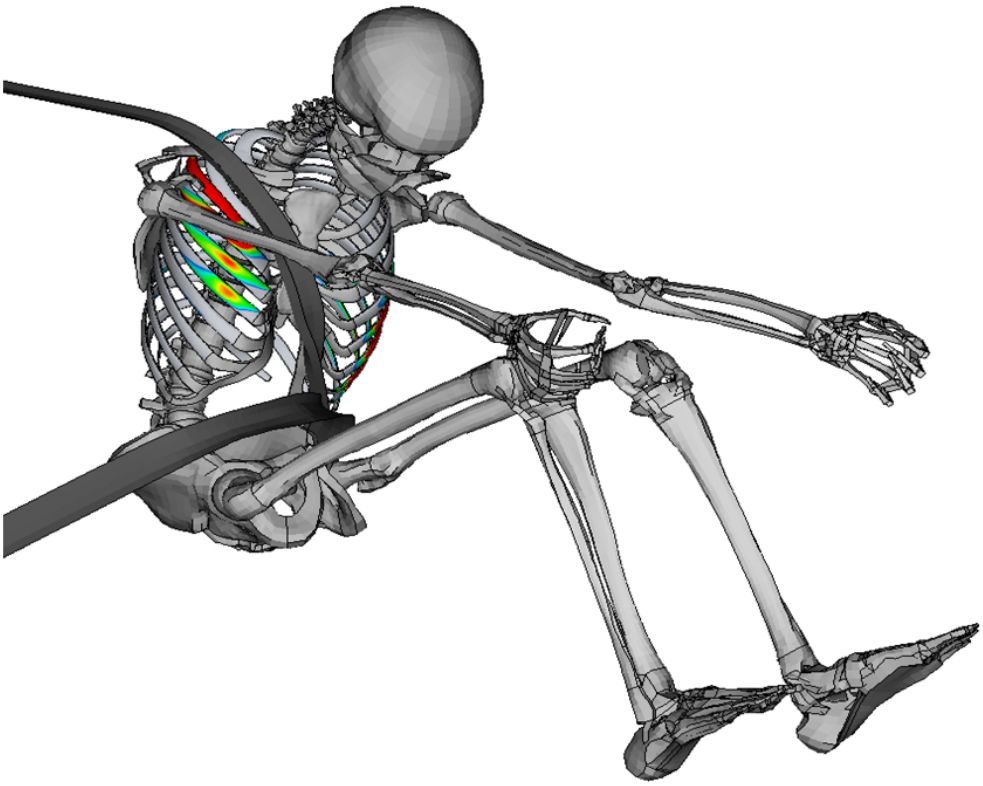
CHALMERS
UNIVERSITY OF TECHNOLOGY



Improvement of Human Body Model Rib Fracture Risk Prediction

Creation of injury risk curves and rib cortical bone regression models for age adjusted risk prediction

Master's thesis in Biomedical Engineering

Amanda Blennow

Department of Mechanics and Maritime Sciences
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2020

Master's thesis 2020:29

Improvement of Human Body Model Rib Fracture Risk Prediction

Creation of injury risk curves and rib cortical bone regression models for age adjusted risk prediction

Amanda Blennow

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Mechanics and Maritime Sciences
Division of Vehicle Safety
Chalmers University of Technology
Gothenburg, Sweden 2020

Improvement of Human Body Model Rib Fracture Risk Prediction
Creation of injury risk curves and rib cortical bone regression models for age adjusted risk prediction
AMANDA BLENNOW

Supervisor: Karl-Johan Larsson, Chalmers
Nils Lübbe, Autoliv
Examiner: Johan Davidsson, Mechanics and Maritime Sciences, Chalmers

Master's Thesis 2020:29
Department of Mechanics and Maritime Science
Division of Vehicle Safety
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Skeletal parts of a belted human body model in frontal crash simulation.

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2020

Improvement of Human Body Model Rib Fracture Risk Prediction
Creation of injury risk curves and rib cortical bone regression models for age adjusted risk prediction
Amanda Blennow
Department of Mechanics and Maritime Science
Chalmers University of Technology

Abstract

In motor vehicle collision, one of the most commonly injured body regions is the thorax and these injuries are in many cases the cause of death among belted, adult drivers. Rib fractures are particularly dangerous among elderly and to enable injury prevention, proper risk assessment is of great importance. In a recent study, the effect of age, sex and strain rate on the rib cortical bone material was investigated, enabling creation of new injury risk curves and rib cortical bone material models.The aim of this thesis was thus to improve the human body model rib fracture risk prediction by creating new rib fracture risk curves and material models of human rib cortical bone for individually adapted rib fracture risk prediction.

Using the methods of parametric survival modelling, injury risk curves for rib fracture risk prediction were developed for prediction using both ultimate strain and total strain energy density. It was found that the new strain based risk curve corresponds to a decrease in ultimate strain of 12 % per decade of life, compared to the 5 % per decade of life in the currently used risk curve. Regression models of the mechanical properties elastic modulus, plastic hardening modulus and yield stress, of human rib cortical bone were created and thus, construction of age dependent stress-strain curves was enabled. The stress-strain curves also model the individual variability, suggesting mechanical parameter values for weak, average and strong individuals of any given age.

The material models were implemented in the SAFER HBM and four post mortem human subject sled tests were reproduced and outputs were evaluated with both the newly created injury risk curves and the injury risk curves that have been used up until now. It was found that both the new injury risk curves and the material models affected the predicted injury risks and particularly among elderly. Due to the large set of data used for the development of injury risk curves, the new strain-based risk curve possesses a higher statistical validity, compared to the old strain-based risk curve, and can therefore be considered as a better alternative. The energy-based risk prediction proved to be less sensitive to the material model used compared to the strain-based risk prediction . This is promising in terms of predicting risks that are less sensitive to individual variability.

Keywords: Injury Risk Curve, Rib Fracture Risk, Risk Prediction, Regression, Rib Cortical Bone, Material Model, Survival Analysis, Human Body Model.

Acknowledgements

During my work with this thesis, I have encountered many that have helped me to turn it in to what it is today. First of all, i would like to thank my supervisors. Thank you, Karl-Johan Larsson for all the time you spent on me and this thesis, thank you for endless discussions on the smallest of details and for all the support. Nils Lübbe, thank you for sharing your knowledge and expertise with me and for pushing me beyond my comfort zone.

Furthermore, I would like to thank Andras Balint for taking the time to discuss statistics with me and helping me understand some of the necessary methods and approaches of this thesis. Lastly, I am very grateful for the warm welcome and including atmosphere from all employees at Autoliv that I have met during my work with my thesis.

Amanda Blennow, Gothenburg, June 2020

Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

AFT Accelerated Failure Time

AIC Akaike's Information Criteria

AIS Abbreviated Injury Scale

ANOVA Analysis of Variance

AUROC Area Under Receiver Operating Characteristics

CV Cross Validation Value

FPR False Positive Rate

HBM Human Body Model

HSR High Strain Rate

IRC Injury Risk Curve

LOOCV Leave-One-Out Cross Validation

LSE Least Squares Estimate

LSR Low Strain Rate

MLE Maximum Likelihood Estimation

MSE Mean Square Error

PMHS Post Mortem Human Subject

ROC Receiver Operating Characteristics

RSS Residual Sum of Squares

SED Strain Energy Density

SED HSR Energy based curve for 0.5 strain/s

SED LSR Energy based curve for 0.005 strain/s

Strain HSR Strain based curve for 0.5 strain/s

Strain LSR Strain based curve for 0.005 strain/s

TPR True Positive Rate

1

Introduction

In motor vehicle collisions, the thorax is one of the most commonly injured body regions. Of all drivers killed in frontal crashes, 84 % sustained an abbreviated injury scale (AIS) thoracic injury of severity level three or greater (AIS3+, serious to fatal) [1] and in approximately 30% of all deaths among belted, adult drivers in frontal collisions, the fatally injured body region was the thorax [2]. The most common thoracic injuries, both for fatal and less severe injuries, are rib fractures [3] and the risk of fatality due to rib fractures increase with age, making rib fractures among elderly particularly dangerous [4, 5]. An increasing number of fractured ribs also increases the risk of breathing issues due to pain, reduced residual volume, hypoxemia and increased respiratory work, as well as the risk of damage and penetration of the underlying soft tissue [4]. This makes the improvement of rib fracture prevention methods crucial.

1.1 Background

To design occupant safety systems that reduce the risk of rib fractures, proper injury risk assessment is of great importance [3]. Today, the injury risk is assessed using both physical dummies, known as anthropometric test devices, and virtual human body models (HBMs). Risk prediction using Anthropometric test devices, uses mainly empirical probabilistic methods [6], where injury occurrance, often in post mortem human subjects (PMHS), is related to measurements from Anthropometric test devices in matched collisions [7]. Forman et al. [6] developed a method for probabilistic rib fracture risk prediction for Finite Element (FE) HBMs. The method for rib fracture risk prediction for HBMs uses the ultimate strain (the strain at failure) to predict rib fractures and includes a method to predict changes in risk with age [6]. The development of the age adjusted risk curves was based on rib cortical bone testing [8, 9] with specimen from a total of 12 PMHS with ages ranging from 18 to 81 years, a sample not large enough to describe the difference in rib cortical bone strength seen in the population over different ages [10]. Instead, to adjust the risk curves for individuals of different age, Forman et al. [6] refer to a study by Carter and Spengler [11] showing that the ultimate strain of cortical bone decreases with an average of 5.1 % per decade of life. The risk curves suggested by Forman et al. are adjusted accordingly to this. However, the study by Carter and Spengler [11] solely bases the value of 5.1 % on femoral cortical bone, classified as long bone,

and it is not stated that the decrease in mechanical properties are the same for ribs, that are classified as flat bones [12]. In the study by Carter and Spengler [11], it can also be seen that the decrease in ultimate strain is more rapid for the younger and the older ages, whereas the decrease in the middle ages is said to be slower. In a more recent study, performed in 2018, by Morgan et al. [13] that reviewed current literature, it was found that the ultimate strain of cortical bone decreases with 10% per decade of life.

Female vehicle occupants are more likely to suffer from rib fractures than male occupants exposed to an identical crash [14] and the difference can partially be described by the structural properties of the ribs [15]. A study performed by Schafman et al. [16] exposed 184 ribs from 93 PMHS to a dynamic bending test. They found that, although it only described a small part of the variance, there is a significant difference in the structural properties of ribs depending on sex. This result corresponds well with a study performed by Agnew et al [17], also finding that sex only explains a small part of the variance in mechanical properties although male subjects tend to have higher tolerances regarding mechanical properties. Albert et al. [18] investigated the contributions to the structural properties of ribs, finding that the rib cortical bone material properties have a significant contribution to the overall structural response of ribs.

In a study published in 2020, Katzenberger et al. [10], that investigated the effects of age, sex and strain rate on the tensile material properties of human ribs. In the study, rib cortical bone samples from 61 PMHS, ranging from 17 to 99 years were tested in two different strain rates, giving a data set that is large enough to represent the variability in humans. This dataset facilitates the creation of sex, age and strain rate dependent injury risk curves (IRCs).

Furthermore, this data set represents a large sample of human rib cortical bone material properties exhibiting both large inter-individual variation and significant age dependency. Human Body Models need material mechanical parameters to represent the rib cortical bone material properties. In the SAFER HBM [19] the rib cortical bone material behavior is represented by a piecewise linear elasto-plastic material model. The mechanical parameters used for this material model represents the average parameters found in [8, 9]. In this thesis, the feasibility to model how the rib material mechanical parameters vary with age and sex, while also accounting for the large inter-individual variation, will be investigated.

1.2 Objectives

The overall aim of this thesis is to improve the HBM rib fracture risk prediction for individuals of different age and sex. To fulfil the aim, three objectives were defined:

1. To investigate how age, sex and strain rate affect the risk of rib cortical bone fracture and to create injury risk curves taking these predictors into account.

2. To investigate how the mechanical properties of rib cortical bone are affected

by age and sex and to create regression models that take these predictors into account. The models should be developed to cover the spread in the population and thus differentiate between stronger and weaker individuals.

3. To apply the models of the rib cortical bone in simulations and to apply the injury risk curves to simulation results to evaluate if implementing the results from objective 1 and 2 improves the accuracy of human body model rib fracture risk prediction.

1.3 Delimitations

The first and second objective regarding the injury risk curves and the material models will be done by the student with some supervision from supervisors and statisticians. This thesis is delimited to only provide the material data for the simulations and to interpret the results from the simulations, not to carry out the simulations. Therefore, the performance of the simulations will not be described in detail but only the preparatory work, risk calculation and data interpretation.

The available data set includes the age, sex, height, weight and BMI of the subject. However, due to the limit of time, the scope of this thesis is delimited to focus on the effect of the covariates age and sex.

The outcome variables measured in the study [10] include yield stress, yield strain, failure stress, failure strain, ultimate stress, elastic strain energy density (SED), plastic SED and total SED. Based on the parameters used for injury risk prediction today, the IRCs will be created for strain and total SED, with the effects of age and sex considered. The regression models for rib cortical bone material properties will be created for elastic modulus, yield stress and tangent modulus.

2

Theory

When studying failure as a function of different parameters, in this case fracture in human cortical bone samples, as a function of age, sex and mechanical load, the theory of survival analysis is an applicable framework to model the data. Thus, the basic principles of survival analysis is presented in section 2.1. As it is of interest to create a parametric representation of the data, both regarding survival analysis and regression models for the mechanical properties of rib cortical bone, coefficients have to be estimated and the fitted models have to be evaluated. For this purpose the theories of coefficient estimation using the least squares estimate and maximum likelihood estimation are presented in section 2.2. For model evaluation and selection, the theories of DFBETAS, Akaike's Information Criteria (AIC) and Cross-validation are presented in section 2.3.

2.1 Survival Analysis

Survival analysis is a statistical regression method that generally can be described as *the time until an event occurs* [20]. This can be compared to logistic regression that only consider binary outcomes. The time in survival analyses is often measured in days, months or years and the event is often referred to as a failure. However, the time can be changed to something else, in biomechanical applications for example, mechanical load is often used rather than time. Another property that differentiate survival analysis from other regression models is its ability to handle censored data [21]. The data can be either, left-, right- or uncensored, where left censored data means that the event occurred before a certain time, but the exact time is unknown. Right censoring is when an event did not occur before a certain time and thus, survival time is greater than the known time. Uncensored data points means that the exact time until an event occurs is known [20].

In a quantative survival analysis, the term most often considered is the survival function, $S(t)$. The survival function describes the probability that the survival time T exceeds a specified time t, or in other words, the probability of being event-free at time t [20], and can be described according to equation 2.1.

$$S(t) = P(T > t) \tag{2.1}$$

Theoretically, the survival function is a continuous function, ranging from time 0 to ∞, where the probability starts at 1 and approaches 0 when t approaches infinity. However, when working with actual data points, the survival function is obtained as a step function that ranges from time 0 to the end of the study, as shown in figure 2.1.

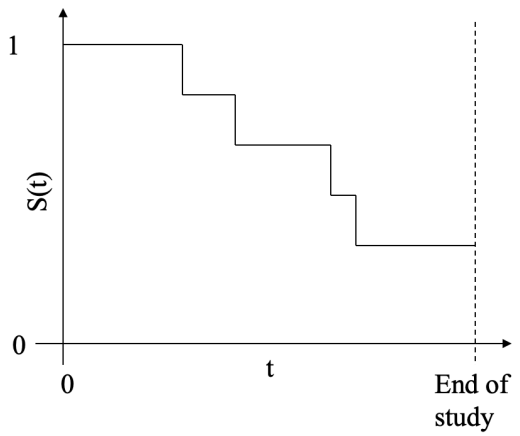


Figure 2.1: Survival curve, illustrated with the Kaplan-Meier estimate. Each step in the curve represents the incidence of an event, causing the risk of no event to decrease for that particular time. In this case, not all subjects under observation experienced an event before the end of the study.

