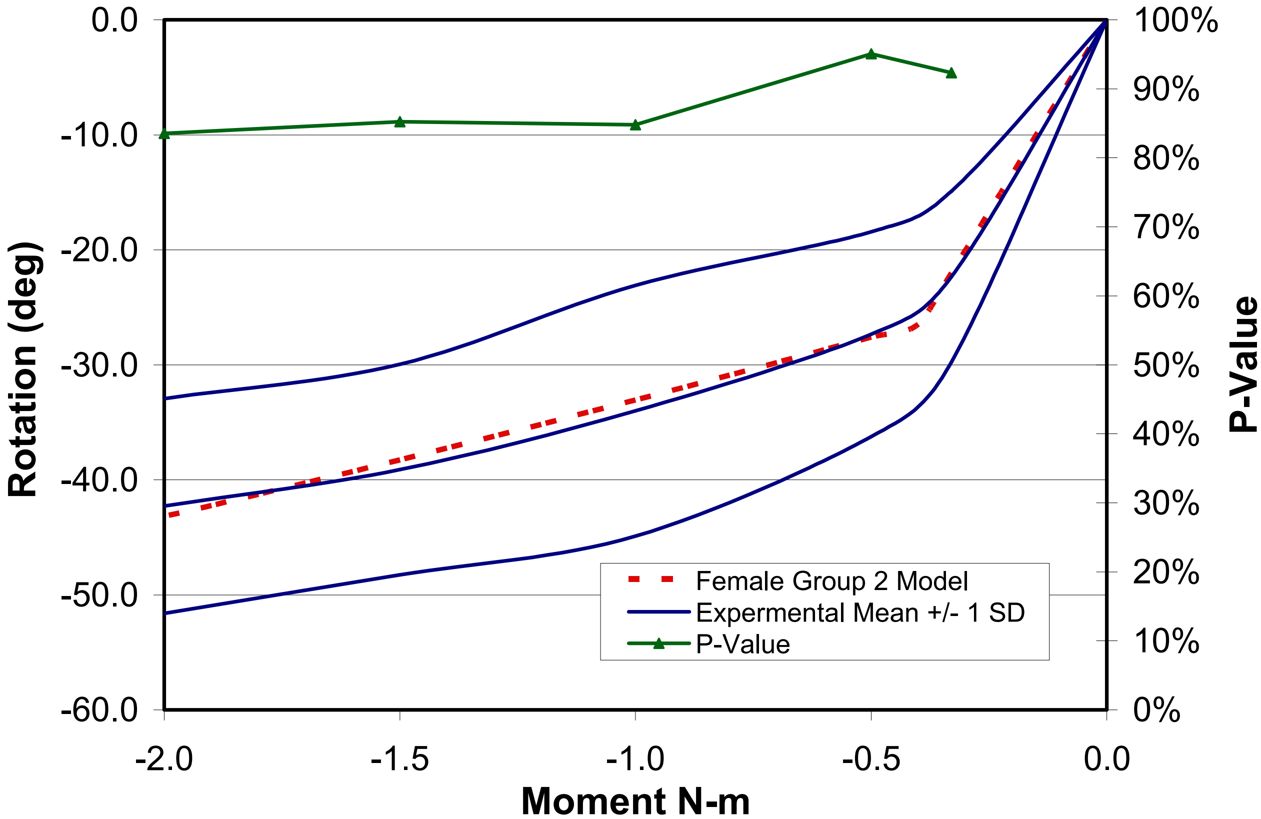


Figure 19: C3 flexion rotation with respect to T1 mean value model response with one standard deviation corridors and mean experimental results

