



# VCU

Virginia Commonwealth University  
VCU Scholars Compass

---

Theses and Dissertations

Graduate School

---

2008

## COMPUTER-AIDED TRAUMA DECISION MAKING USING MACHINE LEARNING AND SIGNAL PROCESSING

Soo-Yeon Ji  
*Virginia Commonwealth University*

Follow this and additional works at: <https://scholarscompass.vcu.edu/etd>



Part of the [Computer Sciences Commons](#)

© The Author

---

Downloaded from

<https://scholarscompass.vcu.edu/etd/1628>

This Dissertation is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact [libcompass@vcu.edu](mailto:libcompass@vcu.edu).

School of Engineering  
Virginia Commonwealth University

This is to certify that the dissertation prepared by Soo-Yeon Ji entitled  
COMPUTER-AIDED TRAUMA DECISION MAKING USING MACHINE LEARNING  
AND SIGNAL PROCESSING has been approved by her committee as satisfactory  
completion of the Dissertation requirement for the degree of Doctor of Philosophy

---

Kayvan Najarian, Ph.D., Committee Chair, School of Engineering

---

Krzysztof Cios, Ph.D., Chair of Computer Science, School of Engineering

---

Lorraine M. Parker, Ph.D., School of Engineering

---

Susan S. Brilliant, Ph.D., School of Engineering

---

Sebastian Ventura Soto, Ph.D., University of Cordoba

---

Kevin R. Ward, M.D., School of Medicine

---

Rosalyn S. Hobson, Ph.D., Associate Dean of Graduate Studies, School of  
Engineering

---

Russell D. Jamison, Ph.D., Dean, School of Engineering

---

F. Douglas Boudinot, Ph.D., Dean, School of Graduate Studies

---

Date

© Soo-Yeon Ji, 2008

All Rights Reserved

COMPUTER-AIDED TRAUMA DECISION MAKING USING MACHINE  
LEARNING AND SIGNAL PROCESSING

A dissertation submitted in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy at Virginia Commonwealth University.

by

SOO-YEON JI

B.S., Kangwon National University (Chunchon, South Korea), 1992

M.S., Kangwon National University (Chunchon, South Korea), 1995

M.SE., University of North Carolina at Charlotte (Charlotte, NC), 2004

Director: KAYVAN NAJARIAN

ASSOCIATE PROFESSOR, DEPARTMENT OF COMPUTER SCIENCE

Virginia Commonwealth University  
Richmond, Virginia  
December, 2008

## Acknowledgments

Foremost, I would like to express my profound gratitude to my advisor, Kayvan Najarian, for his constant guidance from the very early stages of this research and for proving me with lot of chances to have extraordinary experiences throughout the work. His wide knowledge and passion for research have been of great value to me. His encouragement and guidance have provided a good basis for this dissertation.

I wish to express my sincere thanks to my committee, Dr. Krzysztof Cios, Dr. Lorraine M. Parker, Dr. Susan S. Brilliant, Dr. Kevin Ward, and Dr. Sebastin Ventura Soto for their support and constant encouragement. Also, their extensive discussions around my work have been very helpful for improving this dissertation.

My parents, my father Sung Ho Ji and my mother Myung Soon Lim, deserve a special mention for their inseparable support and endless love. Without their encouragement and understanding, it would have been impossible for me to finish this work. I also give my special gratitude to my brothers, my brother-in-law, and my nephews for their love and support. I greatly thank my lovely son, Ji Hun Lee, for his patience and support. There are no words to express my deep gratitude to my family. I am extremely fortunate in having such a great family and a charming son.

I would like to give my special thanks to Rebecca Smith, who helped improve this dissertation. And my special thanks to my friend, Dong-Hyun Jeong, who is giving me endless courage and love to maintain this dissertation.

## Table of Contents

	Page
Acknowledgments . . . . .	ii
List of Tables . . . . .	vii
List of Figures . . . . .	viii
Abstract . . . . .	ix
Executive Summary and Contributions . . . . .	xii
Chapter	
1 Introduction . . . . .	1
1.1 Traumatic Injury . . . . .	1
1.2 Motivation Behind Use of Medical Decision Making Systems . . . . .	2
1.3 Significance of Study . . . . .	3
1.4 Aims of Study . . . . .	5
2 Related work . . . . .	7
2.1 Existing Computer-Aided Decision Making Systems . . . . .	7
2.2 Benefits and Limitations of Logistic Regression and Decision Tree Algorithms . . . . .	10
2.3 Significance of the Machine Learning Approach . . . . .	11
3 Classification Using Significant Features (CUSF) . . . . .	14
3.1 Introduction . . . . .	14

3.2	Method . . . . .	16
3.2.1	Pre-Processing . . . . .	16
3.2.2	Significant Variable Extraction . . . . .	17
3.2.3	Rule Generation . . . . .	19
3.2.4	Rule Testing & Rule Validation . . . . .	20
3.2.5	Machine Learning Algorithms . . . . .	21
4	Application of Computer-Aided Decision-Making System to Traumatic Injury Data . . . . .	25
4.1	Rules for Prediction of Survival and Exact Outcome . . . . .	25
4.1.1	Description of Dataset . . . . .	26
4.1.2	Results . . . . .	28
4.1.3	Discussion . . . . .	31
4.1.4	Conclusion . . . . .	33
4.2	Rules for Prediction of Expected Time to Stay in Intensive Care Unit (ICU) for Airlifted Patients . . . . .	34
4.2.1	Introduction . . . . .	34
4.2.2	Description of Dataset . . . . .	35
4.2.3	Results . . . . .	36
4.2.4	Discussion . . . . .	38
4.2.5	Conclusion . . . . .	39
4.3	Rules for Prediction of Exact Outcome in Traumatic Pelvic Injury . . . . .	40
4.3.1	Introduction . . . . .	40
4.3.2	Description of Dataset . . . . .	41
4.3.3	Model Selection Comparison . . . . .	42
4.3.4	Exact Outcome Prediction . . . . .	44
4.3.5	Discussion . . . . .	46
4.3.6	Conclusion . . . . .	47

5	Testing & Validation Using Non-Medical Datasets . . . . .	49
5.1	Description of Datasets . . . . .	49
5.2	Comparison of Results . . . . .	50
5.3	Discussion . . . . .	51
5.4	Conclusion . . . . .	52
6	Heart Rate Variability (HRV) Analysis Using Signal Processing . . .	53
6.1	Introduction . . . . .	53
6.2	Existing Methods for Heart Rate Variability Analysis . . . .	55
6.3	Specific Aims . . . . .	56
6.4	Description of Dataset . . . . .	58
6.5	Methods . . . . .	60
6.5.1	ECG Segmentation . . . . .	62
6.5.2	Filtering . . . . .	62
6.5.3	QRS Detection . . . . .	63
6.5.4	Feature Extraction . . . . .	66
6.6	Results . . . . .	73
6.6.1	Comparison of LBNP and Exercise Using Wavelet Analysis	74
6.6.2	Comparison of LBNP and Exercise with Traditional Ap- proach . . . . .	75
6.6.3	Classification of Severity of Hemorrhage . . . . .	77
6.7	Discussion . . . . .	81
6.8	Conclusion . . . . .	82
7	Testing & Validation Using Gait in Aging Dataset . . . . .	84
7.1	Description of Dataset . . . . .	84
7.2	Results . . . . .	84
8	Conclusion and Future Work . . . . .	86
8.1	Summary . . . . .	86



8.2 Future Work . . . . .	89
Bibliography . . . . .	91
Appendices . . . . .	104
A Reliable rules for survival prediction . . . . .	105
B Reliable rules for exact outcome prediction . . . . .	107
C Reliable rules for ICU days prediction . . . . .	109
D Reliable rules for exact outcome prediction for pelvic injury . . . . .	111
E Detail non-medical dataset description . . . . .	113

## List of Tables

Table Number		Page
1	On-site dataset collected at the site of accident . . . . .	26
2	Dataset captured after hospital admission . . . . .	28
3	Significant variables of off-site dataset . . . . .	30
4	Performance comparison of five machine learning methods with survival prediction. . . . .	30
5	Performance comparison of five machine learning methods with exact outcome prediction. . . . .	30
6	Detail description of Helicopter dataset . . . . .	37
7	Prediction results for ICU days. . . . .	38
8	Summary of all available input variables . . . . .	42
9	Performance comparison with medical datasets . . . . .	44
10	Accuracy comparison of Pelvic dataset . . . . .	45
11	performance comparison with non-medical dataset . . . . .	51
12	The area under ROC curve comparison with non-medical dataset . . . . .	51
13	Summary of dataset including LBNP and exercise subjects . . . . .	58
14	Summary of statistical comparison between LBNP and exercise with SABWEF . . . . .	74
15	Traditional approach comparisons . . . . .	75
16	Study 1 classification result using ECG . . . . .	79
17	Summary of statistical comparison between Parkinson's disease and healthy subjects with wavelet features . . . . .	85
18	Traditional approach comparisons . . . . .	85

## List of Figures

Figure Number	Page
1 Comparative use of methods in analyzing medical data and trauma injury data . . . . .	10
2 Block diagram for entire approach. . . . .	16
3 Diagram of pre-processing procedures. . . . .	17
4 Reliable rule generation process. . . . .	19
5 Relations between attributes . . . . .	45
6 Detailed schematic diagram of entire process . . . . .	59
7 ECG signal before pre-processing . . . . .	60
8 Detailed schematic diagram of ECG analysis. . . . .	61
9 Example LBNP signal . . . . .	62
10 Diagram of filtering the ECG signal . . . . .	63
11 Example of 60Hz power-line interfering noise . . . . .	63
12 Detailed schematic diagram of QRS wave detection process. . . . .	64
13 Example of QRS detection steps . . . . .	66
14 Results of RR interval detection . . . . .	67
15 Detailed schematic diagram of feature extraction. . . . .	67
16 Detailed schematic diagram of DWT . . . . .	70
17 Diagram for statistical analysis between wavelet and traditional approaches . . . . .	73
18 Pattern comparisons (LBNP vs. Exercise) of wavelet features . . . . .	76
19 Pattern comparisons (LBNP vs. Exercise) of traditional approach . . . . .	77
20 Schematic diagram of classification . . . . .	78
21 Schematic diagram of classification using four signals . . . . .	80

## Abstract

### COMPUTER-AIDED TRAUMA DECISION MAKING USING MACHINE LEARNING AND SIGNAL PROCESSING

Soo-Yeon Ji

A dissertation submitted in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2008

Major Director: Kayvan Najarian  
Associate Professor, Department of Computer Science

Over the last 20 years, much work has focused on computer-aided clinical decision support systems due to a rapid increase in the need for management and processing of medical knowledge. Among all fields of medicine, trauma care has the highest need for proper information management due to the high prevalence of complex, life-threatening injuries. In particular, hemorrhage, which is encountered in most traumatic injuries, is a dominant factor in determining survival in both civilian and military settings. This complication can be better managed using a more in-depth analysis of patient information. Trauma physicians must make precise and rapid decisions, while considering a large number of patient variables and dealing with stressful environments. The ability of a computer-aided decision making system to rapidly analyze a patient's condition can enable physicians to

make more accurate decisions and thereby significantly improve the quality of care provided to patients. The first part of this study is focused on classification of highly complex databases using a hierarchical method which combines two complementary techniques: logistic regression and machine learning. This method, hereafter referred to as Classification Using Significant Features (CUSF), includes a statistical process to select the most significant variables from the correlated database. Then a machine learning algorithm is used to identify the data into classes using only the significant variables. As the main application addressed by CUSF, a set of computer-assisted rule-based trauma decision making system are designed. Computer aided decision-making system not only provides vital assistance for physicians in making fast and accurate decisions, proposed decisions are supported by transparent reasoning, but also can confirm a physicians' current knowledge, enabling them to detect complex patterns and information which may reveal new knowledge not easily visible to the human eyes. The second part of this study proposes an algorithm based on a set of novel wavelet features to analyze physiological signals, such as Electrocardiograms (ECGs) that can provide invaluable information typically invisible to human eyes. These wavelet-based method, hereafter referred to as Signal Analysis Based on Wavelet-Extracted Features (SABWEF), extracts information that can be used to detect and analyze complex patterns that other methods such as Fourier cannot deal with. For instance, SABWEF can evaluate the severity of hemorrhagic shock (HS) from ECG, while the traditional technique of applying power spectrum density (PSD) and fractal dimension (FD) cannot distinguish between the ECG patterns of patients with HS (i.e. blood loss), and those of subjects undergoing physical activity. In this study, as the

main application of SABWEF, ECG is analyzed to distinguish between HS and physical activity, and show that SABWEF can be used in both civilian and military settings to detect HS and its extent. This is the first reported use of an ECG analysis method to classify blood volume loss. SABWEF has the capability to rapidly determine the degree of volume loss from hemorrhage, providing the chance for more rapid remote triage and decision making.

## Executive Summary and Contributions

This work targets two types of complex problems: classification of complex/correlated dataset, and analysis of noisy/complex signals. For the first problem, a novel hierarchical method is developed for classification of highly complex and correlated dataset. This algorithm includes a statistical selection process in which modified version of logistic regression is employed to select the most significant variables in the data. Then a machine learning method is used to classify data into classes using only the identified significant variables. Extensive simulations on a number of databases indicate clear advantages of the proposed method over existing methods for a wide variety of applications.

In the second part of this work, a set of novel wavelet-based features, together with a systematic method to extract them from complex noisy signals, are introduced. These features are proved to be extremely efficient in distinguishing among signals that are very similar to each other. In other words, the introduced features can capture very detailed differences that cannot be detected by traditional signal processing methods such Fourier-based, Fractal-based, and power spectral density based methods. The capabilities of our measures are demonstrated in a number of applications. In all of these applications the novel measures distinguished classes that cannot be separated by any traditional method.

The target application for the developed algorithms is trauma decision making systems. The urgent need for computer-assisted decision making in medical applications is a consequence of the rapid development of novel medical measurement technologies,

fast growth in medical knowledge management, and the need for faster and more accurate decisions. Thus, finding the most relevant variables is critical in developing efficient computer-assisted decision making systems to assist physicians in patient treatment and resource allocation.

In this research, improving feature selection/feature extraction methods are focused in order to identify the most significant variables, considering the relationship of each variable with the outcomes. This will allow the development of systems that can provide physicians with the reasoning behind interesting new information/patterns which may have not been previously observed by experts.

The motivation for developing the decision-making system to analyze traumatic injury data is due to their prevalence and life-threatening nature of these injuries. The mortality rate for trauma patients is high, and the injuries themselves are typically very complex. Also, due to the rapid growth in medical knowledge, physicians may need assistance with exact reasoning when presented with large volumes of complex information. It is therefore increasingly important to develop trauma computer-aided decision-making systems that can improve the quality of patient care.

In this research, two types of algorithms are developed, for two types of medical applications.

- 1) In the first study, predictive computer-assisted rule based trauma decision making systems using a unified computational model that combines statistical regression and machine learning techniques is developed. It was shown that by selecting only the



statistically significant variables, regardless of the specific machine learning method used in the next step, the overall resulting predictive system is not only more accurate and more reliable but also provides predictions and recommendations that are transparent and easier for physicians to understand. Besides this novelty in selection and filtering of variables, in the machine learning step we prove that, despite the hypotheses made in some previous studies, rule-based systems can equal or exceed the performance of non-transparent methods such as neural networks and support vector machines in trauma decision making scenarios, if the models are trained and tested correctly.

- 2) In all types of trauma injury, the existence and severity of hemorrhage is a major factor in determining patient survival. In the second part of this research, novel approaches are proposed to analyze some major physiological signals to detect and evaluate the severity of a patient's hemorrhagic shock (HS). The physiological signals used in this research, such as Electrocardiogram (ECG), can provide invaluable information not visible to human eyes. Specifically, the severity of hemorrhage is often evaluated using heart rate variability (HRV) analysis of the ECG. Currently, power spectral density (PSD) and fractal domain (FD) are the common HRV analysis tools. However, it has been shown that these traditional methods are unable to distinguish between central volume loss (the main indicator of HS) and physical activity (such as typical daily exercise), as they have similar HRV patterns. This is problematic given the desire to use changes in heart rate to detect the presence of acute volume

loss due to hemorrhage. In this research, a new HRV analysis approach is proposed based on wavelet transformation to determine blood loss severity. In addition, these methods can identify the degree of hemorrhage shock. Again, the combined set of algorithms proposed for this study have the to be used for other biomedical and non-biomedical signals.

The following sections provide more details on these new computational approaches/methodologies.

### ***Section 1. Computer-Aided Trauma Decision Making using Machine Learning***

None of the existing trauma decision-making algorithms is in widespread use in trauma centers because: 1) they use non-transparent methods; 2) the performance of these algorithms is typically poor due to the exclusion of relevant attributes and the inclusion of some that are irrelevant to the task at hand. Inclusion of irrelevant variables results in rules that are too complicated to be clinically meaningful.

Predictive computer-assisted rule based decision making system is developed by combining two complementary techniques: logistic regression and different families of machine learning methods. Logistic regression is useful in describing relationships among multiple independent variables and a specific outcome. Furthermore, rule-based methods in machine learning are easy to understand and interpret while also being capable of dealing with categorical variables and missing values. Figure A shows the high-level schematic diagram of the proposed system and it contains four main tasks: pre-processing, extraction of significant variables, rule extraction, and rule testing and evaluation.

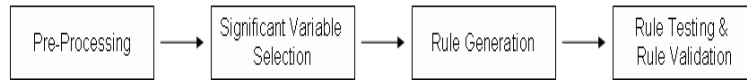


Figure A: Block diagram for entire approach.

It is considered that appropriate feature selection may have a critical impact on prediction accuracy. Therefore, logistic regression is used for selecting significant features and the proposed method (in Chapter 5) is evaluated by comparing the results of machine learning classification using all available variables and only significant variables. In order to examine the significance of the individual variables, logistic regression performs the likelihood ratio significance test. In logistic regression, the stepwise model selection method is commonly used to find the best subset of variables to predict the outcomes, considering all possible combinations. As part of the process, any single predictor variable may be added or deleted. There are major issues with stepwise model selection for medical application. Although the method is designed to find important variables it does not guarantee that the most significant variables are selected, due to the repetition of insertion and deletion. The proposed approach has improved the performance over the stepwise method (Chapter 4 and Chapter 5).

For rule generation, two decision tree algorithms, CART [22] and C4.5 [120, 119, 118], are used to generate the rules. These are then evaluated using 10-fold cross validation. A set of reliable rules is created after rule filtering and assessment by physicians of the clinical merit of the rules. The resulting transparent rule set provides physicians with the exact reasoning behind the system's recommendations and predictions. This research also proves that for a wide range of trauma applications, rule-based methods such as CART and

C4.5 can be trained in such a way that the accuracy, sensitivity, and specificity is either competitive with or even better than non-transparent methods such as support vector machines and AdaBoost. Those machine learning algorithms are described in Chapter 3. This is a finding that, contrary to some reports in the literature, encourages the use of rule-based systems for many medical and non-medical applications. It is hypothesized that one factor in making rule-based systems more reliable and accurate is the use of a “significant feature selection” step as a pre-processing stage, as proposed by this research.

The proposed hierarchical model, CUSF, can not only improve feature selection in order to identify the most significant variables, but also provide physicians with the reasoning behind interesting new information / patterns that may have not been previously observed.

### ***Section 2. Heart Rate Variability (HRV) Analysis Using Advanced Signal Processing***

A new method to analyze HRV is proposed, based on wavelet transformation of Electrocardiogram (ECG) data. The capability of the algorithm to distinguish between ECG for subjects undergoing lower body negative pressure (LBNP) that simulates the loss hemorrhage in human subjects, and subjects doing physical activity is performed. Figure B presents a schematic diagram of the overall method and its comparison with the traditional HRV analysis technique.

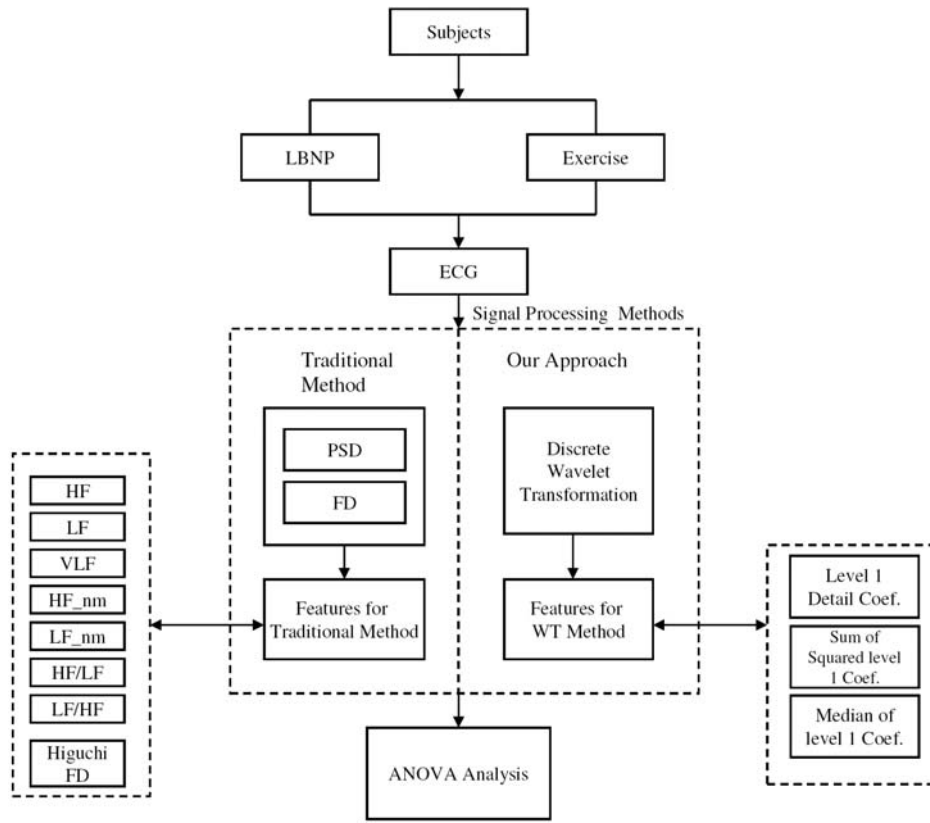


Figure B: Detailed schematic diagram of entire process - multiple tasks are performed on ECG data.

For any HRV analysis, the QRS complex (see Fig 7) must be detected first as it is the most significant waveform in an ECG signal, and the HRV is generally extracted from the ECG recording by detecting RR intervals (Section 5 in Chapter 6). Figure C presents the schematic diagram of the process of analyzing ECG and extracting RR intervals.

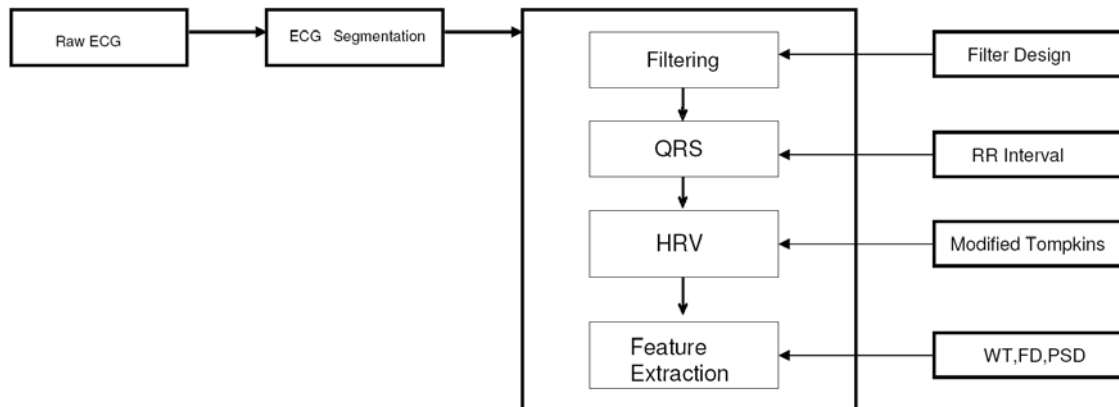


Figure C: Detailed schematic diagram of ECG analysis.

In order to detect the QRS complex, modified version of the Pan-Tompkins algorithm [109] is used. Because subject movement (including minor muscular activity) can cause high frequency noise components or other types of electromagnetic interface, a modified Pan-Tompkins algorithm is used by adding a histogram process step in which the characteristics of the signal's amplitude distribution are analyzed and used for extra filtering of the signal.

Once the HRV is extracted based on the RR interval, the traditional approach (power spectral density and fractal dimension) is applied and compared with the new approach (wavelet transformation). The results with the state-of-the-art in the field are compared and show the advantages of the new method over the existing methods. The traditional method of HRV analysis via power spectral density (PSD) uses an average power with certain ranges of frequencies, and applies the Fast Fourier Transformation (FFT) for calculations. However, FFT cannot reliably be used to process non-stationary signals such as ECG. This renders the traditional HRV analysis method incapable of distinguishing

between the ECG patterns in HS (or LBNP) and physical activity.

In order to overcome this problem, a novel approach, called Signal Analysis Based on Wavelet-Extracted Features (SABWEF), is designed to deal with non-stationary signals. Since wavelet transformation calculates the similarity between the input signal and a mother wavelet, and the Daubechies families of wavelets have the most similar shape to the QRS complex, Daubechies wavelets for ECG analysis are used. The results using Daubechies 4 and Daubechies 32 are compared. Since the results show no significant differences between the two wavelets, Daubechies 4 is used for the full study.

SABWF provides a specialized method of QRS detection, filtering, wavelet decomposition, and feature extraction for ECG analysis. All these techniques as well as their combination can be used to analyze other biomedical and non-biomedical signals.

Unlike traditional HRV analysis, i.e. SABWEF, can differentiate between cases of physical activity and hemorrhage shock. The method may also be useful in quickly determining the degree of volume loss due to hemorrhage. This would prove an invaluable tool in rapid patient triage and decision making.

## CHAPTER 1 Introduction

### 1.1 Traumatic Injury

Many Americans experience minor and major accidents in their daily lives, some of which are serious injuries that can result in death or permanent disability; of these, traumatic injuries are the most prevalent. Patients who survive a head injury may suffer serious consequences, including lifelong paralysis and severe disability [108, 148]. According to the Federal Centers for Disease Control and Prevention (CDC) in 2004, 1.4 million Americans sustain a traumatic brain injury (TBI) each year, and 50,000 of those individuals die as a result of these injuries [26]. Since mild TBI does not affect life expectancy, young people potentially face several decades of disability. Specifically, it is estimated that each year approximately 100,000 children are permanently disabled. TBI accounts for about 29,000 of these cases, and a significant percentage result in neurological impairment [27, 49, 11, 131].

A potentially fatal consequence of traumatic injury is hemorrhagic shock (HS). This is particularly common on the battlefield, where it accounts for nearly 50% of trauma deaths, but is also encountered in civilian life, where it accounts for 39% of trauma deaths. A study of the Israeli military found that in 96% (351 out of 337) of the patients, death due to blood loss occurred within the first four hours after injury. The major causes of the deaths in this group were hemorrhage (50%) and neurological trauma (36%); the rest were



due to severe multiple injuries [19, 134, 158]. The short period between injury and death and the high rate of complications are both associated with a lack of appropriate medical attention and limited evacuation facilities in the field [8, 47].

## ***1.2 Motivation Behind Use of Medical Decision Making Systems***

Due of the rapid growth of medical information and knowledge available at the time of decision making, physicians have become more reliant on assistance from other experts in cases outside of their own area of expertise. Furthermore, physicians who first diagnose a patient must choose from a variety of expensive medical tests to help ensure correct diagnosis and optimal therapeutic management. While these tests provide invaluable information, decision making based on this continuously growing collection of information is a challenging task. Early and effective trauma patient control can improve the chance of survival more than any other measure; however, achieving this is far from simple. In 1996 and 2000 (updated in 2003), the Brain Trauma Foundation published guidelines for the management of severe TBI [54], which were accepted by the American Association of Neurosurgeons and endorsed by the World Health Organization Committee in Neurotraumatology. These guidelines provide valid criteria to identify high-risk patients, with the aim of reducing inappropriate care, controlling geographic variations in practice patterns, and maximizing health care resources. However, the implemetation of practice guidelines was slow due to deficiencies in physician training [101].

