2015

Using the Delta-Model for End-Stage Liver Disease to Improve the Decision-making Process for the Donor Liver System

Joanne Chin
Walden University

Walden University

College of Management and Technology

This is to certify that the doctoral dissertation by

Joanne Chin

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Branford McAllister, Committee Chairperson, Management Faculty
Dr. Thomas Spencer, Committee Member, Management Faculty
Dr. Robert Kilmer, University Reviewer, Management Faculty

Chief Academic Officer
Eric Riedel, Ph.D.

Walden University
2015

Abstract

Using the Delta-Model for End-Stage Liver Disease to Improve the Decision-Making Process for the Donor Liver System

by

Joanne Chin

MS, Hofstra University

BS, New York University

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy in

Applied Management and Decision Science

Walden University

August 2015

Abstract

The purpose of this experimental research was to determine whether using delta-MELD as a criterion for the liver transplant patient selection process could improve the U.S. liver allocation system. This research closed a gap in current literature on the utility of delta-MELD for liver transplant patient selection. The frameworks of systems theory, the analytic hierarchy process, and the Kalman filter contributed to the development of 2 simulation models of the liver allocation system: one that used delta-MELD and one that did not use delta-MELD. The research question examined whether using delta-MELD could improve the liver allocation system by reducing the number of patients dropping off the wait list and lowering the average MELD score. Statistical *t* tests of 2 independent scenarios (allocation with and without delta-MELD), each with 70 runs of 180 simulated days on the liver allocation wait list, did not indicate a significant improvement to the liver allocation system by using delta-MELD for liver allocation. However, observations made from the simulation experiment, such as the median patient wait time being 11 months and delta-MELD being more variable at the end-stage of liver diseases, provided insights into how to improve the model of the liver allocation process. In addition, observations made from the status 1 patient subgroup (patients in ICU with about 7 days to live), which were excluded from this research, suggested including status 1 patients and expanding the simulation timespan from 180 to 360 days to better capture the delta-MELD variability from patients at the end-stage of liver disease. This research provides empirical evidence on the applicability of the delta-MELD criterion for non-status 1 patients, and yields recommendations to include status 1 patients in an improved simulation of the donor liver system while using delta-MELD as criterion.

Using the Delta-Model for End-Stage Liver Disease to Improve the Decision-Making Process for the Donor Liver System

by

Joanne Chin

MS, Hofstra University

BS, New York University

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy in

Applied Management and Decision Science

Walden University

August 2015

Acknowledgement

The UNOS STAR database is the largest and most comprehensive organ transplant database for research in the United States. My dissertation reference of the UNOS STAR database is supported by the U.S. Department of Health and Human Services: Health Resources and Services Administration contract 234-2005-370011c.

In full gratitude, I would like to acknowledge and thank my dissertation committee for guiding me through my dissertation process to completion, Dr. McAllister, Dr. Spencer, and Dr. Kilmer. I have been fortunate to have Dr. McAllister as my dissertation chair, who since the beginning of my program, has encouraged, mentored, and assisted me through my KAM and dissertation writing. I thank Dr. Spencer for his KAM and dissertation guidance. I thank Dr. Kilmer for his dissertation feedback.

I am blessed to have the constant encouragement, inspiration, support, and patience from my husband Steve, our family and friends, that helped me focus on my research and reach my program goals.

Table of Contents

Chapter 1: Introduction 1

Definitions 3

Background 7

Problem Statement 9

Purpose 10

Framework 11

Research Questions 13

Nature of Study 15

Significance of Research 16

Implications for Social Change 17

Scope and Delimitations 18

Assumptions 19

Limitations 20

Acknowledgement 21

Conclusion 22

Chapter 2: Literature Review 24

The MELD Era 25

MELD Variable 25

MELD and MELD-based Models 31

Delta-MELD: Debate and Gap in Literature 36

Literature Review on Time Series Analysis Method 42

Kalman Filter 42
Kalman State Space, Prediction, and Estimation 46
Kalman Error Management 49
Trends in the Liver Allocation System 51
MELD Era Objectives 51
Expanded Criteria Donor and Donor Risk Index 55
Increased MELD Scores and Extended Intensive Care Unit Stays 60
Decision-Making of Multiple Objectives: Urgency, Utility, and Survival 65
Literature on Multiple Criteria and Objective Decision-Making 71
Analytic Hierarchy Process / Analytic Network Process 71
AHP Background and Applications 73
AHP Decision-Making with delta-MELD for Simulation 77
Conclusion 79
Chapter 3: Research Method 81
The Research Design 82
Research Questions and Hypotheses 83
Research Process and Steps 84
Simulation Overview 85
Variables and Parameters in this Research 85
Data Collection through Secondary Data 92
Data Organization for the Simulation 93
Internal and External Validity of the Simulation Components 98