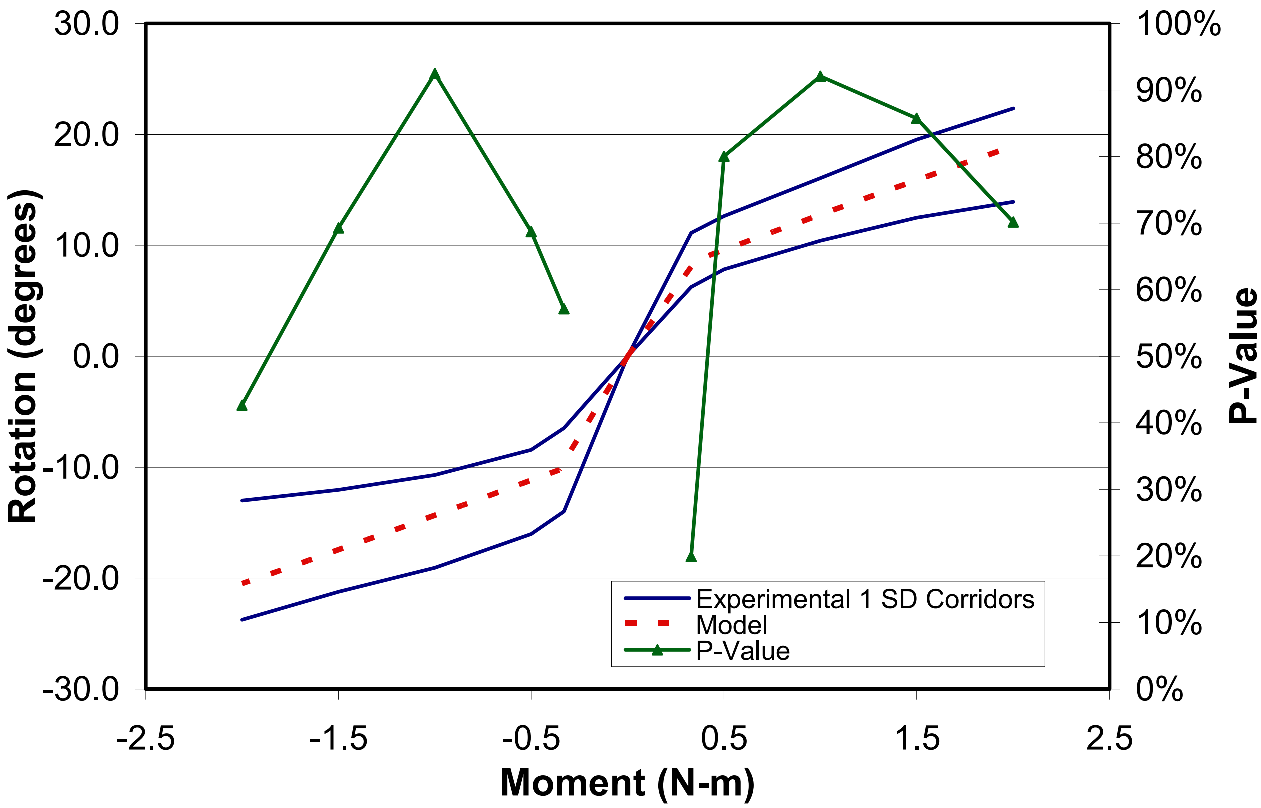


Figure 20: C3 lateral bending rotation with respect to T1 mean value model response with one standard deviation corridors and mean experimental results