In developed countries, the rate of growth of health care expenditures has exceeded that of growth in income for a considerable period of time. For example, in the United

States health care expenditures as a share of gross domestic product (GDP) have tripled since 1950, from 5% then to 15% today [46, 66]. This increase in spending has far exceeded the supply for healthcare revenue and services, and interest in medical decision support systems has risen accordingly due to the need for cost-control and improvement in quality of care.

In a recent study of error in medicine [24], a database was constructed from the 1992 American Hospital Association with 1,116 hospitals participating. The study found 17,338 medication errors that adversely affected patient outcome; furthermore, medication errors occurred in 5.07% of the patients admitted each year, and on average each hospital experienced a medication error every 22.7 hours. Leape [87] found that the average intensive care unit is subject to an even higher error rate - almost two errors per day in each unit- that can have serious and possibly fatal consequences. This illustrates the clear importance of computerized decision making systems as a research issue.

### **1.3 Significance of Study**

The clinical significance of a computer-aided decision making system lies in its ability to improve diagnosis and care by helping physicians make more accurate decisions in a stressful environment, and to provide the reasoning behind all recommendations and predictions. Moreover, by examining the decision-making process using a qualitative methodology, new knowledge can be gained, particularly in the planning of long-term care. An accurate decision-making system may also be of use in rural and remote areas where physicians with extensive trauma experience may not be available, and may be effective in classroom

education of medical students.

In the computer science field, the proposed methods, classification using significant features (CUSF), provide the following novelties and contributions:

1. In the first part of the dissertation, a hierarchical method, CUSF, for classification of highly complex and correlated dataset is presented. This method uses statistical filtering where a modified version of logistic regression is employed to select the most significant variables, and then a machine learning method is used to classify the data into functional classes using only the identified significant variables. Extensive simulations with both medical and non-medical databases prove the superiority of the proposed method over existing methods over a wide variety of applications. We also show that the application of the proposed method, along with rule extraction machine learning methods such as CART, on trauma databases provides rules that can change the way medical decision-making systems are perceived and used by medical community.
2. The second part of the dissertation introduces a set of novel wavelet-based features, signal analysis based on wavelet-extracted features (SABWEF), and a systematic method to extract them from complex noisy signals. These features are proved to be extremely efficient in distinguishing among functional classes that are very similar to each other in signal level, i. e. the features can capture very detailed differences that cannot be detected by traditional signal processing methods such Fourier-based, Fractal-based, and power spectral density based methods. The capabilities of the wavelet-based features are tested, verified, and demonstrated against a number of

applications, in all of which these features distinguished classes that cannot be separated by any traditional method. These applications include the heart rate variability (HRV) analysis that has tremendous impact on patient care. Thus, the proposed wavelet based features may be useful to extract hidden knowledge from highly noisy and complex signals as well as HRV signal.

#### **1.4 Aims of Study**

The main aim of this study is to develop a computer-aided trauma decision making system that integrates all relevant knowledge from medical records, and ultimately generates reliable rules which can support clinicians in applying all information to provide better care to patients. The specific aims of this research are described below:

1. Apply machine learning methods, specifically decision tree algorithms, to extract rules directly from datasets and provide physicians with the reasoning behind them. Show that transparent rule-based systems perform as well as other methods (such as neural networks).
2. Show that developed hierarchical method is well performed on medical dataset as well as non-medical dataset.
3. Analyze heart rate variability (HRV) using wavelet transformation to predict severity of hemorrhage shock (HS). Use advanced signal processing to differentiate between cases of blood loss and physical activity, which present similar HRV patterns in test subjects.
4. Define new features based on discrete wavelet transformation (DWT) of ECG that

can be used to estimate blood loss severity. The features will be defined based on the energy of detail coefficients of Daubechies DWT. Also, the introduced novel features are tested using gait aging signal for further validation.

5. Apply machine learning algorithms to the extracted informative features to predict the severity of blood loss. Statistical analysis is used to validate the extracted features.

The remaining chapters of this thesis are organized as follows. Chapter 2 describes existing work and the benefits and limitations of comparable medical diagnosis techniques. Chapter 3 explains the chosen approaches and methodologies, including the motivation behind their use as well as their significance. Chapter 4 presents application results in predicting patient survival (alive/dead), exact outcome (rehab/home), and pelvic injury severity. In Chapter 5, testing and validation results using non-medical datasets is presented. Another specific approach using signal processing methods for heart rate variability (HRV) analysis is given in Chapter 6. In Chapter 7, the wavelet method features, SABWEF, are tested to gait aging dataset with two conditions (healthy and Parkinson's disease) in order to validate the applicability of our approach. The final chapter presents the conclusions and discussion of this study as well as describing future work and possible enhancements.

## CHAPTER 2 Related work

This chapter describes previous work in developing computer-aided decision-making systems. First, the motivation and background information about such systems are described. Next, machine learning algorithms and their significance in this field are presented. Finally, the benefits and limitations of logistic regression, which is used frequently in analyzing medical data, and decision trees algorithms are described.

### *2.1 Existing Computer-Aided Decision Making Systems*

Several computer-assisted systems exist for decision-making in trauma medicine. For instance, HELP, a hospital information system, has been used at Latter-Day Saints (LDS) Hospital since 1967. Based on an evaluation report of the system [59], HELP has been used by clinical staff because its computerized clinical decision-support provides improvements in patient care. However, the system suffers from several shortcomings: it uses limited patient information, typing information into the system takes time, and the user must wait a significant time for the results. The last disadvantage, i.e. delay, is a major reason why HELP and other decision support systems are not in widespread use [157]. Furthermore, the HELP system is not integrated with the electronic medical records, and the data storage requirements are substantial.

The majority of medical decision-making systems perform a statistical survey of similar cases in trauma database with patient demographic information [53, 65]. As such, they may

not be sufficiently accurate and/or specific for practical implementation. Another issue is the use of artificial neural networks in medical decision-making systems. Even though neural networks have good performance, due to the ‘black box’ nature of neural networks, the knowledge stored in the trained networks is not transparent and the reasoning behind the predictions and recommended decisions is obscured [93, 75].

Bayesian statistical technique is also used by calculating the prior probability of a disease and conditional probabilities of its symptoms. It shows fairly good performance, but has some limitations. Specifically, the assumption of conditional independency of attributes and mutual exclusiveness may not be satisfied in medicine, as overlapping disease categories are common in the real world. Another critical issue is that this method requires a large database to accurately determine all conditional probabilities [121, 135].

As an example of a specific application, studies have been performed to develop an easy and simple predictive model for survival of TBI patients. Signorini [137] presents a simple model that predicts patient survival using age, Glasgow Coma Score (GCS), Injury Severity Score (ISS), pupil reactivity, and the presence of hematoma on CT scans. Although the model is efficient in usage and well-designed, the accuracy and reliability of its rules may be limited due to the small number of variables used.

In general, computer-aided systems have the potential to significantly improve trauma decision making and resource allocation, and it is reported that since trauma injuries generally have specific causes with established methods of treatment, fatal complications and long-term disabilities can be reduced by making less subjective and more accurate decisions in trauma units [108]. In addition, it has been suggested that the cost of trauma care may be significantly reduced by an inclusive trauma system with an emphasis on

computer-aided resource utilization and decision-making [10].

Current guidelines for trauma care in the battlefield also may be significantly improved by continuously observing the patients and their biomedical signals. This can aid the early detection of severe blood loss. Therefore, the most important factors in field care of trauma patients are appropriate training of the medical personnel and sufficient preparation for environmental conditions [44, 83]. Also, early and effective hemorrhage control may improve the chance of survival more than any other measure.

However, even though medical protocols have been developed for the successful management of trauma patients, currently there is no widely used computer-aided system that integrates and processes patient information, compares the current cases with those previously observed, and evaluates the severity of the case to enable faster and more reliable decision making. Therefore all the issues mentioned above should be addressed.

In summary, there are three main reasons why no existing decision-making system is in widespread use in trauma centers:

- 1) The use of non-transparent methods, such as neural networks.
- 2) The lack of a comprehensive database integrating all relevant available patient information.
- 3) Poor performance due to the exclusion of relevant attributes and the inclusion of those irrelevant to the task in hand, resulting in rules that are too complicated to be clinically meaningful.



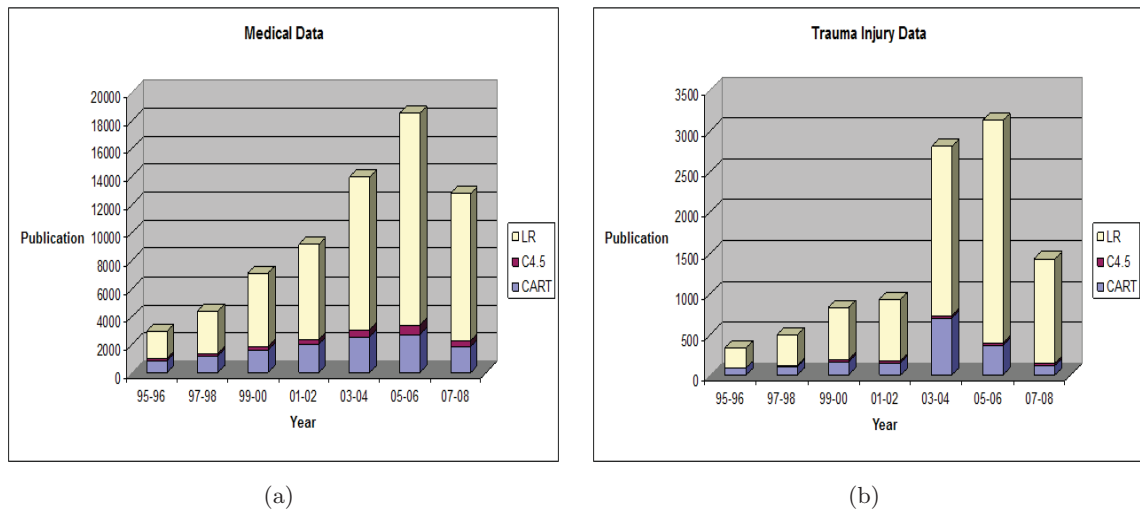


Figure 1: Total number of published papers within two year periods from 1995 to 2008 in PubMed, CiteSeer, and Google Scholar. The charts show how many papers were published related to the analysis of medical data 1(a) and trauma injury data 1(b).

## 2.2 Benefits and Limitations of Logistic Regression and Decision Tree Algorithms

Over the past 20 years, many comparative studies between decision trees and logistic regression (LR) have been performed [92]. Generally, the Classification and Regression Trees (CART) algorithm is compared with LR. Figure 1(a) shows the total number of published papers related to the analysis of medical data, using several algorithms such as CART, C4.5, and LR. Figure 1(b) displays the same information relating specifically to traumatic injury. It is clear that the rate of increase in the use of logistic regression has exceeded that of CART and C4.5.

William [92] compares one of the decision tree algorithms, Iterative Dichotomiser 3 (ID3) algorithm, to logistic regression (LR) using a database of 5,773 cases. His study found that LR outperforms ID3. However, LR analysis is very difficult to use in clinical

applications, particularly when the outcome variable has more than two values. Furthermore, Kuhnert [84] emphasizes that non-parametric methods such as CART can provide more informative variables than LR. His study also states that CART is more useful in the medical field where access to “rule-like” models for decision making is extremely important [91]. However, many previous studies [92, 112, 151, 145, 127, 62] have compared the performance between LR and decision trees, and the results suggest that there is no completely preferable method; the result of comparison depends upon the chosen application.

Although some studies have found that LR outperforms CART, note that there are certain conditions on these comparisons. For instance, LR performs better for smaller datasets and it also provides a concise summary of the relationships between the outcomes and the predictors, which decision trees are unable to do. However, decision trees are easy to understand and the position of a predictor variable at the root shows its dominance compared with other variables. In addition, CART is likely to be more practical in a clinical setting due to its easy interpretation [92, 151], and is useful in decision making for long-term patient care [117]. Therefore, the interdependent use of both techniques may be a promising approach.

### ***2.3 Significance of the Machine Learning Approach***

Machine Learning (ML) [21, 10, 11, 102, 79] is an Artificial Intelligence (AI) technology that has been employed in a variety of statistical, probabilistic and optimization tools to automate complex decision making and problem solving tasks. Many statistical methods

are designed based on multivariate regression or correlation analysis. These approaches, although powerful, assume that the variables are independent and that the data can be modeled using linear combinations of these variables. When the relationships are non-linear and the variables are not independent, these methods cannot be applied. Since many biological systems are fundamentally nonlinear and their parameters conditionally dependent, machine learning algorithms are often more appropriate than statistical methods [138, 139].

There are several reasons why machine learning is popular for medical applications:

- 1) It is possible to build a model from the data which may help the physician, at the first examination of the patient, to decide the severity of the injury or disease, and whether the patient should be admitted to the hospital or could be treated as an outpatient.
- 2) It is possible to extract hidden relationships and correlations among the data.
- 3) Environments change over time. Systems that can adapt to a changing environment would reduce the need for constant redesign.
- 4) Missing values are a common problem in medical applications, and some machine learning algorithms are able to deal with them [90].

Note, however, that the success of machine learning is not always guaranteed. If the dataset is of poor quality, the results may follow suit. A minimum requirement for any machine learning technique is a sufficiently large dataset that can be partitioned

into disjoint training and test sets or subjected to some reasonable form of n-fold cross-validation. Machine learning algorithms that are commonly used in the medical field include support vector machines (SVM) and decision tree algorithms such as Classification and Regression Trees (CART) and C4.5. in Chapter 3, These methods are explained in detail.

## CHAPTER 3 Classification Using Significant Features (CUSF)

This chapter introduces a classification method for highly complex datasets using a hierarchical approach that combines two complementary techniques; logistic regression and machine learning. This method is referred to as Classification Using Significant Features (CUSF). The method incorporates only the most significant variables selected from the potentially correlated input data and uses machine learning algorithms to classify data using only the significant variables.

The rest of this chapter is organized as follows. Section 3.1 introduces the proposed approach. A detailed description of the methods is presented in Section 3.2, including four machine learning algorithms.

### **3.1 Introduction**

Decision tree algorithms, specifically CART and C4.5, are used for rule generation. CART and C4.5 deal effectively with missing values and categorical variables, leading to their widespread use in medical informatics. Kononenko [81] compares some techniques such as Bayes' theorem, neural networks, and decision trees by considering performance, transparency, explanation of reasoning, dealing with missing values, and reliable prediction with small dataset. Note that decision tree algorithms may better satisfy those characteristic of medical diagnosis. However, despite the relatively successful performance of these algo-

rithms in medical applications, they have had limited success in separating and identifying important variables in applications where there are a large number of available attributes. This suggests that combining machine learning with a statistical method to identify the most informative variables can increase our understanding of the patterns in medical data and thus help generate more reliable rules.

Therefore, the use of logistic regression, which provides knowledge of the relationships among the multiple independent variables and the response variable, is considered and therefore it is useful in finding statistically significant variables to model tasks with binary outcomes. Creating rules using all available variables can lead to the inclusion of less relevant or reliable attributes, which can then result in random correlations and the generation of rules which are clinically meaningless. Retaining the less informative variables and/or including highly correlated attributes may also increase the complexity of the rules and make them less suitable for clinical applications. Appropriate feature selection therefore has a significant impact on the accuracy of prediction.

A rule-based computer aided system is proposed to predict the survival (alive or dead), exact outcome (home or rehabilitation), and severity (measured as the number of days stayed at ICU) using traumatic injury dataset. The final rule-based decision making system can provide the reasoning behind its recommendations, and incorporate information from future cases.

The generated rules are also tested against support vector machine (SVM), AdaBoost, and stand alone logistic regression. Though these methods do not generate rules, they are tested in the interest of performance comparison. The intention is to examine whether there are rule-based systems that can compete with or beat non-transparent machine

learning methods in terms of accuracy and performance. Testing other algorithms also validates the accuracy and stability of the rule-based system. Then quantitative measures of the statistical reliability and the accuracy of the resulting predictions and recommendations are tested.

### 3.2 Method

In this section, CUSF methodology, which combines logistic regression and machine learning algorithms, is explained to generate reliable rule based systems. Figure 2 presents a block diagram of the proposed approach. It consists of four steps: pre-processing, significant variable selection, rule generation, and rule testing and validation. Each step is designed to support creation of the most reliable rules.



Figure 2: Block diagram for entire approach.

#### 3.2.1 Pre-Processing

In general, medical datasets contain continuous variables and nominal (categorical) variables. Figure 3 explains the pre-processing procedure in more detail.

As part of the pre-processing step, every nominal variable is replaced with several dichotomous variables. For example, there are seven types of complication, and each type is treated as an individual variable having two levels (Yes/No).

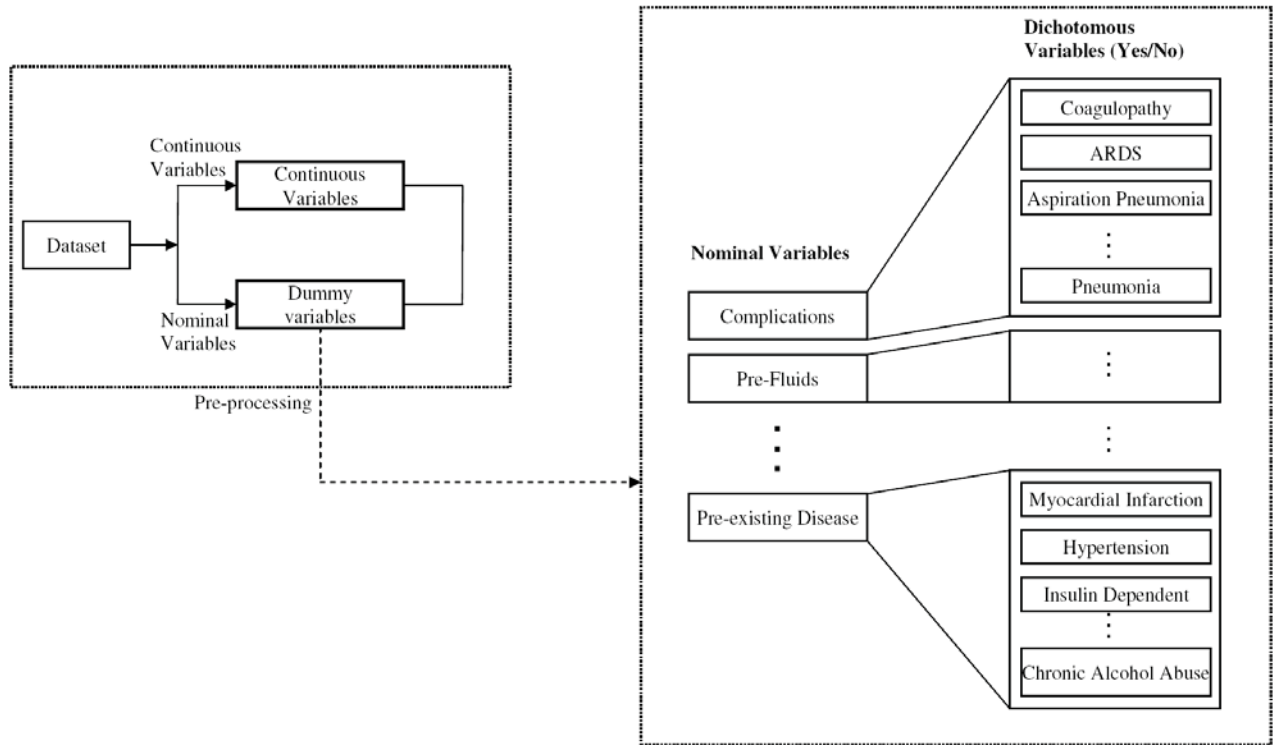


Figure 3: Diagram of pre-processing procedures.

### 3.2.2 Significant Variable Extraction

Logistic regression (LR) is used to extract significant variables. The logistic function is used to calculate the expected probability of a dichotomy as follows:

$$\pi_i = pr(Y = 1|X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots)}} \quad (1)$$

where  $X_i$  are variables with numeric values,  $Y$  is the outcome (dichotomous; 0/1, e.g. Alive/Dead), and the  $\beta$ 's are the regression coefficients that quantify the contributions of the numeric variables to the overall probability [64]. Logistic regression provides knowledge of the relationships and strengths among the multiple independent variables and the response variables. It does not assume any distribution on the independent variables; they



do not have to be normally distributed, linearly related or of equal variance within each group [72, 152]. However, logistic regression does require a linear relationship between the log-odds of outcomes and predictors. The linearity assumption has been checked by categorizing the predictors into a number of categories with equal intervals, then using the Hosmer-Lemeshow (H-L) goodness-of-fit test under the hypothesis that the model set is good. For our dataset, the H-L test, performed using the Statistical Analysis Software (SAS), resulted in a non-significant p-value, indicating that the linearity assumption is valid.

To test the significance of the individual model parameter, logistic regression uses likelihood testing. The likelihood ratio test itself does not tell if any particular independent variables are more important than others. However, the difference between the full model and a nested reduced model which drops one of the independent variables can be analyzed. The difference in -2log-likelihood (-2LL) using maximum likelihood estimation is compared for the full model (L1) and the reduced model (L0):

$$-2 \log\left(\frac{L_0}{L_1}\right) = -2[\log(L_0) - \log(L_1)] = -2(L_0 - L_1) \quad (2)$$

A non-significant difference indicates no effect on performance of the model, hence we can justify dropping the given variable. For this study, only the significant variables (p-value  $\leq .05$ ) are selected. SAS is used to calculate the significance of individual attributes.

Note that stepwise model selection is also available to discover the significance of variables. In the study of statistical regression, the stepwise method is commonly used to find the best subset of variables for outcome prediction, considering all possible combinations of variables. However, the stepwise approach may not guarantee that the most significant

variables are selected due to the repetition of insertion and deletion. For example, age may not be selected as an important variable [144, 150]; however, physicians may believe that patient age is very important in deciding treatment.

The fact that appropriate feature selection may have a critical impact on prediction accuracy is postulated. Thus, the results generated using all available variables and those generated using only significant variables are compared.

### 3.2.3 Rule Generation

Two decision tree algorithms, CART and C4.5, are used to generate the rules using only significant variables. SVM, AdaBoost, and logistic regression are also tested in the interest of performance comparison. Figure 4 describes the rule generation procedure in more detail.

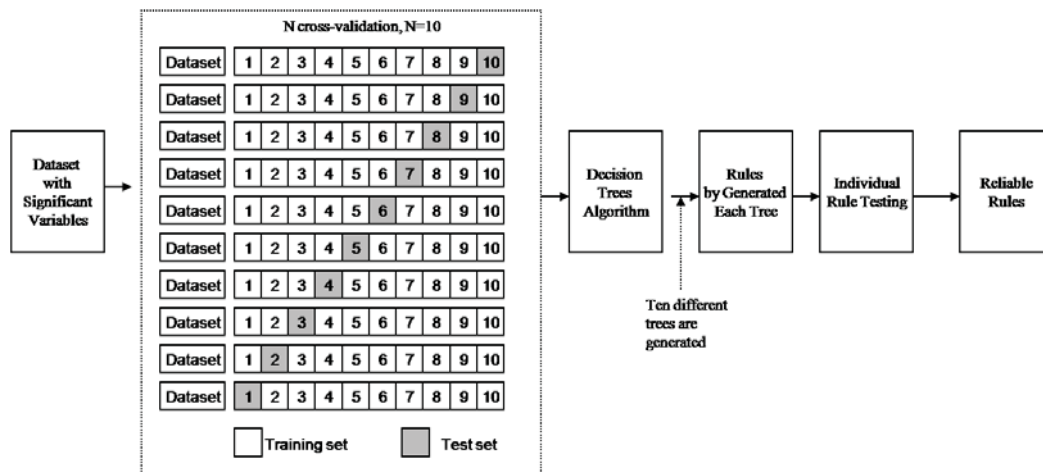


Figure 4: Reliable rule generation process.

Ten-fold cross validation is performed to measure the quality and scalability of the rules. This involves partitioning the data into ten subsets and testing whether the subsets

have a similar outcome distribution. The datasets are divided into  $N$  mutually exclusive subsets, and at each step one is used as the validation set and the other  $N-1$  form the training set (in Figure 4,  $N=10$ ). This is repeated  $N$  times, so each of the subsets is used as the validation set exactly once. The results are then averaged to generate a final estimate. Repeating the analysis multiple times has a considerable computational cost, but the advantage is that it does not matter how the data set is partitioned; every data value will be in a testing set once, and in a training set nine times. The accuracy when using all available variables and using only the significant variables is compared with the machine learning algorithms using 10-fold cross validation.

#### 3.2.4 Rule Testing & Rule Validation

Once rules are generated, each individual rule is tested to measure individual rule accuracy. Then, all rules are evaluated by physicians. Consequently, only rules with both high accuracy and a sufficiently large number of supporting examples are used to form the rule base. The filtered rules therefore both confirm to the physicians' existing knowledge and enable them to analyze new interesting patterns that may reveal new facts. The existing knowledge may therefore be improved.

The next step is to measure rule sensitivity and specificity. Let  $FP$ ,  $TP$ ,  $FN$ , and  $TN$  be the number of false positives, true positives, false negatives and true negatives, respectively. The following measures are calculated:

$$Sensitivity = \frac{TP}{(TP + FN)} \quad (3)$$

$$Specificity = \frac{TN}{(FP + TN)} \quad (4)$$

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (5)$$

Non-medical dataset is also tested to validate this feature selection approach in Chapter 5.

### 3.2.5 Machine Learning Algorithms

In this section two decision tree algorithms, CART and C4.5, are described. Also, two other machine learning algorithms, AdaBoost and SVM, are explained.

#### C4.5

C4.5 [120, 119, 118] extends Quinlan's basic ID3 decision tree algorithm [118]. It is more successful in avoiding overfitting, is able to handle continuous variables, and is more computationally efficient. The original ID3 algorithm calculates information gain when choosing attributes:

$$Info(S) = - \sum_{i=1}^n p(k_i, S) \cdot \log_2 p(k_i, S) \quad (6)$$

where  $p(k_i, S)$  is the relative frequency of examples in  $S$  that belong to class  $k_i$ .

However, ID3 is biased when an attribute has many values. C4.5 uses gain ratio to select attributes. Gain ratio is a modification of the information gain that reduces bias, calculated as:

$$GainRatio(S) = \frac{Info(S)}{- \sum_{i=1}^n \frac{p(k_i, S)}{|S|} \log_2 \frac{p(k_i, S)}{|S|}} \quad (7)$$

where  $|S|$  is the size of  $S$ .

To generate rules, C4.5 uses a divide-and-conquer algorithm to split training data into disjoint regions of the variable space, according to pre-assigned target labels [128]. At

each step, C4.5 splits on the best attribute according to the gain criterion. This criterion is based on entropy, i.e. the randomness of the class distribution in the dataset. The criterion is the greatest difference in entropy of the class probability distribution of the current subset  $S$  and the subsets generated by the split.

The best split is the one that most reduces this value. The output of the algorithm is a decision tree, which can be easily represented as a set of symbolic IF-THEN rules.

#### *Classification and Regression Tree (CART)*

CART, designed by L. Breiman [22], applies information-theoretic concepts to create a decision tree. This allows for the capture of rather complex patterns in data, and their expression in the form of transparent grammatical rules [91]. CART's nonlinear extensions are widely used in data mining and machine learning due to the algorithm's efficiency in dealing with multiple data types [57] and missing data. For missing values, CART simply uses a substitution value, having the most similar split with them [22]. In addition, CART supports an exhaustive search of all variables and split values to find the optimal splitting rules for each node. CART uses the Gini index in order to choose attributes:

$$Gini(S) = 1 - \sum_{i=1}^n p_i^2 \quad (8)$$

where  $p_i$  is the relative frequency of class  $i$  in  $S$ . After splitting  $S$  into two subsets  $S_1$  and  $S_2$  with sizes  $M_1$  and  $M_2$ , the Gini index of the split data is defined as:

$$GiniSplit(S) = \frac{M_1}{|S|} Gini(S_1) + \frac{M_2}{|S|} Gini(S_2) \quad (9)$$

where  $|S|$  is the size of  $S$ . Thus, the smallest Gini split is chosen as the split node.