The Simulation Model 103
Waitlist Entry 104
Donor Liver Arrival 104
Disease Progression 107
Waitlist Patient Management 108
Simulation Inputs and Processing 108
Input Data 108
Input Data for Simulation without delta-MELD 109
Input Data for Simulation with delta-MELD 110
Simulation Outputs 114
Experiment and Sample Size 114
Hypothesis Testing 116
Pilot Testing and Scenario Runs 120
Summary 121
Chapter 4: Simulation Results and Analysis 124
Research Questions and Hypotheses 124
Pilot Testing and Verification of the Simulation Model 126
Waitlist Entry and Waitlist Patient Management Processes 128
Disease Progression Process 131
Donor Liver Arrival Process 134
Weekly Reports 137
Statistics from One Simulation Run 138

Experimental Outcome 139
Results 143
Research Question One 144
Research Question Two 144
Summary 144
Chapter 5: Research Conclusion 146
A Summary of the Research 146
Explanation of Simulation Results 148
Recommendations for Future Studies 151
Conclusion 157
References 159
Appendix A: Kalman Filter and Error Ellipse 168
Appendix B: Analytic Hierarchy Process (AHP) 174
Appendix C: Simulation Programming Notes 182
Appendix D: Statistics Notes 199

List of Tables

Table 1. Research Process and Steps 84

Table 2. Donor Risk Factors 88

Table 3. Recipient Risk Factors 89

Table 4. Hazard Ratios based on MELD 108

Table 5. AHP Weights and Ranking without delta-MELD 110

Table 6. AHP Weights and Ranking with delta-MELD 111

Table 7. Number of Region 9 Liver Transplants from Deceased Donors 113

Table 8. Simulation Variables and Parameters 127

Table 9. Waitlist Entry and Waitlist Patient Management Processes Verification..... 129

Table 10. Disease Progression Process Steps for Verification 131

Table 11. Donor Liver Arrival Process Steps for Verification 134

Table 12. $MELD_{mean}$ without delta-MELD..... 141

Table 13. $MELD_{mean}$ with delta-MELD..... 141

Table B1. Paired Comparison: Value-Description 176

Table B2. AHP Table of Weights of the Criteria..... 176

Table B3. AHP Table of Weights of Alternatives according to MELD 178

Table B4. AHP Table of Weights of Alternatives according to Blood Type 179

Table B5. AHP Table of Weights of Alternatives according to Body Structure..... 180

Table B6. AHP Table Ranking of Alternatives 181

Table D1. Chi-square Goodness of Fit test for Normality..... 203

List of Figures

Figure 1. Hierarchy of MELD levels for liver allocation 98

Figure 2. Analytic hierarchy process structure of objectives and criteria.....102

Figure 3. Donor liver processing based on liver quality.....106

Figure 4. Overview of simulation data processing 112

Figure 5. Waitlist Entry and Waitlist Patient Management processes.....130

Figure 6. Disease Progression process-1 132

Figure 7. Disease Progression process-2 133

Figure 8. Donor Liver Arrival process-1 135

Figure 9. Donor Liver Arrival process-2 136

Figure 10. Weekly reports: Disease progression 137

Figure 11. Statistics from one simulation run 138

Figure 12. Experimental outcome.....140

Figure C1. Initialize data panel 182

Figure C2. Setup scenario(s) panel 185

Figure C3. Run scenario(s) panel.....187

Figure C4. View disease progression panel.....191

Figure C5. View liver arrival panel 193

Figure C6. View *t* tests of two independent populations' parameters panel 194

Figure C7. Simulation progress panel 197

Figure D1. Goodness of fit-test for normality 202

Chapter 1: Introduction

Due to dissatisfaction with the existing liver allocation system and to improved predictive models, the U.S. liver allocation system was revised in 2000 in an effort to better balance the urgency and utility tradeoffs (Freeman, 2009). Bernardi, Gitto, and Biselli (2011) found that the model for end-stage liver disease (MELD) scoring system was originally developed by the Mayo Clinic in Rochester, Minnesota to predict the risk of death in patients undergoing liver transplant and was validated as a reproducible and reasonably accurate predictor of mortality in patients with chronic liver disease. From an urgency point of view, the MELD score offered multiple advantages for prioritizing waiting liver transplant candidates. Freeman (2009) further explained that outside of the liver allocation role which the MELD score supports, many liver transplant researchers have reported that changes in the MELD score over time (delta-MELD) have been associated with increased waiting list mortality. However, the most significant changes tend to occur very late in the course of the disease, which could limit the prognostic usefulness of the delta-MELD measurement. On the other hand, Foxton et al. (2006) and Young et al. (2006) suggested that the current system can further be refined by the use of delta-MELD, the change in MELD score over time. However, there have been limitations regarding how delta-MELD should be interpreted and computed as a predictor of disease progression and waiting list death.