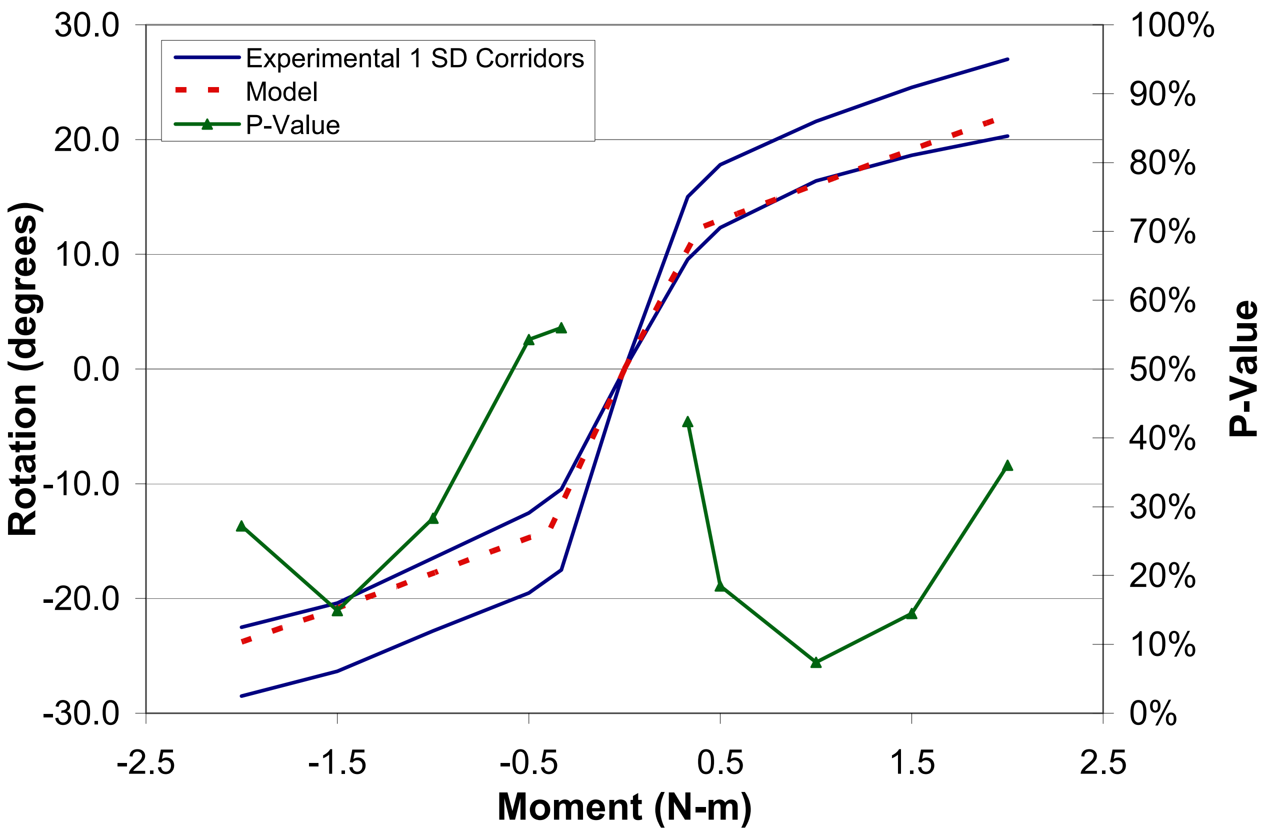


Figure 21: C3 axial rotation with respect to T1 mean value model response with one standard deviation corridors and mean value experimental results

3.5 Full Model (Skull –T1) Validation

A FE representation of the atlas and axis was created from the CT scans and combined with the small female C3-T1 finite element model. A model of the skull was also created using rigid shell elements and added to the model. Stiffness values for the ligaments of the upper cervical spine complex were obtained from Ref. [9]. All ligaments were modeled using spring elements with the exception of the transverse ligament which was modeled using 4 node shell elements. Sliding contact surfaces were used for the atlanto-occipital joint, the facet joints and between the atlas and axis. Head mass and center of gravity (CG) were taken from the literature [10]; the inertia properties were taken from Ref. [11].

Hill type muscles were included in the full cervical spine model (Figure 22). The muscle origin and insertion points as well as muscle physiological cross section area (PSCA) were taken from the literature [12]. For the muscle activation levels and timing, schemes were adapted from Ref. [13]. Peak muscle forces were determined by assuming a peak muscle stress of 50 N/mm^2 [14] and scaling by the PCSA for each muscle group.