The splitting stops at the pure node containing the fewest examples [60].

#### *Adaptive Boost (AdaBoost)*

AdaBoost, introduced by Freund and Schapire [55], is an algorithm that constructs a robust classifier as a linear combination of weak classifiers. Adaboost repeatedly calls a given weak learning algorithm in a set of rounds  $t = 1, \dots, T$ . A distribution of weights is maintained over the training set, such that  $D_t(k)$  is the distribution's weight for training example  $k$  on round  $t$ . The aim of weak learning is to find a good weak hypothesis  $h_t : X \rightarrow \{-1, +1\}$  for the distribution  $D_t$ , where goodness is measured by the error of the hypothesis with respect to  $D_t$ . Then  $D_t$  is updated such that incorrectly classified examples have their weights increased, it forces the weak classifier to concentrate on the more difficult training examples. Correspondingly, correctly classified examples are given less weight. Adaboost selects some parameter  $\alpha_t$  to denote the importance of  $h_t$ , and after all rounds are complete, the final hypothesis  $H$  is a weighted majority vote of all  $T$  weak hypotheses. It has been shown that, as with other boosting algorithms, if each weak hypothesis is at least slightly better than random, then the training error falls at an exponential rate. However, Adaboost is also able to adapt to the error rates of individual weak hypotheses, so each subsequent classifier is adjusted in favor of examples mislabelled by previous classifiers [52].

#### *Support Vector Machine (SVM)*

SVMs [154] are supervised learning methods used primarily for classification. An SVM treats its input data as two sets of vectors in  $n$ -dimensional space: positive and negative examples. In this space, it constructs an optimal hyperplane that preserves the maximum

distance between the two sets [143]. Since SVM is able to handle large feature spaces it has been used successfully to solve many real world problems such as text categorization, image classification, protein analysis, cancer data classification, and hand-writing recognition [58]. Consider a set of  $N$  labelled training examples  $D = (x_1, y_1), \dots, (x_n, y_n)$  with  $y_i \in \{+1, -1\}$  and  $x_i \in R^n$ , where  $n$  is the dimensionality of the input. Let  $\phi : R^n \rightarrow F$  be the mapping function from the input space to the feature space. If the two classes are linearly separable, the SVM algorithm finds a hyperplane  $(w, b)$  that maximizes the margin

$$\gamma = \min_i \{y_i < w, \phi(x_i) > -b\} \quad (10)$$

where  $b$  is a real number (bias term) and  $w$  and  $\phi$  have the same dimensionality. For an unknown input vector  $x_j$ , classification means finding:

$$f(x_j) = \text{sgn}(y_i < w, \phi(x_i) > -b) \quad (11)$$

It can be shown that this minimum occurs when  $w = \sum_i \alpha_i \gamma_i \phi(x_i)$ , where  $\alpha_i$  is a positive real number that represents the strength of training point  $x_i$  in the final classification decision. The subset of points where  $\alpha_i$  is non-zero consists of the points closest to the hyperplane, and these are the support vectors. Since SVM is able to handle large feature spaces, it is frequently used in many real world problems, even though it is computationally expensive [58].

## CHAPTER 4 Application of Computer-Aided Decision-Making System to Traumatic Injury Data

This chapter presents the results of predicting survival (alive/dead), exact outcome (home/rehab), and ICU (Intensive care unit) days via rule generation. Results indicate that the rule-based system can help physicians to make accurate decisions and explain the reasoning behind them; this is expected to enhance patient care, and help derive new knowledge from complex data patterns.

The chapter is organized as follows. Rules for predicting outcomes are generated using three different trauma datasets. In Section 4.1 describes the dataset and examines rules for survival and exact outcome. Section 4.2 focuses on predicting the number of days that a trauma patient transported to hospital via helicopter will spend in the intensive care unit (ICU). The last section follows prediction of exact patient outcome (home/rehab) via classification of pelvic injuries, which is a novel clinical use of data. The importance of filtering in the selection of significant variables is also addressed in this section.

### **4.1 Rules for Prediction of Survival and Exact Outcome**

In this section, several important variables are introduced. The significance of complications and pre-existing disease in predicting the outcomes by comparing two different datasets is explained. The performance comparisons between machine learning algorithms are also presented, plus conclusion and discussion of the findings.



#### 4.1.1 Description of Dataset

Two different datasets are used in the study: on-site and off-site. The on-site dataset collected contains data captured at the site of the accident; the off-site dataset contains data at the hospital after the patient is admitted. The on-site and off-site datasets are used to predict patient survival (dead/alive) and final outcome (home/rehab). The datasets are provided by the Carolinas Healthcare System (CHS) and the National Trauma Data Bank (NTDB).

*On-site dataset:* When making decisions based on the variables available at the accident scene, one has to consider the unavailability of important factors such as pre-existing conditions (comorbidities). Decisions must therefore be made without knowledge of these factors. Some physiological measurements are also excluded because they are only collected after arrival at the hospital. Table 1 presents the variables collected for this dataset, which consist of four categorical and six numerical attributes.

Table 1: On-site dataset collected at the site of accident (\* indicates the categorical variables).

Variable	Possible Values
Gender*	Male, Female
Blunt*	Blunt, Penetrating
ChiefComp*	MVC, Fall, Pedestrian, Motorcycle Crash, etc
Position*	Passenger, Driver, Cyclist, Motorcycle Passenger, etc
Age	$0 \leq \text{Age} \leq 90$
FSBP (Initial Blood Pressure)	$0 \leq \text{FSBP} \leq 300$
GCS (Glasgow Coma Score)	$3 \leq \text{GCS} \leq 15$
ISS (Total Injury Severity Score)	$0 \leq \text{ISS} \leq 75$
Pulse	$0 \leq \text{Pulse} \leq 230$
Respiration Rate	$0 \leq \text{Respiration} \leq 68$

*Off-site Dataset:* The off-site dataset includes information on comorbidities and complications, and contains some other physiological variables. A total of 1589 cases are included in the databases: 588 fatal and 1001 non-fatal. The inputs include both categorical and numerical attributes. The predicted outcomes are defined as patient survival, i.e. alive or dead, and the exact outcome for surviving patients, i.e. rehab or home. For the exact outcome prediction, a total of 628 rehab cases and 213 home cases are used. Table 2 presents the variables for the off-site dataset. Among the categorical variables, “prexcomor” represents any comorbidities that may negatively impact a patient’s ability to recover from the injury and any complication. Abbreviated Injury Scale (AIS) scores for head, thorax, and abdomen are provided, as well as an overall score for patients with multiple injuries. The range of AIS score values in our database is 1(minor injury) to 6 (fatal injury). Injury severity score (ISS) is the most widely used measure of injury severity in patients with trauma and the range of possible values for ISS is 0 to 75. The range of FURR (First Unassisted Respiratory Rate In ED) is between 0 and 99. EDRT indicates a revised trauma score and its range is between 0 and 8. EDEYE (Lowest Glasgow Eye Component in ED) is the value for eye response and ranges from 0 (eyes not opening) to 4 (eyes opening spontaneous ). EDVERBAL (Lowest Glasgow Verbal Component in ED) is the value of patient’s verbal response, which ranges from 1 (no response) to 5 (oriented). Glasgow coma score (GCS) is initially used to assess the patient’s level of consciousness after trauma injury, and the score is used by first aid staff such as the emergency medical services (EMS) and the physicians initially admitting all acute medical and trauma patients. It is also used in patient monitoring, for instance in intensive care. The maximum GCS score is 15; a score higher than 13 is generally treated as a minor case, and a score

less than 8 is considered a severe case.

Table 2: Dataset captured after hospital admission (\* indicates the categorical variables).

Variable	Alive	Dead	Rehab	Home
Cases	1001	588	628	213
Male*	704 (70.3%)	404 (68.7%)	443 (70.5%)	150 (70.4%)
Female*	297 (29.7%)	184 (31.3%)	185 (29.5%)	63 (29.6%)
Age	41.2 ± 19.6	49.2 ± 24.1	39.6 ± 19.3	37.2 ± 16.6
FSBP	126 ± 33.4	119.3 ± 45.6	125.3 ± 31.6	124.5 ± 34.1
FURR	15.3 ± 10.9	13.9 ± 11.9	14.4 ± 11.1	18.2 ± 10.5
GCS	8.7 ± 5.3	27.5 ± 5.2	7.9 ± 5.2	10.5 ± 5.1
ISS	30.5 ± 12.8	35.3 ± 14.7	32 ± 13.2	27.1 ± 11.7
EDEYE	2.4 ± 1.4	2.1 ± 1.4	2.2 ± 1.4	2.8 ± 1.4
ED Verbal	2.7 ± 1.8	2.3 ± 1.7	2.4 ± 1.8	3.3 ± 1.8
EDRT	4.6 ± 3.2	3.8 ± 3.3	4.1 ± 3.3	5.7 ± 2.89
Head AIS	3.0 ± 1.6	3.6 ± 1.6	3.1 ± 1.8	2.5 ± 1.4
Thorax AIS	2.3 ± 1.7	2.4 ± 1.8	2.3 ± 1.8	2.4 ± 1.7
Abdomen AIS	1.1 ± 1.5	1.1 ± 1.6	1.0 ± 1.5	1.5 ± 1.7
Intubation*	Yes/No			
Prexcomor*	17 values: Acquired Coagulopathy, Chronic Alcohol Abuse, Chronic Obstructive Pulmonary Disease, Congestive Heart Failure, Coronary Artery Disease, Coumadin Therapy, Documented History of Cirrhosis, Gastric or Esophageal Varices, Hypertension, Insulin Dependent, Myocardial Infarction, Non-Insulin Dependent, Obesity, Pre-existing Anemia, Routine Steroid Use, Serum Creatinine > 2 mg (on Admission), Spinal Cord Injury			
Complications*	Acute Respiratory Distress Syndrome (ARDS), Aspiration Pneumonia, Bacteremia, Coagulopathy, Intra-Abdominal Abscess, Pneumonia, Pulmonary Embolus			
Safety*	Seat Belt, None Used, Air Bag Deployed, Helmet, Other, Infant/Child Car Seat, Protective Clothing			

#### 4.1.2 Results

The average accuracy of survival prediction, without any knowledge of pre-existing conditions, is 73.9%, rising to 75.8% when this knowledge is included. It was discovered that knowledge of these conditions appears at the highest level of the tree when using CART

and C4.5, indicating their potential importance in the decision-making process. In particular, coagulopathy (a bleeding disorder), which can result in severe hemorrhage, may be among the most important factors to consider in patients with Trauma Brain Injury (TBI). Therefore, the off-site dataset, which contains comorbidity information, was used for further prediction tests.

Since the total number of examples used for training is rather small and some low accuracy rules may have been generated using the small number of examples, only rules with at least 85% prediction accuracy on the testing sets are included in the rule base. This threshold is chosen based on the recommendations made by trauma experts. All rules are presented in Appendix A.

### *Significant Variables*

In order to improve the rule quality and accuracy, it is essential to identify the key variables in the dataset. In addition, shorter rules that are based on fewer, more significant variables are more clinically useful for physicians. Logistic regression was used to extract these key variables from the off-site datasets; the results are shown in Table 3. It can be seen that nine important variables are identified. Mean and standard variation of the critical variables and p values for the significance of independent variables are given.

### *Measuring Performance*

The prediction results of five different machine learning methods are compared in Table 4. The performance for all algorithms is clearly superior when only significant variables are used.

The exact outcome prediction results of five different machine learning methods when

Table 3: Significant variables of off-site dataset (\* indicates categorical variables). Cg stands for Coagulopathy; MI for Myocardial Infarction; ARDS for Acute Respiratory Distress Syndrome; ID for Insulin Dependent; EDRTS for Emergency Department Revised Trauma Score; ISS for Injury Severity Score.

Variable	P-value	Mean $\pm$ S.D.
AIS Head	<.0001	3.25 $\pm$ 1.64
AIS Thorax	0.003	2.33 $\pm$ 1.78
ID*	0.02	-
MI*	<.0001	-
ARDS*	<.0001	-
Cg*	<.0001	-
Age	<.0001	44.15 $\pm$ 21.70
EDRTS	0.03	12.10 $\pm$ 16.03
ISS	0.01	15.82 $\pm$ 19.03

Table 4: Performance comparison of five machine learning methods with survival prediction.

	Logistic	AdaBoost	C4.5	CART	SVM
All Variables	69.4%	70%	68%	75.6%	73%
Significant Variable only	72.9%	73%	75.2%	77.6%	79%

using only significant variables are compared in Table 5. For this prediction of Table 5, no attempt was made to use all available variables, since the survival prediction test has already confirmed improved performance when using only significant variables.

Table 5: Performance comparison of five machine learning methods with exact outcome prediction.

	Logistic	AdaBoost	C4.5	CART	SVM
Significant Vars. only	74.6%	73%	72%	75.6%	72.6%

Discussion with physicians revealed that all generated recommendations and predictions should ideally be transparent in their reasoning; our system therefore uses CART

and C4.5 are used to predict patient survival. If physicians understand the reasoning behind decisions and it follows their own, their confidence in the system may be increased. If the system's reasoning is clinically meaningless, they can choose to disregard the recommendation; however, if the reasoning has some clinical merit, this may alert them to previously hidden factors affecting patient outcome. In summary, the reasoning can confirm the physicians' initial conclusion as well as alert them to interesting new information that may improve patient care.

The most reliable rules ( $\geq 85\%$ ) and supporting rules (between 75% and 85%) are presented in Appendix A (for survival) and Appendix B (for exact outcomes). There are two reasons for including rules with accuracy between 75% and 85%. Firstly, the accuracy of a rule may be low due to the lack of a truly complete database, rather than a flaw in the rule itself. Secondly, even though a rule may have low accuracy, it might include knowledge of hidden relationships between variables. For example, most trauma experts consulted that a patient with an ISS score over 25 would have little chance of survival. However, the survival probability might be higher for a patient with a high ISS score, and lower head and thorax AIS score, an appropriate and prompt treatment is provided. Therefore, those rules with accuracy between 75% and 85% are used as additional "supporting rules" in suggesting possible treatment. This issue is addressed further in the section 4.1.3. In this case, the algorithms were tested using only significant variables.

#### 4.1.3 Discussion

A computer-aided rule base was developed using significant variables selected via logistic regression, and it is shown that this filtering step increases rule quality. The intention was

to extract and formulate medical diagnostic knowledge into an appropriate set of transparent decision rules that can be used in a computer-assisted decision making system. By comparing the performance of five machine learning algorithms - AdaBoost, C4.5, CART, SVM, and logistic regression - using all available variables and significant variables only, it was found that using only the most significant variables provides a considerable improvement in performance. All five methods show improvement using significant-variables-only, indicating that the proposed feature selection method is robust and efficient.

The performance of individual rules was also measured. Reliable rules were identified as those with accuracy above 85%. In addition, all selected rules were considered reliable if the number of cases in the dataset matching the rule was higher than a specified threshold. Rule sensitivity and specificity were also measured, and the average sensitivity and specificity for the two outcome pairs (alive/dead, home/rehab) of over 85% accuracy rules are 87.4% and 88.4%, respectively. Also, average sensitivity and specificity of rules between 75% and 85% of them are 82.6% and 80.3%. This attests to the successful performance of the proposed method. Some additional improvements may be needed to improve rule quality. In particular, large and well balanced datasets across all outcome classes could improve overall quality, as well as sensitivity and specificity measures.

One important issue in rule selection is how to deal with rules having accuracy below 85%. When using only the over-85% rules, some medical knowledge or interesting patterns in the database may be ignored. The accuracy of a rule may be low due to the lack of “database completeness”, rather than a flaw in the rule itself. Therefore, rules with less than 85% accuracy cannot be completely removed from the rule based system.

Some rules were found which had surprising implications. For example, according to

trauma experts, patients with a high ISS score ( $>25$ ) are least likely to survive. One of the “counterintuitive” rules pointed to the fact that there are 52 alive cases (3.3%) with high ISS scores ( $\geq 38$ ). Of these 52 patients, 33 (63.5%) have high AIS head scores ( $\geq 4$ ), and 38 patients (73%) are male. Considering the above conditions, surviving patients have lower thorax (average score=2.61) and lower abdomen AIS scores (average score = 1.03) than fatal cases. The fatal cases typically have a higher head AIS score (average score=5.08) than surviving patients (average head score=3.90).

In addition, it was found that none of the surviving patients had complications such as coagulopathy, and only a few had a pre-existing disease (in particular, Insulin Dependency and Myocardial Infarction). While only Acute Respiratory Distress Syndrome (ARDS) is usually considered an impact factor in predicted survival, according to the created rules, Insulin Dependency, Myocardial Infarction, and Coagulopathy also have significant impact.

#### 4.1.4 Conclusion

The results provide a framework to improve physicians’ diagnostic accuracy with the aid of machine learning algorithms. The resulting system is effective in predicting patient survival, and rehab/home outcome. A method has been introduced that creates a variety of reliable rules that make sense to physicians by combining CART and C4.5 and using only significant variables extracted via logistic regression. The resulting computer-aided decision-making system has significant benefits, both in providing rule-based recommendations and in enabling optimal resource utilization. This may ultimately assist physicians in providing the best possible care to their patients. The diagnosis of future patients may also be improved by analyzing all possible rules associated with their symptoms.



## **4.2 Rules for Prediction of Expected Time to Stay in Intensive Care Unit (ICU) for Airlifted Patients**

This section focuses on predicting the number of days that a TBI patient, who is transported to hospital via helicopter, will spend in the intensive care unit (ICU). Based on the literature, many airlifted patients leave the hospital and/or the ICU on the same day the injury occurs. This puts financial burden on patients and/or the healthcare system as such patients do not truly need helicopter transportation. The ability to successfully predict the injury severity further emphasizes the impacts of computer-aided decision making systems on the care provided to patients as well as the cost effectiveness of trauma care.

This section organized as follows. First, the significance of ICU prediction for the helicopter dataset is explained in Section 4.2.1 and the dataset itself is explained in Section 4.2.2. Experiment results are presented in Section 4.2.3, followed by the discussion and conclusion.

### *4.2.1 Introduction*

It has been reported that traumatic brain injuries are the most expensive affliction in the United States, with an estimated cost of \$224 billion per year [2]. Since the treatment of traumatic brain injuries is extremely time-sensitive, it is widely believed that the predicted length of stay in the ICU should be an important consideration when deciding on the patient transport method (i.e. ambulance or helicopter). Critically injured patients are expected to spend more time in the ICU, and also stand to benefit the most from helicopter transport. Studies have emphasized the impact of helicopter transportation on trauma mortality rates, since the speed of ambulance transport is limited by road and weather

conditions, and may also be constrained by traffic congestion. Cunningham [36] attempts a comparison based on the outcome of the treatment given to trauma patients. Based on this study, patients in critical condition are more likely to survive if transported via helicopter. However, the high cost of helicopter transport remains a major problem [128]. Comparison of ground and helicopter transportation and the corresponding care provided to the patients is a challenging task. While helicopter transportation can provide potential benefits for time-dependent patients, the cost and the risk of an accident are both lower for ground transportation [13, 73].

Gearhart [61] evaluated the cost-effectiveness of helicopter transport for trauma patients, and found that the average cost of helicopter transport is approximately \$2,214 per patient, and \$15,883 for each additional survivor [61]. Ultimately, the cost is almost \$61,000 per survivor for trauma patients. However, Eckstein [43] states that 33% of patients who are transported by helicopter are discharged for home from the emergency department, rather than being sent to ICU. This indicates that a significant number of trauma patients transported by helicopter actually have relatively minor injuries. This emphasizes the necessity of a well-identified transportation policy based on the patient's condition and predicted outcome.

#### 4.2.2 Description of Dataset

This dataset, provided by the Carolinas Healthcare System (CHS), is based on the records of patients who were transported to the CHS hospitals by helicopter. The variables are age, gender, blood pressure, airway (all types of device used to assist patients with breathing), pre-fluids (the amount of blood provided to the patients), GCS, heart rate, respiration

rate, ISS (Injury Severity Score), and ISS-Head&Neck. Age, blood pressure, GCS, pulse rate, ISS-Head&Neck, ISS, and respiration rate are classified as numerical variables. The final outcome is the number of days spent in ICU, as this is considered the most informative measure when deciding the means of transport to hospital.

The use of a relatively small dataset with so many outcomes may result in a complex model that is hard to explain and understand. Pfahringer [114] addresses the benefits of using discretization of continuous attributes including significant improvements of smaller sizes of trees, and improved prediction accuracy. Thus, in this study, the dataset was categorized into two groups. The non-severe group contains patients who stayed in the ICU less than 2 days (ICU stay  $\leq 2$  days). The severe group consists of patients who stayed in the ICU more than 2 days (ICU stay  $\geq 3$  days). This threshold on the number of ICU days, as a criterion of injury severity, was chosen based on discussion with trauma experts. In total, the dataset contains 497 cases: 196 severe and 301 non-severe. Table 6 describes the helicopter dataset in more detail.

As mentioned in Chapter 3, all categorical variables are replaced by multiple dummy variables with two levels (Yes/No). Since a small dataset with many attribute levels may result in bias, the twelve-value airway variable is divided into two groups: not needed (i.e. not performed) and needed (all other levels).

#### 4.2.3 Results

This section presents the results when using only significant variables, and explains the quantitative measures used.

Table 6: Detail description of Helicopter dataset (\* indicates categorical variables).

Variable	Severe (ICU stay > 2 days)	Non-Severe (ICU stay ≤ 2 days)
Cases	301	196
Male	201 (66.8%)	132 (67.3%)
Female	100 (33.2%)	64 (32.7%)
Age	30.6 ± 16.6	32.9 ± 17.2
FSBP	137.7 ± 23.2	127.6 ± 28.0
GCS	11.7 ± 4.87	6.47 ± 5.01
ISS	14.2 ± 8.1	23.7 ± 9.47
Pulse	101.4 ± 22.3	108.2 ± 26.6
Resp. Rate	15.6 ± 9.44	6.45 ± 1.6
ISS-HN	2.83 ± 0.86	3.46 ± 0.91
Airway*	need/no need	
pre-fluids*	<500,500-1000,>2000	

### *Significant Variables*

Out of the ten variables included in the helicopter dataset, only three are identified as significant: age (p-value <0.0001), blood pressure (p-value=0.0078), and ISS (p-value=0.0034). The above significant variables are found using the feature selection process in CUSF described in Chapter 3. Since the scale of the data is small, and ISS is not typically measured at the scene, other variables were included in the model based on the suggestions made by participating physicians. These physicians selected age, GCS, blood pressure, pulse rate, respiration rate, and airway as important factors.

### *Measuring Performance*

Table 7 presents the accuracy of the predicted number of days spent in the ICU for the helicopter dataset. In this case, only the significant variables were used. No attempt was made to use all available variables, since the survival prediction model in Chapter 4.1 already confirmed that using only the significant variables improves accuracy and

performance.

Table 7: Prediction results for ICU days.

Method	Logistic	AdaBoost	C4.5	CART	SVM
Days in ICU	80.6%	78.7%	77.1%	77.4%	80.1%

According to the Table 7, logistic regression, AdaBoost, and SVM are outperformed CART and C4.5. However, as mentioned in Chapter 2, these methods lack the facilitation of understanding for medical professionals. Physicians are more likely to believe in rules when they can understand the reasoning. SVM does not provide the rules and does not have the ability to handle missing values. Thus, CART and C4.5 are used for rule generation, and rules with over 85% accuracy are selected as reliable rules; this threshold was determined via discussion with physicians. The most reliable generated rules ( $\geq 85\%$ ) are presented in Appendix C. The sensitivity and specificity of the rules are 90.6% and 91% respectively. High sensitivity and specificity values indicate that the rules are well classified for predicting the ICU days.

#### 4.2.4 Discussion

Providing rapid care to critically injured patients is essential to improve their chance of survival. Many studies support the use of helicopters and state that the use of an air ambulance can reduce patient mortality. However, helicopter transportation is expensive and risky. Thus, effective usage of air ambulances requires a policy that takes into account the severity of a patient injury. Through the analysis and classification of patient data, the

developed rule-based system may assist in predicting which patients should be transported by air.

Airway status (needed/not needed) was identified as a primary factor in predicting the number of ICU days for patients transported via helicopter. Note that 74.6% of patients spent 2 days or less in the ICU. Only 25.4% of patients stayed more than 2 days, and only 2.9% of those were in the ICU for more than 20 days. This reinforces Eckhart's [43] point that many patients are transported via helicopter unnecessarily.

Therefore, the use of rules for accurate prediction of ICU length of stay may improve the efficiency of helicopter transport, in terms of both cost effectiveness and critical patient care.

#### *4.2.5 Conclusion*

Even though helicopter transport has many advantages, a policy needs to be devised for the use of helicopter transport. Therefore, a rule-based system was designed that predicts the expected number of days a patient will spend in the ICU. This system predicts expected days in ICU specifically for trauma patients. To validate the system, quantitative analysis is performed to calculate the accuracy and significance.

While comprehensive testing using a larger dataset is still required, the method was shown to be capable of extracting accurate and meaningful rules which support making fast and reasonable decisions using the available patient information. Finally, the validity of the rules has been confirmed via discussion with physicians.

### **4.3 Rules for Prediction of Exact Outcome in Traumatic Pelvic Injury**

This section focuses on prediction of exact patient outcome (home/rehab) via classification of pelvic injuries, which is a novel and clinically useful data set. The importance of filtering in the selection of significant variables is also addressed.

The rest of this section is organized as follows. Section 4.3.1 provides a brief introduction, including the significance of studying pelvic injuries. The dataset is described in Section 4.3.2. Comparison results are then presented in Section 4.3.3. This section concludes with discussion of the results.

#### *4.3.1 Introduction*

Traumatic pelvic injuries are often associated with severe, life-threatening hemorrhage, and immediate medical treatment must therefore be provided to reduce mortality and the risk of serious complications. Pelvic injuries are critical injuries because they are associated with a number of complications that often require extensive rehabilitation. Although pelvic fractures are uncommon injuries which affect approximately 3%-8% of trauma patients, they have a relatively high mortality rate (5%-20%) due to the risk of severe hemorrhage [115]. Pelvic fractures therefore represent a significant contribution to mortality rates, [96, 113], where most deaths are caused by complications, such as hemorrhaging and secondary multiple organ failure, rather than the fracture itself [42, 67].

Most serious pelvic injuries occur due to high-speed impact, including motor vehicle-pedestrian accidents, crash injuries, and falls, with an average patient age of 31.5 years [110]. Rapid assessment and diagnosis are important to improve patient survival. However, it is

sometimes difficult for physicians to decide on appropriate treatment, due to the similarity and complexity of the various types of injury. A rule-based computer-aided system was developed to predict the exact outcome (home or rehabilitation) for pelvic trauma patients.

#### 4.3.2 Description of Dataset

The traumatic pelvic injury dataset used in this study was created in collaboration with Carolinas Healthcare System and Virginia Commonwealth University. The database contains not only demographics and physiological data but also trauma scores. These can be taken either at the site of the accident or at the hospital. All medical tests and procedures (e.g. blood transfusion) performed at any stage are also included in the database. Table 8 presents all the available input variables. In Table 8, GCS refers to Glasgow Coma Score, BP refers to blood pressure, and ED is used to note values collected at the emergency department rather than the scene of the accident.

As Table 8 shows, there are twenty variables in the pelvic injury dataset, including all possible scene and revised emergency department (ED) values. The variables are fed to the model as the input, and the output of the model is treated as the exact outcome of the treatment. The outcomes are therefore grouped into two classes: rehabilitation or home. A total of 681 surviving cases were used in this study, consisting of 381 patients sent home and 300 sent to rehabilitation. As described in Chapter 3, every categorical variable is replaced with several dummy variables.

Many researchers have identified Glasgow Coma Score (GCS) and Injury Severity Score (ISS) as important factors used by experts to decide on treatment [42, 39, 89, 126]. Therefore, those two factors are included as significant variables by default.



Table 8: Summary of all available input variables (\* represents categorical variables and mean SD are represented for numerical variables).