In this research, I investigated the utility of the delta-MELD parameter for refining the MELD-based allocation system. Currently, the MELD scores can be used consistently across all types of patients with chronic liver diseases regardless of the

country or region, and regardless of the biases clinical observers may have in assessing liver disease severity. Hence, any improvement to the existing MELD-based scoring system would tend to be MELD-based (Bernardi et al., 2011).

Bambha et al. (2004) described the MELD score as a formula of three parameters to indicate the level of liver disease severity. The MELD score is calculated using serum creatinine, serum total bilirubin, and the international normalized ratio (INR) according to the following formula as is currently used by the United Network for Organ Sharing (UNOS) organization.

$$MELD = 9.57 * log_e\ creantinine\ (mg/dL) + 3.78 * log_e\ bilirubin\ (mg/dL) + 11.20 * log_e\ INR\ (mg/dL)\ + 6.43 \quad (1)$$

Bambha et al. (2004) also provided the definition of delta-MELD as the difference between the current-MELD score and the lowest of all serial-MELD scores in the preceding 30-day window. Thus, delta-MELD is defined as the maximum change in MELD score over a 30-day period. Current-MELD is defined as the most recent MELD score available for each patient. The timing of the current-MELD score depends upon the time lag used in the model. Bambha et al. concluded that the predictive value of delta-MELD is limited, and that further studies based on prospectively collected laboratory data in which the frequency of MELD measurements are controlled could address this issue more definitively. However, other researchers have suggested that delta-MELD can be beneficial towards the improvement of the MELD system (for

example Foxton et al., 2006; Young et al., 2006). This is the conflict and gap that was investigated in this research.

In Chapter 1, I describe the background, problem statement, purpose, theoretical framework, research questions, and nature of study of this research. The limitations, assumptions, and social implication of this research are also provided. Chapter 1 also contains a content description of Chapter 2 and Chapter 3.

In Chapter 2, I provide the literature review with explanations of the investigation of the delta-MELD parameter, background on the objectives of the liver allocation system, data collection methods, and decision-making methods that would address the research questions. The literature review contains an evaluation of multiple objectives of the liver allocation system, the criteria for each of the objectives, and the statistical and decision-making methodologies proposed for the simulation model.

In Chapter 3, I describe the implementation of the simulation model, the simulation experiment, the data to be collected, the computations, and the meaning of the output data. Furthermore, in Chapter 3, I provide details of the analytic hierarchy process (AHP) parameters and the AHP decision process, and the functions behind the four processes of the research simulation. I also provide description on the reliability and validity of the research data collection, data processing, and data analysis of the simulation scenarios with and without delta-MELD as a decision-making criterion.

Definitions

ABO: A, B, AB, and O blood types and their subtypes when used for allocation (Organ Procurement and Transplantation Network, 2014).

Analytic hierarchy process (AHP): A structured technique for organizing and analyzing multiple criteria for decision-making based on mathematics and psychology (Saaty, 1996).

Acute liver failure (ALF): Acute liver failure is a medical condition that includes the rapid loss of liver function, in a matter of days or weeks, usually in a person with no pre-existing liver disease (O'Grady, 2012).

Cold ischemic time (CIT): Cold ischemic time is the amount of time, usually about 12–18 hours, after a donor liver is harvest for transplantation. Reducing the cold ischemia time would improve the quality of the transplanted liver and CIT can be lowered by lowering the logistical and transportation time (Burr & Shah, 2010).

Current-MELD: The most recent MELD score available for each patient (Bambha et al., 2004).

Delta-MELD: The calculated difference between the *current-MELD* score from the lowest of all *serial-MELD* scores in the preceding 30-day window. Thus, delta-MELD is defined as the maximum change in MELD over the 30-day period (Bambha et al., 2004).

Donor risk index (DRI): Donor risk index is a measurement of the donor liver quality based on nine factors (Foxton et al., 2010).

Expanded criteria donor (ECD): Organ Procurement Organizations consider certain conditions of a donor to be expanded criteria donor (ECD) for a liver transplant and the patient has to give informed consent to accept the liver. These conditions may include a donor's age of 70 years or above, a donor who is age 60 years with significant

medical history, or a donor with a history of hepatitis B or hepatitis C (Rodrique, Hanto, & Curry, 2011).