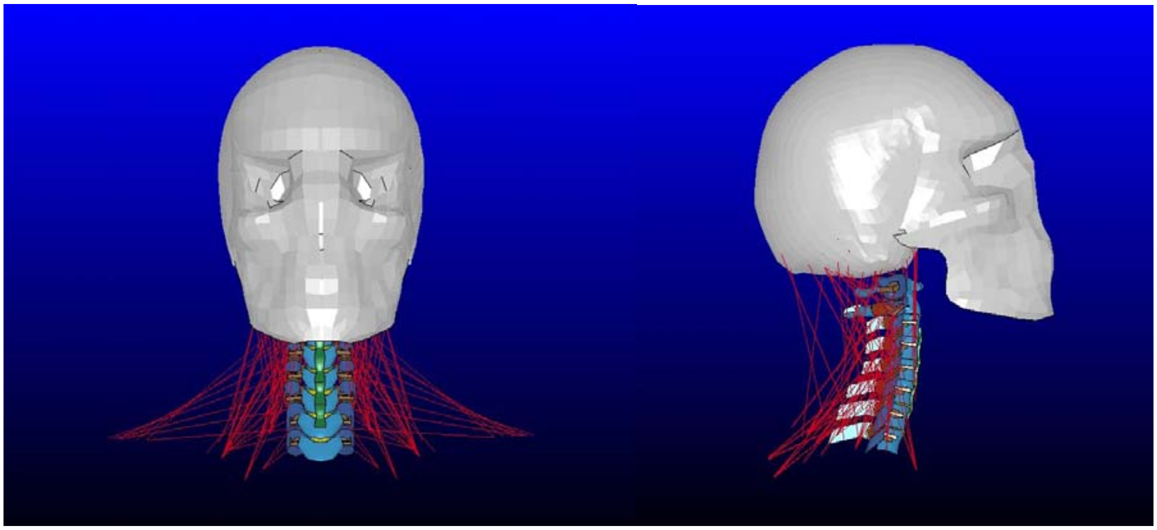


Figure 22: Full cervical column model from the skull to T1 with muscles

Unpublished data from the Naval Biodynamics Laboratory (NBDL) was used to validate the model. Data was provided by NBDL of low G ejection experiments on 12 female subjects. The T1 displacements were recorded during these experiments along with the subject head rotation. The recorded displacements were used to drive the FE model while the skull rotation was monitored at regular intervals. The model head rotation was then compared to the calculated one standard deviation corridors from the female subject tests for +5Gz and +7Gz tests. A 100 sample LHS probabilistic analysis was performed on the model to calculate the variance of the response throughout the ejection simulation.

The results for the small female model of the full cervical spine from the skull to T1 under both +5Gz loading and +7Gz loading is shown in Figures 23 and 24 respectively. These results also include one standard deviation corridors for head rotation versus time for 12 female subject experiments along with the probabilistic FE model mean response and one standard deviation bars. Figure 23 shows that the model response is within the one standard deviation corridors of the experimental. However, the model predicts a variance that is significantly larger than the experimental corridors. The +7Gz results shown in Figure 22 more closely match the mean of the experimental head rotation and the probabilistic model variance prediction is much closer to the experimental than the +5Gz simulation. It would be expected that as the applied G level increases, the variance of the subject's response would also increase.

The p-values shown in both Figures 23 and 24 are significantly lower than the previous validation cases. In several cases the p-values are close to zero, indicating the means are not equal. The BIAS and RMS metrics would be a more useful comparison, and will be implemented when additional experimental data is available.