Using the survival function, another useful summary of survival data can be expressed, the cumulative incidence function, $F(t)$. It can be defined as the probability that the survival time is less than, or equal to t, or equivalently, the probability that he event has occurred by time t and it can be referred to as the *risk function*. The cumulative incidence function can be defined as the complement of the Kaplan-Meier estimate of the survival function as

$$F(t) = 1 - S(t). \tag{2.2}$$

Survival analysis also has the power of dealing with several covariates, which gives the possibility to investigate how different factors, such as age, influence the survival time. The covariates are often referred to as predictors and can be of different types; binary predictors (for example male or female), multilevel categorical predictors (for example different treatments) and continuous predictors (for example age) [21]. How the predictors are dealt with to predict the survival depends on the method used. Either, a parametric modelling can be used, further presented in the next section, or the Cox model can be used. The Cox model is a semi-parametric method since it does not require a parametric model of the baseline but only of the regression parameters. The Cox model will not be further explained within this thesis.

2.1.1 Parametric Survival Models

In parametric modelling one makes the assumption that the outcome follows some kind of distribution. The exact distribution is not known until the values of the parameters are known. If the survival time can be assumed to follow a distribution it is a *parametric survial model*. Examples of distributions commonly used for survival analysis are the Weibull, the exponential, the log-logistic and the lognormal [20, 22]. Typically the parametric survival models are explained with two parameters, a shape parameter that is held fixed throughout the regression and a second parameter that is reparametrized in terms of prediction variables and the corresponding regression coefficients.

The accelerated failure time (AFT) model is a fully parametric approach to survival analysis. The AFT model acts under the assumption that the effect of the covariates are proportional to the survival time, or in this case, the mechanical load required for fracture. This relationship can be expressed as

$$S(t) = S(\gamma t_0), \tag{2.3}$$

where t_0 is the baseline mean survival time. The factor γ is called the *acceleration factor* and denotes how much one, or several covariates affect the expected survival time. In the case of a mechanical load replacing the survival time, the acceleration factor can be interpreted as the factor by which the load for 50% risk of fracture varies. If $\gamma < 1$, the risk is decreased and if $\gamma > 1$, the risk is increased. From equation 2.3 we have that $t = \gamma t_0$. Say we have a covariate that yields $\gamma = 2$, then we can see that $t = 2t_0$. That means that the particular covariate doubled the expected survival time. This is illustrated in figure 2.2.

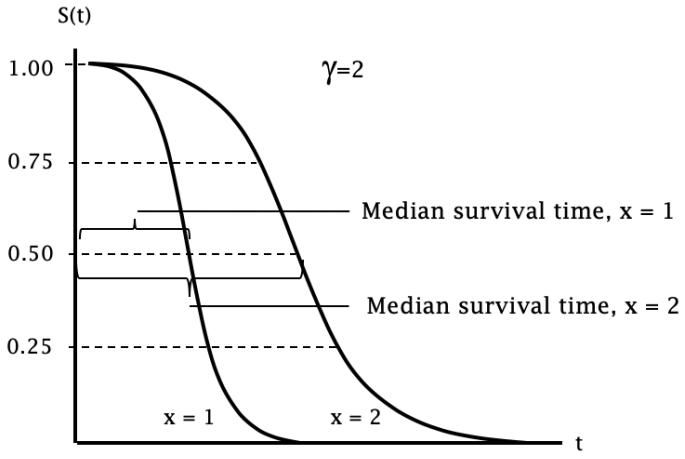


Figure 2.2: The figure illustrates the survival model S(t) with acceleration factor γ and the binomial covariate x. As seen in the figure the median survival time is doubled for x = 2, compared to x = 1.

2.2 Coefficient Estimation

When fitting a parametric model to data, coefficients have to be estimated. There are several different ways to perform coefficient estimation and what method that is the best depends on the model to be fitted. Here, two common methods for coefficient estimation, least squares estimate, that was used for the regression modelling, and maximum likelihood estimation, that was used for the survival analysis, are described.

2.2.1 Least Squares Estimate

When estimating the coefficients for a linear model, the least squares estimate (LSE) is commonly used and it involves minimizing the least squares criterion. A linear model is written as $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$, where $\hat{y}_i$ is the ith prediction of Y - all predicted values - based on the ith observation in X - the corresponding values used to predict Y. Then, the ith error, or *residual*, can be written as $e_i = y_i - \hat{y}_i$ and represents the difference between the ith observed response and the ith predicted response [23]. Using the residuals, the *residual sum of squares* (RSS) can be defined as

$$\text{RSS} = e_1^2 + e_2^2 + ... + e_n^2. \tag{2.4}$$

By inserting the definition of the residual and the equation for the linear model [24], equation 2.4 can be rewritten as

$$\text{RSS} = (y_1 - \hat{\beta}_0 + \hat{\beta}_1 x_1)^2 + (y_2 - \hat{\beta}_0 + \hat{\beta}_1 x_2)^2 + ... + (y_n - \hat{\beta}_0 + \hat{\beta}_1 x_n)^2. \tag{2.5}$$

To find the values for $\hat{\beta}_0$ and $\hat{\beta}_1$ that minimize the RSS, the least squares problem is solved by setting the gradient, with regards to β_j, to zero. This gives the minimizers

$$\hat{\beta}_1 = \frac{\Sigma_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\Sigma_{i=1}^n (x_i - \bar{x})^2}, \tag{2.6}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}, \tag{2.7}$$

where $\bar{x}$ and $\bar{y}$ are the sample means. In figure 2.3, the estimation $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$, found using the least squares, is plotted together with the residuals e_i, represented by the thin lines, and the observations (x_i, y_i), represented by the circles. For multiple linear regression with p covariates, the predictions can be written as

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + ... + \hat{\beta}_p x_{ip}. \tag{2.8}$$

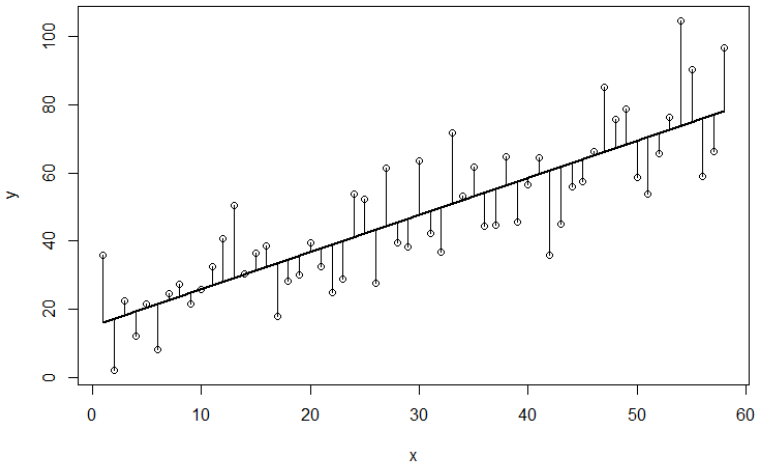


Figure 2.3: The diagonal line represents the estimation using the linear model and the circles are the actual observations. The vertical lines illustrates the residuals that describe the difference between the true and the estimated value of Y based on X.

The coefficient estimations are made by using the same least squares approach as for the simple linear regression. For non-linear models such as a quadratic or cubic model, the same method as for multiple linear regression can be used , rewriting the quadratic and cubic terms as $X_1 = X$, $X_2 = X^2$ and $X_3 = X^3$ [23].

The RSS value is closely related to the mean square error (MSE), that is another measure of the quality of the estimate. The MSE is defined as

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N} e_i^2 \tag{2.9}$$

where e is the residual defined previously.

2.2.2 Maximum Likelihood Estimation

The maximum likelihood estimation (MLE) is a method to estimate the coefficients of a probability distribution so that under the assumed statistical model, the observed data is the most probable. The maximum likelihood estimation is a very general method and the least squares approach presented in section 2.2.1, is actually a special case of the MLE method [23].

The MLE is a method, that given a specific distribution, finds the coefficient estimations θ so that the likelihood that the observations are produced by the given distribution is maximised. To find the maximum likelihood, the function to maximise is the *likelihood function*,

$$\text{lik}(\theta) = \prod_{i=1}^{n} f(X_i|\theta), \tag{2.10}$$

where the function $f(X_i|\theta)$ is the joint density function and the expression for f depends on the distribution under examination. To find the maximum value is often complicated and either, numerical methods are required or, instead of maximising the likelihood itself, it is easier to maximise its natural logaritm. This can be done since the natural logaritm is a monotonic function, meaning that it has the same maxima and minima as the original function, and the function that are to be maximised is then called the *log likelihood* [25].

2.3 Model evaluation

When choosing a particular model for a regression problem, it has to be evaluated. Both with the purpose to understand if the model is suitable for the data set and to compare one or several models to each other to ensure the best model is chosen.

2.3.1 DFBETAS

The DFBETAS statistics is used to identify overly influential observations and represents the change of the parameter estimation when deleting one observation at a time and is defined as

$$\text{DFBETAS}_j = \frac{b_i - b_{ij}}{S_i\sqrt{(X'X)^{-1}_{jj}}}, \tag{2.11}$$

where the numerator indicates the difference between the parameter value b_i, for predictor i, before and after excluding observation j from the full sample. The denominator represents the standard error of the estimated parameter b_{ij}, where S_i is the standard deviation withouth the ith observation and X is the predictor matrix [26]. Ususally a threshold is set based on the sample size, n, to $2/\sqrt{n}$ [27]. If the DFBETAS for a certain observaiton exceeds the threshold it can be considered as an overly influential observation. If an observation is overly influential it could be be of interest to omit it from the sample. The sample should, however, be examined before being omitted. In chapter 3, DFBETAS is used to identify potentially erroneous observations.

2.3.2 Akaike's Information Criteria

Akaike's information criterion (AIC) is a method for evaluating a model in the terms of simplicity and goodness of fit. The AIC value is defined as

$$\text{AIC} = -2\ln\hat{L} + 2k, \tag{2.12}$$

where $\hat{L}$ is the maximum likelihood estimation and k is the number of indepently adjusted parameters within the model [28]. For normally distributed errors, which is assumed to be the case, $\hat{L}$ is maximised when the residual sum of squares is minimised and thus, the AIC is more negative for a better fitting model [28]. The second term, called the penalty term, is smaller the lower the number of parameters is and therefore the best model, in terms of goodness-of-fit and simplicity, is the model with the minimum AIC value. For different models to be comparable using AIC, they need to model the same data set, using the same response variables.

2.3.3 Cross-Validation

When fitting and selecting a regression model to a set of observations, statistical learning is a powerful method when used correctly [29]. When assessing the quality of the fitted model, the MSE is often used. As the MSE was introduced in section 2.2.1, it was computed using the same data that was used to fit the model. This is called the *training error*. However, how well the model fits to the training data is not of real interest and instead, what is of interest is what is called the *test error*. The test error tells us how well the model fits to new, previously unseen data, and just as the training error, it is often assessed using MSE [23]. In some cases there might be a set of validation data available but often, that is not the case. In the absence of a large test data set, the observations can be divided in to a set of training data, on which the model is trained, and a set of validation data, on which the accuracy of the model is assessed, as seen in figure 2.4. This is called the *validation set approach*. Although the validation set approach may be useful, there are two potential drawbacks. Firstly, the estimated test error rate is highly variable, depending on exactly which observations that are chosen for the training and test data set respectively. Secondly, it is well established that statistical models tend to perform worse when fitted to a smaller number of observations, dividing the data set, as done in the validation set approach, might lead to an overestimation of the test error rate for a model fit to the entire data sample [23]. A refinement of the validation set approach that addresses these two issues is the *cross-validation method*.

Figure 2.4: The full data set can be divided into one training set, illustrated in dark grey, and one validation set, illustrated with light grey.

2.3.3.1 Leave-One-Out Cross-Validation

Just as for the validation set approach, the method of leave-one-out cross-validation (LOOCV) implies that the data set is divided into two, a training set and a validation set [23]. The difference is that in the LOOCV, the validation set contains only one observation, (x_i, y_i) and the training set, on which the model is fit, contains the remaining $\{(x_1, y_1), ..., (x_{i-1}, y_{i-1}), (x_{i+1}, y_{i+1}), ..., (x_n, y_n)\}$ observations. The MSE can then be computed as

$$\text{MSE}_i = (y_i - \hat{y}_i)^2, \tag{2.13}$$

where $\hat{y}_i$ is the point estimate for x_i. This estimte gives an MSE with a very high variability, which makes it a poor estimate although it is unbiased for the test error. To decrease the variability of the MSE, the procedure can be repeated n times, where the validation set is choosen according to figure 2.5, computing the cross validation estimate (CV) as

$$CV_{(n)} = \frac{1}{n}\sum_{i=1}^{n} \text{MSE}_i. \tag{2.14}$$

Figure 2.5: The data set is divided into a training set, containing n-1 observations and a validation set containing 1 observation. The procedure is repeated n times until all observations have acted as the validation set.

Using this method, the training set is substantially larger than the training set in the validation set approach and consequently, the LOOCV tends to not overestimate the test error rate to the same extent.

2.3.4 Reciever operator characteristics

When evaluating how well a diagnostic method works, the terms *sensitivity* and *specificity* are often used. The sensitivity is a measure of how many of the real positive outcomes that are correctly identified and the specificity is a measure of the proportion of the real negative outcomes that are correctly identified [30]. To assess the usefulness of a test method, both the sensitivity and the specificity need to be known and the best possible test would have both sensitivity and specificity close to 100 %. However, a diagnostic method that has a high sensitivity may as well have a low specificity and the other way around. Therefore, to assess the quality of a test or to select between different diagnostic tests, a method that can help selecting between models with different combinations of sensitivity and specificity is desired. In many test settings, for example risk prediction, the diagnostic result is a probability, rather than a simple "yes" or "no" and a threshold has to be set to

translate the continous prediction outcome into a binary classifier. What threshold that is set affects the sensitivity and specificity.

One useful method for evaluation and comparison that deals with the issues of thresholds and sensitivity - specificity trade-off is the *reciever operating characteristics* (ROC) curve [30]. For the ROC, the threshold is allowed to vary. Starting at 100%, the threshold decreases to every possible threshold. At every threshold level, an analysis of the sensitivity and the specificity is performed. The sensitivity can also be referred to the *true positive rate* (TPR), and is computed as

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{P}} \tag{2.15}$$

where TP is the number of true positive diagnoses and P is the number of real positive outcomes. The *false positive rate* (FPR) is a measure of 1-specificity and is computed as

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{N}} \tag{2.16}$$

where FP is the number of false positive diagnoses and N is the number of real negative outcomes. These values are computed for each threshold level, resulting in a number of different coordinates, (FPR, TPR) on the interval [0,1].