Graphical user interface (*GUI)*: Graphical user interface is the interface that provides text-based or graphical information to the user via a computer interface.

Hepatocellular carcinoma (*HCC)*: Hepatocellular carcinoma is a liver cancer and it is also known as malignant hepatoma.

Model for end-stage liver disease (*MELD)*: Model for end-stage liver disease score is used to quantify the severity of end-stage liver disease for liver transplant planning (Bernardi, 2011).

Organ Procurement Organization (*OPO)*: Organ Procurement Organization is an organization accepted as a Member and is authorized by the Centers for Medicare and Medicaid Services (CMS) to procure organs for transplantation. For each OPO, CMS defines a geographic procurement territory within which the OPO concentrates its procurement efforts. No OPO is limited to or granted exclusive procurement right to procure organs in its territory (Organ Procurement and Transplantation Network, 2014).

Organ Procurement and Transplantation Network (OPTN): Organ Procurement and Transplantation Network is an organization governed by the U.S. Department of Health and Human Services and is formed by multiple committees to develop organ transplantation policies (Organ Procurement and Transplantation Network, 2014).

Pediatric end-stage liver disease (*PELD)*: PELD score is the pediatric version of the MELD score for the purpose of liver transplant planning (Organ Procurement and Transplantation Network, 2014).

Scientific Registry of Transplant Recipients (SRTR): Scientific Registry of Transplant Recipients is a national database of transplant statistics. (Scientific Registry for Transplant Recipients, 2012).

Serial-MELD: The MELD scores collected over the 30 day window serially (Bambha et al., 2004).

Standard criteria donor (SCD): Standard criteria donor liver comes from a decease donor who is brain dead, but still has a beating heart, albeit may be supported by a respirator. Unless the donor liver has been evaluated to have certain risk factors, it is a SCD liver for liver transplant (Rodrique et al., 2011).

Survival outcomes following liver transplantation (SOFT): SOFT score is based on the MELD score and other risk factors for assessment of overall survival outcomes in order to consider waitlist mortality against posttranplant mortality (Rana, et al., 2008).

Transplant center: A hospital that is a member in which transplants are performed. It is the responsibility of the transplant surgeon of the transplant center receiving the organ to offer the surgeon's candidate to ensure medical suitability of donor organs for transplantation into the potential recipient according to the candidate's blood type and subtype (Organ Procurement and Transplantation Network, 2014).

Transplant program: A transplant center, or hospital, may have one or more transplant programs. Each program oversees transplantation of one or more organ types (Organ Procurement and Transplantation Network, 2014).

United Network for Organ Sharing (UNOS): United Network for Organ Sharing is a private, non-profit organization that manages the nation's organ transplant system, under contract with the federal government (United Network for Organ Sharing, 2014).

Waiting list: This list is a computerized list of candidates who are waiting to be matched with specific donor organs in hopes of receiving transplants. Waiting list candidates are registered on the Waiting list by member transplant centers. The candidate's transplant program would be responsible for ensuring the accuracy of candidate ABO data on the waiting list (Organ Procurement and Transplantation Network, 2014).

Background

Malinchoc et al. (2000) from the Mayo Clinic in Rochester, Minnesota, devised a mathematical model to prioritize patients for liver transplantation based on medical urgency, named the Mayo end-stage liver disease score. This model was proved to accurately predict the probability of death within three months after the procedure. Subsequently, the model name was changed from Mayo end-stage liver disease to model for end-stage liver disease (MELD) and it was successfully validated in patients with different liver disease severity. Bernardi et al. (2011) explained that because of this, the time has come for a sickest-first policy to be reliably fulfilled, and the MELD score became the means to allocate donor livers for medical transplant in the United States from February 2002. Bernardi et al. further explained that MELD has several features of an ideal prognostic model to predict the probability of survival. It incorporates objective variables readily determined in all laboratories and each of these variables is weighted

according to the influence on prognosis.

Bernardi et al. (2011) further elaborated that the MELD score is not a time-dependent model, because it is computed by a single measurement of laboratory parameters. In an attempt to weigh the time-related changes, the delta-MELD which is defined as the difference between the MELD score calculated at two time points has been proposed. Bernardi et al. noted that studies showed this new score was able to predict the mortality risk of patients more accurately than standard MELD score alone. However, Bernardi et al. also noted that there have also been other studies debating its usefulness in predicting survival on the waiting list.

Having a time-dependent variable as a criterion could be beneficial to the liver allocation system because a time-dependent parameter such as delta-MELD could help align the MELD scores more accurately when assessing patients' status upon an arrival of donor liver. This is because pretransplant patients would not have the same time-stamps of their latest MELD scores, and their MELD scores could vary due to liver deterioration when an actual donor liver is made available.