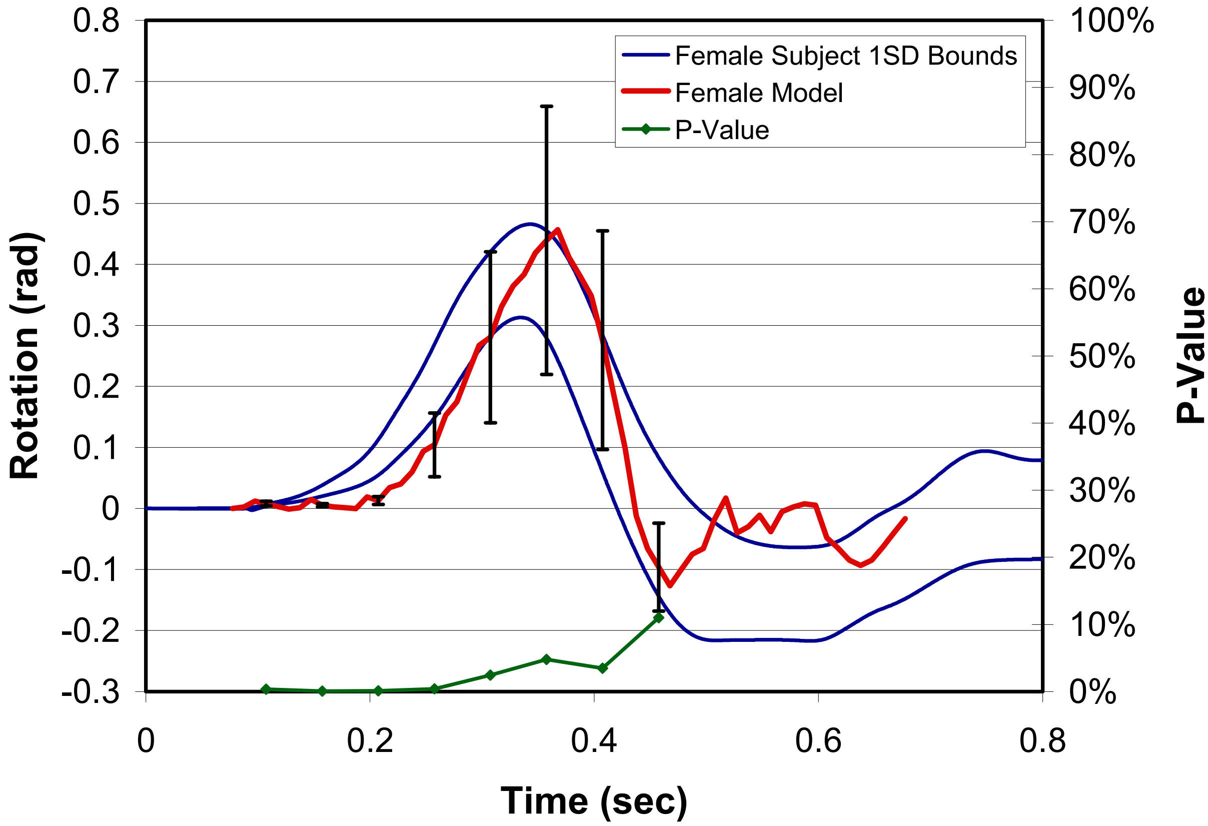


Figure 23: +5Gz ejection simulation head rotation for the full cervical spine probabilistic FE model. The one standard deviation corridors are shown for the experimental results as well as the probabilistic model response