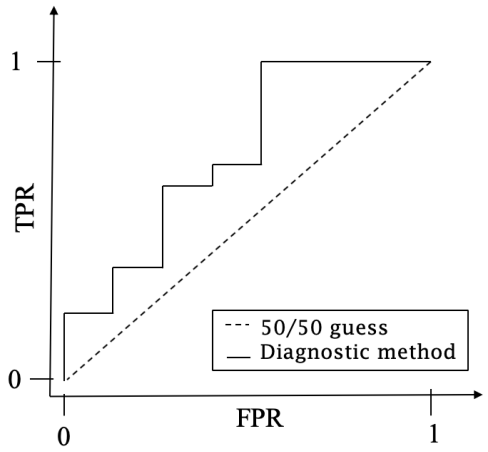


Figure 2.6: An ROC curve together with a random guess for a diagnostic test.

By plotting the coordinate values, a curve as the one in figure 2.6 is obtained, showing the relationship between the sensitivity, or TPR, and 1-specificity, or FPR. The dashed line in the figure is the result of a 50% guess. As stated earlier, a perfect diagnostic model would have both specificity and a sensitivity of 1 and produce an ROC curve that is one for any non-zero value of FPR. A good test would thus be somewhat close to this. When the ROC is created, the performance of a diagnostic

test can be quantified by calculating the Area Under the ROC curve (AUROC). Per definition, the ideal curve would have an AUROC of 1 and a random guess, as the one illustrated by the diagonal line in figure 2.6, would have an area of 0.5. The AUROC is therefore a powerful method for comparing two diagnostic methods to each other where the test with the highest AUROC can be considered as the better one, with increased capability of classifying the true outcomes.

3

Construction of Injury Risk Curves

This chapter handles the creation of the injury risk curves, which was listed as the first objective. The creation of the IRCs followed a method proposed by Petitjean et al. [22] and resulted in four different curves for ultimate strain and total strain energy density (SED) and high strain rate (HSR) and low strain rate (LSR). The resulting IRCs were compared to the previously developed IRCs by Forman et al. [6] and a larger influence of age on fracture risk was found in the new risk curves.

3.1 Materials and Methods

This section explains the available data and the methods used for the creation and evaluation of the IRCs.

3.1.1 Available Data

The study was performed using data from Katzenberger et al. [10]. In the study, the mechanical properties were measured in testing of human rib cortical bone. The tests were performed using rib cortical bone samples from 61 PMHSs, 32 males and 29 females, ranging from 17 to 99 years of age. The average age was 56.4 $\pm$ 26.2 years. From the PMHSs, dog-bone shaped coupons were extracted, from rib level 3 to 7, and the spongy bone was removed, leaving only the cortical bone. The coupons were placed in a servo-hydraulic high-rate Material Testing System and exposed to a uniaxial tensile test. For each subject, one coupon was tested to failure in tension at a targeted strain rate of 0.005 strain/s and another coupon was tested to failure in the higher strain rate of 0.5 strain/s. The tension in the cortical bone was measured using an extensometer, the data was filtered at SAE CFC 180 and used to calculate elastic modulus, yield stress, yield strain, failure stress, failure strain, ultimate stress, elastic SED, plastic SED and total SED. For a test to be considered as successful, the coupon had to fracture within the gauge length. In the cases where two successful tests on specimen at the same strain rate were performed, the results from the two tests were averaged [10].

For the validation of the obtained injury risk curves, human body model rib strain

predictions from real-life accident reconstruction simulations were available. Also, the corresponding strain-based rib fracture risk predictions obtained by the Forman et al. [6] probabilistic method and the Forman 2012 fracture risk curves were available.

3.1.2 Creation of Injury Risk Curves

IRCs were created both for failure strain and for total SED separately, as well as for high and low strain rate. Thus, each step in the following procedure was repeated four times; for failure strain at both high and low strain rate and for total SED at both high and low strain rate.

The IRCs were created using a method proposed by Petitjean et al. [22], with a few exceptions to make the method suitable for the available set of data. The method by Petitjean et al. follows a 12 step procedure, starting with collecting relevant data and thereafter, to assign censoring status. The third step is to check for dual injury mechanism and after that, in the cases where there are evidence of multiple injury mechanisms, the data should be separated in terms of mechanisms. The fifth step is to estimate the distribution parameters, using survival analysis with the maximum likelihood method. Step six is to identify overly influential observations using the DFBETAS statistics and step seven is to check the distribution assumptions. Based on the AIC, the distribution is chosen in the eigth step and the ninth step is to check the validity of the predictions againt existing results. The tenth step is to calculate the 95 % confidence intervals and to assess the relative size of the confidence intervals. Step eleven is to assess the quality index of the injury risk curves based on the relative size of the confidence interval. Lastly, one injury risk curve per body region should be recommended in order to avoid inconsistent injury risk predictions.

Based on the available set of data and the intended outcome, the method used for IRC creation for available set of data were as follows; first, relevant data was collected by analysing the data set [10]. The tests that were not deemed as successful were excluded from the analysis. The data set contained measures of engineering stress, σ_{eng} and engineering strain, ε_{eng}. For use in simulations, true stress, σ_{true} and true strain, $\varepsilon_{\text{true}}$, is required and thus, the values were recomputed according to equation 3.1 and 3.2 respectively and the total SED was recomputed with new values for true stress and true strain, as the area under the stress-strain curve [31].

$$\sigma_{\text{true}} = \sigma_{\text{eng}} \cdot (1 + \varepsilon_{\text{eng}}) \tag{3.1}$$

$$\varepsilon_{\text{true}} = \log(1 + \varepsilon_{\text{eng}}) \tag{3.2}$$

The analysis also investigated the effect of age and sex on Ultimate strain and total SED by conducting a one-way analysis of variance (ANOVA) and estimating the p-value. Second, the censoring status was assigned to each data point as exact data since all tests used in the analysis contained data for when the fracture occurred.

According to Petitjean et al. [22], the third and fourth step is to check for dual injury mechanisms and to separate data by injury mechanism, respectively. Since all tests in the study by Katzenberger et al. [10] were performed with identical setups, the fractures were all of the same type and none of the material responses showed multi-modality, the data points were, thus, not separated.

The distribution parameters were estimated through survival analysis with the MLE method, using the R software, version 3.6.3, with the flexsurv package, version 1.1.1. Regression models of the ultimate strain versus age and total SED versus age were performed to identify suitable distributions describing the trends in the failure strain and total SED. It was found that both the ultimate strain and the Total SED decreased log-linearly with regards to age and thus, it was assumed that the survival curves follow the AFT model and the parameters were estimated according to the method of maximum likelihood, presented in section 2.2.2. Since the risk of fracture at zero stimulus is zero, the distributions log-normal, log-logistic and Weibull are considered since they possess this property.

To find data points that might be erroneous and affect the modelling negatively the DFBETAS statistics was used. As recommended by Besley et al. [27], a threshold for the absolute value of DFBETAS was set to $2/\sqrt{n}$, where n is the number of observations in the sample. If an observation had a DFBETAS with a value above the threshold, that specific observation was examined. If there was no obvious difference in the examined observation compared to the other observations, no action was required and the observation was kept in the sample.

The distribution assumptions were checked using a Q-Q plot where the percentiles of the distributions were plotted against the percentiles of the sample. If an appropriate distribution had been chosen, the plot should follow a straight line through the origin with slope one. From the appropriate distributions, the one with the best fit was identified using the AIC value.

The 95% confidence intervals for the risk curves were determined under the assumptions of an asymptotic normal distribution of the maximum likelihood estimates of parameters β_0 and β_1 and a normal distribution of the error ϵ. By determining the width of of the 95% confidence interval at any given injury risk, called the relative size of the confidence intervals, the quality indices of the injury risk curves were assessed. The relative sizes of the confidence intervals were computed at 5%, 25% and 50% risk of injury for the curves adjusted for 25, 50 and 75 years of age and the quality indices were determined according to Table 3.1 [22].

Table 3.1: Quality index based on the relative size of the 95 % confidence interval.

Quality Index	Relative size of the 95 % confidence interval
Good	from 0 to 0.5
Fair	from 0.5 to 1
Marginal	from 1 to 1.5
Unacceptable	over 1.5

The recommendation of one curve per set of data was performed to ensure consistency in injury risk prediction. The recommendations were based on the AIC value and the quality index. First, the difference in AIC, Δ_{AIC}, between the current model and the model with the lowest AIC was evaluated according to the rule of thumb, presented by Burham and Anderson [32]. $\Delta_{\text{AIC}} < 2$ was considered as essentially no difference between the models, $4 < \Delta_{\text{AIC}} < 7$ was considered as a slight difference in models and $\Delta_{\text{AIC}} > 10$ was considered as a substantial difference between the two models compared. Thereafter, the Quality indices were taken into consideration under the criteria that a quality index as low as possible was desired and no injury risk curve with an "unacceptable" quality index should be recommended.

The newly developed IRCs should be checked for validity. However, since available simulation results only include ultimate strain and the higher strain rate, only one of the four IRCs can be compared at this stage. The new strain based IRC for high strain rate, Strain HSR, and the Forman 2012 IRC [6] were applied to rib strains obtained in simulations reconstructing real-life crashes with the SAFER HBM [33]. The maximum strain measured in each rib from the simulations were used for evaluating the risks for one or more (1+), 2+ and 3+ fractured ribs in every specimen. This was compared to the outcome of the real life vehicle crashes to see whether the Strain HSR risk curve predicts more accurate risks for rib fractures. To give an estimate of if, and how, the Strain HSR IRC improve rib fracture risk prediction, an ROC was performed on the Forman 2012 and Strain HSR risk predictions. Since available simulation results only include maximum strain and not total SED, the IRCs based on total SED were not compared to previous results.

3.2 Results

The collection and grouping of the data used for the creation of the rib fracture IRCs resulted in two sets of data, tests performed in high and low strain rate. After excluding three tests that were non-successful, the distribution of the specimen in each of the groups deviated slightly from that of the complete test group. In Table 3.2, the number of observations, mean age and the number of men and women are presented. The p-values for age and sex are also included in the table. Due to the high p-values resulting from the ANOVA, sex was not further considered as a covariate when constructing the injury risk curves since including non-essential variables increases the risk of overfitting in the model [24]. Regarding age, the ANOVA indicated that age should kept as a covariate. In Figure A.1-A.4, the

effects of age and sex on Ultimate strain and total SED are illustrated as plots .

Table 3.2: Properties of the used data set

Strain rate	Number of samples	Mean age	Male / Female	Measure	P-value age	P-value sex
0.5	58	56.2 ± 26.1	31 / 27	Ultimate strain	<0.0001	0.213
				Total SED	<0.0001	0.266
0.005	58	55.8 ± 25.8	32 / 26	Ultimate strain	<0.0001	0.554
				Total SED	< 0.0001	0.807

The distribution parameters were estimated for the weibull, log-normal and log-logistic distributions and the estimated parameter values are presented in Table 3.3.

Using the DFBETAS statistics to identify erroneous observations, it was found that a number of observations could, according to the DFBETAS, be considered as overly influential. In Figure 3.1, an example of one of the observations that were distinguished with the DFBETAS (shown in red) is shown together with stress-strain curves of individuals of the same age. As seen in the figure, there was no obvious difference between the red curve and the others. After examining all the observations, it was judged that there was no obvious difference in any of the test performance, tested specimen and test result in any of the identified observations. All observations were therefore kept.

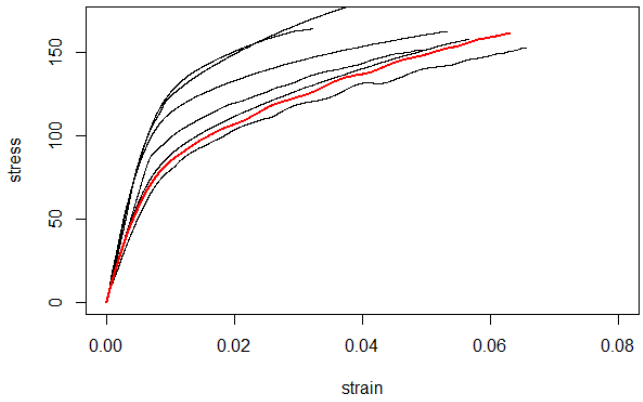


Figure 3.1: Stress-strain curves for the observation of a 18 year old that was distinguished by DFBETAS and other observations of indviduals of the same age.

The QQ-plots were evaluated for each of the distributions under investigation and they all showed good correspondence between the data and the chosen distributions. The AIC values for the distributions were determined according to equation 2.12 and the values are presented in Table 3.3. The lowest AIC values are marked with bold font.

Table 3.3: Scale and shape parameters, as well as AIC value for Weibull, log-logistic and log-normal distribution for the different data types.

Simulation Measurement	Distribution	0.5 strain/s				0.005 strain/s			
		Shape	β_0	β_1	AIC	Shape	β_0	β_1	AIC
Ultimate strain	Weibull	3.3561	-2.9236	-0.0114	-389.499	3.5840	-2.9530	-0.0120	-400.889
	Log-logistic	5.6986	-2.9802	-0.0133	-397.09	5.6671	-2.9924	-0.0139	-402.477
	Log-normal	0.3026	-2.9802	-0.0130	**-399.3**	0.3062	-3.0241	-0.0134	**-404.273**
Total SED	Weibull	2.4088	2.0775	-0.0194	166.35	2.4196	1.9483	-0.0211	145.900
	Log-logistic	3.8083	2.1086	-0.023	**165.27**	3.8037	1.8445	-0.0229	**145.816**
	Log-normal	0.4793	2.1370	-0.0245	167.04	0.4682	1.8084	-0.0226	145.912

The relative sizes of the 95 % confidence intervals were determined for 5%, 25% and 50% risk of injury and used to calculate the quality indices by analysing the relative size of the confidence intervals in accordance with the values in Table 3.1. All distributions were assigned with quality index good except the Weibull distribution for total SED that only scored quality index "Fair". In Table B.2 values for the relative sizes of the confidence intervals and the assigned quality indices are presented.

A comparison between the distributions were performed for each of the strain rates and data types. The values for Δ_{AIC} are presented in Table 3.4. For a strain rate of 0.5 strain/s and a measure of Ultimate strain, Strain HSR, the log-normal distribution has a $\Delta_{\text{AIC}} > 2$, and was therefore recommended. For a strain rate of 0.5 strain/s and a measure of total SED, the log-logistic distribution is the best according to the AIC value. However, compared to both the log-normal and the Weibull distribution, $\Delta_{\text{AIC}} < 2$ and according to Burham and Anderson [32], this implies that the difference between the models is essentially none. The models were thus compared in quality index. The quality index of the log-logistic distribution was classified as good for all tested CI points, whereas the quality index for the Weibull distribution only reached the second best level, "Fair", for the CI at 5% risk of injury. Both the log-logistic and the log-normal distribution had a good quality index but the log-logistic distribution had a lower AIC. Therefore, the Log-logistic distribution is recommended for the SED-based IRC at a strain rate of 0.5 strain/s, called SED HSR. For the curve for ultimate strain at the low strain rate of 0.005 strain/s, Strain LSR, the log-normal distribution was recommended, both due to the lowest AIC value and the good quality index but also out of simplicity since the log-normal distribution also was recommended for the Strain HSR IRC. For the same reasons, the log-logistic distribution was recommended for the SED based risk curve at 0.005 strain/s, SED LSR. The parameter values for the recommended IRCs are presented in Table 3.5. The parametric expressions of the injury risk, F, using the log-logistic and the log-normal distributions, are presented in equation 3.3 and 3.4, respectively.

$$F(\text{SED}, \text{Age}) = 1 - \frac{1}{e^{-(\beta_0+\beta_1 \cdot \text{Age})\text{Shape}} \cdot \text{SED}^{\text{Shape}}} \tag{3.3}$$

$$F(\alpha, \text{Age}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{\frac{-t^2}{2}} \mathrm{d}t \tag{3.4}$$

Where the upper integral limit, α is computed as

$$\alpha(\text{Strain}) = \frac{\text{Log}(\text{Strain}) - (\beta_0 + \beta_1 \cdot \text{Age})}{\text{Shape}} \tag{3.5}$$

Table 3.4: Δ_{AIC} for the different distributions, computed with regards to the distribution with the lowest AIC value, as presented in Table 3.3

Simulation Measurement	Distribution	0.5 strain/s	0.005 strain/s
		Δ_{AIC}	Δ_{AIC}
Ultimate strain	Weibull	9.801	3.384
	Log-logistic	2.21	1.796
	Log-normal	0	0
Total SED	Weibull	1.08	0.084
	Log-logistic	0	0
	Log-normal	1.77	0.096

Table 3.5: Distribution and parameter values of the chosen distributions.