Young et al. (2006) acknowledged that the allocation of donor livers through the MELD score, implemented in the United States on February 2002 by United Network for Organ Sharing (UNOS), has resulted in a fall in waiting list deaths in the United States. In addition, liver transplant centers in the United States are able to transplant a sicker population of patients with no deterioration in results. Foxton et al. (2006) suggested that the current system can further be refined by the use of delta-MELD, the change in MELD score over time. However, there have been limitations regarding how delta-MELD

should be interpreted as a predictor of disease progression and waiting list death.

Young et al. (2006) also concluded that there is value in using delta-MELD score for decision-making regarding the allocation of donor livers. However, there has been a problem calculating delta-MELD due to the various collections of MELD data. Hence, the data collected and studied by Young et al. has been considered biased. Young et al. concluded that a study designed to minimize data collection bias was needed to fully clarify the role of delta-MELD in liver allocation. Young et al. also concluded that using MELD and delta-MELD in allocation decision-making could possibly improve overall outcomes by allocating livers more efficiently to reduce waiting list deaths.

Rahman and Hodgson (2001) divided acute hepatic failure (AHF) into three categories, which are hyperacute, acute, and subacute. Hyperacute is when encephalopathy is developed within 7 days after the onset of jaundice. Acute is when encephalopathy is developed in 8 to 28 days after the onset of jaundice. Subacute is when encephalopathy is developed in 5 to 26 weeks after the onset of jaundice. These classifications of hepatic failure suggest that not all liver diseases deteriorate at the same rate or that the MELD score alone is an indicator of the most urgent patient in need of a liver transplant. In the current donor liver system, when multiple matching recipients have the same MELD scores, the patient who waited longest rather than the patient with the faster deteriorating disease, will get the transplant.

Problem Statement

While there has been elaborate research on the MELD-based topic, there is a gap in scholarly literature clarifying whether the delta-MELD parameter used as a

recipient criterion in addition to the MELD score, together as primary criteria, could refine the liver allocation system. Specifically, no research has been conducted in current literature to analyze how using the delta-MELD parameter in addition to the MELD score as a liver transplant selection criterion would affect the number of liver patients saved by liver transplants. Overcoming a data collection bias of delta-MELD, as described by Young et al. (2006), through statistical time series analysis techniques can depict an accurate account of disease progression and in predicting patient waiting list outcomes.

Purpose

The purpose of this research was to address the lack of scholarly understanding about the utility of the delta-MELD criterion by investigating whether using the delta-MELD criterion can improve and refine the liver transplant patient selection process of the U.S. liver donor allocation system. Cholangitas and Burroughs (2012) stated that an ideal donor liver allocation model should not only be able to allocate according to the highest probability of dying before liver transplant, but also be able to predict which patients have the lowest post-liver transplant mortality in order to improve utility (i.e. a survival benefit system). In this research, I investigated whether the utilization of the delta-MELD parameter could help reduce patients from dropping off of the waiting list due to being too sick to undergo liver transplant as well as reduce the average MELD score among pretransplant patients waiting for liver transplant.

O'Grady (2012) noted that patients undergoing liver transplantation with more advanced MELD scores are more likely to have acute liver failure and have longer stays in intensive care environments. Although this may not affect the outcome of post liver

transplant, O'Grady noted that it is still an important medical transplant practice to understand the different patterns of disease progression and to be able to assess prognosis based on recognized prognostic models. For this reason, this research refers to delta-MELD that is based on the Kalman estimation reflecting MELD progression.

There has been a problem in calculating delta-MELD due to the various collection methods of MELD data (Foxton et al., 2006; Young et al., 2006). Overcoming a data collection bias of delta-MELD through statistical time series analysis techniques could depict a real-life account of disease progression as well as predict patients' waiting list MELD outcomes.

Framework

The theoretical frameworks that shaped the simulation model and framed the research questions were based on multi-criteria decision-making techniques of the operations research discipline, and time series forecasting and estimation methods from mathematical statistics. The research framework also has the objective of the U.S. Federal law that mandates a sickest first system that would be employed for ranking candidates for liver transplantation based on medical urgency (Freeman et al., 2009). These methods, techniques, and objective were the conceptual and theoretical frameworks for determining whether the liver allocation system could be improved upon. They are mentioned here and discussed in greater detail in Chapter 2 and Chapter 3.

A multi-criteria decision-making tool for the selection of the most suited and sickest patient was useful because there would be multiple objectives and criteria to weigh in the selection consideration among the many patients waiting for a donor liver.