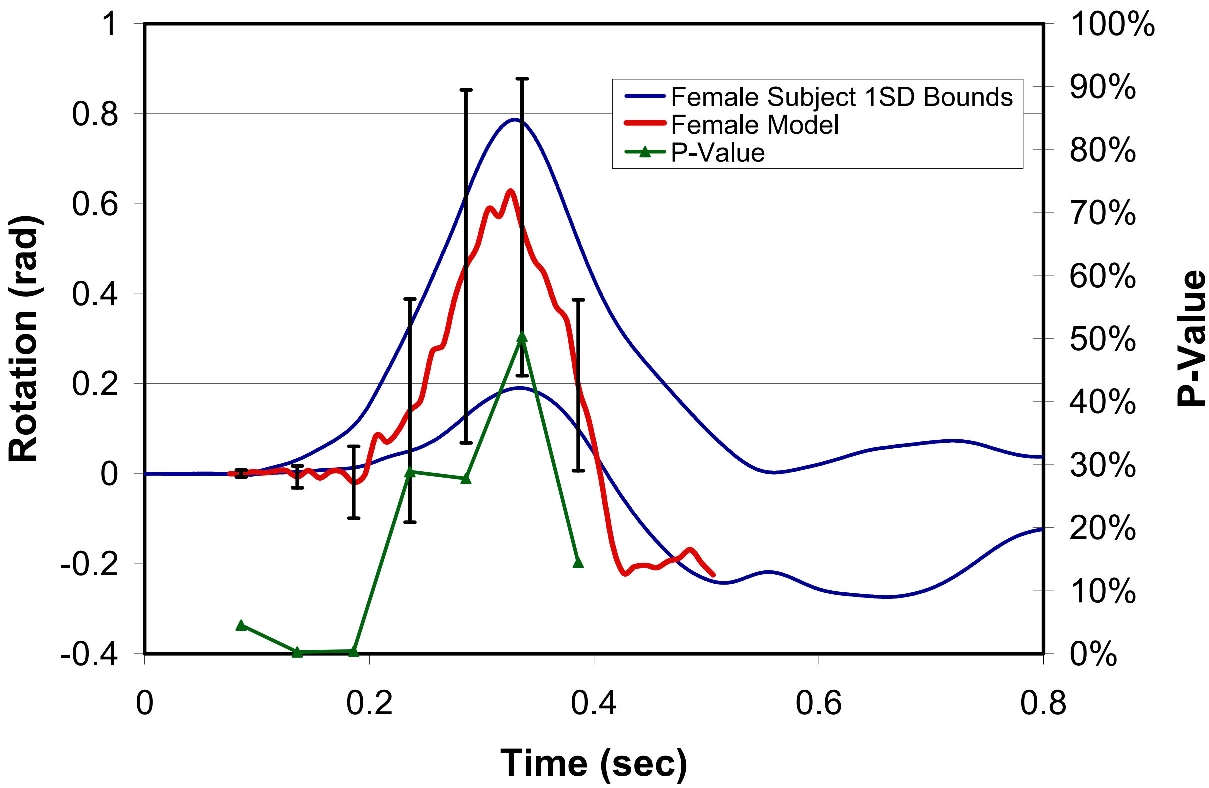


Figure 24: +7Gz ejection simulation head rotation for the full cervical spine probabilistic FE model. The one standard deviation corridors are shown for the experimental results as well as the probabilistic model response

4.0 SUMMARY AND CONCLUSIONS

A hierarchical model development and validation process was performed to create a predictive tool that could be used to quantify the risk of injury to military aviators. Fundamental model parameters such as the ligament and IVD material model constants were calibrated to quasi-static and dynamic experiment data. Once calibrated, these model input parameters were fixed and the FE model predicted responses at increasing levels of complexity were compared to experimental data and showed excellent agreement.

The goal of model V&V is to quantify confidence in the predictive capability of the model. The approach to validation assessment is to measure the agreement between model predictions and experimental data from appropriately designed and conducted experiments. Agreement is measured, for example, by quantifying the difference (error) between the experimental data and the model output. Uncertainty in both model output and experimental data will confound estimate of the error. Consequently, agreement should be expressed as a statistical statement, for example, as the expected error with associated confidence limits.

As demonstrated in this paper, uncertainty quantification plays a key role in model V&V. Nondeterminism refers to the existence of errors and uncertainties in the outputs of computational simulations due to inherent and/or subjective uncertainties in the model. Likewise, the measurements that are made to validate these simulation outputs also contain errors and uncertainties. This paper presented the main concepts involved in model V&V with specific focus on the critical role of hierarchical model development and uncertainty quantification via nondeterministic analysis.

There are many open issues in the definition and practice of model V&V. From a practical standpoint, the up-front costs associated with conducting a high quality V&V program will certainly be formidable. Therefore, the long-term benefits of using a model to supplement testing must be balanced against the model development and V&V costs. It is certain that V&V will remain an issue of high importance in the field of computational mechanics, and further developments are needed to develop recommended practices for performing model V&V.