Variable	Outcomes	
	Home	Rehab
Total Dataset	381 Cases	300 Cases
Male	238 (62.5%)	180 (60%)
Female	143 (37.5%)	120 (40%)
Age	35.57 ± 13.98	46.56 ± 17.54
Scene BP	125.1 ± 27.53	125.49 ± 32.98
Pre-Hospital Eye score	3.77 ± 0.68	3.30 ± 1.20
Scene GCS	14.01 ± 2.51	12.26 ± 4.47
Scene Motor	5.68 ± 1.01	4.97 ± 1.88
Scene Pulse	98.25 ± 22.71	102.08 ± 22.11
Scene Respiration	20.58 ± 4.92	19.22 ± 7.92
ED Blood Pressure	129.83 ± 22.51	127.10 ± 27.25
ED GCS	13.57 ± 3.57	11.75 ± 4.97
ISS	16.52 ± 9.10	21.68 ± 9.97
Chief Complaint*	MVC, Pedestrian, Mtrcycle, Crash, Accident, Fall, Struck, GSW, Machine, Bicycle, Assault, ATV, Animal, Aircraft	
Position	Driver, Pedestrian, Passenger, Pedal Cyclist, Motorcycle Driver, Motorcycle Passenger	
Pre-Hospital Fluids*	<500, IVF Unk. Amount, Not Performed, 500-2000, >2000, IVF Attempted	
Scene Needle Thorax*	Performed, Not performed	
ABD. U/S*	Positive, Negative, Not performed	
ABD CT*	Positive, Negative, Not performed	
Chest CT*	Positive, Negative, Not performed	
Head CT*	Positive, Negative, Not performed	
Diagnosis type*	Open fracture of pubis, Closed fracture of ilium, Open fracture of ilium, etc	

#### 4.3.3 Model Selection Comparison

This section compares the significant variable selection methods. Specifically, stepwise model selection method and the feature selection process in CUSF are quantitatively compared. The stepwise model selection method, described in Chapter 3, is commonly used in medical data analysis to find models incorporating the best subset of input variables that

successfully predicts the outcomes, considering all possible combinations of independent variables. As part of the process, any single predictor variable may be added or deleted and the starting subset is typically the empty set.

There are major issues with using stepwise model selection for rule based decision-making application. Although the method is designed to find important variables, it may not guarantee choosing the most relevant attributes due to its repetition of the add and delete steps, as mentioned in Chapter 3 [100].

To compare the two model selection methods, CUSF and stepwise, the outcome, length of stay in ICU as a measure of severity for the pelvic injury patient dataset. The outputs are grouped into two classes: non-severe cases (2 or fewer days spent in ICU days) and severe cases (3 or more days spent in ICU). A total of 764 cases are used: 491 non-severe and 273 severe.

Using the stepwise method, age (p-value=0.0026), ED GCS (p-value < 0.0001), and ISS (p-value < 0.0001) are selected as a significant variables. When using the feature selection of proposed approach, CSUF, age (p-value=0.0003), gender (p-value=0.0279), ED GCS (p-value=0.0160), and ISS (p-value=0.0165) are selected as the most significant variables.

In terms of performance, the stepwise model selection has 83% training accuracy and 68.8% testing accuracy, whereas the same values for CUSF are 83.8% and 70%, respectively.

The two model selection methods were compared using the traumatic brain injury dataset given in Table 2 and Table 8. Table 9 presents the results of comparison between CUSF and the stepwise model selection.

Table 9: The performance comparison between the proposed approach (CUSF) and stepwise.

	C4.5		CART		SVM		AdaBoost	
	CUSF	Stepwise	CUSF	Stepwise	CUSF	Stepwise	CUSF	Stepwise
Survival Dataset	71.2%	68.7%	71.7%	69.6%	73%	69.9%	73.2%	70.9%
Pelvic Dataset	67.4%	65.7%	69.7%	65.2%	68.6%	67.6%	68%	66.6%

The results show that the proposed feature selection, CUSF, has better performance than stepwise selection. Non-medical datasets are also tested in Chapter 5 to validate the results. Also, statistical test, specifically Friedman test, will perform to test the significance of the difference between the two methods on medical and non-medical dataset in Chapter 5.

#### 4.3.4 Exact Outcome Prediction

It was found that age (p-value < 0.0001), pre-fluids (p-value=0.0317), and chest CT (p-value=0.0311) are all significant variables for predicting the exact patient outcome (home/rehab). Figure 5(a) shows the relationship between patient age, ED GCS, and ISS for the two outcomes. Figure 5(b) provides the same information for age, ED BP, and ISS.

Figure 5 (a) shows good separation of outcomes (home/rehab), but the lack of separation in Figure 5 (b) suggests that ED BP is not a significant feature for our application. Based on Figure 5, it is apparent that patients who are discharged to rehab are typically much older and have a higher ISS than those who are sent home. Also, when ED GCS values are the same, rehab patients have a higher ISS score than home outcome patients in 41% of cases. These observations support the filtering done by the significant variables

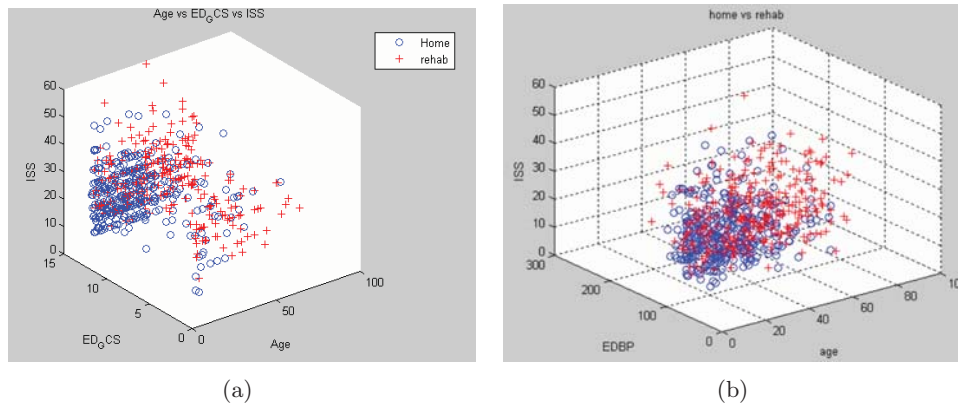


Figure 5: 5(a) Relation between age, ED GCS, and ISS 5(b) Relation between age, ED BP, and ISS .

selection process.

Table 10 compares the results of each machine learning algorithm except logistic regression on the testing sets. According to Table 10, none of the machine learning methods offers a significantly higher accuracy than the others. This indicates that the variable selection method, as described in Chapter 3, provides an efficient and appropriate process in filtering the input space. This observation also suggests that the variable selection step may have a significant impact on predicting the outcome. Table 10 shows that logistic regression outperforms the other methods. Based on this result, it was found that CART has marginally higher accuracy than other machine learning algorithms in predicting outcomes for the pelvic dataset.

Table 10: Performance comparison using pelvic dataset to predict exact outcome (home/rehab).

Method	Logistic	AdaBoost	C4.5	CART	SVM
Accuracy	71.8%	68%	65.3%	69.7%	68.6%

Rules with over 85% accuracy are selected as reliable rules; this threshold was determined via discussion with physicians. Only these rules are incorporated into the final rule base. In total, twenty-seven CART and seven C4.5 reliable rules were generated and the average rule accuracy is 90.4%. The sensitivity and specificity of the rules were calculated as 87.5% and 90.7% respectively. The most reliable generated rules ( $\geq 85\%$ ) are presented in Appendix D. High sensitivity and specificity values indicate that the rules constructed using the pelvic injury dataset are well classified. The final rules are transparent to physicians; as discussed previously, this is a desirable quality, as it allows them to understand the reasoning behind the predictions made by the computer-aided system.

#### 4.3.5 Discussion

This work is intended to solve the problem of formalizing medical diagnostic knowledge so it may be transferred into an appropriate computer system, using transparent methods that can be easily understood by physicians. This method also addresses a current problem in the medical application domain, specifically, how to deal with the many predictors that make variable selection difficult.

Interestingly, gender proved to be a significant variable in predicting the exact outcome (home/rehab). In the dataset, 415 out of 681 patients are male (61%) with an average age of 40. Also, comparing the average intensive care unit (ICU) days for home and rehab outcomes, the length of stay for rehab outcome patients is five times longer than the stay for those patients sent home (home:  $1.77 \pm 5.06$ , rehab:  $8.63 \pm 11.1$ ). However, 16% of home outcome patients stayed more than 2 days in ICU and 43% of rehab outcome patients stayed less than 2 days, contrary to expectations. This illustrates the need for computer-

aided decision making in both treating patients and in considering cost effectiveness.

Based on the ANOVA test, two values - GCS at the scene (p-value < 0.0001) and EDGCS (p-value < 0.0001) - are important factors in predicting the exact outcome (home/rehab). However, when using logistic regression it was found that only the EDGCS value is significant. In addition, a statistical t-test was performed to compare the scene and ED GCS scores. The resulting p-value of 0.04 is significant, as it is relatively high when using 5% as the confidence level. This indicates that the GCS measurements at the two different locations are typically quite different. It should be noted that for 352 cases out of 681 (51.7%) the GCS is not measured at the scene and for 16 cases out of 329 (4.9%) there is a large difference ( $\geq 7$ ) between scene GCS and ED GCS (typically the value at the scene is higher than the value at the ED). This reinforces the importance of filtering in the selection of significant variables.

A limitation of this study is the small size of the dataset used to generate the rules. However, the model goodness-of-fit test indicates that despite its size, the dataset is still suitable.

#### 4.3.6 Conclusion

In this section, a computer-aided system has been designed which aids physicians in making fast and accurate decisions for trauma injury cases. The system uses machine learning algorithms, specifically decision tree algorithms, to assist physicians in the decision-making process. Using this rule-based system, the most similar existing case to a new patient can be found in order to recommend a suitable treatment. As a next step, it is intended to integrate all available heterogeneous patient information during the decision-making

process by extracting diagnostically significant features from images and signals. These are very important measures such as the exact radius of the pelvis ring after the injury which are not easily visible to human eyes.

## CHAPTER 5 Testing & Validation Using Non-Medical Datasets

This section presents the results of testing and validation of CUSF using important applications. These tests further validate the applicability of CUSF for modeling of complex datasets. Decision tree algorithms are used as machine learning methods to produce and compare results on three different datasets, as described below.

### 5.1 Description of Datasets

#### *Census-income Dataset*

The income dataset was obtained from the UCI machine learning Repository [15]. The data consists of a binary classification to determine income level, with 33 input attributes and 1400 examples. Of these, 433 cases have under 50K income, and 966 have over 50K income. Of the attributes, 12 are numerical, including age, wage per hour, gross income, and dividends from stock and 21 are categorical, including gender, education, and own business or self employed. Also, a few variables are missing. A detailed description about the attributes are presented in Appendix E. For reliable estimation, 10-fold cross validation was used in testing.

#### *University Dataset*

The university dataset was obtained from the UCI machine learning Repository [15]. The data consists of a binary classification to predict the number of applicants, with 14 input



attributes including 10 numerical such as verbal SAT score, math SAT score, and expenses and 4 categorical variables such as major and control. Among the attributes, university name, state, and location are removed for testing. Detailed descriptions of the attributes are presented in Appendix E. A total of 269 examples are used. Some missing values were included. Again, 10-fold cross validation was used.

#### *Credit Approval Dataset*

This dataset was obtained from the UCI machine learning Repository [15]. The data consists of a binary classification concerning credit card approval, with 15 input attributes consisting of 6 numerical and 9 categorical variables. In the original database, all attribute names and values have been changed to meaningless symbols to protect the confidentiality of the data. A total of 690 examples are tested with 10-fold cross validation. The class distribution of the data contains 307 positive examples and 383 negative examples. Approximately 5% of values are missing.

## **5.2 Comparison of Results**

This section presents the results when using the non-medical datasets to test the proposed approach, CUSF. Table 11 compares the results of CUSF and the stepwise feature selection method.

According to Table 11, CUSF is slightly better than stepwise feature selection. Table 12 presents the results for the area under the ROC curve.

Considering the accuracy and the area of ROC curve analysis, the CUSF presents better performance than the stepwise with decision tree algorithm. Non-parametric statistical

Table 11: Performance comparison results between CUSF approach and stepwise.

Method	C4.5		CART	
	CUSF	Stepwise	CUSF	Stepwise
income	83.3%	82.4%	84.3%	82.6%
university	81.8%	81.1%	82.5%	79.2%
credit approval	87%	84.7%	85.5%	85.2%

Table 12: The area under ROC curve analysis between CUSF approach and stepwise.

Method	C4.5		CART	
	CUSF	Stepwise	CUSF	Stepwise
income	0.87	0.86	0.88	0.85
university	0.82	0.81	0.82	0.79
credit approval	0.9	0.89	0.89	0.86

test, Friedman test, is performed on both medical and non-medical dataset (i.e. the results of CART and C4.5 in Table 9 and Table 11). Since insufficient non-medical datasets are used for testing, the medical dataset is also included when performing the Friedman test. Based on Friedman test ( $p$  value=0.002), it was found that the difference between the performance of CUSF and stepwise selection method is significant.

### 5.3 Discussion

This section describes the validation of CUSF. Three non-medical datasets are used, with 10-fold cross validation for testing. In contrast to the medical dataset, the non-medical datasets have less than 5% missing values and fewer categorical variables.

In many fields of study, the stepwise method is commonly used for selecting the best

subset of variables for outcome prediction. However, it requires additional computation to find the best model. Furthermore, in the medical field it may not include some important variables which physicians believe critical to determine the outcome, such as age. This is due to its add and delete approach to variable selection, which can lead to any variable being dropped from the model based on the subset model significance. The proposed method may especially useful for medical datasets due to the nature of medical data - such as the higher proportions of missing values and categorial variables. However, according to the comparison results of the Section 3 in Chapter 4 and this chapter, the proposed feature selection method may also be useful for non-medical datasets.

#### **5.4 Conclusion**

The results show that the proposed approach (CUSF) is slightly better than stepwise selection. The proposed approach may be better, when combining it with decision tree algorithms.

## CHAPTER 6 Heart Rate Variability (HRV) Analysis Using Signal Processing

This chapter introduces heart rate variability analysis (HRV), which can provide useful information in identifying the degree of bleeding using signal processing methods. Since blood loss and physical activity have similar HRV responses in human subjects, differentiating between the two conditions has become of increasing importance and interest, particularly in military applications. This chapter presents a signal processing approach, specifically using wavelet transformation analysis, which can be employed to distinguish the two conditions.

The rest of this chapter is organized as follows. First, background and a description of a previous heart rate variability study are provided in Section 6.1 and 6.2 respectively. Then the objective of this study is presented in Section 6.3. Section 6.4 describes the dataset, and methodology is presented in Section 6.5, starting with ECG segmentation. The results are presented in Section 6.6. Finally, conclusions and discussion are provided at the end of the chapter.

### 6.1 Introduction

Hemorrhage shock is the most critical life-threatening factor in battle injuries, and the relationship between hemorrhage and outcome for trauma patients has long been a topic of interest for study. Failure to promptly identify hemorrhage shock (HS) cases can prove

fatal. In one study of the Israeli military, 96% (351 out of 337) of patient fatalities occurred in the first four hours, typically due to blood loss [134, 159]. Both these rapid deaths and the many complications associated with the injury result from a lack of appropriate medical attention and limited evacuation facilities in the field [8]. 90% of combat field deaths occur before the patient can reach medical care. Thus, the treatment of battlefield injuries is extremely challenging, and fatalities within the first four hour of wound occurrence are highly dependent on battlefield conditions [164]. Moreover, the likelihood of death depends on a number of factors, such as the severity of hemorrhagic shock, the time delay, and the type of treatment provided. Thus, monitoring the health status of combatants using easily obtained signals such as heart rate variability remains a critical challenge.

Heart rate variability (HRV) is a non-invasive measurement of cardiovascular autonomic regulation. Recently, analysis of HRV from electrocardiography (ECG) recording has become a popular method for assessing activity of the autonomic nervous system. Several papers over the past 20 years have emphasized the significant relationship between HRV and cardiovascular mortality [30, 32, 33, 35, 133, 45]. Monitoring heart beat fluctuations appears to provide valuable information concerning cardiovascular and neurological diseases, as well as physical and mental stress. In particular, heart variability in cardiovascular activity, such as RR interval (described in Section 6.5.3), has been widely studied as a measure of cardiovascular function that can be used in both risk estimation and diagnosis of cardiac events.

## 6.2 Existing Methods for Heart Rate Variability Analysis

There are two main traditional approaches for HRV analysis; time domain analysis of HRV for standard deviation of normal to normal intervals, and frequency domain analysis for power spectrum density (PSD) using simple electrocardiograms (ECG). Previous studies [7, 125, 104] have demonstrated that PSD analysis is a good non-invasive tool for examination of the cardiovascular system, and it is currently the most popular linear technique used for studying HRV signals [7, 116, 160]. PSD analysis provides three bands: high frequency (HF: 0.15-0.5 Hz), low frequency (LF: 0.04-0.15 Hz), and very low frequency (VLF: 0.0033-0.04 Hz). However, PSD estimation methods are unsuitable for analyzing series whose characteristics change rapidly [125, 98, 104].

The importance of biological time series analysis in describing complex patterns has been studied over a long period. The nonlinear dynamic techniques used are used based on the concept of chaos theory and have been applied to many areas, including medicine and biology [7, 130]. Thus, the physiological phenomena of HRV have been characterized by fractal properties and prior studies have emphasized fractal dimension (FD) analysis, a useful tool in the identification of complex biological systems under different conditions [78, 95, 163, 155]. Previous studies state that FD analysis can reliably identify heart disease, as the irregularity of HRV causes abnormal cases to have greater fractal complexity than normal cases [48, 76, 4].

Lower body negative pressure (LBNP) is widely used as a human demonstration model for studying acute hemorrhage analysis [33, 23]. Several preliminary studies [16, 18, 31, 34, 32, 33, 140] have found that HRV becomes lower and more persistent with an increase

in negative pressure and that LBNP is a useful model to simulate acute hemorrhage in humans, since both induce similar physiological responses. Also, Convertino [31] and Stevens [140] have confirmed that lower body negative pressure (LBNP) is a useful technique to study cardiovascular activities and hemodynamic effects associated with severe hemorrhage shock in humans, in particular in combat settings. Comparisons between physiological response of LBNP and blood loss have demonstrated that some amount of blood loss and LBNP cause a similar physiological reactions.

However, several studies [140, 129, 106, 51, 68, 123, 124] of physical activity have found that exercise and LBNP generate similar physiological effects. Heidi et. al [68] studied heart activity during different activity states, and found that RR intervals also decrease significantly during exercise and other vigorous activity.

Therefore, the differentiation between two conditions, blood loss and exercise, is vital for better early identification of bleeding, especially in military applications. Currently, power spectral density (PSD) and fractal domain (FD) are being studied and advocated as a means to detect sensitive HRV changes due to HS. Unfortunately, traditional HRV analysis appears to be unable to distinguish between central volume loss and exercise. This is problematic given the desire to use changes in heart rate to detect the presence of acute volume loss due to hemorrhage.

### **6.3 Specific Aims**

This study proposes a new method (based on wavelet transformation) to distinguish HRV between different electrocardiogram (ECG) data for LBNP and physical activity. The

hypothesis under test is that wavelet transformation analysis can distinguish between LBNP and physical activity subjects, whereas fractal dimension (FD) and traditional power spectral density (PSD) analysis cannot. Therefore, the specific aims for this study are as follows:

- 1) To propose an algorithm based on a set of novel wavelet features to analyze complex signals, such as Electrocardiograms (ECGs), that can provide important information typically invisible to human eyes. This wavelet-based method is referred to as Signal Analysis Based on Wavelet-Extracted Features (SABWEF). The algorithm extracts information that can be used to detect and analyze complex patterns.
- 2) To compare traditional methods of signal processing and SABWEF for discovering hidden patterns from complex signals using statistical analysis.
- 3) To assess the success of SABWEF in analyzing ECG, in particular to apply SABWEF, combined with machine learning algorithms, to estimate of the severity of blood loss.
- 4) To assess the success of SABWEF in analyzing multiple signals, in particular, to apply SABWEF, combined with machine learning algorithms, to four signals, ECG, arterial blood pressure (ABP), and impedance signals (IZT, and DZT), to predict the severity. Also, to assess the improvements in prediction of severity of blood loss by analyzing four signals rather than analyzing only ECG.
- 5) To assess the success of SABWEF in analyzing signals other than ECG, such as gait aging signal.



## 6.4 Description of Dataset

The dataset comprises fifty-nine subjects from the US Army Institute of Surgical Research, including forty-six LBNP subjects and thirteen exercise subjects. LBNP testing is done using a chamber in which the subject is exposed to varying levels of negative pressure. All measures are either continuously monitored or semi-continuously monitored at regular intervals. If the subject becomes distressed, he can request termination of the test, or end the test himself using a switch located on the LBNP device. Table 13 presents the detail dataset information containing LBNP and exercise subjects as well as the monitoring timeframe for the LBNP protocol.

Table 13: Summary of dataset including LBNP and exercise subjects 13(a), and monitoring schedule for lower body negative pressure (LBNP) protocol 13(b).

(a)				(b)		
LBNP		Exercise		LBNP Protocol	Stage	Time
Collapse Stage	Subject Number	Collapse Stage	Subject Number			
3	5(10.8%)	5	2(15.4%)	0 mmHg	Baseline	5 min
4	18(39.1%)	6	11(84.6%)	-15 mmHg	Stage 1	5 min
5	12(26.1%)	-	-	-30 mmHg	Stage 2	5 min
6	9(19.6%)	-	-	-45 mmHg	Stage 3	5 min
7	2(4.4%)	-	-	-60 mmHg	Stage 4	5 min
Total	46 subject		13 subject	-70 mmHg	Stage 5	5 min
				-80 mmHg	Stage 6	5 min
				-90 mmHg	Stage 7	5 min
				-100 mmHg	Recovery	5 min

The LBNP protocol consists of a 5-minute rest period (0 mm Hg) followed by 5 minutes of chamber decompression of the lower body to -15, -30, -45, and -60 mm Hg and additional increments of -10 mm Hg every 5 minutes until the onset of cardiovascular collapse. Cardiovascular collapse is defined for LBNP as the stage at which the experiment is terminated due to physical or physiological signs, or symptoms of distress. The collapse state for exercise is defined as the stage of exercise resulting in the same heart rate at

which LBNP was terminated for a particular subject recovery stage occurs after removing the negative pressure from the subject. The exercise protocol consists of 5 minutes of baseline followed by 5 minutes of exercise at gradually increasing workloads. All ECGs were sampled at 500 Hz.

Figure 6 presents a detailed schematic of overall method. Two condition datasets, LBNP and exercise, are used as input, and both traditional heart rate analysis and the proposed analysis, SABWEF, are applied. Then statistical analysis between the SABWEF and traditional analysis methods is performed.

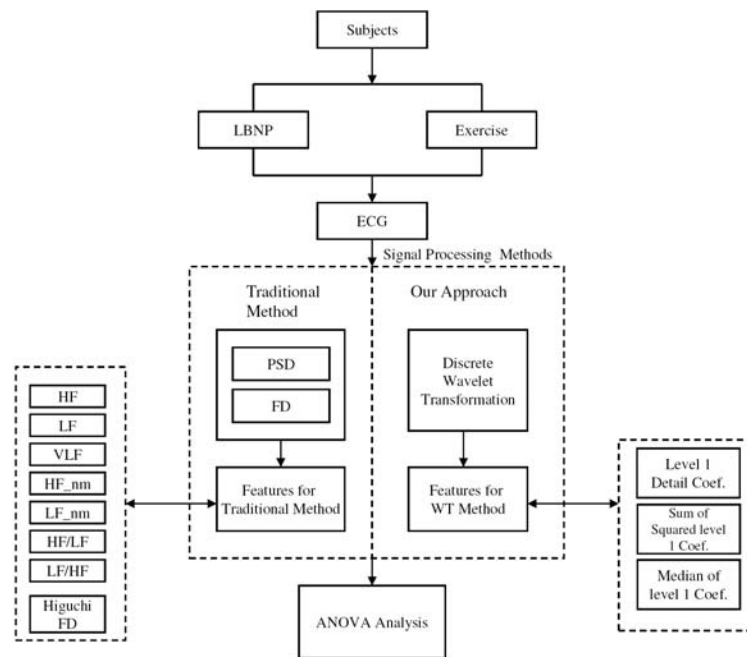


Figure 6: Detailed schematic diagram of entire process - multiple tasks are performed on ECG data.

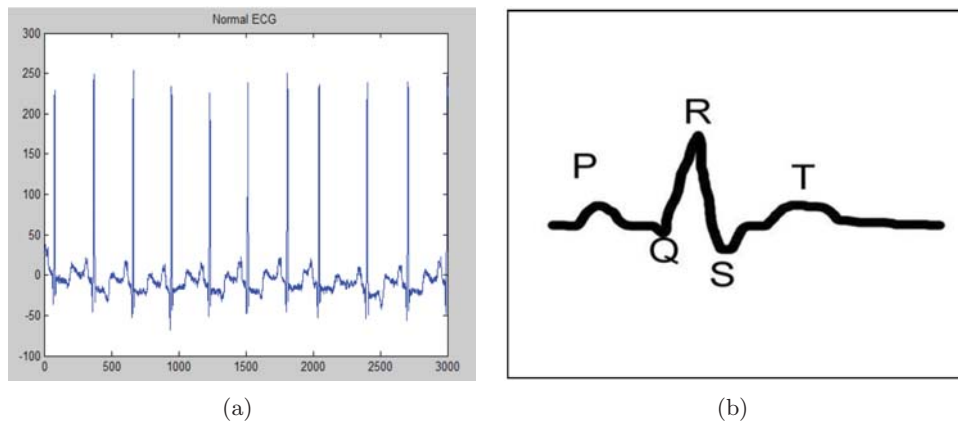


Figure 7: ECG signal before pre-processing 7(a) and description of P wave, QRS, and T wave 7(b).

### 6.5 Methods

Physiological data, such as electrocardiographs (ECG), are commonly used to record patient condition using a waveform which is suitable for monitoring. The ECG signal is an electrical signal generated by the beating of the heart, and can be used as a non-invasive diagnostic tool in examining heart function. The accurate recognition of ECG signals, which is referred to as QRS detection, is an important task for diagnosis. Figure 7 displays a standard ECG with characteristic P, R, and T waves. The P wave represents the spread of electrical activity over the atrium, and it usually has 0.08 to 0.12 sec (80 to 120 ms) duration. The T wave represents the recovery of the ventricles. Since the R wave has the highest amplitude in the QRS complex, its accurate detection is important in detecting the heart function. The interval between QRS complexes is called the R to R interval and the variation of the intervals is referred to as HRV. Note that QRS complex is the most significant waveform within the ECG signal to get HRV response information and the HRV is generally extracted from the ECG recording by detecting RR intervals.

In order to correctly assess the clinical meaning of the displayed information, it is highly desirable that the waveforms accurately reflect the measured physiological data regardless of the monitor type, pixel resolution, and the size of the window. The original electrical signal sample used in the dataset is recorded at a high data rate of 500 frames per second (also described as 500 Hz); an individual heartbeat consists of 400 to 700 samples (corresponding to heart rates between 70 beats per minute and 50 beats per minute). This is a significant volume of data to consider, and may therefore be re-sampled at a lower frequency such as 250Hz or 125Hz, a process known as down-sampling [69].

During the tests, electrodes are used to record a subject's ECG. The processing includes three parts; obtaining the raw ECG, ECG segmentation, and QRS detection. As mentioned earlier, QRS detection is important to examining heart function. Figure 8 gives an overview of the process of ECG analysis for this study. Each step will now be explained in more detail.

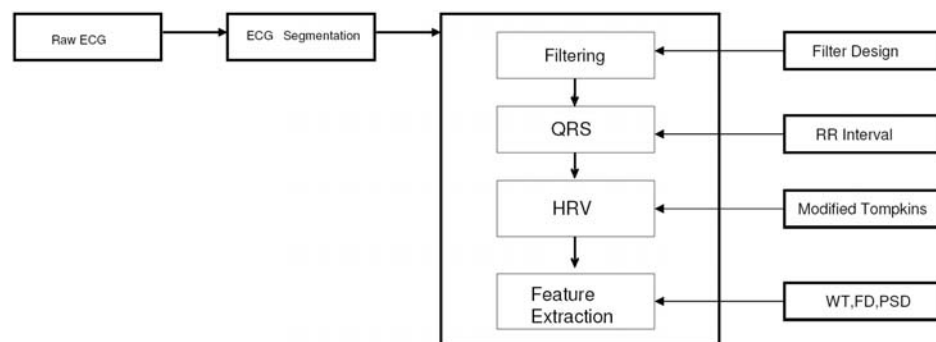


Figure 8: Detailed schematic diagram of ECG analysis.