Measure	Curve name	Distribution	Shape	β_0	β_1
Ultimate strain, 0.5 strain/s	Strain HSR	Log-normal	0.32026	-2.9802	-0.0130
Total SED, 0.5 strain/s	SED HSR	Log-logistic	3.8083	2.1086	-0.023
Ultimate strain, 0.005 strain/s	Strain LSR	Log-normal	0.03602	-3.0241	-0.0134
Total SED, 0.005 strain/s	SED LSR	Log-logistic	3.8037	1.8445	-0.0226

In figure 3.2 and 3.3, the recommended strain based IRCs are plotted, together with their 95% confidence intervals for a 25 year old and a 75 year old. The IRCs based on total SED are plotted in figure 3.4 and 3.5 for the same ages.

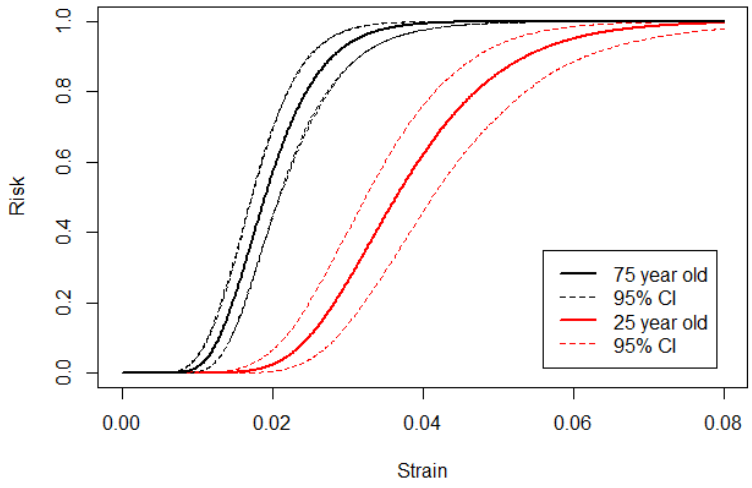


Figure 3.2: Recommended risk curve for Strain HSR, estimating risk based on ultimate strain for 0.5 strain/s

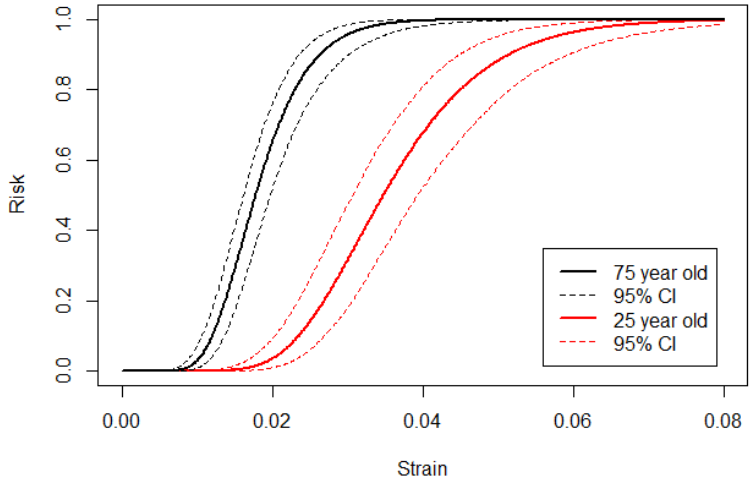


Figure 3.3: Recommended risk curve for Strain LSR, estimating risk based on ultimate strain for 0.005 strain/s

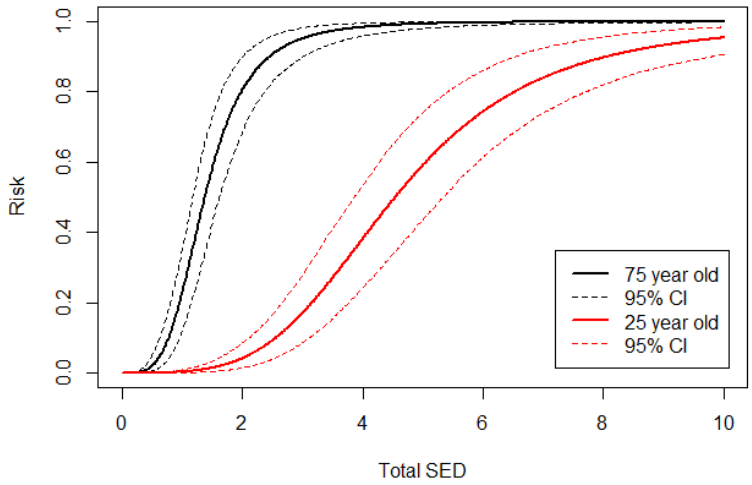


Figure 3.4: Recommended risk curve for SED HSR, estimating risk based on total SED for 0.5 strain/s

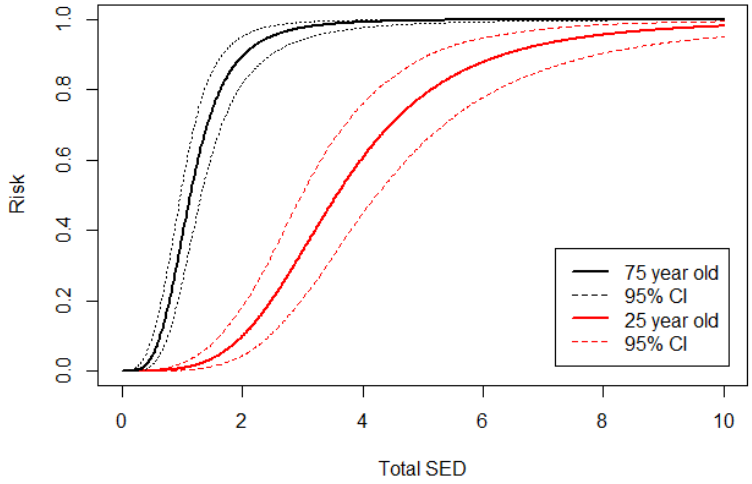


Figure 3.5: Recommended risk curve for SED LSR, estimating risk based on total SED for 0.005 strain/s

The risks estimated with the new strain based IRCs, for 0.5 strain/s, based on the human body model accident reconstruction simulation results are presented in Table 3.6 together with the Forman 2012 risks and the actual number of fractured ribs. In figure 3.6 and 3.7, the resulting plots from the ROC are shown for one or more and two or more rib fractures respectively. The AUROC values, representing the area under the ROC curve, were computed and the values are presented in Table 3.7. The Strain HSR risk curve results in a slightly higher AUROC and is therefore better, according to the analysis.

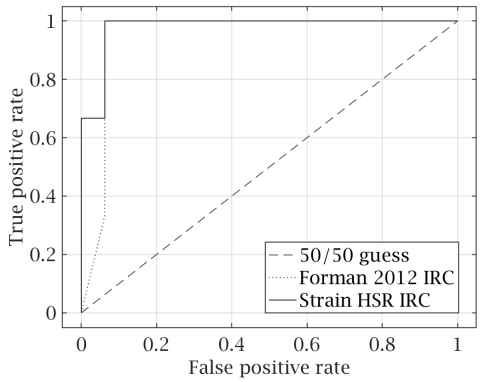


Figure 3.6: ROC for one or more fractured ribs, comparing the Forman 2012 and the Strain HSR IRCs

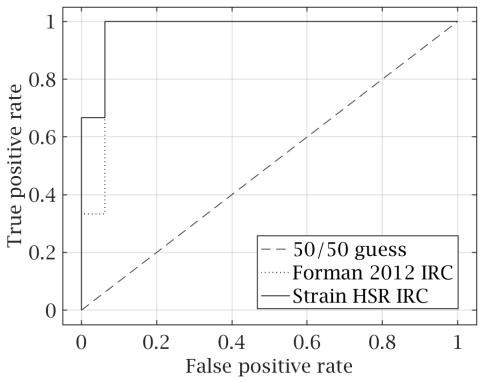


Figure 3.7: ROC for two or more rib fractures, comparing the Forman 2012 and the Strain HSR IRCs

Table 3.6: Rib fracture risks evaluated for 1+, 2+ and 3+ fractured ribs with Strain HSR risk curves for high strain rate, and Forman 2012 risk curves.

Age	Sex	Actual outcome	Forman 2012 risk			Strain HSR risk		
			1+ ribfx (%)	2+ ribfx (%)	3+ ribfx (%)	1+ ribfx (%)	2+ ribfx (%)	3+ ribfx (%)
81	M	0	24.75	2.72	0.16	36.64	5.43	0.33
82	F	0	67.57	25.42	5.13	90.33	58.6	21.77
44	M	0	100.0	99.7	91.98	99.68	95.17	74.32
44	M	0	97.99	86.74	62.07	93.74	71.05	38.7
67	M	1-2 (AIS2)	99.97	99.39	95.56	99.99	99.88	98.93
62	F	1 (AIS1)	88.22	56.45	23.36	93.06	67.82	33.14
85	F	4 (AIS3)	100.0	100.0	99.98	100.0	100.0	100.0
79	M	0	7.73	0.27	0.01	3.33	0.04	0.0
42	M	0	62.54	21.65	4.3	42.02	7.45	0.57
19	M	0	91.33	47.51	12.4	52.21	7.31	0.36
42	M	9 (AIS3)	99.32	92.78	71.95	96.09	77.17	43.76
46	M	0	2.57	0.03	0.0	0.01	0.0	0.0
52	F	0	36.91	6.05	0.53	26.39	21.13	0.06
39	M	0	32.97	5.2	0.47	10.64	0.37	0.0
37	F	0	63.76	13.53	1.3	41.95	2.61	0.06
38	F	0	13.81	0.78	0.02	2.08	0.01	0.0
34	M	0	12.81	0.59	0.01	2.05	0.01	0.0
42	M	0	69.11	29.94	8.39	45.92	10.71	1.41
22	M	0	34.83	5.78	0.53	5.07	0.09	0.0

Table 3.7: Area under curve for the ROC presented in figure 3.6 and 3.7 comparing the Forman 2012 risk curve and the Strain HSR risk curves for 1+ and 2+ rib fractures.

	AUROC	
	1+ fractures	2+ fractures
Forman 2012 IRC	0.9083	0.9583
Strain HSR IRC	0.9500	0.9792

3.3 Discussion

The analysis of the data showed that the effect of sex on Ultimate strain and Total SED was not significant. Thus, it was left out from the analysis. Therefore, it is suspected that the increased rib fracture risk for females in vehicle accidents is not related to differences in the rib material, but rather the structure of the ribs. For example, it is possible that females on average build more slender ribs than similar sized males which could be a contributing factor to the lower female rib bending stiffness identified by Agnew [17]. The average thickness of the cortical bone layer can possibly differ between males and females, which could be another contributing factor to the increased female rib fracture risk that cannot be revealed in this study.

When constructing the IRCs, the method of accelerated failure time was used for the survival analysis, which results in a proportional relationship between the survival time and the covariate, or in this case between age and the mechanical load required for fracture. The acceleration factors, computed as $\gamma = \exp(\beta_1 x)$, where x is the age, are presented in Table 3.8. From the formula, it can be seen that the acceleration factor indicates a decrease that is, in percent, constant over all ages. This type of non-linear decrease is thus the same as found by Carter and Spengler [11] and Morgan et al. [13].

The values in Table 3.8 for the decrease over one decade of life can be compared to the decrease in ultimate strain of 5.1 % per decade of life that Forman et al. [6] used for age adjustment of their IRCs.

Table 3.8: Acceleration factor for the different injury risk curves.

Simulation	0.5 strain/s		0.005 strain/s	
Measurement	Acceleration factor	Decrease per decade	Acceleration factor	Decrease per decade
Ultimate strain	0.9871	12.19 %	0.9867	12.54%
Total SED	0.9773	20.55 %	0.9777	20.23 %

In figure 3.8-3.10, the risks computed by Forman et al. [6] have been plotted together with the Strain HSR IRC presented above. In figure 3.8, the estimated risks for a 75 year old is plotted, showing a higher estimation of the risk using the Strain HSR IRC. Regarding the risk of fracture for a 45-year old, in figure 3.9 the risk is rather similar, however with a slightly lower risk using the Strain HSR IRC. The risk of rib fracture for the 25-year old, as seen in figure 3.10, differs remarkably with a lower risk using the Strain HSR IRC. As seen in the plot in figure 3.8, the two IRCs are rather similar for the lower strains. This makes intuitively sense since the method used by Forman et al. [6], adjusted each sample individually using the scaling of 5.1 % decrease in ultimate strain per decade of life. In the cases where the adjusted data point belonged to a specimen that was close to the adjusted age, the difference is thus smaller and the effect of age scaling does not affect the movement of the IRC as much. Since the specimen that fractured at the low strains are likely to be older, these data point did not need much re-scaling, whereas the specimen with the higher strains are more likely to be younger, requiring more scaling and thus, also deviate more from the Strain HSR IRC that assume a larger decrease in ultimate strain. In the same way, by looking at figure 3.9, it can be seen that the two curves are closer for the mid values, where the specimen are likely to be middle-aged and thus not requiring as much scaling.

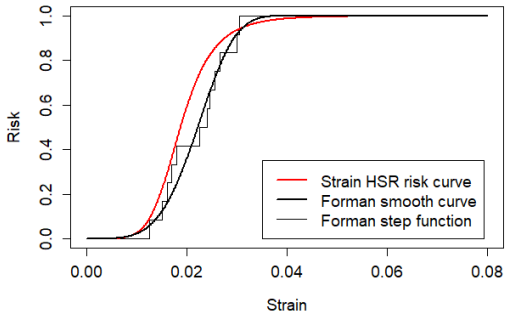


Figure 3.8: Comparison of rib fracture risk using Forman 2012 risk curves and the Strain risk curves for 75 year olds

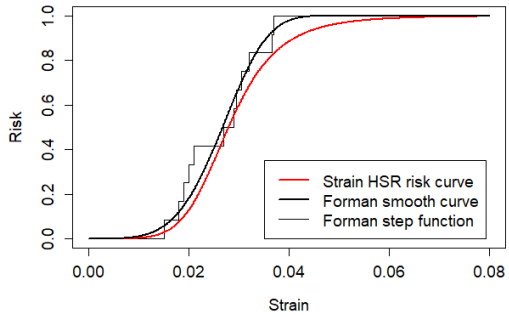


Figure 3.9: Comparison of rib fracture risk using Forman 2012 risk curves and the Strain risk curves for 45 year olds

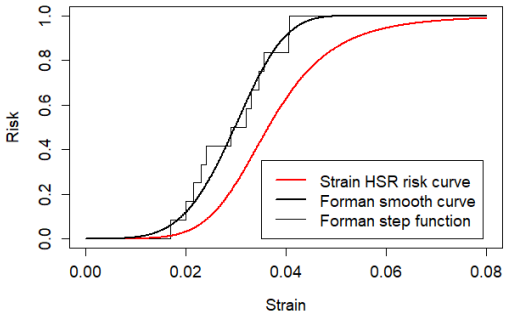


Figure 3.10: Comparison of rib fracture risk using Forman 2012 risk curves and the Strain risk curves for 25 year olds

The ROC comparison, which was showed in Figures 3.6 and 3.7 and Table 3.7, was performed for one or more and two or more fractures. In both cases it can be seen that the Strain HSR IRC performs slightly better. However, both the Forman 2012 and the Strain HSR IRCs performs well in the ROC analysis. The ROC only shows the relationship between true positive and 1 - false positive. Since the data used for the comparison contained very few real positive outcomes, already when only one positive outcome is identified, the true positive rate will be high.