5.0 ACKNOWLEDGMENTS

Development and validation of the cervical spine injury model was supported by the Naval Air Warfare Center Aircraft Division (NAWCAD). Dr. Barry Shender serves as the technical monitor.

6.0 REFERENCES

[1] American Society of Mechanical Engineers (2006), *Guide for Verification and Validation of Computational Solid Mechanics*, ASME V&V 10-2006, New York, NY.

[2] Nicolella, D., Francis, W., Bonivtch, A., Thacker, B., Paskoff, G., & Shender, B. (2006) *Development, verification, and validation of a parametric cervical spine injury prediction model.* 47th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference.

[3] Weiss, J., Maker, B., & Govindjee, S., (1996) *Finite element implementation of incompressible transversely isotropic hyperelasticity*. Comp. Meth. Appl. Mech. Eng. 135: 107-128.

[4] Yoganandan, N., Kumaresan, S., & Pintar, F. (2000) *Geometric and mechanical properties of human cervical spine ligaments*. Journal of Biomechanical Engineering 122: 623-629

[5] Lucas, S., Bass, C., Salzar, R., Shender, B., & Paskoff, G.. (2006) *High rate viscoelastic properties of human cervical spinal intervertebral discs.* American Society of Biomechanics Annual Meeting. Blacksburg, VA.

[6] Lucas, S. (2004) *Viscoelastic characterization of cervical spinal ligaments.* MS Thesis, Biomedical Engineering, University of Virginia.

[7] Maurel, N., Lavaste, F., & Skalli, W. (1997) *A three-dimensional parameterized finite element model of the lower cervical spine study of the influence of the posterior articular facets.* Journal of Biomechanics 30(9): 921-931

[8] Wheeldon, A., Pintar, A., Knowles, S., & Yoganandan, N. (2006) *Experimental flexion/extension data corridors for validation of finite element models of the young, normal cervical spine,* Journal of Biomechanics 39(2): 375-380.

[9] Halldin, P., Brolin, K., Kleiven, S., von Holst, H., Jakobsson, L., & Palmertz, C. (2000) *Investigation of conditions that affect neck compression-flexion injuries using numberical techniques.* 44th Stapp Car Crash Conference Proceedings

[10] Plaga, J., Albery, C., Boehmer, M., Goodyear, C., & Thomas, G. (2005) *Design and development of anthropometrically correct head forms for joint strike fighter ejection seat testing.* US-AFRL Report AFRL-HE-WP-TR-2005-0044.

[11] Becker, E. (1972) *Measurement of mass distribution parameters of anatomical segments.* SAE Number 720964.

[12] Chancey, V., Nightingale, R., Van Ee, C., Knaub, K., & Myers, B. (2003) *Improved estimation of the human neck tensile tolerance: reducing the range of reported tolerance using anthropometrically correct muscles and optimized physiologic initial conditions.* Stapp Car Crash Journal 47

[13] Brolin, K., & Halldin, P., (2004) *Development of a finite element model of the upper cervical spine and parameter study of ligament characteristics.* Spine 29(4): 376-385.

[14] Winters, J., & Stark, L. (1988) *Estimated mechanical properties of synergistic muscles involved in movements of a variety of human joints*. J Biomechanics 21(12): 1027-1041.

Paper No. 26

Discusser's Name: I. Celik

Question: (1) What was the contribution of discretization errors on the uncertainty bars? (2) What was the most dominant uncertainty parameter?

Author's Reply: (1) Discretization errors were not included in the plots shown. (2) The probabilistic sensitivity values were not slown, but were computed at part of our analysis. The top three parameters were head mass, head CG and initial position.

Discusser's Name: M. Hemsch

Question: How do you estimate the uncertainty for some system response quantity (SRQ) from a given simulation/prediction (e.g. the spine "breaks" for this loading scenario)?

Author's Reply: Our primary SRQ is currently "rotation". This is predicted using both Monte Carlo and Latin Hypercube simulation. The "tolerance" level is subject/scenario dependent and was not shown.