### 6.5.1 ECG Segmentation

ECG segmentation is applied to the raw signal to distinguish the stages based on Table 13(b). Figure 9 presents an example of an LBNP signal.

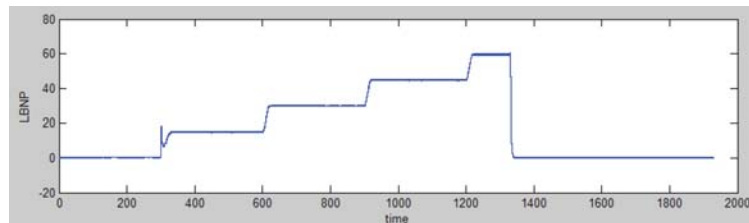


Figure 9: Example LBNP signal - The x-axis represents time and the y-axes represent LBNP values. The original ECG and LBNP signals must be scaled down for clear visualization and to help differentiate the stages.

The graph in Figure 9 displays increasing sections representing transient states, and flat horizontal lines representing steady states. Therefore, the ECG signal is segmented based on the pressure levels of LBNP in Table 13(b). This study observes only the steady state sections at each specific stage. A signal is collected for each stage.

The application of a pre-processing filter is necessary in order to attenuate signal noise. To extract the main ECG features, three main pre-processing steps are performed: filtering, QRS detection, and feature extraction.

### 6.5.2 Filtering

The purpose of filtering is to remove noise caused by a 60Hz power-line interface, as well as to attenuate noise such as motion artifact and baseline drift, which is generally caused by amplifiers [147]. This noise has a considerable influence on the quality of signal analysis; therefore filtering must be done prior to processing the ECG signal. A diagram of the

filtering process is presented at Figure 10.

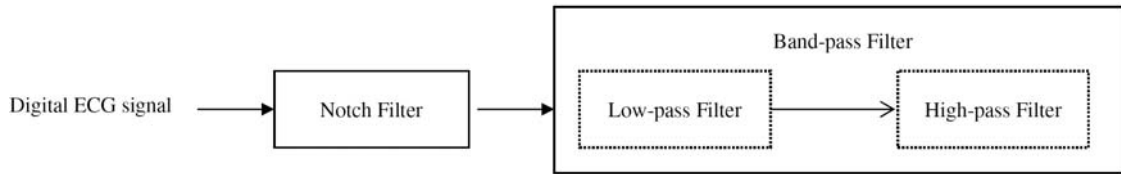


Figure 10: The schematic diagram of filtering methods on ECG signal.

Both a notch filter and a band-pass filter are used. The notch filter is applied first, to remove the 60 Hz power-line noise from the ECG. The notch filter can be seen in Figure 11 as the simultaneous application of high-pass and low-pass filters. This notch filter blocks only a specific predefined frequency. This study uses a 60Hz notch filter, as this is the frequency associated with the power noise [17].

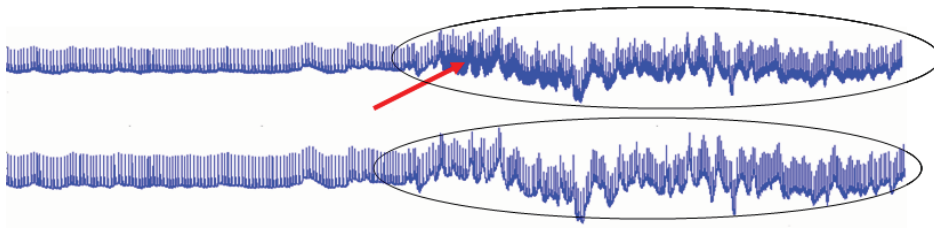


Figure 11: Example of 60Hz power-line interfering noise before (upper) and after (bottom) filtering. The arrow highlights the effect of the 60Hz power-line noise in the original signal.

### 6.5.3 QRS Detection

To extract heart rate variability based on RR intervals (see Figure 14), QRS detection must be performed first. The QRS complex is the most distinguishable component in the analysis because of its spiked nature and high amplitude. Since the P and T wave occur before and after the QRS respectively, they are difficult to distinguish without knowledge of

the QRS location [109, 74]. There are several techniques for detecting the QRS complex using a variety of techniques and approaches. For example, QRS detection algorithms based on signal derivatives are described in [6, 80, 71], algorithms based on a digital filter are described in [109, 142], algorithms based on wavelet transform are discussed in [88, 77], algorithms based on neural network are investigated in [162, 156], and algorithms based on mathematical morphology are developed in [28, 149].

In order to detect the QRS complex, a modified version of the Tompkins algorithm is formed and used. The Pan-Tompkins algorithm [109, 6] is the most commonly used real-time QRS detection algorithm, which is based on an analysis of the slope, amplitude, and the width of QRS complexes. Additional procedures such as histogram analysis, thresholding process, and RR interval checking are added to the original Tompkins algorithm.

Figure 12 illustrates the detection process.

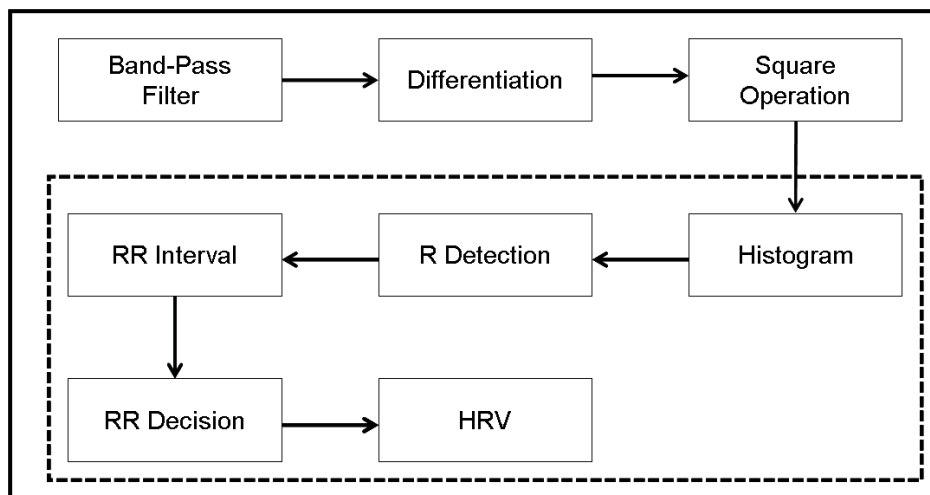


Figure 12: Detailed schematic diagram of QRS wave detection process.

First, the differentiation step is applied to remove the low-frequency components, such

as P and T waves, retaining the high-frequency components associated with the high slopes of the QRS complex. Next, the squaring operation makes the result positive and emphasizes large differences resulting from QRS complexes. In other words, it emphasizes the higher frequency component nonlinearly and attenuates the lower frequency component. This suppresses small differences caused by the P and T waves. Then, a moving average window is used as a further smoothing filter, reducing the high-frequency noise. A moving average is used over the squared signal to obtain the smooth pulse corresponding to the QRS complex. The moving window size is defined based on half of the sampling rate in this study. After this, a threshold is selected with the smoothed (averaged) signal to determine the presence of the QRS complex in the waveform. However, ECG measurements can also include noise introduced by muscle activity, which can cause high frequency noise, or from other sources such as electromagnetic interface. These may impede detection of the true RR interval. In order to overcome this, the result of the moving average signal is subtracted from the original ECG signal. Additionally, histogram analysis is performed over the moving average signals. The histogram accurately describes characteristics of the amplitude distribution across the signals. Thus, the histogram procedure removes unexpected noise by removing small frequency parts of the signal. Then, an adaptive threshold value,  $Thresh = mean(x_i) + max(x_i) * \alpha$ , where  $i = 1, \dots, n$  ( $n$  is the length of the signal) and  $x_i$  is the signal, is applied. After testing different values of  $\alpha$ , it was observed that  $\alpha=0.4$  provides the best performance. To avoid False Positives (FP) due to high T-wave detection, the acceptable heart beat range is set to between 30 bpm and 200 bpm, and the RR interval is checked as well. Any single RR interval is compared using previous RR intervals with sliding small window ( $=8$ ). At this time, median value of previous RR intervals is used. If



the RR interval value is greater than a certain range of interval, new RR interval is added. Let  $I = \{I_1, I_2, I_3, \dots, I_n\}$  is a set of RR interval at each stage and  $n$  is a length of RR interval,  $l$  indicates an added new RR interval. For adding new RR interval, the following rules are applied. If  $\omega_0 \leq I_i \leq \omega_1, i = 9, \dots, n \implies l = I_i$ . If  $\omega_1 \leq I_i \leq \omega_2, i = 9, \dots, n \implies l$  is added and where  $l = \text{median}(\{I_{i-8}, I_{i-7}, \dots, I_{i-1}\})$  is added RR interval and  $I_{i+1} = I_i - l$ . If  $\omega_2 \leq I_i \leq \omega_3 \implies$  two  $l$  is added, where  $l = \text{median}(\{I_{i-8}, I_{i-7}, \dots, I_{i-1}\})$  is added RR interval and  $I_{i+2} = I_i - 2 \times l$ , where  $\omega_0 = 0.89 \times m, \omega_1 = 1.29 \times m, \omega_2 = 2 \times m$ , and  $\omega_3 = 3 \times m$ , where  $m$  is the median value of the previous eight RR intervals. Based on this process, final heart rate variability is calculated.

Figure 13 shows the processing steps 1 through 3. The two circles in (a) show that a trend movement problem caused an incorrect RR interval detection. Figure 13 (c) shows how this algorithm handles the trend movement problem to correctly detects QRS. Figure 14 shows a portion of the RR detection result from Figure 13-(c).

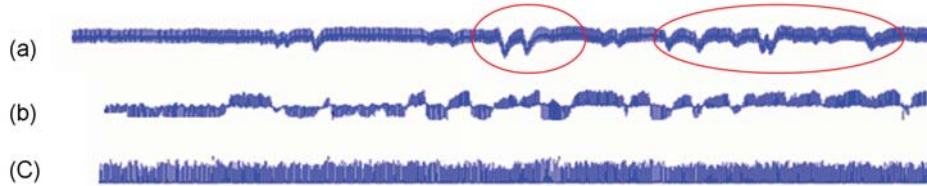


Figure 13: Example of QRS detection steps - (a) original signal (b) result signal after applying the moving average filter (c) result signal after applying the histogram procedure.

#### 6.5.4 Feature Extraction

Once the HRV is extracted based on RR intervals, three methods are applied to extract features. Figure 15 describes the feature extraction step. A total of fifty-seven features are obtained using discrete wavelet transform (DWT), PSD, and FD methods. Feature

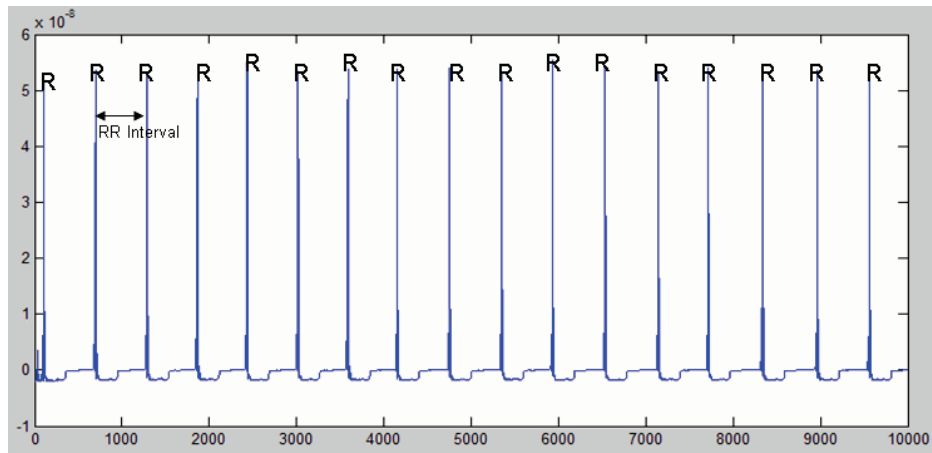


Figure 14: Results of R peak detection and RR interval. R indicates R peak and RR interval means time between R peak

extraction contains three steps: HRV normalization, signal processing analysis, and feature analysis. For this step, HRV is considered as an input signal.

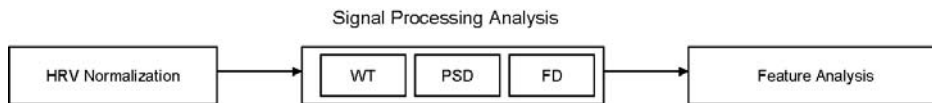


Figure 15: Detailed schematic diagram of feature extraction.

- a) *HRV normalization*: Before calculating discrete wavelet transform (DWT), power spectral density (PSD), and fractal dimension (FD) for the HRV, normalization is performed because different subjects have different baselines. First the mean at the baseline stage (HR first mean) is calculated and then all values at the rest of the stages are divided by the HR first mean. These values, after division by the HR first mean, are then used as input to further analysis such as the DWT. The HR first

mean is calculated as follows:

$$hr_m = \frac{1}{N} \sum_{i=1}^N x'_i \quad (12)$$

where  $x'_i$  is the HRV signal of the baseline and N is the length of HRV signal of the baseline. HRV normalization is then performed as follows:

$$y_i = \frac{x_i - hr_m}{hr_m}, \quad i = 1, \dots, N \quad (13)$$

where  $y_i$  is the new signal after normalization and  $x_i$  is the original HRV signal of each stage. Using the new normalized signal calculated, features are extracted.

- b) *Signal Processing Analysis:* Three signal processing methods, wavelet transformation, PSD and fractal dimension (FD) are described.

**Wavelet Transformation Analysis:** Wavelet transformation analysis (WT) acts like a mathematical microscope which allows one to zoom in and discover the detailed structure of a signal, or alternatively to reveal large scale structures by zooming out. WT is a promising technique for time-frequency analysis, providing several features not supported by Fourier transformation analysis [94, 141, 99, 111, 153].

The Fourier Transform (FT) uses sine and cosine base functions that have infinite span and are globally uniform in time. However, the frequency of an ECG signal changes over time; the QRS complex is a high frequency wave while the T wave contains low frequency components. Therefore, it is critical to maintain the correspondence of ECG frequency information to location in time [88]. However, FT does not provide easy access to a signal's time domain information, and is therefore unsuitable for preserving location information. FT is also limited to measuring

non-stationary signals such as electrocardiogram (ECG) and blood pressure (BP) [103].

Unlike FT, wavelet transformation can provide easy access to both frequency and location information. This combination of time and frequency resolution makes wavelet transform potentially very valuable so it is commonly used for many practical applications in the field of biology and medicine [94, 4]. There are no absolute rules for choosing a wavelet function, and therefore careful testing of different wavelets and their efficiency is needed. Since the Daubechies wavelet family is similar in shape to the QRS complex [37], discrete wavelet transform (DWT) with the Daubechies wavelet, well suited to local analysis of fast time varying and non-regular signals, is applied [146].

DWT not only captures the frequency content of the input, by examining it at different scales, but also investigates the times at which these frequencies occur. It was developed as an alternative to the Short Time Fourier Transform (STFT), to overcome problems related to its frequency and time resolution properties [82, 153]. Figure 16 shows the detail processing steps of calculating DWT, where H is a high-pass filter and G is a low-pass filter associated with H (For more details see [105]).

**Power spectral density (PSD):** Power spectral density (PSD) is described as the distribution of energy with frequency. In other words, it shows at which frequencies variations are strong and at which frequencies variations are weak. Mathematically:

$$\Omega(w) = \left| \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iwt} dt \right|^2 \quad (14)$$

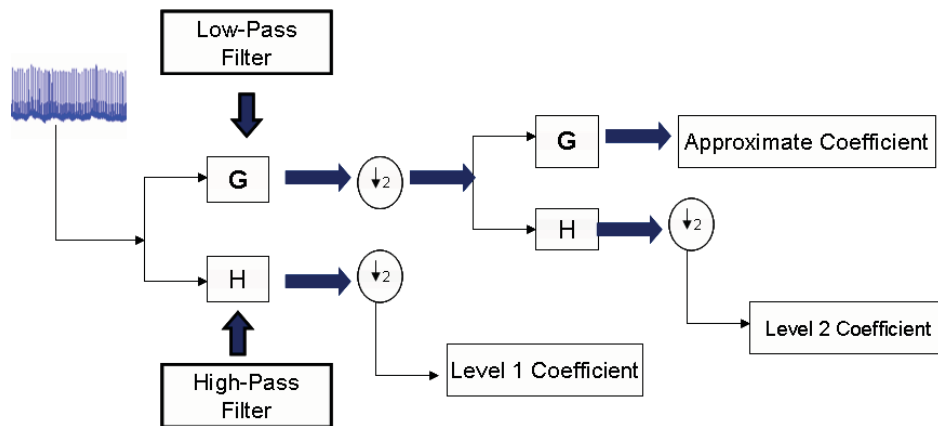


Figure 16: Detailed schematic diagram of Discrete Wavelet Transformation with level 2.

where  $f(t)$  is the correlation function of the signal. The traditional way to analyze HRV using PSD is to use an average power of high frequency, low frequency, very low frequency, HF\_normalize, LF\_normalize, ratio of HF to LF, and ratio of LF to HF [116, 86]. The PSD is calculated from RR intervals using linear interpolation, then re-sampled at 5Hz, and a low-pass filter is applied with cut-off frequency of 0.5Hz. The Fourier transformation with Hanning window is employed to obtain the power spectra. The spectral power is exhibited as the integrated area of HF, LF, and VLF. The power between 0.15 Hz and 0.4Hz is considered as the power of the high frequency (HF), between 0.04Hz and 0.15Hz as the low frequency (LF), and between 0.003Hz and 0.04Hz as very low frequency (VLF) range. Also, HF\_normalize and LF\_normalize are measured by normalizing HF and LF respectively by the difference between total average power and VLF:

$$HF_m = \frac{HF}{(TAP - VLF)} \quad (15)$$

$$LF_m = \frac{LF}{(TAP - VLF)} \quad (16)$$

where TAP is the total average power of RR interval [45].

**Fractal Dimension (FD):** Fractal Dimension (FD) analysis is helpful in understanding complex biological signals such as ECG. Fractals have the characteristic that each subset is similar to the whole set, and fractal dimension (FD) is a measure of this self-similarity [163, 76, 4, 136]. The Higuchi FD, as explained below, is applied for this study because it is easy to use. This method first re-generates the original signal as a finite time series based on a pre-defined window size. For this study, first the window sizes of 8 and 15 are applied, based on the values reported in the literature [4, 14, 132], and then the results of these two window sizes are compared. For a given input signal  $x(1), x(2), \dots, x(N)$ , the new finite time-series,  $x_k^m$ , is constructed as follows:

$$x(m), x(m+k), x(m+2k), \dots, x(m + [\frac{N-m}{k} \cdot k]), \quad m = 1, 2, \dots, k \quad (17)$$

where “[ ]” denotes the floor function, that is, the greatest integer that is less than or equal to the value, and both  $k$  and  $m$  are integers representing the initial time and interval. Then the length of the curve  $L_m(k)$  is defined as follows

$$L_m(k) = \frac{1}{k} \left\{ \left( \sum_{i=1}^{[\frac{N-m}{k}]} |x(m+ik) - x(m+(i-1) \cdot k)| \right) \frac{N-1}{[\frac{N-m}{k}] \cdot k} \right\} \quad (18)$$

where  $\frac{N-1}{[\frac{N-m}{k}] \cdot k}$  represents the normalization factor for the curve length and  $N$  is the total length of the signal.  $L(k)$  is defined as the length of the curve for the time series

$k$  and  $L_m(k)$  is denoted as the average value over  $k$ . Thus, if  $L(k) \propto k^{-D}$ , then the curve has the dimension  $D$ . In other words, FD identifies the slope of the best fit-line at the log-log plot for  $\log(L(k))$  versus  $\log(k)$  [4, 14, 48, 70, 132]. Theoretically, the FD for a signal should be between 1 and 2.

c) *Feature Analysis*: For DWT, mother wavelets Daubechies 4 (db4) and Daubechies 32 (db32) are applied and compared in this study. The standard deviation of the detail coefficients for each level is calculated and used as one of the features of each stage. The set of features using both db4 and db32 coefficients include:

- Standard deviation of coefficients at each level, i.e.  $sd_1, sd_2, sd_3, sd_4$ , and standard deviation of approximate coefficients, i.e.  $sa$
- Sum of square of coefficients at each level, i.e.  $sqd_1^2, sqd_2^2, sqd_3^2, sqd_4^2$ , and standard deviation of approximate coefficients, i.e.  $sqa^2$
- Median of the twenty highest coefficients at each level,
- Coefficient right before median at level 1 ( $d_{1\_before}$ ), and coefficient right after median at level 1 ( $d_{1\_after}$ ), and coefficient for middle ( $d_{1\_middle}$ ) when median is formed using the twenty highest detail coefficients,

### *Statistical Analysis*

Analysis of variance (ANOVA) is performed using the statistical software tool SAS to compare the HRV response over 4 stages (baseline to stage 4) of LBNP and exercise subjects. Comparison of LBNP and exercise using wavelet analysis and traditional way are performed respectively using ANOVA. Figure 17 illustrates the process of comparing

the results for the LBNP and exercise datasets using the traditional and wavelet analysis approaches.

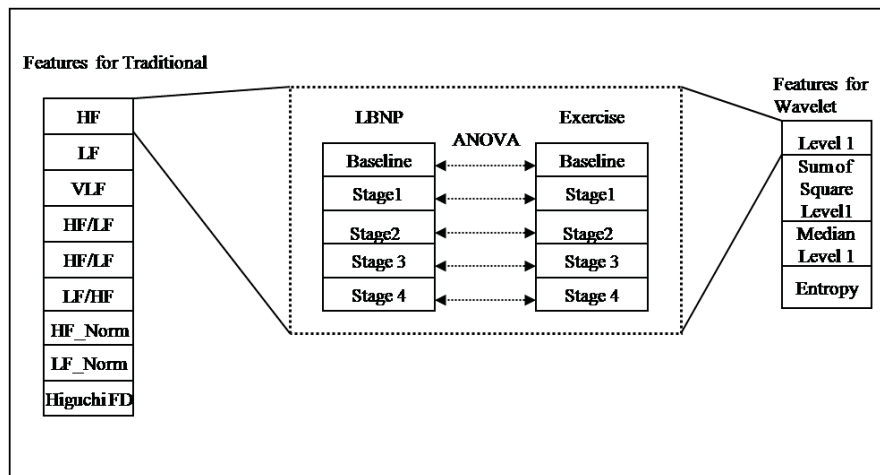


Figure 17: Comparisons between LBNP and exercise conditions with traditional and wavelet analysis.

## 6.6 Results

A total of forty-five features are extracted, thirty-six features for DWT (db 32) and DWT (db4), seven for PSD, and two features for Higuchi FD.

The data consists of the cardiovascular response over 3 stages of LBNP and 4 stages of exercise. In Study 1, ANOVA analysis is performed using the proposed approach (SAB-WEF). In study 2, ANOVA analysis is performed using the traditional approach. In study 3, classification is performed to predict the current stage of the LBNP subjects using machine learning methods. The original sampling rate is 500Hz. A 125Hz down-sampling rate is also applied to the ECG signal to examine HRV at low frequency.



### 6.6.1 Comparison of LBNP and Exercise Using Wavelet Analysis

Based on repeated two-way factor ANOVA considering subject condition (LBNP/exercise) and blood loss, it was found that the standard deviation of wavelet coefficients at level 1 (p-value=0.0348), sum of squared of wavelet coefficients at level 1 (p-value < .0001), and median of wavelet coefficients at level 1 (p-value=0.0524) are significant to distinguish between LBNP and exercise. However, the Higuchi FD (p-value=0.4377) is not significant. The results of the repeated measure ANOVA tests are presented in Table 14. Each column shows the p-value. The features based on db4 have very similar p-values as those of db32 and are significant.

Table 14: The summary of statistical comparison between LBNP and exercise with SABWEF. Results for 500 Hz and 125 Hz sampling rates are presented here. P values inside parentheses are for 125Hz and without parentheses are for 500 Hz. (SS stands for sum of squared).

Stage	Level 1 500Hz (125Hz)	SS level 1 500Hz (125Hz)	level 1 median 500Hz (125Hz)	Entropy 500Hz (125Hz)
Baseline	0.2078 (0.2775)	0.3473 (0.5148)	0.4243 (0.4050)	0.9448 (0.9637)
Stage 1	0.0214 (0.0063)	0.0065 (0.0134)	0.0112 (0.1580)	0.1181 (0.2409)
Stage 2	0.0025 (0.0085)	0.0001 (0.0007)	0.0021 (0.0039)	0.0133 (0.0233)
Stage 3	0.0235 (0.0338)	0.0109 (0.0162)	0.0175 (0.0385)	<0.0001 (<0.0001)
Stage 4	0.0572 (0.0746)	0.0209 (0.0392)	0.0805 (0.0468)	<0.0001 (<0.0001)

Based on this statistical analysis there is sufficient evidence to claim that SABWEF accurately differentiates between LBNP and exercise subjects.

Figure 18 shows the average pattern and standard deviation of some of the above-mentioned features for LBNP and exercise groups at different stages. These patterns include the standard deviation of wavelet coefficients at level 1 using db4 and the median

of wavelet coefficients at level 1 using db4. Pattern comparisons from baseline to 5 stages of LBNP and exercise are shown in Figure 18.

### 6.6.2 Comparison of LBNP and Exercise with Traditional Approach

This study also compares the results of the traditional methods, which measures the average power of HF, LF, VLF, ratio of LF to HF, and ratio of HF to LF at specific frequency bands using PSD and Higuchi FD. As mentioned previously, the RR interval is used to measure PSD with RR interpolation. Table 15 shows the comparison results across the LBNP and exercise groups. Based on Table 15, the features extracted using the traditional approach may not be sufficient to distinguish between the LBNP and exercise subjects. Thus, the traditional method may not prove effective in differentiating LBNP from exercise subjects.

Table 15: Statistical comparisons for HRV traditional approach. Results for 500 Hz and 125 Hz sampling rates are presented here. P values for 125 Hz are inside parenthesis.

Stage	HF 500Hz (125Hz)	LF 500Hz (125Hz)	VLF 500Hz (125Hz)	HF_nm 500Hz (125Hz)	LF_nm 500Hz (125Hz)	LF/HF 500Hz (125Hz)	HF/LF 500Hz (125Hz)	Higuchi 500Hz (125Hz)
Baseline	0.5897 (0.9973)	0.1460 (0.2487)	0.8536 (0.6934)	0.6180 (0.7511)	0.5561 (0.4247)	0.5067 (0.4084)	0.1657 (0.1003)	0.0165 (0.2393)
Stage 1	0.9062 (0.6057)	0.0656 (0.0714)	0.1905 (0.1693)	0.5825 (0.6365)	0.7617 (0.8003)	0.9436 (0.9296)	0.4306 (0.4082)	0.1401 (0.1580)
Stage 2	0.6735 (0.7401)	0.8286 (0.6759)	0.5750 (0.3316)	0.8781 (0.7915)	0.3133 (0.5642)	0.7966 (0.7829)	0.4045 (0.3599)	0.1001 (0.3122)
Stage 3	0.3680 (0.7152)	0.3911 (0.1466)	0.6291 (0.3358)	0.9361 (0.5476)	0.1179 (0.1378)	0.9152 (0.5619)	0.4584 (0.4832)	0.9841 (0.8495)
Stage 4	0.6054 (0.7082)	0.1331 (0.325)	0.7488 (0.1199)	0.3554 (0.1900)	0.3504 (0.0616)	0.1683 (0.1639)	0.1635 (0.0965)	0.1304 (0.1953)

These results are supported by a recent study reporting that HF/LF measure is not sufficient to distinguish across LBNP and exercise at 500Hz [123]. It was found that low sampling rates, e.g. 125Hz, cannot distinguish between LBNP and exercise.