According to the method suggested by Petitjean et al [22], only one injury risk curve per body region should be recommended to ensure consistent injury risk predictions. Although the aim with the different risk curves is to find the risk curve that best predicts the risk for rib fractures, the selection between maximum strain and total SED based risk curves could not be made with the tools suggested by Petitjean et al. The method of using AIC for choosing between different models are not consistent when the models are based on different variables. Instead the methods of computing

Δ_{AIC} and the quality index was used to ensure the validity of the selected model for each of the four risk curves. The new suggested IRCs are thus, based on both ultimate strain and total SED for both 0.5 strain/s and 0.005 strain/s. To evaluate how the effect of using total SED for risk prediction instead of using ultimate strain, new simulations with total SED as output have to be performed. Then the risks obtained with strain as a measure can be compared with the risks using SED as a measure.

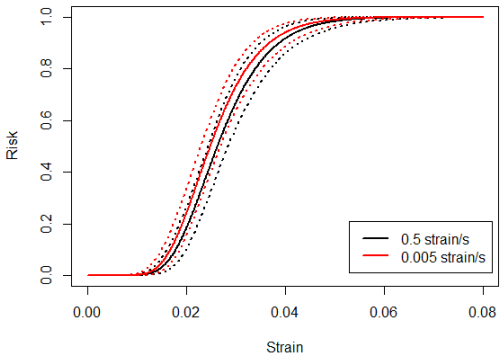


Figure 3.11: Comparison of Strain HSR (black) and Strain LSR (red) and their confidence intervals. The example shows the risk curves for a 50 year old.

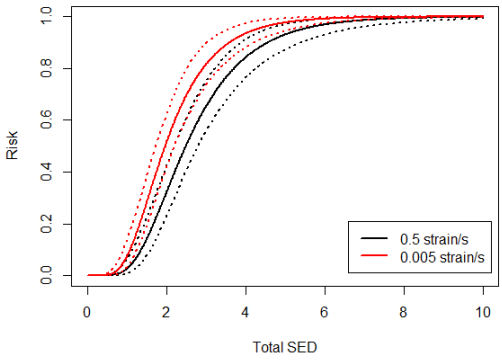


Figure 3.12: Comparison of SED HSR (black) and SED LSR (red) and their confidence intervals. The example shows the risk curves for a 50 year old.

Regarding strain rate, it can be seen in figure 3.11 that there is no significant differences in the risk when measuring strain. When measuring total SED, there is a slightly larger difference, however, looking at the plots in figures 3.11 and 3.12, there is no significant difference in neither strain nor SED based risks, since the confidence intervals in both cases are overlapping. Since motor vehicle collisions are high-speed events, the lower strain rate is of less interest in the setting of injury risk in crash simulation. If the results would have shown a large difference in risk based on strain rate, it could have been of interest to model the risk using strain rate as a covariate rather than as a separate curve, to capture the effect of strain rate. In the current situation however, the difference in risk that the different strain rates, that could occur in a motor vehicle collision, would give rise to would be negligible.

3.4 Conclusion

From the analysis of the data presented by Katzenberger et al. [10], it can be concluded that the only predictor for rib cortical bone fracture risk with statistical significance regarding ultimate strain and total SED is age. Thus, the function to determine the age-specific risk of rib cortical bone fracture, based on strain, is

$$F(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{\frac{-t^2}{2}} \mathrm{d}t, \tag{3.6}$$

$$\alpha(\text{MaxStrain, Age}) = \frac{\text{Log}(\text{MaxStrain}) - (-2.9802 - 0.0130 \cdot \text{Age})}{0.32026}. \tag{3.7}$$

The function to determine the age-specific risk of rib cortical bone fracture, based on total SED, is

$$F(\text{SED, Age}) = 1 - \frac{1}{e^{-(2.1086 - 0.023 \cdot \text{Age})3.8083} \cdot \text{SED}^{3.8083}}. \tag{3.8}$$

4

Creation of Regression Models for Material Parameters

This chapter handles the creation of regression models for the mechanical properties of human rib cortical bones, thus fulfilling the second objective of this thesis. Due to the large individual variability in mechanical properties, also within a given age span, equations that capture this variability was constructed to provide the opportunity to evaluate how this large individual variability affects human body model rib strain and energy predictions.

4.1 Materials and Methods

This section presents the materials and methods used to create the regression models. First the set of available data is presented and thereafter follows the method of data analysis, parameter estimation and model evaluation used for creation of regression models.

4.1.1 Available Data

The available data set contained data from experiments performed by Katzenberger et al. [10], the same data set as used in chapter 3, however, only the tests performed with the high strain rate were used. For each subject in the study, the strain and stress in the dogbone shaped coupon were calculated from force and elongation data collected at 40 kHz.

The available data that was used for the creation of the regression models included experimental stress-strain curves, given as engineering stress and engineering strain. The elastic modulus from the study by Katzenberger et al. [10] was computed based on the engineering values. The values were thus recomputed to true stress and true strain based on the experimental values from the study by Katzenberger et al.

4.1.2 Creation of Regression Models

The purpose was to create regression models to estimate mechanical properties of rib cortical bone for individuals of different age and sex. The predicted mechanical

properties were to be used in simulations for rib fracture risk prediction in LS-Dyna. The material used for the ribs in the simulation setup is a piecewise linear material, called MAT_024 [34]. For simplicity, the material model, that can contain up to eight linear segments, was simplified to a bilinear material model and thus consisting of one linear elastic curve and one linear plastic curve. Therefore, the aim was to create regression models that together have the capability to construct the age adjusted, bilinear approximations of the stress-strain curve as well as the upper and lower 80 % prediction intervals in order to simulate a weaker or stronger individual. Therefore, it was decided that regression models should be constructed for elastic modulus, yield stress and plastic hardening modulus separately. As shown in Figure 4.1 the elastic and plastic modulus define the inclination of the two lines and the yield stress defines the intersection.

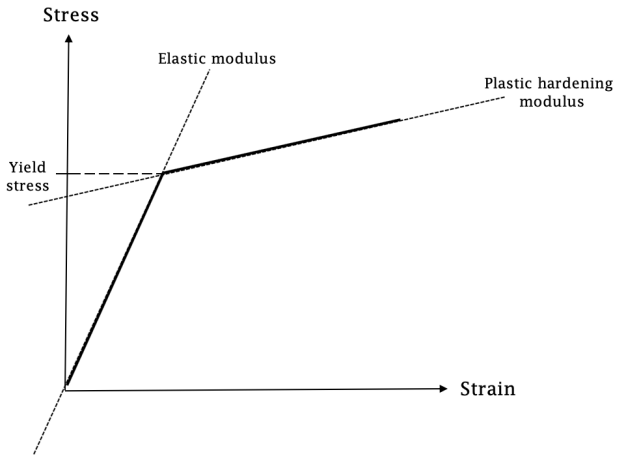


Figure 4.1: The bilinear approximation of the stress-strain curve can be constructed by the inclinations of the elastic and plastic part of the curve and the yield stress that defines their intersection.

The data was first analysed and non-successful tests, where the fracture had not occurred within the gauge length, were removed from the sample. For every specimen the engineering stress and engineering strain was recomputed to true stress and true strain according to equations 3.1 and 3.2. Thereafter the elastic modulus was computed , according to [10], as the gradient of the elastic portion of the stress-strain curve, between 10 % and 50 % of the yield point and the plastic hardening modulus was computed as the gradient of the plastic portion of the stress-strain curve, between the yield point and the failure point. The first step to investigate the effect of sex and age on the mechanical properties was to conduct an ANOVA and estimate the p-value. Only statistically significant predictors ($p< 0.05$) were kept in the analysis.

If neither sex nor age was found to be of statistical significance, no regression model was created. Instead, the mechanical property was assumed to be normally distributed and the lower and weaker individual was estimated as the lower and upper 10 % tail from the normal distribution. The assumption of the normal distribution

was controlled with a QQ-plot.

For the mechanical properties with statistically significant predictors, regression models were fitted. After plotting the data points and visually estimating the regressions it was decided that linear polynomial and log-linear polynomial models, equation 4.1 and 4.2 respectively, should be fitted to the data and evaluated for the best fit.

$$\hat{\mathbf{Y}} = \beta_0 + \beta_1 \mathbf{X} + ... + \beta_n \mathbf{X}^n, \tag{4.1}$$

$$\mathrm{LN}(\hat{\mathbf{Y}}) = \beta_0 + \beta_1 \mathbf{X} + ... + \beta_n \mathbf{X}^n, \tag{4.2}$$

for $n = 1, ..., 5$. Models were fitted using the LSE method for coefficient estimation in a LOOCV setting. To enable comparison in the cross validation value, CV, between linear polynomial regression and log-linear polynomial regression the estimate $\hat{y}_i$ was transferred back to the linear scale before being compared to the CV of the linear polynomial regression models.

Although the model with the lowest CV is the one that corresponds best with the data points, it might conflict with existing empirical knowledge on how age affects cortical bone mechanical properties. Therefore, theoretical behaviour of cortical bone was taken into consideration when choosing the most appropriate model. Also, due to the limited amount of data, a simple model was desired to avoid overfitting and in the cases where only a small difference in CV could be seen, a simpler model was selected.

To enable representation of the weaker and stronger individuals in the experimental data, the prediction intervals for each of the regression models were calculated. The prediction intervals was computed at 80 % confidence level, meaning that by using the lower boundary in simulation, 90 % of the population of that age will be stronger than the simulated individual and by taking the upper boundary, only 10 % of the population will be as strong.

From the regression models and the prediction intervals, equations were constructed so that given a specific predictor, the desired estimation was given.

4.2 Results

From the data set, three specimen were removed due to non-successful tests, resulting in a data set of 58 observations, a mean age of 56.2 ± 26.1 years, divided between 31 males and 27 females. The p-values for each of the measures are presented in Table 4.1. For none of the measures, sex has a statistical significance and was therefore omitted from the creation of the regression models.

Regarding the plastic hardening modulus, neither age nor sex has a statistical significance and therefore, a normal distribution was fitted to the data, using a mean of 2.52 MPa and a standard deviation of 0.69 estimated from the sample. In Figure 4.2, a histogram of the plastic hardening modulus together with the fitted normal distribution is plotted. The 10 % and the 90 % percentile was computed in order to represent the weaker and the stronger individual respectively.

Table 4.1: Properties of the used data set for creation of regression models

Measure	P-value age	P-value sex
Yield stress	<0.0001	0.803
Elastic modulus	<0.0001	0.960
Plastic hardening modulus	0.422	0.792

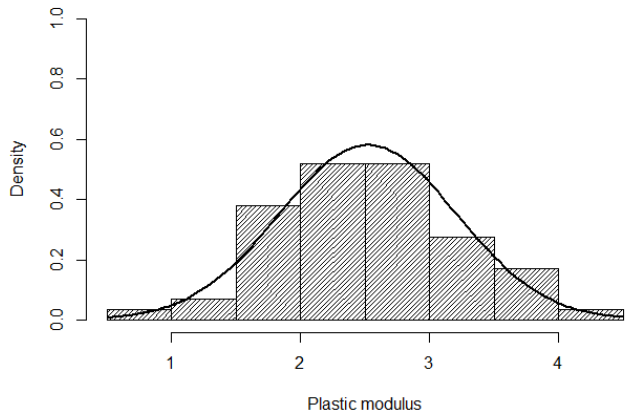


Figure 4.2: Histogram of the plastic hardening modulus together with the fitted normal distribution with mean = 2.52 MPa and standard deviation = 0.69

The cross-validation was performed for the different material parameters and the different CVs are presented in Table 4.2.

Table 4.2: Cross validation values for polynomials of degree 1-5 in linear-linear scale and log-linear scale

	Linear-linear					Log-linear				
	1st	2nd	3rd	4th	5th	1st	2nd	3rd	4th	5th
Yield stress	450.0	450.3	461.2	464.0	482.4	472.1	458.8	470.7	473.8	489.5
Elastic modulus	9.32	9.48	9.39	8.72	8.98	9.56	9.49	9.34	8.86	9.11

Regarding yield stress, the first order polynomial in the linear-linear scale has the lowest CV and was therefore chosen. For the Elastic modulus, the fourth degree polynomial in the linear-linear scale has the lowest CV. However, by plotting the

data together with the model it could be seen that the model suggests a very irregular behaviour of the elastic modulus with age. This behaviour of the material due to ageing is not supported in theory and is likely due to overfitting of the data. The same goes for the two models with the second and third lowest CV (fourth order log-linear and fifth order linear-linear models) they deviate from the expected behaviour due to overfitting. Instead, the model with the fourth lowest CV was selected, that is the first order model in the linear-linear scale. In Figure 4.3 and 4.4, the chosen models together with their 80 % prediction intervals are shown.

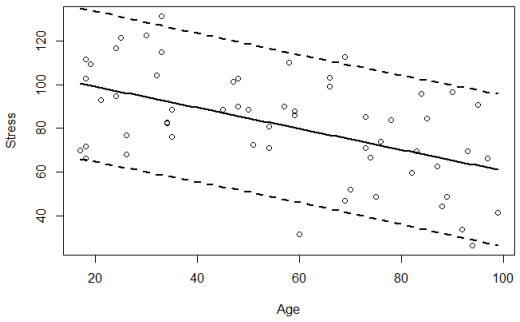


Figure 4.3: Linear regression model of yield stress as a function of age together with 80 % prediction intervals.

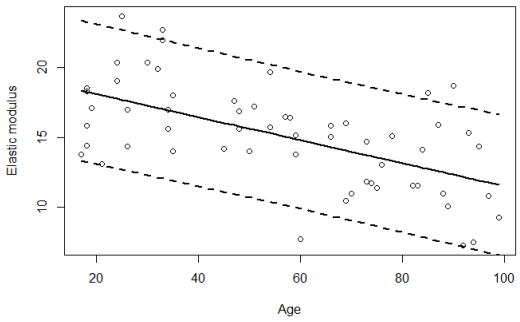


Figure 4.4: Linear regression model of elastic modulus as a function of age together with 80 % prediction intervals.