Winston (2004) described that multi-criteria decision-making could be a complex process because when multiple objectives are important to a decision-maker, it could be difficult to choose among the many alternatives. By using a multi-criteria decision-making tool, the AHP, the simulation would be an ideal method for reflecting the selection of the most suitable patient.

Time series analysis is an arm of mathematical statistics that provides tools useful for estimation and forecasting of time series values. Box, Jenkins, and Reinsel (2008) explained that an intrinsic feature of a time series is that, typically, adjacent observations can be dependent. The nature of this dependence among observations of a time series is of considerable practical interest. Time series analysis is concerned with techniques for the analysis of this dependence. This research applies time series analysis techniques to determine and predict patients' MELD and delta-MELD parameters.

Kalman filter is a technique for forecasting and estimating time series values and it has its strength in the observability and controllability of time series data (Brockwell & Davis, 2006). Asemoto (2010) described the Kalman filter as a statistical algorithm that enables certain computations to be carried out for a model cast in state space form. The Kalman filter is also known for its simplicity and straightforwardness of its algorithm. This research employed a time series prediction of liver disease progression in order to aid in the simulation model in computing the delta-MELD parameter for the selection of compatible transplant recipients.

In summary, the theoretical frameworks that shaped the simulation model and framed the research questions were the AHP technique and the Kalman filter. And the

conceptual framework was to meet the objective of the U.S. Federal law that mandates a sickest first system that would be employed for ranking candidates for liver transplantation based on urgency (Freeman et al., 2009). These methods contributed to the development of a simulation experiment to determine whether the current liver allocation system could be improved upon.

Research Questions

Gotthardt et al. (2009) argued that their study was not only an analysis of the number of deaths on the waiting list but also an analysis of the number of removals from the waiting list due to patients' poor condition. These combinations of numbers more accurately reflect the natural history of liver diseases. Gotthardt et al. also stated that while their data do not argue against the use of MELD scores to be taken to prioritize patients during the initial period on the waiting list, their study showed that for patients with longer times on waiting list, additional factors for assessment of patient prognosis could assist in the development of a new scoring system for allocation. Hence, the following were the research questions for this research.

1. Does a simulation model using the additional parameter of delta-MELD as a patient selection criterion reduce the number of pretransplant patients who dropped off of the waiting list?

The null and alternative hypotheses are as follows for the first research question:

H_o: There is no difference in the number of patients who dropped off of the waiting list between simulation models with and without delta-MELD along with the MELD score as primary criteria for patient selection in donor liver allocation.

H_a: There is a difference in the number of patients who dropped off of the waiting list between simulation models with and without delta-MELD along with the MELD score as primary criteria for patient selection in donor liver allocation.

Quante, Benckert, Thelen, and Jonas (2012) stated that the implementation of the MELD score system in Europe affected a change not only by the reduction in waitlist mortality among pretransplant patients, but also in the average MELD score that increased among pretransplant patients. This trend is also reflected in the Eurotransplant Annual Report 2010, which describes a 24% increase in the number of high-MELD recipients within the total population of liver-graft recipients in 2010 compared with 2009. Hence, the null and alternative hypotheses are as follows for the second research question.

2. Does a simulation model using the additional parameter of delta-MELD as a patient selection criterion lower the average MELD score among pretransplant patients?

H_o: There is no difference in the average MELD score among pretransplant patients between simulation models with and without delta-MELD along with the MELD score as primary criteria for patient selection in donor liver allocation.

H_a: There is a difference in the average MELD score among pretransplant patients between simulation models with and without delta-MELD along with the MELD score as primary criteria for patient selection in donor liver allocation.

Nature of Study

This quantitative simulation experiment investigated the utility of the delta-MELD parameter for decision-making in liver transplant patient selection. This research used a simulation model that was unlike any previous research on the study of the delta-MELD parameter because of its unique approach to compute part of the MELD scores based on the Kalman estimation. The Kalman filter was also used to generate and simulate the disease progression of patients that are on the waiting list for a donor liver. In addition, the technique of AHP was used to simulate the decision-making process of the existing and proposed patient selection criteria of the liver allocation according to OPO hierarchy of priority.

The experiment simulated two scenarios, one that would reflect the current system's utilization of the MELD score as a primary criterion, and another scenario that utilized both MELD and delta-MELD scores as primary criteria in determining medical urgency. The two scenarios, utilizing and not utilizing delta-MELD as criterion, counted the number of patients who dropped off of the waiting list and computed the average MELD over a 180-day period. This quantitative study generated the values of patients removed and average MELD from the two scenarios for statistical comparison.