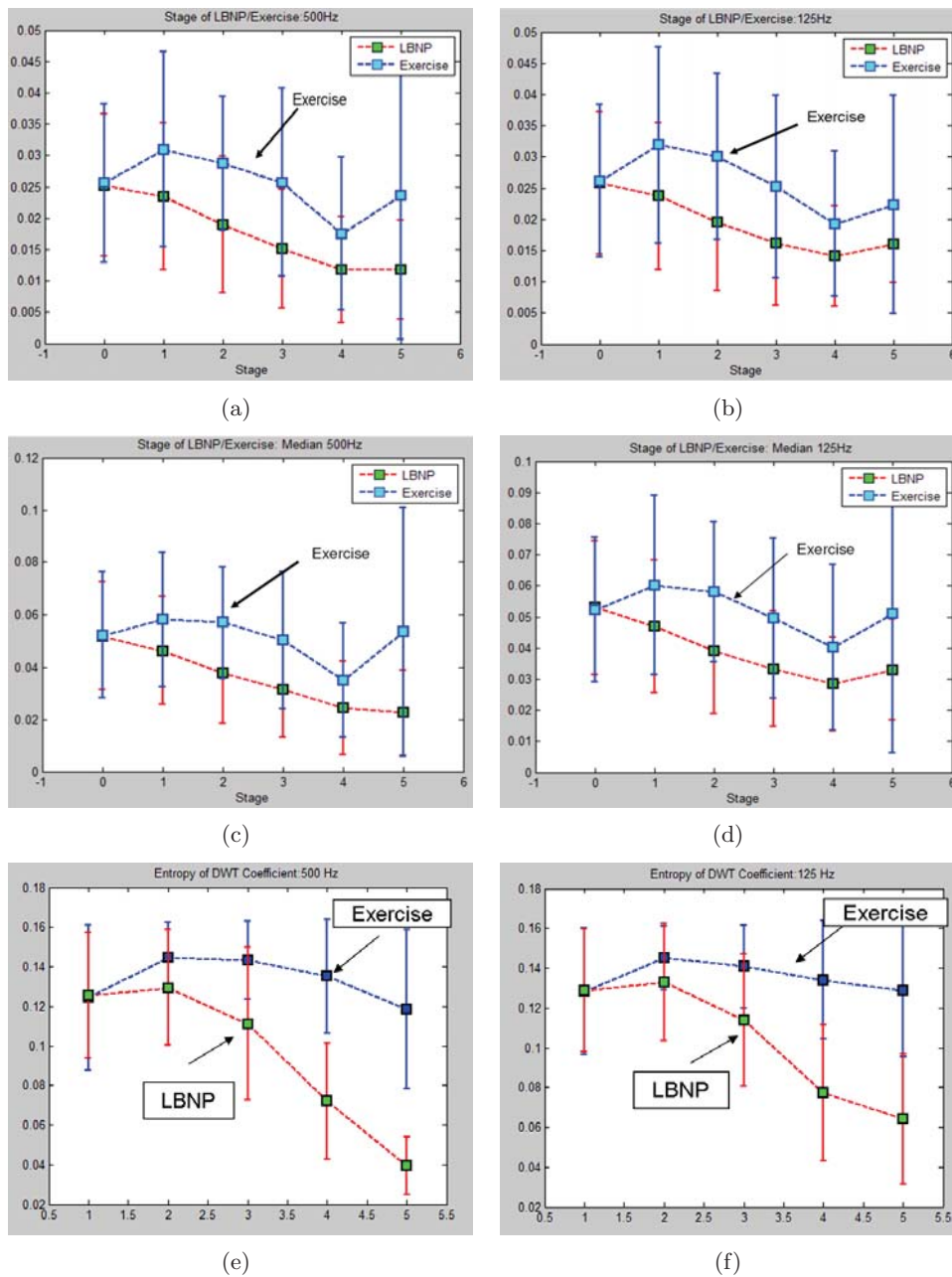


Figure 18: Pattern comparison using wavelet features: Standard deviation of wavelet coefficients at level 1 using db4 pattern of LBNP and exercise 500 Hz 18(a) and 125 Hz 18(b) Median of wavelet coefficients at level 1 using db4 of LBNP and exercise at 500 Hz 18(c) and 125Hz 18(d) Entropy of wavelet coefficients at level 1 using db4 of LBNP and exercise at 500 Hz 18(e) and 125Hz 18(f) .

Figure 19 shows the patterns of traditional approach for LBNP and exercise groups at different stages. The HF and ratio between HF and LF patterns are presented in here.

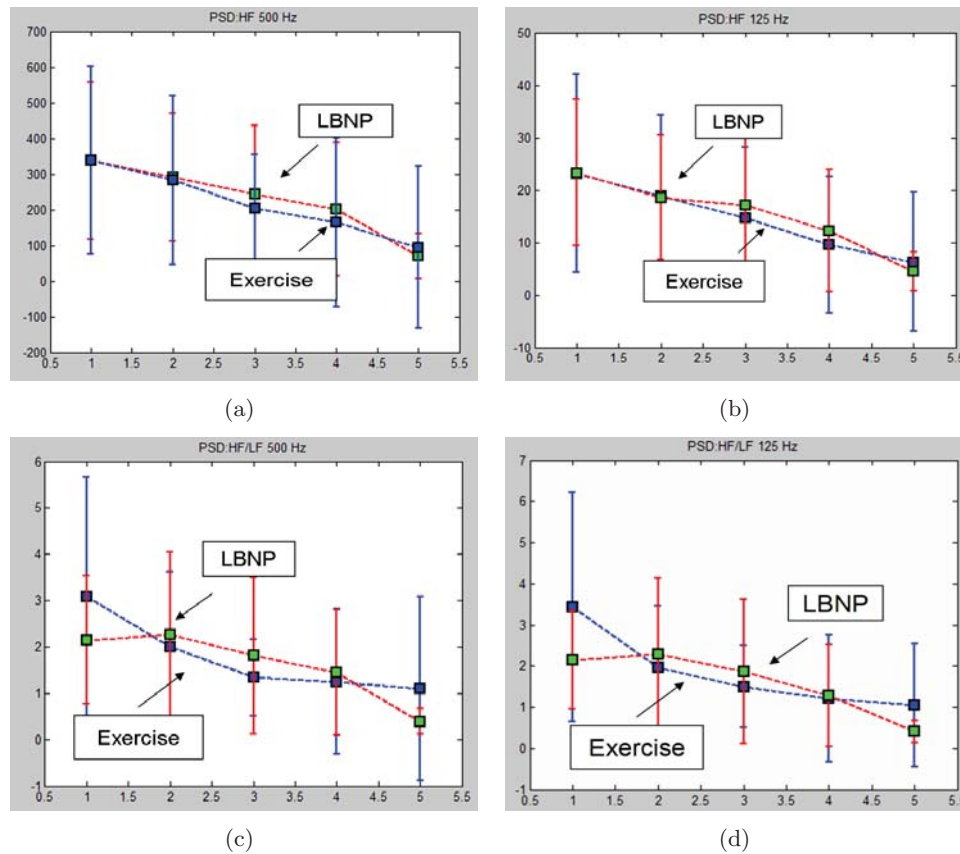


Figure 19: Pattern comparison using PSD features: HF using PSD of LBNP and exercise 500 Hz 19(a) and 125 Hz 19(b) HF/LF using PSD of LBNP and exercise at 500 Hz 19(c) and 125Hz 19(d) .

### 6.6.3 Classification of Severity of Hemorrhage

This section presents the classification of hemorrhage severity using the LBNP human model. In order to predict blood volume loss, it is divided into 3 classes (mild: -15 to -30 mmHg; moderate: -45 to -60 mmHg; severe: over -60 mmHg). These levels correspond to estimated blood losses of 400-550 cc, 500-1000 cc and greater than 1000 cc respectively.

Two types of classification study are performed. In the first, only the ECG is used to predict the severity of hemorrhage. In the second four signals, ECG, arterial blood pressure (ABP), and impedance signals (IZT, and DZT), are used to predict the hemorrhage severity using LBNP. Figure 20 shows the schematic diagram of the classification process for study 1.

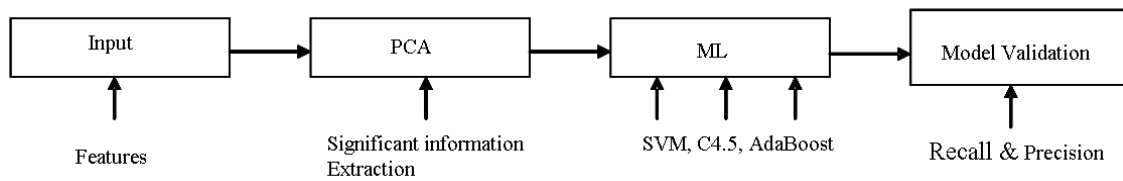


Figure 20: Schematic diagram of classification (Study 1).

The input ECG used for processing includes the baseline ECG as well as the ECG signal of the stage to be classified. From each of the ECG signals, all features listed in Section 6.3 are extracted. Since these features may have some level of correlation with each other, principal component analysis (PCA) is used in order to eliminate potential redundancy. The resulting PCA features are then used as input to the machine learning algorithms (SVM, C4.5, and AdaBoost). Once the features are classified by the machine learning algorithm, precision and recall of the algorithm are calculated to validate the model and testing is performed using 10-fold cross validation.

Note that a subject that collapses at stage 7 contributes to the data in all classes, i.e. this subject produces a set of input-output data for class 1, a set for class 2, and another set for class 3. Thus, the number of samples produced for this study is much larger than the number of subjects. A total of 219 samples are considered for classification into three

classes using 10-fold cross validation. Specifically, class 1 has 92 subjects, class 2 has 88 subjects, and class 3 has 39 subjects.

Table 16 describes the classification results of study 1. Precision and recall are used for assessing model performance. Precision is the probability of correctly predicting the experiment result and recall (i.e. sensitivity) indicates the probability that the experiment prediction is correct.

Table 16: Classification result using ECG signal only. TP indicates a True Positive

Method	C4.5	AdaBoost	SVM
Accuracy	74.4%	69%	77.2%
TP for class 1 (Mild)	82/92	91/92	89/92
TP for class 2 (Moderate)	58/88	45/88	62/88
TP for class 3 (Severe)	30/39	15/39	18/39

According to the comparison results, SVM has the higher prediction accuracy, 77.2%. SVM's average precision and recall (sensitivity) for all three classes are 71.4% and 79.3%, respectively. However, C4.5 seems to have more reliable results as it correctly classifies 30 out of 39 cases in the severe class, which is an important factor for clinical decision making. C4.5 also has an accuracy of 74.4% and the average precision and recall of 77.4% and 76.1%, respectively.

In study 2, four signals, ECG, ABP, and impedance signals (IZT and DZT), are used to predict the severity of hemorrhage into three class as described above. Figure 21 presents the schematic diagram of the classification process using four signals for study 2. In this study, ECG, ABP, and impedance (IZT and DZT) are considered as input signals in order to predict the severity of blood loss. Unlike study 1, approximate coefficients with db4

level 6 are used for this study. Let  $c_i$  be an approximate coefficient of a signal, and  $d_i^s$  (where  $s$  indicates each level) is a detail coefficient of the signal. Then the features are defined as follows:

$$\lambda_1 = \frac{1}{N} \sum_i^N (c_i)^2 \quad (19)$$

$$\lambda_2 = \text{median}(c_i)^2 \quad (20)$$

$$\lambda_3 = -\lambda_1 \log_2(\sum d^1 + d^2 + d^3 + d^4 + d^5 + d^6 + \lambda_1) \quad (21)$$

where  $N$  is the length of coefficient and  $d^s = \frac{1}{N} \sum_{i=1}^N d_i^s$ , where  $s = 1, \dots, 6$ . Thus, a set of features,  $\alpha = \{\lambda_1, \lambda_2, \lambda_3\}$ , are extracted from each signal and used as features set.

Testing is done with 10-fold cross validation.

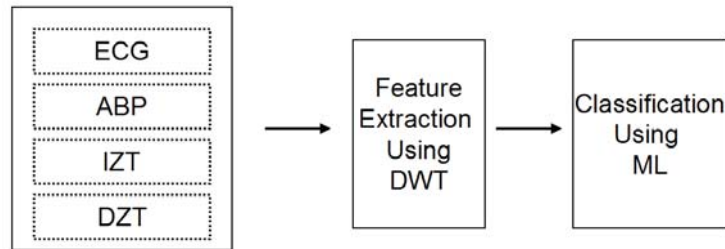


Figure 21: Schematic diagram of classification using four signals (study 2).

According to this study using four signals, SVM has the highest prediction accuracy, 83%. Based on this result, adding more information (signals) may provide better performance than using a single signal. Also, the two studies of classification results show that the features calculated via the wavelet method are useful in predicting severity of hemorrhage. Verification that these features are applicable to the analysis of other signals is tested in Chapter 7.

## 6.7 Discussion

Diagnosis and assessment of hemorrhage based on low level physiologic signals such as heart rate, including determining the severity of hemorrhage, remains a challenge. Heart rate variability (HRV) contains significant information regarding cardiovascular activities, and can provide additional information about autonomic control of the heart rate. This can be used to evaluate the degree of hemorrhage shock, and assist in assessing the effects of treatment before cardiovascular collapse occurs. Studying the effects of HRV may help improve the quality of medical care in cases of hemorrhage shock. Several previous studies have examined the use of HRV detection in reducing the mortality rate of patients in the field. Rapid response and early detection are potential factors in improving the chance of survival from severe blood loss. In a combat environment, the differentiation of sensitive HRV response due to blood loss or physical activity is essential for determining appropriate treatment.

Using an LBNP model of hemorrhage, it has been shown that wavelet analysis can differentiate between hemorrhage and exercise based on heart rate variability. Although preliminary studies indicate that FD analysis has a superior performance in distinguishing normal and pathological subjects, this study has shown that it may not be sufficient to differentiate physically active soldiers from a bleeding soldier in a combat situation. Even though preliminary results indicate that the traditional approach, PSD, is a good noninvasive tool for studying HRV, it cannot effectively differentiate between volume loss and exercise subjects. Contrary to previous FD and PSD analysis results, wavelet anal-



ysis may well be capable of informing users of the physical condition of the combatant. Level 1, sum of squared level 1, median of level 1, entropy, and approximate coefficient using wavelet coefficient are significantly more effective in distinguishing hemorrhage and physical activity. In particular, the entropy measure can differentiate well between the two conditions, as well as different stages of blood loss.

Another important issue in ECG analysis is noise. Notch and band pass filters are well suited to remove the noise caused by power lines and subject movement. Subject movement can significantly affect RR interval detection and generate incorrect results; it must therefore be dealt with carefully.

Despite some limitations in the dataset, signal analysis based on wavelet-extracted features (SABWEF) has been confirmed as useful in differentiating between hemorrhage and physical activity. According to the result of study 2, information extracted from more signals may provide an even better prediction of hemorrhage severity. Therefore, in future continuation of this study, additional signals such as mean arterial pressure (MAP) and respiratory rate will be added and analyzed to predict blood loss.

## **6.8 Conclusion**

As mentioned previously, since rapid response and early detection of HS are important to improving the likelihood of survival in the battlefield, a combat medic must be able to differentiate between seriously injured soldiers with blood loss and those undergoing physical activity in order to make accurate remote triage decisions. Therefore, the proposed method, signal analysis based on wavelet-extracted features (SABWEF), to improve

the quality of monitoring will provide the more accurate detection of the HS condition in battlefield conditions. The development of accurate and fast methods for real-time electrocardiogram (ECG) analysis is vital, and can be combined with automated monitoring devices for soldiers and other high-risk individuals for early detection and evaluation of HS. Our method is suitable for long-term ECG monitoring with low computational costs and a low sampling rate. Thus, this research initiative represents an important step in improving care for both civilians and combat soldiers.

## CHAPTER 7 Testing & Validation Using Gait in Aging Dataset

This chapter presents the results for further test and validate the capabilities of the signal analysis based on wavelet-extracted features (SABWEF) method. In Section 7.1, the gait in aging dataset is described. Comparison results are presented in Section 7.2.

### 7.1 Description of Dataset

This gait in aging dataset was obtained from the Physio Bank [3]. Data was collected from healthy subjects and from subjects with Parkinson's disease by measuring the stride interval signal, the time between successive heel strikes of the same foot. the dataset for five subjects with each condition were used for this study. The signal was sampled at 300 Hz. The statistical results between our defined features (level 1 wavelet coefficient, sum of square of level 1 coefficient, and median of level 1 coefficient) and traditional features using PSD were compared using ANOVA.

### 7.2 Results

This section presents the comparing results of the wavelet features and the traditional approach.

Table 17 presents the statistical comparison between Parkinson's disease and healthy subjects using wavelet features. A small p-value ( $<0.005$ ) indicates that the wavelet fea-

tures introduced can be used to distinguish the two types of subject.

Table 17: The summary of statistical comparison between Parkinson's disease and healthy subject with wavelet features.

Wavelet	Level 1	SS level 1	level 1 median
Disease vs healthy	0.0444	0.0120	0.0233

Table 18 presents the difference between Parkinson's disease and healthy subject when using traditional features. The larger p-value ( $>0.005$ ) indicates that the traditional approach with PSD cannot distinguish between the two subjects.

Table 18: Statistical comparisons for between Parkinson's disease and healthy subjects with traditional features.

Traditional Features	HF	LF	VLF	HF_nm	LF_nm	LF/HF	HF/LF
Disease vs. Healthy	0.3401	0.3518	0.2796	0.1629	0.4228	0.0697	0.3456

As shown in Table 17 and Table 18, it was found that SABWEF is applicable in differentiating healthy subjects and those with the Parkinson's disease.

This chapter has shown that the wavelet-based method may be useful for other signals as well as HRV analysis. Statistical analysis based on a small dataset shows that wavelet features can identify between the gait of healthy and diseased subjects in a dataset of aging patients. A larger dataset is needed for further testing.

## CHAPTER 8 Conclusion and Future Work

A computer assisted rule-based system is designed to help trauma physicians make faster and more precise decisions. Due to the complex nature of medical data, feature selection is crucial in increasing the prediction accuracy as well as transparency of the decisions recommended by the system. A rule-based system using only significant features may not only help trauma experts better predict outcomes such as the likelihood of survival, but also provide more transparent recommendations to experts. Also, by introducing a novel set of wavelet-based features, the critical issue of extracting hidden information from a highly complex signal which may not be identified by traditional method such as Fourier Transform and fractal analysis is addressed. This chapter is organized as follows. First, a summary of the work is described in Section 8.1. Then possible future directions are presented in Section 8.2.

### 8.1 Summary

Among causes of death and permanent disability, traumatic injuries are the most prevalent in both civilian and military settings. Also, trauma injuries, which are often associated with bleeding, are the main risk factors in determining the chance of survival. Since decisions in trauma care giving have to be made quickly, developing a system for computer-assisted trauma decision making has gained interest amongst medical researchers.

However, developing this kind of system is not easy for the following reasons: patient

information is complex in nature, the trauma unit is a stressful environment, the injury types are very diverse, and the need for rapid decisions underlies every aspect of trauma medicine. In addition, the most critical issue in decision making is the ability to integrate all available complex clinical information and recommend a reasonable course of treatment.

Therefore, the main goal of the first part of this research is to design computer aided trauma decision making based on rules which are transparent to experts, easy to understand, and have the capability of being used for diagnosis of new patients based on faster and more accurate decisions. As a result, two complementary techniques, logistic regression for finding significant variables and machine learning for rule extraction, are used to provide accurate and understandable recommendations for physicians. Logistic regression is useful for describing relationships between multiple independent variables and a specific outcome. On the other hand, decision tree algorithms have the benefits of being easy to understand and interpret, being capable of using categorical variables, and having the ability to deal with missing values. The combination of statistical analysis and decision tree algorithms, specifically CART and C4.5, and using only significant variables extracted via logistic regression can create high quality rules that make sense to physicians. Also, this combination approach was found to be useful for non-medical as well medical decision-making.

The proposed computer-aided rule based system has significant benefits: 1) It provides rule-based recommendations and enables optimal resource utilization. This may assist physicians in providing the highest possible standard of care to patients with traumatic injuries. 2) A previously recorded case that resembles that of a new patient can be identified in order to recommend a suitable treatment. The diagnosis of future patients may be

improved by analyzing all possible rules associated with their symptoms. 3) The system also may deliver potential guidance to the physicians based patient characteristics which may improve efficiency of expenditures. 4) The system could be used as a teaching tool to educate students about treatment of trauma patients.

In the second study, another algorithm is developed whose main application is an important problem in trauma care. Since hemorrhage shock (HS) is a consequence of serious trauma injuries, diagnosis and assessment of HS, including determination of the severity of hemorrhage based on low level physiological signals such as heart rate, remains a challenge. Heart rate variability (HRV) contains significant information regarding cardiovascular activities, and can provide additional information about abnormality of the heart function. This can be used to evaluate the degree of the blood loss and assist in assessing the effects of treatment before cardiovascular collapse occurs. Studying the effects of HRV may help improve the quality of medical care in the case of hemorrhage shock, and several previous studies have examined the use of early HRV change detection as a potential factor in improving the chance of survival in the event of severe blood loss.

The differentiation of sensitive HRV response due to blood loss or physical activity is essential in determining appropriate treatment and improving the survival rate in a combat environment, given that most deaths in combat occur in the first four hours after a soldier is injured. Preliminary studies indicate that fractal dimension (FD) and power spectral density (PSD) analysis may not be sufficient to differentiate physically active soldiers from bleeding soldiers. The second part of research proposes that analysis of novel wavelet features that may be capable of making this distinction. Features such as level 1 coefficients, sum of squared of coefficients at level 1, relative entropy of coefficients

at level 1, and median of coefficients at level 1 are calculated using detail coefficient of wavelet transformation. This approach was shown to be significantly more effective than FD and PSD analysis in distinguishing hemorrhage from physical activity.

Since rapid response and early effective hemorrhage control has obvious benefits and can improve the chance of survival more than any other measure, the novel wavelet features proposed by this research will improve the quality of monitoring by providing more accurate detection of the HS condition in civilians as well as provide valuable information for making remote triage decisions on the battlefield. Note that the signal analysis based on wavelet-extracted features (SABWEF) also may contribute to predicting early detection of severe blood loss before it has occurred. Also, the proposed approach may be useful for analyzing other complex signals, as shown in this research.

## **8.2 Future Work**

The main future direction in developing reliable rule-based trauma decision making is adding more information such as signal features and image features to extract more significant features from all integrated information. Other valuable signals such as respiratory rate and transcranial Doppler (TCD), measured using ultrasonography, will help improve the prediction accuracy of factors such as the volume of blood loss. More signal processing features, extracted from ECG, such as P duration time, T duration time, ratio S wave to T wave, and ratio P wave to R wave, will also help improve predictions. Such a prediction system will provide invaluable information in estimating early risk factors before hemorrhage shock actually occurs. It is also likely that such analysis will prove to be useful for



determining the states of other critical illness and injury including sepsis, cardiac arrest, prediction of sudden death, the ability to defibrillate, and many others.

## Bibliography

## Bibliography

- [1] Office of public affairs, u.s. department of transportation, 2003.
- [2] The coalition for american trauma center, 2005.
- [3] physiologic signal archives for biomedical research, 2008.
- [4] R. Acharya, P.S. Bhat, N. Kannathal, A. Rao, and C.M. Lim. Analysis of cardiac health using fractal dimension and wavelet transformation. *ITBM-RBM*, 26:133–139, 2005.
- [5] Advanced Trauma Life Support (ATLS) Course for Physicians. *American College of Surgeons Committee on Trauma*, 1997.
- [6] V.X. Afonso, W.J. Tompkins, T.Q. Nguyen, and S. Luo. Ecg beat detection using filter banks. *IEEE Transactions on Biomedical Engineering*, 46(2):192–202, 1999.
- [7] S. Akselrod and D. Gordon. Power spectrum analysis of heart rate fluctuation: A quantitative probe of beat-to-beat cardiovascular control. *Amer. J. Physiol.*, 249:H867–H875, 1985.
- [8] R.H. Alexander and H.J. Proctor. Committee on trauma resource document: trauma prevention, in advanced trauma life support, 1993.
- [9] T. Allan, J.D. Pryor, S.J. Clark, and W.A. Miller H.R. Warner. Help: A computerized system for medical decision making. *IEEE Computer magazine*, 8(1):34–38, 1975.
- [10] J.R. Anderson, S. Amarel, R.S. Michalski, J.G. Carbonell, and T.M. Mitchell. *Machine Learning: An Artificial Intelligence Approach*. Morgan Kaufmann, 1983.
- [11] R.N. Anderson, A.M. Minino, L.A. Fingerhut, M. Warner, and M.A. Heinen. Deaths: Injuries, 2001. *National Vital Statistics Reports*, 52(21):1–87, 2001.
- [12] P.J. Andrews, D.H. Sleeman, P.F. Statham, A. McQuatt, V. Corruble, P.A. Jones, T.P. Howells, and C.S. Macmillan. Predicting recovery in patients suffering from traumatic brain injury by using admission variables and physiological data: a comparison between decision tree analysis and logistic regression. *Journal of Neurosurg.*, 97(2):326–336, 2002.
- [13] C.L. Arfkbn, M.J. Shapiro, P.Q. Besse, and B. Littenberg. Effectiveness of helicopter versus ground ambulance services for interfacility transport. *Journal of Trauma-Injury Infection & Critical Care*, 45(4):785–790, 1998.

- [14] Y. Ashkenazy. The use of generalized information dimension in measuring fractal dimension of time series. *Physica A*, 271(3):427–447, 1999.
- [15] A. Asuncion and D.J. Newman. UCI machine learning repository, 2007.
- [16] L.J. Badra, W.H. Cooke, J.B. Hoag, A.A. Crossman, T.A. Kuusela, K.U.O. Tahvanainen, and D.L. Eckberg. Respiratory modulation of human autonomic rhythms. *Am J Physiol Heart Circ Physiol*, 280:H2674–H2688, 2001.
- [17] Y.W. Bai, W.Y. Chu, C.Y. Chen, Y.T. Lee, Y.C. Tsai, and C.H. Tsai. Adjustable 60hz noise reduction by a notch filter for ecg signals. In *Instrumentation and Measurement technology Conference (IMTC)*, pages 1706–1711, 2004.
- [18] A.I. Batchinsky, W.H. Cooke, T. Kuusela, and L.C. Cancio. Loss of complexity characterizes the heart rate response to experimental hemorrhagic shock in swine. *Crit Care Med*, 35(2):519–525, 2007.
- [19] R.F. Bellamy. The causes of death in conventional land warfare: implications for combat casualty care research. *Mil Med*, 149(2):55–62, 1984.
- [20] C.M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [21] A.L. Bluma and P. Langley. Selection of relevant features and examples in machine learning. *Artificial Intelligence*, 97(1-2):245–271, 1997.
- [22] L. Breiman. *Classification and Regression Trees*. Chapman & Hall, Boca Raton, 1993.
- [23] C.M. Brown, M.J. Hecht, M.J. Hilz, and B. Neundorfer. Effects of lower body negative pressure on cardiac and vascular responses to carotid baroreflex stimulation. *Physiological Research*, 52:637–645, 2003.
- [24] C.A. Bond, C.L. Raehl, and T. Franke. Medication errors in united states hospitals. *Pharmacotherapy*, 21(9):1023–1036, 2001.
- [25] C.J. Carrico, J.B. Holcomb, and I.H. Chaudry. Post resuscitative and initial utility of life saving efforts: scientific priorities and strategic planning for resuscitation research and life saving therapy following traumatic injury. *Acad Emerg Med.*, 9:621–626, 2002.
- [26] Centers for Disease Control and Prevention. *Traumatic Brain Injury*, 1997.
- [27] Centres for Disease Control and Prevention. *Facts about Concussion and Brain Injury*, 2004.