The expressions for the regression models of yield stress, elastic modulus and plastic modulus are presented in equation 4.3, 4.4 and 4.5 respectively. The variable $t_{(1-\alpha)/2}$ is the t-value for the desired confidence level under the assumption that the error follows a t-distribution. For an 80% prediction interval, $t_{(1-80)/2} = 1.28$.

$$\text{Yield stress} = 108.8 - 0.482 \cdot \text{Age} \pm t_{(1-\alpha)/2}\sqrt{423.6 + 0.0123 \cdot (\text{Age} - 56.16)^2} \quad (4.3)$$

$$\text{Elastic modulus} = 19.74 - 0.082 \cdot \text{Age} \pm t_{(1-\alpha)/2}\sqrt{8.891 + 0.00026 \cdot (\text{Age} - 56.16)^2} \quad (4.4)$$

$$\text{Plastic modulus} = 2.522 \pm 0.884 \quad (4.5)$$

The term to the left of the $\pm$ sign represents the average individual of the certain age, and the term to the right represent the prediction interval, and thus either the weaker or the stronger individual.

4.3 Discussion

The ANOVA showed that there was little or no correlation between sex and either of yield stress, elastic modulus or plastic modulus and sex was therefore omitted in the construction of the regression models. This is consistent with the result presented by Katzenberger et al. [10]. The difference in rib fracture risk between males and females is thus not likely to depend on the mechanical properties of rib cortical bone but rather on other properties, such as the shape of the ribs and ribcage and the bone porosity, as suggested by Schafman et al. [16]. The ANOVA also showed little or no correlation between age and plastic modulus. Thus, plastic modulus for the weaker, average and stronger individual was estimated based on the normal distribution. This approximation seems well suiting to the data based on the QQ-plot shown in Figure 4.5.

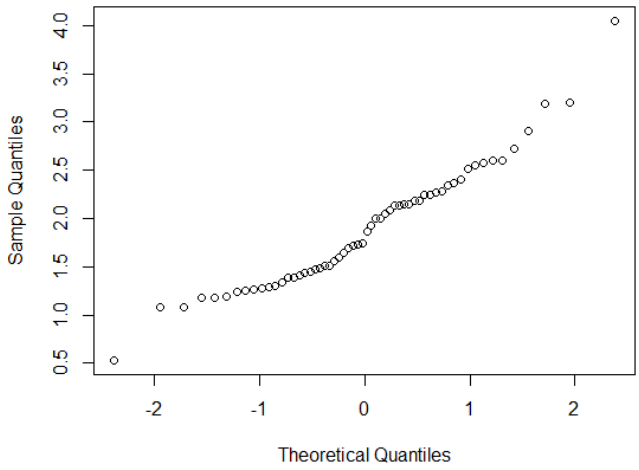


Figure 4.5: QQ-plot showing the teoretical quantiles against the sample quantiles for the distribution of the plastic hardening modulus

The choice of creating regression models for yield stress, elastic modulus and plastic modulus to estimate the mechanical properties of human rib cortical bone was based on the desire to enable prediction of weaker and stronger individuals. To visually make the approximations of the bilinear stress-strain curves, as presented in figure 4.6 and 4.7, the average value and the 80 % prediction boundaries were computed for the elastic modulus, plastic hardening modulus and yield stress. The elastic and plastic hardening modulus were plotted so that they intercepted at yield stress value for the desired material model; weak, average or strong.

As seen in Figure 4.6 and 4.7, the lower 80 % boundary is below the lowest true stress-strain curve. That is because the prediction intervals are computed for the three parameters separately and for all ages, not a specific age as shown in the figures.

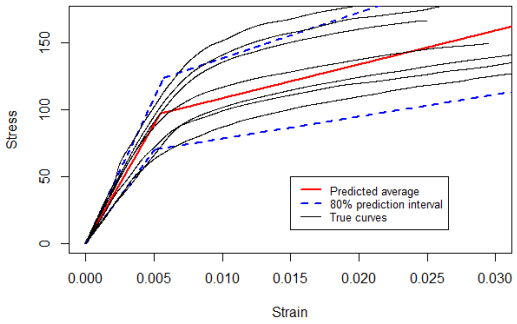


Figure 4.6: Predicted stress-strain curve and 80 % prediction intervals for 25 years olds together with the true stress strain curves for individuals of the same age ± 5 years.

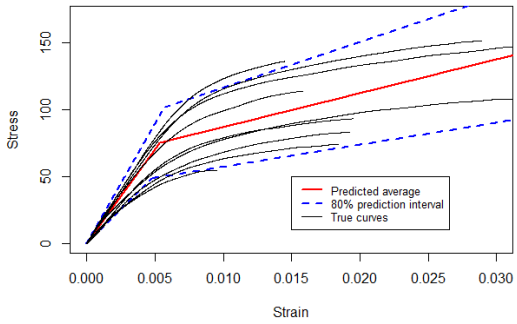


Figure 4.7: Predicted stress-strain curve and 80 % prediction intervals for 70 years olds together with the true stress strain curves for individuals of the same age ± 5 years.

When constructing the upper and lower prediction boundary, the 10th percentile plastic hardening modulus was selected for the lower boundary and the 90th percentile was selected for the upper boundary. This was since that would make the lower boundary the weakest and the upper the strongest. However, if it actually is so that the individuals with a lower elastic modulus also possess a lower plastic hardening modulus and the stronger individuals possess a higher plastic hardening modulus is not investigated and the opposite relationship is possible.

The prediction level was set to 80 %, giving that the lower boundary ensures that 90 % of all individuals are stronger. This level can however be changed by using a different t-value if it instead is desired to, for example, cover 80 % or 100 % of the population of a certain age.

4.4 Conclusion

The second objective was to create regression models that allowed the construction of individually adapted stress-strain curves of the human rib cortical bone. Therefore, models were created for yield stress, elastic modulus and plastic hardening modulus. To capture the inter-individual variation measured in the experimental results, the 80 % prediction intervals were estimated. It was seen that age was the only statistically significant predictor for yield stress and elastic modulus and the plastic hardening modulus was not affected by neither age nor sex.

For the yield stress, the equations for the average, the weaker 10 % and the stronger 90 % are as follows;

$$\text{Yield stress}_{\text{average}} = 108.8 - 0.482 \cdot \text{Age} \quad [\text{MPa}] \tag{4.6}$$

$$\text{Yield stress}_{\text{weak}} = 108.8 - 0.482 \cdot \text{Age} - 1.28\sqrt{423.6 - 0.012 \cdot (\text{Age} - 56.16)^2} \quad [\text{MPa}] \tag{4.7}$$

$$\text{Yield stress}_{\text{strong}} = 108.8 - 0.482 \cdot \text{Age} + 1.28\sqrt{423.6 + 0.012 \cdot (\text{Age} - 56.16)^2} \quad [\text{MPa}] \tag{4.8}$$

For the elastic modulus, the equations for the average, the weaker 10 % and the stronger 90 % are as follows;

$$\text{Elastic modulus}_{\text{average}} = 19.74 - 0.082 \cdot \text{Age} \quad [\text{GPa}] \tag{4.9}$$

$$\text{Elastic modulus}_{\text{weak}} = 19.74 - 0.082 \cdot \text{Age} - 1.28\sqrt{8.891 + 0.0003 \cdot (\text{Age} - 56.16)^2} \quad [\text{GPa}] \tag{4.10}$$

$$\text{Elastic modulus}_{\text{strong}} = 19.74 - 0.082 \cdot \text{Age} + 1.28\sqrt{8.891 + 0.0003 \cdot (\text{Age} - 56.16)^2} \quad [\text{GPa}] \tag{4.11}$$

The plastic hardening modulus was estimated with the normal distribution and the values are 2.522 GPa, 1.638 GPa and 3.406 GPa for the average, weaker and stronger respectively.

5

Evaluation of Developed Material Models and Risk Curves

This Chapter adresses the third objective; to evaluate the effect of individually adapted mechanical properties of rib cortical bone in human body models in simulation for rib fracture risk prediction. The chapter also aims to investigate the validity and effect of predicting rib fractures from simulation results, using total SED rather than maximum strain.

5.1 Materials and Methods

In this section, the results from chapter 3 and 4 are used in combination with simulations in LSDyna and SAFER HBM. The material used in the simulations to represent rib cortical bone is a piecewise linear plastic material, in LSDyna called MAT_024 [34]. The simulations were performed with a bilinear elastic-plastic material, created with two linear approximations. The material model is defined by the parameters E, EPS and ES. E is youngs modulus, or the elastic modulus, EPS is the effective plastic strain and ES is the corresponding stress value. In MAT_24, up to eight values can be used for EPS and ES. For a bilinear material, two values are used for EPS and ES, EPS1, EPS2, ES1 and ES2.

The elastic modulus, E, was obtained from equation 4.4, for the desired age and prediction level (weak, average or strong). For ES1, the yield stress was used, obtained for the specific age and prediction level, according to the regression model in equation 4.3. EPS1, that is the corresponding effective plastic strain at yield stress, per definition set to 0. ES2 and EPS2 were identified using the plastic hardening modulus identified according to equation 4.5. EPS2 was set to 0.1 to ensure the interval covers the failure point. Thereafter, ES2 was computed according as

$$\text{ES2} = \text{ES1} + (\text{plastic modulus} \cdot \text{EPS2}). \tag{5.1}$$

For the evaluation, four different simulations with four different material models - weak, average, strong, and original - were performed. The simulations were reconstructions of previously performed tests with post mortem human subjects. The

SAFER HBM used in the simulations was morphed to better resemble the occupant in the crash [35]. The simulations were selected to represent individuals of different ages and to simulate crashes with varying outcomes in terms of number of fractured ribs. The characteristics of each chosen simulation is presented in table 5.1. The mechanical properties for rib cortical bone were estimated as presented above for the age corresponding to the age of the simulated individual. All parameter values for each of the different simulation setups are presented in table 5.2.

Table 5.1: Details of the simulated individuals

Simulation ID	Age	Sex	Impact Direction	Impact Speed [km/h]	No Fractured Ribs
S1	65	F	Frontal	30	4
S2	34	M	Lateral	29	1
S3	83	F	Lateral	50*	12
S4	57	F	Frontal	47	17

*Sled test simulates intrusion and lateral acceleration from 50 km/h moving deformable barrier to vehicle impact test.

Table 5.2: The parameters estimated from the regression models for an individual with weak, average and strong rib cortical bone material properties, used in the simulations.

Simulation ID	Age	Level	Material Parameter				
			E [GPa]	EPS1 [-]	ES1 [GPa]	EPS2 [-]	ES2 [GPa]
S1	65	Weak	10.57	0.00	0.051	0.10	0.215
		Average	14.39	0.00	0.077	0.10	0.330
		Strong	18.21	0.00	0.104	0.10	0.442
S2	34	Weak	13.09	0.00	0.066	0.10	0.230
		Average	16.94	0.00	0.092	0.10	0.345
		Strong	20.78	0.00	0.119	0.10	0.230
S3	83	Weak	9.051	0.00	0.042	0.10	0.206
		Average	12.91	0.00	0.069	0.10	0.321
		Strong	16.76	0.00	0.095	0.10	0.436
S4	57	Weak	11.23	0.00	0.055	0.10	0.219
		Average	15.05	0.00	0.081	0.10.	0.334
		Strong	18.86	0.00	0.108	0.10	0.448

From the simulations, both total SED and maximum strain for each of the elements in the ribs were extracted. The maximum strain and maximum SED were identified for each of the ribs and the risks of rib fractures were evaluated using the Strain HSR, SED HSR and the Forman 2012 injury risk curves and probabilistic method for risk prediction from Forman et al. The risks were evaluated for the risk of 1+, 2+, and 3+ fractured ribs.

To evaluate whether the rib fracture risk prediction was improved using the new injury risk curves and age adapted material models, an ROC analysis was performed. ROC curves were computed for the Forman 2012, Strain HSR and SED HSR IRCs and the four different material models; weak, average, strong and original.

5.2 Results

The rib maximum strains and maximum total SED were extracted from the simulation results. For total SED, in some ribs one or a few elements obtained noisy and increasing levels of SED, which was not seen to be followed by neighboring elements, that instead followed a smooth trend of increasing and decreasing SED as external load was applied and relieved. The results from the noisy elements are believed to be caused by an unknown numerical error, and were therefore removed from consideration. Instead the peak SED was selected from the element in each rib exhibiting the largest SED, where the time history of SED was smooth and similar to that of surrounding elements. In Figure 5.1, a plot of the output energies from one of the simulations are shown. As seen in the figure, three elements reach energies much higher than the surrounding elements and the original max value, marked in orange, was updated with the new max value, marked in blue. In Figure 5.2, the strains for the same rib in the same individual is shown. As seen in the plot the strain output follows a smooth trend, compared to the disturbed results in the SED output.

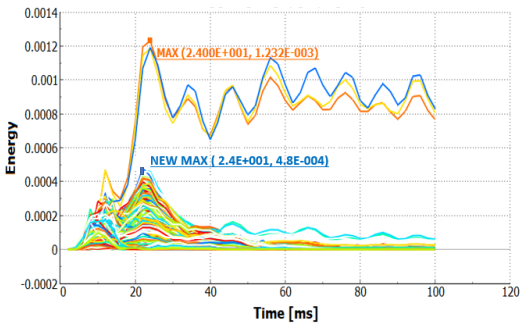


Figure 5.1: Total energy density for each element of one rib of the strong 34 year old. The max value marked in orange is the old max value and the blue text indicates the new max value

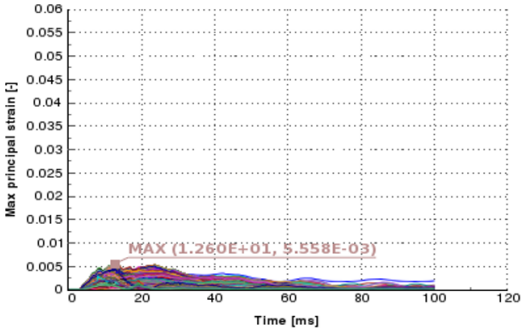


Figure 5.2: Maximum strain for each element of one rib of the strong 34 year old. The plot shows the strains in the same rib that is illustrated in the plot to the left, Figure 5.1.

Using the maximum energies and and maximum strains, the risks for fractures were computed and the resulting risk for 1+, 2+ and 3+ rib fractures for each simulation are presented in Table 5.4. As seen in the table, the risks differ between the different simulation and evaluation combinations as could be expected. It can be seen that for individuals below the average age of the test specimen, in this case the 34 year old, S2, the Strain HSR injury risk curve predicts lower risks than Forman 2012 injury risk curves. For individuals above the average age of the test specimen, with an exception of the 65 year old, S1, the Strain HSR injury risk curve predicts higher risks than the Forman 2012 risk curves.

The ROC analysis was performed and the AUROC is presented in table 5.3. As seen in the table, there are no values for the AUROC for the risk of one or more rib fractures. This is due to the fact that for one or more fractures, there are no true negative outcomes, making the definition of the false positive rate undefined.