The data used in this simulation model include secondary data from the United Network of Organ Sharing Organization's Standard Transplant Analysis and Research (UNOS STAR) database. The data consisted of both patient and donor liver data. Patient data included age group, gender, race, primary cause of disease, transplant history, blood type, MELD scores, date of MELD scores, time on wait list, and status. Donor liver data

included donor age, donor height, donation after cardiac death donors, split liver donors, race, donor's cause of death from cerebrovascular accident, regional sharing, local sharing, and cold ischemia time. Patient data supported the patient and disease progression simulation portion of the scenarios. Multiple and sequential MELD scores records taken over time determined the patients' delta-MELD scores.

The simulation model is a software program that is comprised of four components or processes. These four components included patient waitlist entry, donor liver arrival which performs scoring processing for liver recipient selection, patient disease progression, and waitlist patient management.

This quantitative study addressed the research questions by formulating a time series estimation of patients' illness based on known and estimated MELD, and delta-MELD scores. This research applied a multi-criteria decision-making process consisting of the proposed new criterion, and performed statistical *t* tests of two independent populations for comparison of the system utilizing delta-MELD scenario against the existing system's scenario.

The elements crucial to the simulation included MELD and delta-MELD patient data. I performed *t* tests of two independent populations (with and without delta-MELD criterion) to determine whether the delta-MELD parameter is useful for patient selection.

Significance of Research

In this research, I investigated the delta-MELD parameter as a transplant patient selection criterion. The delta-MELD parameter was helpful for its predictive attributes when assessing pretransplant patients' prognosis because it helped to align and predict

patients' health statuses upon the arrival of an available donor liver. Gotthardt et al. (2009) stated that an effort that could improve the MELD system would involve analyzing the change in MELD scores over time, bearing in mind that this dynamic variable would reflect the dynamic of the disease. Gotthardt et al. further explained that several attempts have been made where some studies concluded that the delta-MELD score had better prediction ability for mortality than the baseline MELD score, while other studies concluded that delta-MELD was not as predictive compared to the most updated MELD score. This suggested that an in-depth study still needs to be conducted but with different supporting methodologies. This research could potentially improve upon the current allocation system by including the use of the delta-MELD parameter as a criterion for patient selection in order to reduce the number of patients from dropping off of the waiting list and to reduce the average MELD score of waiting list patients. Strategies used in this study are a methodology for forecasting and estimating liver disease progression and the use of a multi-criteria decision-making process, while incorporating the multiple processes of the liver allocation system. This research provided further understanding on the usefulness of the delta-MELD parameter.

Implications for Social Change

The implication of positive social change is the potential of saving more lives through an improved decision-making system for allocating donor livers to transplant patients. Time series prediction technique could be applied to other areas of health care for better control and management of disease progression. This research helped to promote using knowledge of disease progression into the decision-making refinements

for the donor liver allocation system.

Quante et al. (2012) noted that, in December 2006, the MELD score system was implemented as the basis for new liver allocation system in many countries within the Eurotransplant area. Cholongitas and Burroughs (2012) also noted that the MELD system adopted by Eurotransplant helped to allocate organs in seven countries of central Europe: Austria, Belgium, Croatia, Germany, Luxemburg, the Netherlands, and Slovenia. Since then, there was a significant reduction in waiting-list mortality in Europe. This affected a social change in the reduction in waiting-list mortality among pretransplant patients. A potential refinement to the liver donor system by including delta-MELD to patient selection criteria could benefit the transplant community in the United States and in Europe in two ways. The donor allocation system could be more effective in achieving the sickest first policy, and thus, would provide a more fair system for the recipients.

Scope and Delimitations

This research was limited to the investigation of the delta-MELD. Although the donor risk index (DRI) was considered in the multi-criteria decision-making, and this research referenced the DRI, the DRI values were the same for both before and after (utilizing delta-MELD) simulation scenarios. The delta-MELD was the sole interest of this study. The prediction of MELD scores was limited only to supporting the simulation and was not meant to be used for prognosis. The propagated MELD scores used in the simulation were intended to project the MELD scores in future time and in accordance with secondary data.

In this research, I included data from the UNOS STAR database. I used a simulation model to simulate and fill in additional MELD scores that were estimated MELD scores in order to support a sequence of MELD scores so that the derivation of delta-MELD values was possible. This was done because secondary data may not have enough MELD values to formulate delta-MELD values throughout a 180-day timeline.

The 2010 American Association for the Study of Liver Diseases guidelines advised against the use of prognostic models in an individual patient (Siciliano et al. 2012). The simulation used de-identified patient and donor liver data. The simulation's processing and output data of MELD average and number of patients dropped from waitlist cannot be used to single out any individual patient. In addition, the research was limited to adult liver donations, adult transplant patients, non-HCC patients, and non-status 1 patients.