- [28] C.-H.H. Chu and E.J. Delp. Impulsive noise suppression and background normalization of electrocardiogram signals using morphological operators. *IEEE Transactions on Biomedical Engineering*, 36(2):262–273, 1989.
- [29] D.V. Cicchetti. Neural networks and diagnosis in the clinical laboratory: state of the art. *Clin Chem.*, 38:9–10, 1992.
- [30] A. Coats, J. Conway, J. Isea, G. Pannarale, P. Sleight, and V. Somers. Heart rate variability and its association with mortality in prehospital trauma patients. *Journal of Physiology*, 413:289–298, 1989.
- [31] V.A. Convertino, W.H. Cooke, and John B. Holcomb. Arterial pulse pressure and its association with reduced stroke volume during progressive central hypovolemia. *The Journal of TRAUMA Injury, Infection, and Critical Care.*, 61:629–634, 2006.
- [32] W.H. Cooke, J.Salinas, V.A. Convertino, D.Hinds D.A. Ludwig, J.H. Duke, F.A. Moore, and J.B. Holcomb. Heart rate variability and its association with mortality in prehospital trauma patients. *The Journal of TRAUMA Injury, Infection, and Critical Care*, 60(2):363–70, 2006.
- [33] W.H. Cooke, K.L. Ryan, and V.A. Convertino. Lower body negative pressure as a model to study progression to acute hemorrhagic shock in humans. *J Applied Physiology*, 96(4):1249–1261, 2004.
- [34] W.H. Cooke, J. Salinas, V.A. Convertino, D.A. Ludwig, D. Hinds, J.H. Duke, F.A. Moore, and J.B. Holcomb. Systemic and forearm vascular resistance changes after upright bicycle exercise in man. *The Journal of TRAUMA Injury, Infection, and Critical Care*, 60:363–370, 2006.
- [35] P.B. Corr, K. Yamada, and F.X. Witkowski. *Mechanisms controlling cardiac autonomic function and their relationships to arrhythmogenesis*. In: *The Heart and Cardiovascular System*. Raven Press, New York, 2006.
- [36] P. Cunningham, R. Rutledge C.C. Baker, and T.V. Clancy. A comparison of the association of helicopter and ground ambulance transport with the outcome of injury in trauma patients transported from the scene. *Journal of Trauma*, 43(6):940–946, 1997.
- [37] I. Daubechies. *Ten Lectures on Wavelets(CBMS-NSF Regional Conference Series in Applied Mathematics)*. SIAM: Society for Industrial and Applied Mathematics, Philadelphia, 1 edition, 1992.
- [38] F.T. deDombal, D.J. Leaper, and J.R. Staniland. Computer aided diagnosis of acute abdominal pain: a multicentre study. *BMJ*, 293(6550):800–804, 1986.

- [39] S.C. Ducrocq, P.G. Meyer, G.A. Orliaguet, S. Blanot, D. Renier, and P.A. Carli. Epidemiology and early predictive factors of mortality and outcome in children with traumatic severe brain injury: Experience of a french pediatric trauma center. *Pediatric Critical Care Medicine*, 7(5):496–497, 2006.
- [40] R.O. Duda and P.E. Hart. *Pattern Classification and Scene Analysis*. Wiley, New York, 1973.
- [41] R.O. Duda and P.E. Hart. *Pattern Classification and Scene Analysis*. John Wiley & Sons Inc, New York, 1973.
- [42] B.J. Eastridge, A. Starr, J.P. Minei, G.E. OKeefe, and T.M. Scalea. The importance of fracture pattern in guiding therapeutic decision making in patients with hemorrhagic shock and pelvic ring disruption. *Journal of Trauma-Injury Infection and Critical Care*, 53(3):446–451, 2002.
- [43] M. Eckstein, T. Jantos, N. Kelly, and A. Cardillo. Helicopter transport of pediatric trauma patients in an urban emergency medical services system: A critical analysis. *The Journal of trauma, injury, infection, and critical care*, 53(2):340–344, 2002.
- [44] J.H. Eichom. Standards for patient monitoring during anesthesia at harvard medical school. *Journal of the American Medical Association*, 258(8):1017–1020, 1986.
- [45] Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology. Heart Rate Variability : Standards of Measurement, Physiological Interpretation, and Clinical Use. *Circulation*, 93(5):1043–1065, 1996.
- [46] A.C. Enthoven and N.A. Wilson. Health-care financing and technology assessment. In *Medical Informatics: Computer Applications in Health Care*. Addison-Wesley Longman Publishing Co., Inc., Boston MA. USA, 1990.
- [47] T.J. Esposito, N.D. Sanddal, J.D. Hansen, and S. Reynolds. Analysis of preventable trauma deaths and inappropriate trauma care in a rural state. *Journal of Trauma*, 39(5):955–962, 1995.
- [48] R. Esteller, G. Vachtsevanos, J. Echauz, and B. Litt. A comparison of waveform fractal dimension algorithms. *IEEE Transactions on Circuits and Systems: Fundamental Theory and Applications*, 48(2):177–183, 2001.
- [49] Expert Working Group, Atlanta, GA. *Traumatic Brain Injury in the United States: Assessing Outcomes in Children*, 2000.
- [50] T.C. Fabian, J.H. Patton, M.A. Croce, G. Minard K.A. Kudsk, and F.E. Pritchard. Blunt carotid injury: Importance of early diagnosis and anticoagulant therapy. *Annals of Surgery*, 223(5):513–525, 1996.

- [51] L. Fagraeus and D. Linnarsson. Autonomic origin of heart rate fluctuations at the onset of muscular exercise. *Journal of Applied Physiology*, 40:679–782, 1976.
- [52] W. Fan, S.J. Stolfo, and J. Zhang. The application of adaboost for distributed, scalable and on-line learning. In *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 362–366, 1999.
- [53] J.M. Fitzmaurice, K. Adams, and J.M. Eisenberg. Three decades of research on computer applications in health care: Medical informatics support at the agency for healthcare research and quality. *Journal of the American Medical Informatics Association*, 9:144–160, 2002.
- [54] The Brain Trauma Foundation. The american association of neurological surgeons: The joint section on neurotrauma and critical care. *J Neurotrauma*, 17:457–53, 2000.
- [55] Y. Freund and R. Schapire. A short introduction to boosting. *Journal of Japanese Society for Artificial Intelligence*, 14(5):771–780, 1999.
- [56] J.H. Friedman. A recursive partitioning decision rule for nonparametric classification. *IEEE Transactions On Computers*, 26:404–408, 1977.
- [57] C.Y. Fu. Combining loglinear model with classification and regression tree (cart): an application to birth data. *Computational Statistics & Data Analysis*, 45(4):865–874, 2004.
- [58] T.S. Furey, N. Cristianini nd N. Duy, D.W. Bednarski, M. Schummer, and D. Hausler. Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics*, 16(10):906–914, 2000.
- [59] R.M. Gardner. Monitoring of physiological data in a clinical environment. *Annual Review of Biophysics and Bioengineering*, 1:211–224, 1972.
- [60] M. Garzotto, T.M. Beer, R.G. Hudson, L. Peters, Y.C. Hsieh, E. Barrera, T. Klein, and M. Mori. Improved detection of prostate cancer using classification and regression tree analysis. *Journal of Clinical Oncology*, 23(19):4322–4329, 2005.
- [61] P.A. Gearhart, R. Wuerz, and A.R. Localio. Cost-effectiveness analysis of helicopter ems for trauma patients. *Annals of emergency medicine*, 30(4):500–506, 1997.
- [62] H.M. Guo, Y.I. Shyu, and H.K. Chang. Combining logistic regression with classification and regression tree to predict quality of care in a home health nursing data set. *Stud Health Technology Inform.*, 122:891, 2006.
- [63] B. Hasan, D.B. Alam, and J.A. DaCorta. Hemorrhage control in the battlefield: Role of new hemostatic agents. *Military Medicine*, 170:1–63, 2005.

- [64] J. Hasford, H. Ansari, and K. Lehmann. Cart and logistic regression analyses of risk factors for first dose hypotension by an ace-inhibitor. *Therapie.*, 48(5):479–482, 1993.
- [65] P.J. Haug and R.M. Gardner and K.E. Tate. Decision support in medicine: examples from the help system. *International Journal of the Computers and Biomedical Research*, 27(5):396–418, 1994.
- [66] Health and Medical Service. *US Industry Trade Outlook*, 1998.
- [67] W.L. Heap, C. Werken, R.W.Keunen, and R.J. Goris. Unstable fractures and dislocations of the pelvic ring : results of treatment in relation to the severity of injury. *Netherlands Journal of Surgery*, 37(5):148–152, 1985.
- [68] A.K. Heidi, H.W. Robert, and M.A. Welsch. Vagal modulation of the heart and central hemodynamics during handgrip exercise. *American journal of Heart and circulatory physiology*, 279(5):1648–1652, 2000.
- [69] L. Hejjel and E. Roth. What is the adequate sampling interval of the ecg signal for heart rate variability analysis in the time domain. *Physiol. Meas.*, 25:1405–1411, 2004.
- [70] T. Higuchi. Approach to an irregular time series on the basis of the fractal theory. *Physica*, 31:277–283, 1998.
- [71] W.P. Holsinger, K.M. Kempner, and M.H. Miller. Qrs preprocessor based on digital differentiation. *IEEE Transactions on Biomedical Engineering*, 18(3):121–217, 1971.
- [72] D. Hosmer and S. Lemeshow. *Applied Logistic Regression*. Wiley, New York, 1989.
- [73] R. Hotvedt, I.S. Kristiansen, O.H. Fade, J. Thoner, S.M. Almdahl, G. Bjosvik, L. Berge, A.C. Magnus, K. Mamen, T. Sparr, and K. Ytre-Arne. Which groups of patients benefit from helicopter evacuation? *Lancet*, 347(9012):1362–1366, 1996.
- [74] H.S. Hsuing, C.K. Cheng, and G.J. Jan. A real-time qrs detection and delineation algorithm based on the inflection angle of ecg waveform. In *Proceedings of the 11th Annual International Conference of IEEE Engineering in Medicine & Biology Society*, pages 138–139, 1989.
- [75] I. Jagielska. Linguistic rule extraction from neural networks for descriptive data mining. In *International Conference on Knowledge-Based Intelligent Electronic Systems*, pages 89–92, 1989.
- [76] T.T. Jartti, T.A. kuusela, J.K. Timo, and U.O. Kari. The dose-response effects of terbutaline on the variability, approximate entropy and fractal dimension of heart rate and blood pressure. *Journal of Cardiovascular Pharmacology*, 45:277–285, 1998.



- [77] S. Kadambe, S. Murray, and G.F. Boudreaux-Bartels. Wavelet transform-based qrs complex detector. *IEEE Trans. Biomed. Engineering*, 46:838–848, 1999.
- [78] J.W. Kantelhardt, S.A Zschiegner, E. Koscielny-Bunde, S.Havlin, and A. Stanley S. Bunde. Multifractal detrended fluctuation analysis of non-stationary time series. *Physica A: Statistical Mechanics and its Applications*, 316:87–114, 2002.
- [79] A. Kelemen, Y. Liang, and S. Franklin. *A Comparative Study of Different Machine Learning Approaches for Decision Making*. WSEAS Press, Piraeus, Greece, 2002.
- [80] B.U. Kohler, C. Hennig, and R. Orglmeister. The principles of software qrs detection. In *IEEE Engineering in Med. and Biol.*, pages 42–57, 2002.
- [81] I. kononenko. Machine learning for medical diagnosis: history, state of the art and perspective. *Artificial Intelligence in medicine*, 23(1):89–109, 2001.
- [82] J. Kovacevic and I. Daubechies. Special issue on wavelets and signal processing. *IEEE Trans. Signal Processing*, 41(4):507–509, 1996.
- [83] M.M. Krausz. Combat fluid resuscitation original articles. *Journal of Trauma-Injury Infection & Critical Care*, 54(5):39–42, 2003.
- [84] P.M. Kuhnert, K. Do, and R. McClure. Combining non-parametric models with logistic regression: an application to motor vehicle injury data. *Comput. Stat. Data Anal.*, 34(3):371–386, 2000.
- [85] J.F. Kurtzke and L.T. Kurland. *The epidemiology of neurologic disease*. Clinical neurology, New York, 1989.
- [86] P. Laguna, G.B. Moody, and R.G. Mark. Power spectral density of unevenly sampled data by least-square analysis: Performance and application to heart rate signals. *IEEE Transactions on Biomedical Engineering*, 45(6):698–715, 1998.
- [87] L.L. Leape. Error in medicine. *Journal of the Americans Medical Associations*, 272(23):1851–1857, 1994.
- [88] C. Li, C. Aheng, and C. Tai. Detection of ecg characteristics points using wavelet transforms. *IEEE Trans. Biomed. Engineering*, 42(1):21–28, 1995.
- [89] Y.C. Li, L. Li, W.T. Chiu, and W.S. Jian. Neural network modeling for surgical decisions on traumatic brain injury patients. *International Journal of Medical Informatics*, 57(1):1–9, 2000.
- [90] W.Z. Liu, A.P. White, S.G. Thompson, and M.A. Bramer. Techniques for dealing with missing values in classification. In *Second International Symposium on Advances in Intelligent Data Analysis*, pages 527–536, 1997.

- [91] W.Y. Loh and N. Vanichsetakul. Tree-structured classification via generalized discriminant analysis. *Journal of the American Statistical Association*, 83(403):715–725, 1998.
- [92] W.J. Long, J.L. Griffith, and H.P. Selker. A comparison of logistic regression to decision-tree induction in a medical domain. *Computers in Biomedical Research*, 26(1):74–97, 1993.
- [93] P.S. Maclin, J. Dempsey, and J. Brooks. Using neural networks to diagnose cancer. *Journal of Medical Systems*, 15(1):11–19, 1991.
- [94] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1998.
- [95] B.B. Mandelbrot. *The Fractal Geometry of Nature*. Freeman, New York, 1983.
- [96] J.R. Mccarroll, P.W. Braunstein, W. Cooper, M. Helpert, M. Seremetis, P.A. Wade, and S.B. Weinberg. Fatal pedestrian automotive accidents. *Journal of American Medical Associations*, 180:127–133, 1962.
- [97] R. McMurtry, D. Walton, D. Dickinson, J. Keilam, and M. Tile. Pelvic disruption in the polytraumatized patient : a management protocol. *Clinical or Orthopaedics and related research*, 151:22–30, 1980.
- [98] J.C. McNames, C. Crespo, M. Aboy, J. Bassale, L.Jenkins, , and B.Goldstein. Harmonic spectrogram for the analysis of semi-periodic physiologic signals. *Engineering in Medicine and Biology*, 1:143– 144, 2002.
- [99] Y. Meyer. *Wavelets Algorithms and Application*. SIAM, 1993.
- [100] A.J. Miller. *Subset Selection in Regression*. Chapman & Hall, 1990.
- [101] F.H. Miller. The legal ramifications of the nccn practice guidelines. *Journal of Oncology*, 10:35–39, 1996.
- [102] T. Mitchell. *Machine Learning*. McGraw Hill, 1997.
- [103] P.A. Morettin. From fourier to wavelet analysis of time series. In *Proceedings in Computational Statistics*, pages 111–122, 1996.
- [104] M.Pagani, F. Lombardi, and S. Guzzetti. lower spectral analysis of heart rate and arterial blood pressure as a marker of sympatho-vagal interaction in man and conscious dog. *Circ. Res.*, 59:17–193, 1986.
- [105] K. Najarian and A. Darvish. Neural networks and their applications in biomedical engineering. In M Akay, editor, *Wiley Encyclopaedia of Biomedical Engineering*. John Wiley & Sons, Hoboken, NJ, 2006.

- [106] Y. Nakamura, Y. Yamamoto, and I. Muraoka. Autonomic control of heart rate during physical exercise and fractal dimension of heart rate variability. *The American Physiological Society*, 74(2):875–8819, 1993.
- [107] Nation Center for Injury Prevention and Control, Atlanta, GA. *Traumatic brain injury in the United States: emergency department visits, hospitalizations, and deaths*, 2006.
- [108] NIH Consensus Statement Online. *Rehabilitation of Persons with Traumatic Brain Injury*, 1998.
- [109] J.T. Pan and J. Willis. A real-time qrs detection algorithm. *IEEE Transactions on Biomedical Engineering*, 32(3):230–236, 1985.
- [110] G.F. Pennal, M. Tile, J.P. Waddell, and H. Garside. Pelvic disruption: assessment and classification. *Clin Orthop*, 151:12–21, 1980.
- [111] D. Percival and A.T. Walden. *Wavelet Methods for Time Series Analysis*. Cambridge University Press, 2000.
- [112] C. Perlich, F. Provost, and J.S. Simonoff. Tree induction vs. logistic regression: a learning-curve analysis. *The Journal of Machine Learning Research*, 4:211–255, 2003.
- [113] J.F. Perry and R.J. Mccellan. Autopsy findings in 127 patients following fatal traffic accidents. *Surg., Gynec., and Obstet.*, 19:586–590, 1964.
- [114] B. Pfahringer. Compression-based discretization of continuous attributes. In *Conference Machine Learning*, pages 456–463, 1995.
- [115] T. Pohlemann, H. Tscherne, F. Baumgartel, H.J. Egbers, E. Euler, F. Maurer, M. Fell, E. Mayr, W.W. Quirini, W. Schlickewei, and A. Weinberg. Pelvic fractures: epidemiology, therapy and long-term outcome. overview of the multicenter study of the pelvis study. *Unfallchirurg.*, 99(3):160–167, 1996.
- [116] B. Pomeranz, R.J.B. Macaulay, M.A. Caudill and I. Kutz, D. Adam, and K.M. Kilborn. Assessment of autonomic function in humans by heart rate spectral analysis. *American journal of Physiology*, 248:151–153, 1995.
- [117] M. Quartararo, P. Glasziou, and C.B. Kerr. Classification trees for decision making in long-term care. *Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 50(6):298–302, 1995.
- [118] J.R. Quinlan. Learning efficient classification procedures and their application to chess end games. In *In Machine Learning: An Artificial Intelligence Approach*, pages 463–482, 1987.

- [119] J.R. Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann, 1993.
- [120] J.R. Quinlan. Improved use of continuous attributes in c4.5. *Journal of Artificial Intelligence Research*, 4:77–90, 1996.
- [121] J. A. Reggia. Computer-assisted medical decision making:a critical review. *Biomedical Engineering*, 9:605–619, 1981.
- [122] D.P. Rice and E.J. MacKenzie. Cost of injury in the united states: A report to congress. In *National Highway Traffic Safety Administration and U.S. Department of Health and Human Services*, 1989.
- [123] A.C. Rickards, L.K. Ryan, H.W. Cooke, A.S. Romero, and A.V. Convertino. Combat stress or hemorrhage? evidence for a decision-assist algorithm for remote triage. *Aviation, Space, and Environmental Medicine*, 79:670–676, 2008.
- [124] C.A. Rickards, D.C. Kenneth, L.L. Bergeron, B.L. Burton, P.J. Khatri, C.T. Lee, K.L. Ryan, W.H. Cooke, D.F. Doerr, and V.A. Convertino. Cerebral blood flow response and its association with symptoms during orthostatic hypotension. *Aviation, Space, and Environmental Medicine*, 78(7):635–658, 2007.
- [125] M.D. Rienzo, P. Castiglioni, G. Mancina, G. Parati, and A. Pedotti. Sequential spectral analysis of 24-hour arterial blood pressure and pulse interval in free-moving subject. *IEEE Transactions on Biomedical Engineering*, 36(11):1066–1075, 1989.
- [126] A. Rovlias and S. Kotsou. Classification and regression tree for prediction of outcome after severe head injury using simple clinical and laboratory variables. *Pharmacotherapy*, 21(7):886–893, 2004.
- [127] S.M. Rudolfer, G.Paliouras, and I. Peers. A comparison of logistic regression to decision tree induction in the diagnosis of carpal tunnel syndrome. *Computers and Biomedical Research*, 32:391–414, 1999.
- [128] S. Ruggieri. Efficient c4.5. *IEEE Trans. on Knowl. and Data Eng.*, 14(2):438–444, 2002.
- [129] M. Saito, T. Mano, and S. Iwase. Changes in muscle sympathetic nerve activity and calf blood flow during static handgrip exercise. *Eur J Appl Physiol*, 60:277–281, 1990.
- [130] M. Sarkar and T.Y.Leong. Characterization of medical time series using fuzzy similarity-based fractal dimensions. *Artificial Intelligence in Medicine*, 27:201–222, 2003.

- [131] A. Sauaia, F.A. Moore, E.E. Moore, K.A. Moser, R. Brannan, R.A. Read, and P.T. Pons. Epidemiology of trauma deaths; a reassessment. *Journal of Trauma*, 38:185–93, 1995.
- [132] H.E. Schepers, B. James, and B. waighe. Four methods to estimate the dimension from self-affine signals. *IEEE Engineer in Medicine and Biology*, 71(11):57 – 64,, 1992.
- [133] P.J. Schwartz and S.G. Priori. Sympathetic nervous system and cardiac arrhythmias. *Journal of Cardiovascular Electrophysiology*, 4(3):338–355, 2007.
- [134] A. Scope, M. Lynn M, A. Abargel, and A. Eldad. Mortality epidemiology in low-intensity warfare: Israel defense forces' experience. *Injury*, 32(1):1–3, 2007.
- [135] E.H. Shortliffe, B.G. Buchanan, and E.A. Feigenbaum. Knowledge engineering for medical decision making: A review of computer-based clinical decision aids. *Proceedings of the IEEE*, 67:1207– 1224, 1979.
- [136] S. Shoupeng and Q.Peiwen. A fractal-dimension-based signal-processing technique and its use for nondestructive testing. *Russian Journal of Nondestructive Testing*, 43(4):270–80, 2007.
- [137] D.F. Signorini, P.J.D. Andrews, P.A. Jones, J.M. Wardlaw, and J.D Miller. Predicting survival using simple clinical variables: a case study in traumatic brain injury. *Journal Neurol Neurosurg Psychiatry*, 66:20–25, 1999.
- [138] R.J. Simes. Treatment selection for cancer patients: application of statistical decision theory to the treatment of advanced ovarian cancer. *Journal Chronic Dis*, 38:171–186, 1985.
- [139] G.W. Snedecor and W.G. Cochran. *Statistical Methods*. Iowa State University Press, 1989.
- [140] P.M. Stevens and L.E. Lamb. Effects of lower body negative pressure on the cardiovascular system. *Am. Journal Cardiol.*, 16:506–515, 1966.
- [141] Z.R. Struzik. Determining local singularity strengths and their spectra with the wavelet transform. *Fractals*, 82:163–179, 2000.
- [142] S. Suppappola and Y. Sun. Nonlinear transforms of ecg signals for digital qrs detection: A quantitative analysis. *IEEE Trans. Biomed. Engineering*, 41:397–400, 1994.
- [143] Joachims T. Text categorization with support vector machines: learning with many relevant features. In *10th European Conference on Machine Learning*, pages 137–142, 1998.

- [144] K. H. Tan, N. Sulke, N. A. Taub, E. Watts, S. Karani, and E. Sowton. Determinants of success of coronary angioplasty in patients with a chronic total occlusion: a multiple logistic regression model angioplasty in patients with a chronic total. *Heart*, 70:126–131, 1993.
- [145] J.H. Teng and K.C. Lin. Application of classification tree and logistic regression for the management and health intervention plans in a community-based study. *Journal of Evaluation in Clinical Practice*, 13(5):741–748, 2007.
- [146] N.V. Thakor and Y.S. Zu. The american association of neurological surgeons: The joint section on neurotrauma and critical care. *J Neurotrauma*, 38(5):185–194, 1991.
- [147] K.W. Thomas, C.S. Dayton, and M. W. Peterson. Applications of adaptive filtering to ecg analysis: noise cancellation and arrhythmia detection. *IEEE Transactions on Biomedical Engineering*, 1:457–553, 1999.
- [148] D.J. Thurman, C.M.S. Clinton, and K. Dunn. Traumatic brain injury in the united states: A public health perspective. *Journal of Head Trauma Rehabilitation*, 14(6):602–615, 1999.
- [149] P.E. Trahanias. An approach to qrs complex detection using mathematical morphology. *IEEE Trans. Biomed. Engineering*, 40:201–205, 1993.
- [150] G. Tricot, T. Spencer, J. Sawyer, D. Spoon, R. Desikan, A. Fassas, A. Badros, M. Zangari, N. Munshi, and E. Anaissie. Predicting long-term (5 years) event-free survival in multiple myeloma patients following planned tandem autotransplants. *British Journal of Haematology*, 116:211–217, 2002.
- [151] F. Tsien and K. Long. Using classification tree and logistic regression methods to diagnose myocardial infarction. In *MEDINFO98*, pages 493–497, 1998.
- [152] J.V. Tu. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcome. *Journal Clinical Epidemiol*, 49(11):1225–1231, 1996.
- [153] M. Unser and A. Aldroubi. A review of wavelets in biomedical applications. *Journal Clinical Epidemiol*, 84(4):626–638, 1996.
- [154] V. Vapnik. *Statistical Learning Theory*. Wiley, 1989.
- [155] L. vargas, L. Guzman, and F. Angulo-Brow. Simple model of the aging effect in heart interbeat time series. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 67:052901.1–052901.4, 2003.
- [156] G. Vijaya, V. Kumar, and H.K. Verma. Ann-based qrs-complex analysis of ecg. *Journal Med. Eng.*, 22:160.167, 1998.

- [157] M.C. Vissers, A.H. Lindenb, and J. Cees. Consultation behavior of residents supported with a protocol processing system (protoview) at the emergency ward. *International Journal of Bio-Medical Computing*, 38:181–187, 1995.
- [158] A. L. Wade, J.L. Dye, C.R. Mohrle, and M.R. Galarneau. Head, face, and neck injuries during operation iraqi freedom ii: Results from the us navy-marine corps combat trauma registry. *Journal of Trauma-Injury Infection & Critical Care.*, 63(4):836–840, 2007.
- [159] The Washington Post. *Number of wounded in action on the rise: Iraq toll reflects medical advances, resistance troops face*, 2003.
- [160] M.W. Weissman, J.S. Markowitz, R. Ouelette, S. Greenwald, and J.P. Hahn. Panic disorder and cardiovascular/cerebrovascular problems. *Am J. Psychiatry*, 147(4):1504–1508, 1990.
- [161] S.H. Woolf. Practice guidelines, a new reality in medicine. ii. methods of developing guidelines. *Arch Intern Med*, 152:946–52, 1992.
- [162] Q. Xue, Y.H. Hu, and W.J. Tompkins. Neural-network-based adaptive matched filtering for qrs detection. *IEEE Trans. Biomed. Engineering*, 39:317–329, 1992.
- [163] V.K. Yeragani, K. Srinivasan, S. Vempati, R. Pohl, and R. BalonN. Fractal dimension of heart rate time series, an effective measure of autonomic function. *Journal of Applied Physiology*, 75(6):2429–2438, 1993.
- [164] R. Zajtchuk and G.R. Sullivan. Battlefield trauma care: Focus on advanced technology. *Military Medicine*, 160(1):1–7, 1995.

## Appendix A Reliable rules for survival prediction

Reliable rules are defined as those with accuracy greater than 85% and rules between 75% and 85%. Cg stands for coagulopathy; MI for myocardial infarction; ARDS for Acute Respiratory Distress Syndrome; EDRTS for Emergency Department Revised Trauma Score; ISS for Injury Severity Score; ID for Insulin-Dependent.