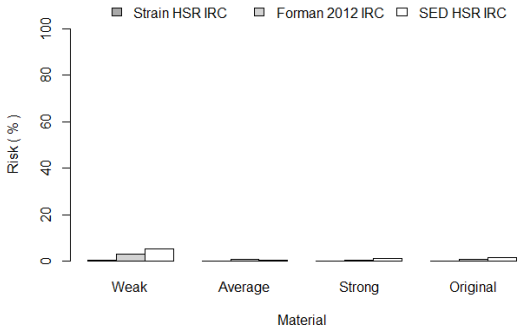


Figure 5.3: Predicted risks for 1+ rib fractures, for S1 with the four different material models and the Forman 2012, Strain HSR and SED HSR IRCs

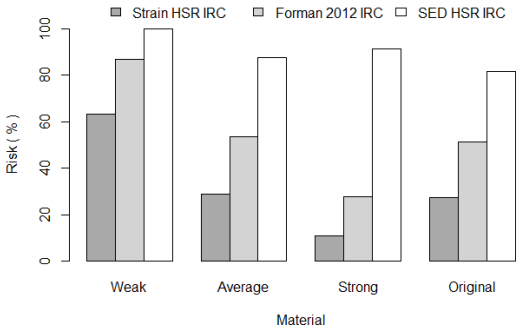


Figure 5.4: Predicted risks for 1+ rib fractures, for S2 with the four different material models and the Forman 2012, Strain HSR and SED HSR IRCs

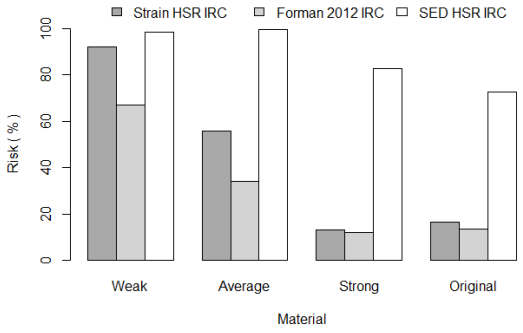


Figure 5.5: Predicted risks for 1+ rib fractures, for S3 with the four different material models and the Forman 2012, Strain HSR and SED HSR IRCs

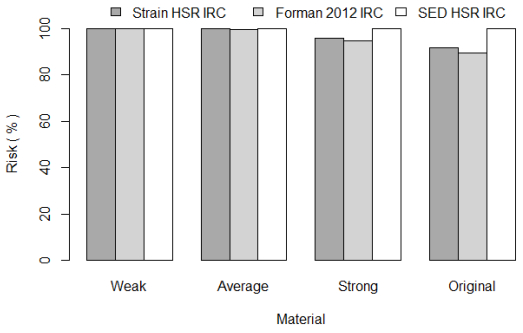


Figure 5.6: Predicted risks for 1+ rib fractures, for S4 with the four different material models and the Forman 2012, Strain HSR and SED HSR IRCs

Table 5.3: AUROC evaluated with both the Forman 2012 injury risk curves and the new injury risk curves based on results from simulations using both the new material models for weak, average and strong individuals as well as the old material models.

Injury level	Injury Risk Curve	Material			
		Weak	Average	Strong	Original
2+	Strain HSR	0.67	0.67	0.67	0.33
	Forman 2012	0.33	0.33	0.33	0.33
	SED HSR	0.33	0.67	0.33	0.33
3+	Strain HSR	0.67	0.67	0.83	0.67
	Forman 2012	0.33	0.33	0.33	0.33
	SED HSR	0.33	0.67	0.33	0.33

5.3 Discussion

When evaluating the injury risk curves and the rib cortical bone material models using FE-HBM simulations, there are many factors affecting the simulation results. For example, one problem consists of too small chest deformations, which in turn may result in underprediction of the rib fracture risk. Therefore, to be able to investigate the effect of the new injury risk curves and rib cortical bone material models, what we are most interested in is the relative difference between the different risk curves and material models. The reason that only the risks for 1+, 2+, and 3+ has been evaluated and presented is that the main goal of the evaluation was to investigate the effect of applying the new material models and the constructed injury risk curves and not to investigate if the new risk curves and material models predicts the correct number of rib fractures.

In Table 5.4 it can be seen that, for S1, all predicted risks are close to zero, which is a clear underestimation of the real outcome of four rib fractures. This implies that the woman is very weak and lies outside of the 80 % prediction interval. This is likely due to porosity that is not properly captured in the testing and material modelling in a macroscopic scale. It can also be seen that the Strain HSR curve predicts lower risks than the Forman 2012 curve although the individual is older than the mean age. This can be explained by the shape of the injury risk curve. The log-normal distribution used for the Strain HSR curve is closer to zero for the lowest strains compared to the Weibull distribution used for the Forman 2012 curve.

For simulation S2, the true test outcome was 1 rib fracture. By setting a threshold of 50 %, it can be seen that the Strain HSR curve underpredicts slightly compared to the Forman 2012 curve. If only looking at the weaker individual, both strain-based IRCs performs well. The energy-based risk curves do, on the other hand, overpredict the risk when regarded with a 50 % threshold. Simulation S3 represents an individual that, at an age of 83 years, sustained 12 rib fractures. Using the method of the 50 % threshold, both the Strain HSR and the Forman 2012 risk curves underpredicts the risks. The Strain HSR IRC tend to perform slightly better but not as good as the SED HSR IRC.

The simulation denoted S4 is a recreation of a high energy crash with a severe outcome of 17 rib fractures. By using the 50 % threshold again, it can be seen that all risk curves perform good but the old material does not result in satisfying risks, instead the original material results in underprediction of the risks. An exception is the energy-based risk curve that predicts 100 % risk for all material models and up to 3+ rib fractures.

In a comparison between the different IRCs it can be seen that all of them show clear difference between the different material models. For both of the strain based IRCs, the weaker individual is exposed to higher risk than the average individual, that in turn is exposed to lower risk than the stronger individual. The energy based risks show a deviation from this behaviour. As seen in table 5.4, and Figure 5.3-5.6, the risks evaluated with the SED based risk curves follow a different pattern.

For S1 and S2, the risks are higher for the stronger individual than for the average individual and for S3 and S4, the risks are higher for the average individual than for the weaker individual. It can be seen in Figure 5.3-5.6 that the SED-based risks are less affected by the different material models than the strain-based risks are. It is however worth to mention that the SAFER HBM only is validated for strain-based risk prediction.

An other factor that might affect the reliability of energy based risk prediction is that the energy values used for the risk estimation were, in many cases, manually adjusted since some elements in the FEM-model had substantially higher energy values than their surrounding elements. Due to the obvious errors, it can also be assumed that there are errors in the energy output from the simulation that are less obvious and therefore not manually corrected.

Regarding the ROC analysis, it can be concluded that the available data and the limited amount of performed simulations were not ideal for an ROC analysis. Since there was no real negative outcomes for 1 or more rib fractures, the analysis was not even possible to perform for 1+ rib fractures. However, in the performed ROC analysis, the Strain HSR IRC always performs better than the Forman 2012 IRC with an exception of risk for 2+ rib fractures with the original material model where the Strain HSR and the Forman 2012 IRCs perform equally. From Table 5.3 it can be seen that, in the example simulations performed, the Forman 2012 IRC actually performs worse than a random guess ($0.33 < 0.50$). For the Strain HSR IRC, it is necessary to also use the new material models to perform more accurate prediction than a random guess. Regarding the SED HSR curve, it only performs better when used with the average material model.

5.4 Conclusion

The third objective was to evaluate how the new injury risk curves and the individually adapted material models affect the injury risk prediction. It has been shown that both the new IRCs and the new material models affect the injury risk prediction. Both the Strain HSR and the SED HSR curves have a higher sensitivity to age compared to the Forman 2012 IRC, and particularly the SED HSR curve. How the material models affect the risk of rib fractures depends on the IRC used for risk evaluation. Using both the strain-based IRCs, the material models have a large impact, where a weaker material model gives higher risk and vice versa. For the SED HSR, the effect of the material model is smaller and a weaker material model does not imply a higher risk for rib fractures.

Table 5.4: Risks of rib fractures, based on the different simulation results evaluated with both Forman 2012 and Strain HSR injury risk curves.

ID	Age	Actual nfr	NFR	Strain HSR % fracture risk				Forman 2012 % fracture risk				SED HSR % fracture risk			
				Weak	Average	Strong	Original	Weak	Average	Strong	Original	Weak	Average	Strong	Original
S1	65	4	1+	0.45	0.01	0.00	0.02	2.99	0.77	0.49	1.00	5.47	0.58	1.48	1.51
			2+	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.03	0.00	0.00	0.00
			3+	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S2	34	1	1+	63.28	28.97	11.08	27.55	86.78	53.64	27.63	51.34	100.00	87.42	91.15	81.52
			2+	17.35	1.59	0.12	1.29	47.47	11.08	2.22	9.56	97.02	53.80	62.13	41.16
			3+	2.26	0.02	0.00	0.02	15.06	1.06	0.07	0.80	69.36	20.56	26.73	12.21
S3	83	12	1+	92.02	55.90	13.34	16.68	66.95	34.06	12.03	13.78	98.31	99.49	82.83	72.55
			2+	66.31	15.34	0.66	1.01	26.64	5.39	0.63	0.82	88.18	92.41	46.40	31.40
			3+	32.75	2.12	0.01	0.03	6.35	0.47	0.02	0.03	63.78	68.75	15.93	7.72
S4	57	17	1+	100.00	99.71	95.89	91.78	100.00	99.52	94.53	89.39	100.00	100.00	100.00	100.00
			2+	100.00	97.29	79.22	67.91	100.00	95.96	75.04	62.47	100.00	100.00	100.00	100.00
			3+	99.99	88.65	50.75	37.39	100.00	84.64	45.51	32.07	100.00	100.00	100.00	100.00

6

Discussion

This chapter aims to discuss the combined results from Chapter 3, 4 and 5 and how they answear the aim and objectives of this thesis.

The first objective was to develop injury risk curves for rib fracture risk prediction. The developed risk curves accounted for age and risk curves were developed for high and low strain rate and were based on both strain and total SED. However, it was seen that there was no significant difference between the risks from high and low strain rate. In chapter 3, where risks evaluated with both the Strain HSR injury risk curve and the Forman 2012 curve, were compared, it was seen that the Strain HSR IRC performed better in human body model reconstructions of real-life accidents, both in terms of AUROC and in terms of over- and underpredicting based on a 50 % threshold. In Chapter 5, The AUROC showed similar result, an improvement in risk prediction using the new strain based injury risk curves. Also when looking at a 50 % threshold, there was an improvement, although small. Thus, there are no indications that the Strain HSR risk curve performs worse than the Forman 2012 risk curve. The Strain HSR injury risk curve does also capture a larger effect of ageing, which is supported by Morgan et al. [13]. The new injury risk curves are also based on a larger set of data, including more individuals compared to the Forman 2012 risk curve, which makes it more statistically reliable.

The second objective was to develop models of the mechanical parameters of human rib cortical bone. When these material models were implemented into the simulations in Chapter 5, they proved to be more consistent with the actual test outcome. This is consistent with the findings of Albert et al. [18], that the rib cortical bone mechanical properties have a significant effect to the overall structural response of ribs. In particular did the material model resembling the weaker individual improve the risk prediction. However, if it is enough to use only the weaker material model or not can not be determined without knowing the reason why the risk prediction is performed. If the goal of the risk prediction is to identify how dangerous a crash is in the worst case scenario, then only the weaker material model needs to be considered. If it, on the other hand, is of interest to know the risk for an average or strong individual of a certain age, the other material models could be of interest.

Regarding the SED based risks, it was seen that the predicted risks did not follow the strength of the material model. This could perhaps be because the total SED,

that is the area under the stress-strain curve, is accounting for both the stress and the strain that is experienced by the ribs. This makes the dependence of the material smaller than for the strain-based risk and it can be seen in the plots in Figure 5.3 - 5.6. In the figures it can be seen that the difference between the different models are smaller than for the strain-based risks. As seen by Katzenberger et al. [10], total SED is more dependent on age than ultimate strain, which corresponds with the acceleration factors of the risk curves. Therefore, it would be of interest to further investigate the possibilities of using SED-based injury risk curves. The first steps in such an investigation would be to perform more simulations of real tests and to validate the SAFER HBM rib cage for prediction of SED.

Regarding the difference in rib fracture risk among men and women, it has been seen that there is no significant difference in the material properties, neither when estimating the risks based on strain or energy, nor when constructing material models of the rib cortical bone on a macroscopic level. Yet, we know there is a difference [16] and the reason behind it could be either the shape of the ribs, that females build slender ribs, and the shape of the rib cage, for example the rib angle.

The aim of this thesis was to improve the human body model rib fracture risk prediction. The combination of the new material models and the new injury risk curves gives the possibility to an improved prediction. Among the middle-aged, average and strong individuals, the difference in risk prediction is very small but for the elderly, weak individuals, the new possibilities for injury risk prediction are an improvement compared to risk prediction with baseline material models and the Forman 2012 curve. It has been seen that the Strain HSR and the SED HSR IRCs perform better than the Forman 2012 IRC in most situations and based on the performed tests, there are no evidence that the new IRCs would result in worse risk prediction than the Forman 2012 IRC. Regarding the material models, the ROC analysis shows little or no improvement in the risk prediction when evaluated only with the Forman 2012 IRC, but a clear improvement compared to the original model when evaluated with any of the new IRCs.

7

Conclusion

In this thesis, I have created new injury risk curves for rib fracture risk prediction and models of the mechanical properties of rib cortical bone. In Chapter 3 it was seen that the new injury risk curves, both strain-based and energy-based, had a higher sensitivity to age compared to the Forman 2012 risk curves. This was confirmed when the new IRCs were applied to simulation outputs for risk prediction in both Chapter 3 and Chapter 5. The material models developed in Chapter 4 showed that age had a significant effect on the stress-strain curve used in the material model of the rib cortical bone mechanical properties. This effect was visible also in Chapter 5 when the material models were applied to the SAFER HBM.

The real effect of the new IRCs and the material models need further investigation and for the energy-based IRC to be of use, the SAFER HBM rib cage should be validated for energy-based . However, the results from Chapter 3 and Chapter 4, and the implementation in Chapter 5 indicates that, if used correctly, the combination of an individually adapted rib cortical bone material and the new age adapted risk curves has the possibility to improve the rib fracture risk prediction, which was the aim of this thesis.

Bibliography

[1] Cuerden R, Cookson R, Massie P, Edward M. A review of the european 40% offset frontal impact test configuration. In: Proceedings of the International Technical Conference on the Enhanced Safety of Vehicles, June. Lyon, France; 2007. p. 18–21.

[2] Kent R, Henary B, Matsuoka F. On the fatal crash experience of older drivers. Annual Proceedings - Association for the Advancement of Automotive Medicine. 2005;(May):371–391.

[3] Mendoza-Vazquez M, Davidsson J, Brolin K. Construction and evaluation of thoracic injury risk curves for a finite element human body model in frontal car crashes. Accident Analysis and Prevention. 2015 12;85:73–82.

[4] Kent R, Woods W, Bostrom O. Fatality risk and the presence of rib fractures. In: Annals of Advances in Automotive Medicine - 52nd Annual Scientific Conference. vol. 52; 2008. p. 73–82.

[5] Carroll J, Adolph T, Chauvel C, Labrousse M, Trosseille X, Pastor C, et al. Overview of serious thorax injuries in European frontal car crash accidents and implications for crash test dummy development. International Research Council on the Biomechanics of Injury - 2010 International IRCOBI Conference on the Biomechanics of Injury, Proceedings. 2010;(September):217–234.

[6] Forman JL, Kent RW, Mroz K, Pipkorn B, Bostrom O, Segui-Gomez M. Predicting rib fracture risk with whole-body finite element models: Development and preliminary evaluation of aprobabilistic analytical framework. In: Annals of Advances in Automotive Medicine; 2012. .

[7] Poplin GS, McMurry TL, Forman JL, Ash J, Parent DP, Craig MJ, et al. Development of thoracic injury risk functions for the THOR ATD. Accident Analysis and Prevention. 2017;106:122–130.

[8] Kemper AR, McNally C, Kennedy EA, Manoogian SJ, Rath AL, Ng TP, et al. Material Properties of Human Rib Cortical Bone from Dynamic Tension Coupon Testing. Stapp Car Crash Journal. 2005;49:199–230.

[9] Kemper AR, McNally C, Pullins CA, Freeman LJ, Duma SM, Rouhana SM. The biomechanics of human ribs: material and structural properties from dynamic tension and bending tests. Stapp car crash journal. 2007;51:235–273.

[10] Katzenberger MJ, Albert DL, Agnew AM, Kemper AR. Effects of sex, age, and two loading rates on the tensile material properties of human rib cortical bone. Journal of the Mechanical Behavior of Biomedical Materials. 2020;102.

[11] Carter DR, Spengler DM. Mechanical properties and composition of cortical bone; 1978.

[12] Tortora GJ, Derrickson B. Introduction to the Human Body. New York: John Wiley and Sons Inc; 2014.

[13] Morgan EF, Unnikrisnan GU, Hussein AI. Bone Mechanical Properties in Healthy and Diseased States. Annual Review of Biomedical Engineering. 2018;20(1):119–143.

[14] Evans L, Gerrish PH. Gender and age influence on fatality risk from the same physical impact determined using two-car crashes. In: SAE Technical Papers; 2001-01-1174. .

[15] Schoell SL, Weaver AA, Vavalle NA, Stitzel JD. Age- and Sex-Specific Thorax Finite Element Model Development and Simulation. Traffic Injury Prevention. 2015;16.

[16] Schafman MA, Kang YS, Moorhouse K, White SE, Bolte JH, Agnew AM. Age and sex alone are insufficient to predict human rib structural response to dynamic A-P loading. Journal of Biomechanics. 2016;49:3516–3522.

[17] Agnew A, Murach MM, Dominguez VM, Sreedhar A, Misicka E, Harden A, et al. Sources of Variability in Structural Bending Response of Pediatric and Adult Human Ribs in Dynamic Frontal Impacts. Stapp car crash journal. 2018;62:119–192.

[18] Albert DL, Kang YS, Agnew AM, Kemper AR. A Comparison of Rib Structural and Material Properties from Matched Whole Rib Bending and Tension Coupon Tests. In: Proceedings of IRCOBI Conference, IRC -17-71; 2017. p. 567–576.

[19] Iraeus J, Pipkorn B. Development and Validation of a Generic Finite Element Ribcage to be used for Strain-based Fracture Prediction. Proceedings of IRCOBI Conference. 2019;p. 193–210.

[20] Kleinbaum D, Klein MG. Survival Analysis: A Self-Learning Text Third Edition; 2005.

[21] Vittinghoff E, Glidden DV, Shiboski SC, McCulloch CE. Regression Methods in Biostatistics. New York: Springer New York; 2005.

[22] Petitjean A, Trosseille X, Praxl N, Hynd D, Irwin A. Injury risk curves for the WorldSID 50th male dummy. Stapp car crash journal. 2012;53.

[23] Hastie T, Tibshirani R, James G, Witten D. An Introduction to Statistical Learning, Springer Texts. vol. 102. Springer; 2006.

[24] Hastie T, Tibshirani R, Friedman J. Springer Series in Statistics. vol. 27; 2009. Available from: `http://www.springerlink.com/index/D7X7KX6772HQ2135.pdf`.

[25] Rice JA. Mathematical statistics and data analysis. 3rd ed. Crockett C, editor. Duxbury: Thomson Brooks/Cole; 2007.

[26] Van der Meer T, Te Grotenhuis M, Pelzer B. Influential cases in multilevel modeling: A methodological comment. American Sociological Review. 2010 feb;75(1):173–178.

[27] Belsley DA, Kuh E, Welsch RE. Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. New Jersey: John Wiley sons & Inc.; 1981.

[28] Akaike H. A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control. 1974;19(6):716–723.

[29] Picard RR, Cook RD. Cross-validation of regression models. Journal of the American Statistical Association. 1984;387:575–583.

[30] Bewick V, Cheek L, Ball J. Statistics review 13: Receiver operating characteristics curves. Critical Care. 2004;8(6):508–512.

[31] LSDyna. From engineering to true strain, true stress;. Available from: `https://www.dynasupport.com/howtos/material/from-engineering-to-true-strain-true-stress`.

[32] Burnham KP, Anderson DR, Huyvaert KP. AIC model selection and multimodel inference in behavioral ecology: Some background, observations, and comparisons. Behavioral Ecology and Sociobiology. 2011;65(1):23–35.

[33] Pipkorn B, Iraeus J, Björklund M, Bunketorp O, Jakobsson L. Multi-Scale Validation of a Rib Fracture Prediction Method for Human Body Models. Proceedings of IRCOBI Conference. 2019;(0):175–192.

[34] LsDyna. Keyword User ' S Manual Volume II. vol. II; 2012.

[35] Larsson KJ. Thesis, Chalmers University of Technology, Evaluation of Morphed Human Body Models for Diverse Occupant Safety Analysis; 2020.

A

Data for Failure Strain and SED

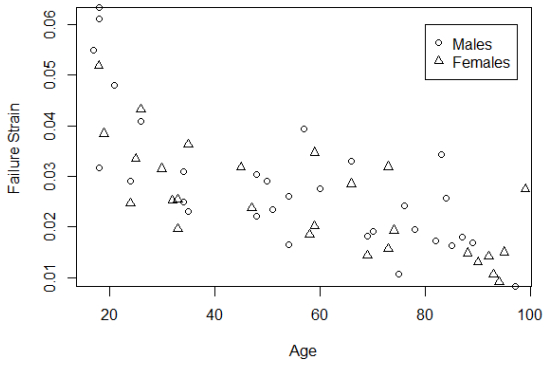


Figure A.1: Data points for failure strain versus Age, for a strain rate of 0.5 strain/s.

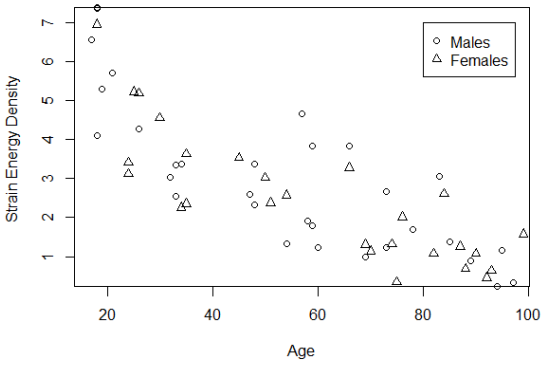


Figure A.2: Data points for total SED versus Age, for a strain rate of 0.5 strain/s.

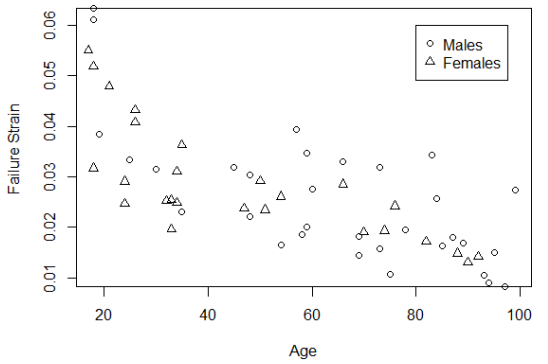


Figure A.3: Data points for failure strain versus Age, for a strain rate of 0.005 strain/s.

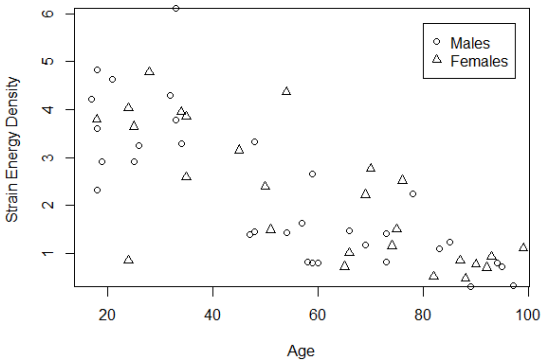


Figure A.4: Data points for total SED versus Age, for a strain rate of 0.005 strain/s.

B

Distribution Characteristics

Table B.1: Confidence intervals and relative size of the confidence intervals for 5%, 25% and 50% risk of injury for individuals of 25, 50 and 75 years of age, tests performed at a strain rate of 0.5 strain/s.

Distribution	Age	Risk of injury	Ultimate Strain					Total SED				
			Quantile location	Lower 95% CI	Upper 95% CI	Reltavie CI size	Quality Index	Quantile location	Lower 95% CI	Upper 95% CI	Relative CI size	Quality Index
Weibull	25	5%	0.0167	0.0133	0.0204	0.423	Good	1.4333	1.0265	1.9098	0.616	Fair
		25%	0.0279	0.0241	0.0319	0.281	Good	2.9326	2.3700	3.5624	0.407	Good
		50%	0.0362	0.0321	0.0408	0.237	Good	4.2246	3.5346	5.04	0.356	Good
	50	5%	0.0125	0.0101	0.0153	0.431	Good	0.883	0.6296	1.1862	0.63	Fair
		25%	0.0210	0.0185	0.0236	0.255	Good	1.8065	1.4806	2.1512	0.371	Good
		50%	0.0273	0.0248	0.0296	0.177	Good	2.6024	2.2626	2.9739	0.273	Good
	75	5%	0.0094	0.0075	0.0116	0.464	Good	0.5439	0.3747	0.7348	0.662	Fair
		25%	0.0158	0.0137	0.0181	0.289	Good	1.1129	0.8957	1.3569	0.414	Good
		50%	0.0205	0.0185	0.0230	0.215	Good	1.6037	1.3618	1.8792	0.323	Good
log-logistic	25	5%	0.0217	0.0183	0.0255	0.334	Good	2.0937	1.6428	2.6101	0.482	Good
		25%	0.0300	0.0265	0.0344	0.260	Good	3.3995	2.83705	4.0355	0.36	Good
		50%	0.0364	0.0326	0.0414	0.253	Good	4.5363	3.8681	5.2884	0.335	Good
	50	5%	0.0156	0.0135	0.0177	0.279	Good	1.153	0.9203	1.4005	0.388	Good
		25%	0.0215	0.0196	0.0235	0.175	Good	1.8721	1.6253	2.142	0.269	Good
		50%	0.0261	0.0241	0.0283	0.158	Good	2.4982	2.2126	2.81331	0.239	Good
	75	5%	0.0111	0.0096	0.0128	0.293	Good	0.635	0.4799	0.8044	0.478	Good
		25%	0.0154	0.0138	0.0173	0.218	Good	1.031	0.8599	1.2254	0.353	Good
		50%	0.0187	0.0169	0.0207	0.196	Good	1.3758	1.1741	1.6092	0.332	Good
log-normal	25	5%	0.0221	0.0189	0.0257	0.283	Good	2.0862	1.6612	2.6632	0.461	Good
		25%	0.0971	0.0260	0.0337	0.238	Good	3.3215	2.754	4.0998	0.387	Good
		50%	0.0364	0.0323	0.0412	0.241	Good	4.589	3.8014	5.6484	0.375	Good
	50	5%	0.0160	0.0140	0.0179	0.248	Good	1.1298	0.9119	1.3472	0.376	Good
		25%	0.0215	0.0197	0.0233	0.181	Good	1.7987	1.5467	2.0550	0.275	Good
		50%	0.0263	0.0243	0.0286	0.153	Good	2.4852	2.1686	2.8135	0.253	Good
	75	5%	0.0116	0.0102	0.0130	0.262	Good	0.6119	0.4845	0.7466	0.427	Good
		25%	0.0155	0.0141	0.0171	0.205	Good	0.9741	0.8191	1.1511	0.326	Good
		50%	0.0190	0.0174	0.0208	0.195	Good	1.3459	1.1519	1.5845	0.313	Good

Table B.2: Confidence intervals and relative size of the confidence intervals for 5%, 25% and 50% risk of injury for individuals of 25, 50 and 75 years of age, tests performed at a strain rate of 0.005 strain/s.

Distribution	Age	Risk of injury	Ultimate Strain					Total SED				
			Quantile location	Lower 95% CI	Upper 95% CI	Reltavie CI size	Quality Index	Quantile location	Lower 95% CI	Upper 95% CI	Relative CI size	Quality Index
Weibull	25	5%	0.0169	0.0136	0.0206	0.415	Good	1.2128	0.8702	1.6343	0.630	Fair
		25%	0.0273	0.0236	0.0314	0.284	Good	2.4734	1.9641	3.0482	0.438	Good
		50%	0.0349	0.0311	0.0395	0.241	Good	3.5574	2.9692	4.2622	0.365	Good
	50	5%	0.0125	0.0102	0.0152	0.398	Good	0.7154	0.5243	0.9537	0.600	Fair
		25%	0.0202	0.0180	0.0227	0.234	Good	1.4590	1.2176	1.7254	0.348	Good
		50%	0.0259	0.0238	0.0281	0.169	Good	2.0984	1.8476	2.3765	0.252	Good
	75	5%	0.0093	0.0073	0.0112	0.421	Good	0.4220	0.2965	0.5796	0.671	Fair
		25%	0.0150	0.0130	0.0170	0.265	Good	0.8606	0.6953	1.0604	0.424	Good
		50%	0.0191	0.0173	0.0211	0.198	Good	1.2378	1.0506	1.4499	0.323	Good
log-logistic	25	5%	0.0210	0.0180	0.0249	0.316	Good	1.6444	1.2791	2.0690	0.480	Good
		25%	0.0292	0.0259	0.0331	0.252	Good	2.6715	2.2087	3.2006	0.371	Good
		50%	0.0354	0.0317	0.0396	0.244	Good	3.5661	2.9915	4.2528	0.354	Good
	50	5%	0.0148	0.0129	0.0168	0.263	Good	0.9271	0.7454	1.1252	0.410	Good
		25%	0.0206	0.0188	0.0225	0.177	Good	1.5062	1.3183	1.7342	0.276	Good
		50%	0.0250	0.0229	0.0271	0.156	Good	2.0106	1.7995	2.2714	0.235	Good
	75	5%	0.0105	0.0090	0.0121	0.296	Good	0.5227	0.4124	0.6379	0.431	Good
		25%	0.0145	0.0130	0.0162	0.225	Good	0.8492	0.7193	0.9937	0.323	Good
		50%	0.0176	0.0160	0.0196	0.205	Good	1.1336	0.9803	1.3236	0.303	Good
log-normal	25	5%	0.0210	0.0179	0.0244	0.308	Good	1.6039	1.2694	2.0097	0.462	Good
		25%	0.0283	0.0247	0.0322	0.263	Good	2.5264	2.0796	3.0603	0.388	Good
		50%	0.0348	0.0306	0.0392	0.249	Good	3.4657	2.8699	4.1452	0.368	Good
	50	5%	0.0150	0.0131	0.0169	0.247	Good	0.9109	0.7331	1.0908	0.393	Good
		25%	0.0202	0.0186	0.0221	0.176	Good	1.4347	1.2458	1.6365	0.272	Good
		50%	0.0248	0.0230	0.0270	0.161	Good	1.9675	1.7351	2.2364	0.255	Good
	75	5%	0.0107	0.0094	0.0122	0.259	Good	0.5173	0.4187	0.6340	0.416	Good
		25%	0.0145	0.0131	0.0159	0.198	Good	0.8148	0.6922	0.9578	0.326	Good
		50%	0.0178	0.0161	0.0196	0.193	Good	1.1173	0.9577	1.3016	0.308	Good