Rules	Test Accuracy	Method
(Cg='Yes') and HEAD<2 and AGE<76.65 Then Alive	29/34(85.3%)	CART
(Cg='No') and (MI='No') and AGE<61.70 and HEAD≤4 and (ARDS='No') Then Alive	334/375(89.1%)	CART
(Cg='No') and (MI='No') and HEAD≥5 and AGE<22.35 Then Alive	55/64(85.9%)	CART
ISS>28 and (Cg='No') and THORAX≤4 and 62.25≤AGE<69.00 and EDRTS≥2.88 Then Alive	10/11(90.9%)	CART
ISS>23 and (Cg='No')and THORAX≤4 and 69<AGE<72.35 Then Alive	13/15(86.7%)	CART
HEAD ≤ 2 and (MI = 'No') and (Cg = 'No') and AGE ≤ 62 Then Alive	182/206(88.3%)	C4.5
(MI = 'Yes') and AGE ≤ 62 and EDRTS > 5.39 and ISS ≤ 25 Then Alive	19/20(95%)	C4.5
THORAX > 3 and HEAD ≤ 4 and (ARDS = 'No') and AGE ≤ 62 Then Alive	126/148(85.1%)	C4.5
THORAX ≤ 2 and EDRTS ≤ 0.87 and ISS > 38 Then Dead	12/13(92.3%)	C4.5
(MI = 'Yes') and AGE > 82.6 Then Dead	16/18(88.9%)	C4.5
(MI = 'Yes') and ISS > 30 Then Dead	45/50(90%)	C4.5
HEAD > 4 and (MI = 'Yes') Then Dead	25/27(92.6%)	C4.5
(Cg = 'Yes') and HEAD ≤ 4 and AGE > 78 Then Dead	12/14(85.7%)	C4.5
(ID = 'Yes') and AGE > 78 and (MI = 'Yes') and HEAD ≤ 4 Then Dead	27/31(87.1%)	C4.5
HEAD > 0 and HEAD ≤ 2 and (ID = 'Yes') and (ARDS = 'No') and AGE ≤ 75.2 Then Alive	107/118(90.7%)	C4.5
(ID = 'Yes') and (MI = 'Yes') and HEAD > 3 Then Dead	43/49(87.8%)	C4.5
(MI = 'Yes') and (ID = 'Yes') and AGE > 78 Then Dead	32/37(86.5%)	C4.5
HEAD > 4 and (MI = 'Yes') Then Dead	25/27(92.6%)	C4.5
(MI = 'Yes') and ISS > 30 Then Dead	45/50(90%)	C4.5
(MI = 'Yes') and AGE > 79.6 and ISS > 12 Then Dead	27/30(90%)	C4.5
(Cg='Yes') and HEAD ≤ 4 and AGE > 79.6 Then Dead	12/14(85.7%)	C4.5
(ARDS = 'No') and (MI = 'No') and (Cg = 'No') and HEAD ≤ 4 and AGE ≤ 62 Then Alive	335/376(89.1%)	C4.5
(MI = 'Yes') and (ID = 'Yes') and AGE > 78 Then Dead	15/16(93.8%)	C4.5
(MI = 'Yes') and HEAD ≤ 4 and ISS > 38 Then Dead	29/34(85.3%)	C4.5
(MI = 'Yes') and AGE ≤ 61.6 and ISS > 27 Then Dead	26/30(86.7%)	C4.5
HEAD ≤ 2 and (MI = 'No') and AGE ≤ 62 and ISS ≤ 38 Then Alive	235/270(87%)	C4.5
THORAX > 0 and (ID = 'Yes') and ISS ≤ 30 Then Alive	13/14(92.9%)	C4.5

### Extracted supporting rules for survival prediction (75% - 85% accuracy).

Though these rules are not reliable enough for practical use; however, they can contain



pattern information which may be of interest to physicians.

Rules	Test Acc.	Method
(Cg='Yes') and $2.5 \leq \text{HEAD} < 3.5$ and $\text{EDRTS} < 6.07$ and $35.65 \leq \text{AGE} < 55.25$ Then Alive	10/12(83.3%)	CART
(Cg='Yes') and $\text{HEAD} \geq 3$ and $\text{EDRTS} \geq 6.07$ and $\text{THORAX} < 1$ Then Alive	33/43 (76.7%)	CART
(Cg='No') and (MI='No') and $\text{AGE} < 61.70$ and (ARDS='Yes') and $\text{HEAD} < 3$ Then Alive	50/59(84.7%)	CART
(Cg='No') and (MI='No') and $\text{ISS} \geq 24$ and $61.70 \leq \text{AGE} < 68.90$ and $\text{HEAD} \leq 3$ Then Dead	11/13(84.6%)	CART
$\text{AGE} < 61.70$ and $\text{HEAD} \leq 4$ and (MI='No') Then Alive	625/793(78.8%)	CART
$\text{HEAD} \geq 5$ and (Cg='No') and $\text{AGE} < 22.85$ Then Alive	60/73(82.2%)	CART
$\text{HEAD} \geq 5$ and (Cg='No') and $\text{EDRTS} < 5.02$ and $22.85 \leq \text{AGE} < 28$ and $\text{ISS} \geq 33$ Then Dead	11/13(84.6%)	CART
$\text{ISS} \geq 23$ and (ID='Yes') and $61.70 \leq \text{AGE} < 80.50$ and (ARDS='No') and (Cg='Yes') Then Dead	12/15(80.0%)	CART
$\text{ISS} \geq 23$ and (ID='Yes') and $\text{AGE} \geq 80.50$ Then Dead	42/51(82.4%)	CART
$\text{AGE} < 61.70$ and (Cg='Yes') and $\text{HEAD} \leq 3$ and $\text{ISS} < 42$ Then Alive	47/56(83.9%)	CART
$\text{AGE} < 61.70$ and (Cg='No') and (MI='No') Then Alive	559/706 (79.2%)	CART
(MI='No') and (Cg='Yes') and $\text{HEAD} \leq 3$ and $\text{AGE} < 60.40$ and $\text{ISS} < 42$ Then Alive	47/56(83.9%)	CART
(MI='No') and (Cg='No') and $\text{ISS} \leq 23$ and $\text{EDRTS} < 6.07$ Then Alive	578/728(79.4%)	CART
$\text{AGE} < 62$ and $\text{HEAD} \leq 4$ and $\text{ISS} \leq 25$ Then Alive	648/822(78.8%)	CART
$\text{HEAD} \geq 5$ and (Cg='No') and $\text{AGE} < 22.85$ Then Alive	60/73(82.2%)	CART
$\text{ISS} \geq 23$ and $\text{AGE} \geq 80.50$ Then Dead	45/55(81.8%)	CART
$\text{AGE} < 61.70$ and $\text{HEAD} \leq 4$ and (MI='No') Then Alive	625/793(78.8%)	CART
$\text{AGE} < 61.60$ and (MI='No') and $\text{ISS} < 42$ and (Cg='Yes') and $\text{HEAD} \leq 3$ Then Alive	47/56(83.9%)	CART
$\text{AGE} < 61.60$ and $\text{HEAD} \leq 4$ and (MI='No') and (Cg='No') Then Alive	421/503(83.7%)	CART
$\text{AGE} \geq 61.60$ and $\text{ISS} \geq 23$ and (Cg='Yes') Then Dead	44/54(81.5%)	CART
$\text{AGE} < 61.70$ and $\text{HEAD} \leq 4$ and ( $\text{ISS} < 27$ ) Then Alive	646/820(78.8%)	CART
$\text{HEAD} \geq 5$ and (Cg='No') and $22.35 \leq \text{AGE} < 25.85$ and (MI='No') and $\text{ISS} \geq 30$ Then Dead	10/13(76.9%)	CART
$\text{ISS} \geq 23$ and $61.70 \leq \text{AGE} < 74.10$ and $\text{EDRTS} < 2.88$ Then Dead	21/28(75.0%)	CART
$\text{ISS} \geq 23$ and $\text{AGE} \geq 74.10$ Then Dead	92/119(77.3%)	CART
(MI = 'No') and $\text{HEAD} \leq 4$ and $\text{AGE} \leq 62$ $\text{ISS} \leq 38$ Then Alive	508/612(83%)	C4.5
$3 < \text{HEAD} \leq 4$ and (MI='No') and (ID='Yes') and (Cg='No') and $\text{ISS} \leq 59$ Then Alive	138/175(78.9%)	C4.5
$\text{HEAD} > 1$ and (MI = 'Yes') and $\text{ISS} > 22$ Then Dead	59/70(84.3%)	C4.5

## Appendix B Reliable rules for exact outcome prediction

Reliable rules for exact outcome (Home/Rehab) are defined as those with accuracy greater than 85%. FSBP represents initial blood pressure; ISS stands for Injury Severity Score; EDGCSTOTAL is the total Glasgow Coma Score recorded in the emergency department; EDRTS is the Emergency Department Revised Trauma Score; ARDS stands for Acute Respiratory Distress Syndrome.

Rules	Test Acc.	Method
HEAD $\leq$ 3 and AGE $<$ 43.45 and FSBP $<$ 143.50 and ISS $\leq$ 33 and EDRTS $<$ 0.87 and THORAX $\geq$ 2 Then Rehab	17/19(89.5%)	CART
EDRTS $<$ 5.36 and HEAD $\leq$ 3 and 33 $<$ FSBP $\leq$ 143 and ISS $\geq$ 33.50 Then Rehab	69/79(87.3%)	CART
HEAD $\geq$ 4 and FSBP $<$ 171 and EDRTS $<$ 2.25 Then Rehab	125/135(92.6%)	CART
2.25 $\leq$ EDRTS $<$ 5.36 and HEAD $\geq$ 4 and FSBP $<$ 171 and AGE $\geq$ 10.90 Then Rehab	45/52(86.5%)	CART
EDRTS $\geq$ 5.36 and AGE $<$ 48.15 and THORAX $\geq$ 1 and ISS $\leq$ 21 Then Home	23/27(85.2%)	CART
EDRTS $\geq$ 5.36 and AGE $\geq$ 48.15 and EDGCSTOTAL $\geq$ 9 and ISS $\geq$ 25 Then Rehab	61/65(93.8%)	CART
EDRTS $<$ 5.02 and HEAD $\leq$ 3 and 11.65 $\leq$ AGE $<$ 24.40 and EDGCSTOTAL $\leq$ 8 and FSBP $\geq$ 108 and THORAX $\leq$ 4 Then Rehab	24/28(85.7%)	CART
EDRTS $<$ 5.02 and HEAD $\leq$ 3 and 26.05 $\leq$ AGE $<$ 37.30 and EDGCSTOTAL $\leq$ 8 and FSBP $\geq$ 108 Then Rehab	22/24(91.7%)	CART
EDRTS $<$ 5.02 and HEAD $\leq$ 3 and AGE $\geq$ 43.30 Then Rehab	50/55(90.9%)	CART
EDRTS $<$ 5.02 and HEAD $\geq$ 4 Then Rehab	179/201(89.1%)	CART
EDRTS $\geq$ 5.02 and AGE $<$ 48.15 and THORAX $\geq$ 1 and ISS $\leq$ 21 Then Home	23/27(85.2%)	CART
EDRTS $\geq$ 5.02 and 24.25 $\leq$ AGE $<$ 48.15 and THORAX $\geq$ 1 and 22 $\leq$ ISS $<$ 28 and FSBP $<$ 146 Then Home	11/12(91.7%)	CART
EDRTS $\geq$ 5.02 and AGE $<$ 48.15 and THORAX $\geq$ 1 and ISS $\geq$ 22 and 74 $\leq$ FSBP $\leq$ 93 Then Rehab	10/11(90.9%)	CART
EDRTS $\geq$ 5.02 and AGE $\geq$ 48.15 and EDGCSTOTAL $\geq$ 9 and ISS $\geq$ 24.50 Then Rehab	104/122(85.2%)	CART
EDRTS $<$ 2.69 and 113 $\leq$ FSBP $\leq$ 170 and HEAD $\leq$ 3 Then Rehab	49/57(86.0%)	CART
EDRTS $<$ 2.69 and FSBP $\leq$ 170 and HEAD $\geq$ 4 Then Rehab	126/137(92.0%)	CART
EDRTS $\geq$ 2.69 and AGE $<$ 45.10 and THORAX $\geq$ 1 and ISS $\leq$ 21 and EDGCSTOTAL $\geq$ 7 Then Home	24/28(85.7%)	CART
2.69 $\leq$ EDRTS $<$ 5.02 and 22.80 $\leq$ AGE $<$ 48.15 and THORAX $\geq$ 1 and ISS $\geq$ 22 and (ARDS='No') Then Rehab	25/28(89.3%)	CART
EDRTS $\geq$ 2.69 and 48.15 $\leq$ AGE $<$ 84.30 and ISS $\geq$ 25 Then Rehab	74/79(93.7%)	CART
EDRTS $<$ 5.02 and 39.25 $\leq$ AGE $<$ 51.35 and HEAD $\leq$ 3 and FSBP $\leq$ 116 and ISS $\leq$ 33 Then Rehab	12/13(92.3%)	CART

Extracted supporting rules for exact outcome prediction (75% - 85% accuracy).

Though these rules are not reliable enough for practical use, they can contain pattern information which may be of interest to physicians.

Rules	Test Acc.	Method
EDRTS $\geq$ 5.36 and EDGCSTOTAL $\geq$ 9 and ISS $\leq$ 24 and THORAX $\leq$ 3 and AGE $\geq$ 53.95 and FSBP $\geq$ 93 Then Rehab	49/62(79.0%)	CART
EDRTS $\geq$ 7.12 and AGE $<$ 47.55 and THORAX $\geq$ 1 and 28 $\leq$ ISS $<$ 35 and 94 $\leq$ FSBP $\leq$ 135 Then Rehab	16/20(80.0%)	CART
EDRTS $\geq$ 2.69 and AGE $<$ 22.80 and THORAX $\geq$ 1 and ISS $\geq$ 22 and 123 $\leq$ FSBP $\leq$ 139 Then Rehab	11/13(84.6%)	CART
EDRTS $\geq$ 7.70 and 22.80 $\leq$ AGE $<$ 45.90 and THORAX $\geq$ 1 and ISS $\geq$ 28 and FSBP $\geq$ 76 Then Rehab	31/39(79.5%)	CART
5.02 $\leq$ EDRTS $<$ 7.12 and AGE $<$ 45.90 and THORAX $\geq$ 1 and 22 $\leq$ ISS $\leq$ 39 Then Rehab	9/12(75.0%)	CART
EDRTS $\geq$ 7.12 and AGE $<$ 48.15 and ISS $\geq$ 25 and HEAD $\leq$ 4 and THORAX $\geq$ 1 and 69 $\leq$ FSBP $<$ 98 Then Rehab	15/19(78.9%)	CART
EDRTS $\geq$ 2.69 and AGE $<$ 47.80 and ISS $\leq$ 24 and HEAD $\leq$ 2 and Then Home	43/56(76.8%)	CART
2.69 $\leq$ EDRTS $<$ 5.02 and 26.75 $\leq$ AGE $<$ 47.80 and ISS $\geq$ 25 and HEAD $\geq$ 1 Then Rehab	28/34(82.4%)	CART
EDRTS $\geq$ 2.69 and ISS $\leq$ 24 and THORAX $\leq$ 3 and HEAD $\geq$ 3 Then Rehab	151/182(83%)	CART
EDRTS $\geq$ 4.75 and AGE $<$ 48.15 and FSBP $\geq$ 94 and THORAX $\geq$ 1 and ISS $\leq$ 21 Then Home	21/25(84.0%)	CART
EDRTS $\geq$ 2.69 and AGE $\geq$ 48.15 and ISS $\leq$ 25 and THORAX $\leq$ 3 and FSBP $\geq$ 80 Then Rehab	59/74(79.7%)	CART
EDGCSTOTAL $\geq$ 7 and 26.75 $\leq$ AGE $<$ 43.00 and ISS $\geq$ 25 and FSBP $\geq$ 138 and HEAD $\geq$ 3 Then Rehab	12/16(75.0%)	CART
EDGCSTOTAL $\geq$ 6 and AGE $\geq$ 50.60 and ISS $\leq$ 25 and THORAX $\leq$ 3 and HEAD $\leq$ 4 and FSBP $\geq$ 74 Then Rehab	61/79(77.2%)	CART
(ID='Yes') and AGE $>$ 44 and (ARDS='Yes') Then Rehab	30/39(76.9%)	C4.5
THORAX $\leq$ 3 and ISS $>$ 18 Then Rehab	342/431(79.4%)	C4.5
(Cg='No') and 18.4 $<$ AGE $\leq$ 59.7 and ISS $>$ 30 Then Rehab	162/199(81.4%)	C4.5

## Appendix C Reliable rules for ICU days prediction

Reliable rules for ICU days are defined as those with accuracy greater than 85%. ED-BP is Emergency Department Blood Pressure; ED-RESP is Emergency Department Respiratory Rate; ED-PULSE is Emergency Department Pulse Rate; ED-GCS is Emergency Department Glasgow Coma Score.

Rules	Test Acc.	Method
(AIRWAY='Need') and $115 \leq \text{ED-BP} < 156$ and $\text{AGE} \geq 47.05$ and Then ICU stay days $\geq 3$	14/15(93.3%)	CART
(AIRWAY='Need') and $115 \leq \text{ED-BP} < 156$ and $\text{ED-RESP} < 18$ and $4.35 \leq \text{AGE} < 14.5$ Then ICU stay days $\geq 3$	12/12(100%)	CART
(AIRWAY='No Need') and $\text{ED-RESP} \geq 21$ and $45 \leq \text{AGE} < 55.85$ Then ICU stay days $\leq 2$	10/11(90.1%)	CART
(AIRWAY='Need') and $\text{ED-BP} < 91$ Then ICU stay days $\geq 3$	14/14(100%)	CART
(AIRWAY='Need') and $93.5 \leq \text{ED-BP} < 156.5$ and $\text{ED-PULSE} \geq 60.5$ and $\text{AGE} \geq 54.2$ Then ICU stay days $\geq 3$	10/10(100%)	CART
(AIRWAY='Need') and $94 \leq \text{ED-BP} < 156$ and $\text{ED-PULSE} \geq 61$ and $\text{ED-RESP} < 19$ and $18.45 \leq \text{AGE} < 44.5$ Then ICU stay days $\geq 3$	60/76(86.6%)	CART
(AIRWAY='No Need') and $\text{AGE} < 52.9$ and $\text{ED-BP} \geq 107$ and $\text{ED-GCS} \geq 11$ Then ICU stay days $\leq 2$	175/192(91.1%)	CART
(AIRWAY='Need') and $\text{ED-BP} < 150.5$ and $\text{ED-RESP} < 19$ and $\text{AGE} \geq 4.9$ and $\text{ED-PULSE} \geq 138$ Then ICU stay days $\geq 3$	18/20(90%)	CART
(AIRWAY='Need') and $\text{ED-RESP} < 19$ and $\text{ED-PULSE} < 138$ and $\text{ED-BP} < 115$ and $10.9 \leq \text{AGE} < 47.3$ Then ICU stay days $\geq 3$	31/33 (93.9%)	CART
(AIRWAY='No Need') and $\text{AGE} < 37.1$ and $\text{ED-GCS} \geq 11$ and $\text{ED-BP} \geq 125$ Then ICU stay days $\leq 2$	89/90(98.9%)	CART
(AIRWAY='No Need') and $\text{AGE} < 37.1$ and $\text{ED-GCS} \geq 11$ and $\text{ED-BP} < 119$ Then ICU stay days $\leq 2$	39/44(88.6%)	CART
(AIRWAY='No Need') and $\text{AGE} < 37.1$ and $\text{ED-GCS} \geq 13$ and $119 \leq \text{ED-BP} < 125$ and $\text{ED-PULSE} \geq 90$ Then ICU stay days $\leq 2$	21/22(95.5%)	CART
(AIRWAY='Need') and $146 \leq \text{ED-BP} < 156$ and $\text{AGE} < 22.5$ Then ICU stay days $\geq 3$	11/12(91.2%)	CART

Rules	Test Acc.	Method
(AIRWAY='No Need') and AGE<37.05 and ED-GCS $\geq$ 9 Then ICU stay days $\leq$ 2	157/172(91.3%)	CART
(AIRWAY='No Need') and 37.05 $\leq$ AGE<46.9and ED-RESP<21 and ED-PULSE<121 Then ICU stay days $\leq$ 2	23/25(92%)	CART
(AIRWAY='No Need') and ED-RESP<21 and ED-PULSE<121 and AGE $\geq$ 49.7 and ED-BP $\geq$ 141 Then ICU stay days $\leq$ 2	12/13(92.3%)	CART
(AIRWAY='No Need') and AGE<37.05 and ED-GCS<10 and 114 $\leq$ ED-BP<142 Then ICU stay days $\leq$ 2	11/12(91.7%)	CART
(AIRWAY='Need') and ED-BP<91.5 Then ICU stay days $\geq$ 3	14/14(100%)	CART
(AIRWAY='Need') and 91 $\leq$ ED-BP<156 and 95.5 $\leq$ ED-PULSE<102.5 Then ICU stay days $\geq$ 3	15/17(88.2%)	CART
(AIRWAY='No Need') and AGE<52.9 and ED-BP $\geq$ 99 and ED-GCS $\geq$ 13 Then ICU stay days $\leq$ 2	177/196(90.3%)	CART

## Appendix D Reliable rules for exact outcome prediction for pelvic injury

Reliable rules for exact outcome for pelvic injury are defined as those with accuracy greater than 85%.

Rules	Accuracy	Method
AGE<46 and EDGCS≤13 and ISS≥4 and (PRHFLU=IVF Unk Amount OR PRHFLU≥ 2000) Then Rehab	11/12(91.7%)	CART
EDGCS≥13 and (CHESTCT='Not Performed') and ISS≤9 and AGE≤28 Then Home	31/31(100%)	CART
36≤AGE≤46 and EDGCS≥14 and (CHESTCT='Not Performed') and ISS≤9 Then Home	16/18(88.9%)	CART
21≤AGE≤46 and EDGCS≥14 and 30≤ISS≤39 Then rehab	10/11(90.9%)	CART
52≤AGE≤68 and EDGCS≥15 and (PRHFLU='IVF Unk Amount' or PRHFLU='Not Performed' or PRHFLU='500-2000' or PRHFLU>2000 or PRHFLU='IVF Attempted') and ISS≥19 Then Rehab	14/15(93.3%)	CART
EDGCS≤13 and ISS≤13 and AGE≤6 Then Home	11/12(91.7%)	CART
EDGCS≤13 and ISS≥14 and AGE≤18 Then Rehab	9/10(90%)	CART
18≤AGE≤47 and EDGCS≤13 and ISS≥14 and (PRHFLU='IVF Unk Amount' or PRHFLU>2000) Then Rehab	11/12(91.7%)	CART
48≤AGE≤66 and ISS≥10 and EDGCS≤11 Then Rehab	18/19(94.7%)	CART
AGE<47 and EDGCS≥14 and ISS≤9 Then Home	75/88(85.2%)	CART
AGE≤48 and EDGCS≤13 and ISS≥20 and (PRHFLU='IVF Unk Amount' or PRHFLU> 2000) Then Rehab	11/12(91.7%)	CART
49≤AGE≤0 and ISS≥12 Then Rehab	14/16(87.5%)	CART
10≤ISS≤18 and 48≤AGE≤68 and (PRHFLU='IVF Unk Amount' or PRHFLU='500-2000' or PRHFLU>2000 or PRHFLU='IVF Attempted') Then Rehab	12/14(85.7%)	CART
AGE≤46 and EDGCS≥14 and ISS≤9 Then Home	75/87(86.2%)	CART
10≤ISS≤24 and 47≤AGE≤52 and EDGCS≥9 and (PRHFLU='performed') Then Rehab	13/14(92.9%)	CART
AGE<47.5 and EDGCS≤13 and (PRHFLU>2000) Then Rehab	10/11(90.9%)	CART
EDGCS≤13 and (PRHFLU='is not > 2000' and AGE<36.50 and ISS<20 Then Home	17/20(85%)	CART
EDGCS≥14 and ISS≤9 and AGE≤28 Then Home	35/35(100%)	CART

Rules	Accuracy	Method
EDGCS $\geq$ 14 and ISS $\leq$ 9 and 29 $\leq$ AGE $\leq$ 35 and (PRHFLU='500-2000' OR PRHFLU $>$ 2000 OR PRHFLU='Not Performed') Then Home	13/15(86.7%)	CART
24 $\leq$ AGE $\leq$ 47 and EDGCS $\geq$ 14 and (CHESTCT='Not Performed') and 26 $\leq$ ISS $\leq$ 29 Then Home	9/10(90%)	CART
29 $\leq$ AGE $\leq$ 47 and EDGCS $\geq$ 14 and ISS $<$ 9.50 and (PRHFLU='is not < 500' and PRHFLU=' is not > 2000') Then Home	34/39(87.2%)	CART
ISS $>$ 10 and 48 $\leq$ AGE $\leq$ 68 and EDGCS $\leq$ 11 Then Rehab	18/20(90%)	C4.5
ISS $\geq$ 10 and EDGCS $\geq$ 12 and 48 $\leq$ AGE $\leq$ 50 Then Rehab	17/20(85%)	C4.5
AGE $<$ 48 and EDGCS $\geq$ 14 and ISS $<$ 7 Then Home	28/31(90.3%)	C4.5
(PRHFLU = Not Performed) and 3 $<$ EDGCS $\leq$ 14 and ISS $>$ 24 Then Rehab	18/19(94.7%)	C4.5
(CHESTCT = 'Not Performed') and AGE $\leq$ 47 and ISS $\leq$ 10 Then Home	74/85(87%)	C4.5
AGE $>$ 43 and EDGCS $\leq$ 11 Then Rehab	51/54(94.4%)	C4.5

## Appendix E Detail non-medical dataset description

A detail description of two non-medical dataset, Census-income and university, is presented.

### Census-Income Dataset Description:

Type of Variable	Variables
Numerical (12)	age, wage per hour, capital gross, capital losses, dividends from stocks, number persons worked for employer, household, family member, weeks worked in year, family members under 18, taxable income amount, year of starting working
Categorical (21)	education, marital status, major occupation, gender,race, labor, member of a labor union, reason for unemployment, full or part time employment status, citizenship, own business or self employed, tax filer status, live in this house 1 year ago, veterans benefits status, industry, enrolled in education institute, country of birth father, country of birth mother, country of birth self, region of previous residence, state of previous residence
Class	Yes/No

### University Dataset Description:

Type of Variable	Variables
Numerical (10)	number of students, ratio male to female, ratio student to faculty, verbal SAT score, math SAT score, expenses, financial aid, admittance, enrolled, quality of life,
Categorical (4)	Control (private, state,city), academic emphasis, academics
Class	Yes/No



## VITA

Soo-Yeon Ji received the her B.S. and M.S. degrees in Mathematics from Kangwon National University at South Korea in 1992 and 1995. She received her M.S. degree in Software Information Systems from the University of North Carolina at Charlotte in 2004. Her research interests include biomedical signal and image processing, computer-assisted decision making, machine learning, and statistical data analysis.

### **List of Relevant Publications:**

S.Y. Ji, R. Smith, T. Huynh and K. Najarian, "A Comparative Analysis of Multi-level Computer-Assisted Decision Making Systems for Traumatic Brain Injuries," To appear: *BMC Medical Informatics and Decision Making*, 2009.

S.Y. Ji, K. Ward, K. Ryan, C. Rickard, V. Convertino, and K. Najarian, "Heart Rate Variability Analysis using Wavelet Transform to Predict Severity of Hemorrhagic Shock," *American Heart Association Resuscitation Science Symposium*, Nov 2008.

S.Y. Ji and K. Najarian, "A Modified Maximum Correlation Modeling Method for fMRI Brain," *International Workshop on Biomedical and Health Informatics (BHI)*, NOV 3-5, 2008.

W. Chen, R. Smith, S.Y. Ji, and K. Najarian, "Automated Segmentation of Lateral Ventricles in Brain CT images," *International Workshop on Biomedical and Health*

*Informatics (BHI)*, NOV 3-5, 2008.

S. Vasilache, R. Smith, S.Y. Ji, K. Najarian and T. Huynh, "Outcome Prediction in Traumatic Pelvic Injuries Using Maximum Similarity and Quality Measures," *Proceedings of IEEE Information Reuse and Integration (IRI)*, July 2008.

S.Y. Ji, K. Najarian, R. Smith, and T. Huynh, "Computer Aided Traumatic Pelvic Injury Decision Making," *Proceedings of Southern Association for Information Systems (SAIS)* March 2008.

S.Y. Ji, T. Huynh, K. Najarian, "A Comparative Medical Informatics Approach to Traumatic Pelvic Injuries," *Proceedings of IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Nov. 2007.

S.Y. Ji, T. Huynh, K. Najarian, "An Intelligent Method for Computer-Aided Trauma Decision Making System," *Proceedings of ACMSE* March, pp.198-202, 2007.

Internal Patent Disclosure:

S.Y. Ji, K. Ward, and K. Najarian, "HRV Analysis Based on Wavelet," Filed April 2008.