This research limited the use of the UNOS STAR database for the simulation to data from the recent five years of 2008-2012, and from one region, Region 9. The purpose for this was to limit the simulation's scope from the need to concern with cold ischemia time and the varying MELD averages of additional regions, by focusing solely on one region from the recent five years, 2008-2012. Region 9 is confined to the area of New York state and western Vermont.

Assumptions

Massie et al. (2011) explained that although MELD was adopted to estimate the short-term (90-day) risk of waitlist mortality, it is believed to underestimate such risk for certain patients with non-normative conditions. Some diseases have low risk of short-

term mortality, but require transplant before progression to the point of irreversible complications. For such cases, additional MELD points can be granted, and these patients receive priority based on the exception MELD rather than the calculated MELD. The Organ Procurement and Transplantation Network (OPTN) policies originally allowed exception points for certain recognized exceptional diagnoses such as hepatocellular carcinoma (HCC) and hepatopulmonary syndrome (HPS).

The two categories of liver diseases are cholestatic and noncholestatic. The assumption is that the exception points are already incorporated in the patient data for the noncholestatic disease of HCC. It is known that HCC patients receive exception MELD points that are not derived from the MELD formula. This research filtered out HCC patients for the simulation to allow the MELD scores to be comparable to other MELD scores without the concern of how the exception points were applied.

Limitations

Simulation limitations, such as using the simulation sampling intervals of 180-day instead of 360-day in duration, along with using a limited data sampling from the OPO Region 9, may have masked the true effect of delta-MELD, and prevented the *t* tests from producing a significant outcome. This may be because while OPO Region 9 has many liver transplant patients, it is also a region of many organ donors. This could be the predominant factor for keeping the waitlist MELD averages low, and possibly reducing the occurrence of waitlist patients from having sizable delta-MELDs in this limited data sample. The delta-MELD values usually have more variability near the end-stage of patients' liver disease (Freeman, 2009), and the simulation model may not have used a

duration that was long enough to allow the patients' disease progression to follow its course to completion in order to encounter the variability in delta-MELD values.

Finally, the simulation employed two main theoretical frameworks, the Kalman estimation for simulating disease progression, and the AHP for simulating decision-making. The AHP was used for its sophistication and straightforwardness in applying the same decision criteria to all patients. However, in actuality, the selection of patients is likely based on physicians' medical experience, expertise, and knowledge of patients' medical history (Bernardi et al., 2011), while weighing other factors in addition to the AHP criteria. Many times, transplant physicians are knowledgeable of their patients' medical history, allowing them to see subtle changes to their patients' conditions (Schiano, 2012), which this simulation model or any simulation model may not be able to replicate, given only the UNOS STAR database to work with.

Acknowledgement

By using the UNOS STAR database to support the simulation model, it is required by the UNOS organization to specify that this work was supported by Health Resources and Services Administration contract 234-2005-370011c. The content of this work was the responsibility of the authors alone and did not necessarily reflect the views or policy of the Department of Health and Human Services, nor did the mention of trade names, commercial products, or organizations imply endorsement by the United States Government (Organ Procurement and Transplantation Network, 2014).

Conclusion

The supporting methodologies of this research, AHP and Kalman filter, may be new to the health care field and may be new to the donor liver allocation scoring system. The research question was to determine whether the parameter of delta-MELD should be used, in addition to the MELD score together as primary criteria, for patient selection. As a patient selection criterion, delta-MELD could potentially improve the liver allocation system by reducing the number of pretransplant patients from dropping off the waiting list and by lowering the average MELD score among pretransplant patients.

The conceptual framework included the delta-MELD and MELD being primary criteria for patient selection decision-making process for selecting the most urgent patient in need of a liver transplant. The underlying theoretical frameworks for the simulation were applying the Kalman filter and the AHP technique into the scenarios of the existing and proposed liver allocation system.

By using the conceptual and theoretical frameworks, the decision-making process could potentially be refined and equitably judged among the patients waiting for a liver transplant. An additional strategy for the simulation was taking a systems perspective of the liver allocation system.

In Chapter 2, I present a literature review to explain why the investigation of the delta-MELD parameter is needed and how the chosen methodologies could bring further understanding regarding delta-MELD. In the literature review, I examine and conclude the need to bring forth a different set of methodologies which include statistical

prediction, criteria and decision-making development, and a systems perspective into the simulation to address the research problem.

Chapter 2: Literature Review

In this introduction, I outline and describe the structure of the literature review. In the first section, I describe the MELD of the donor liver allocation system, its problem and gap in research, and the methodology that was used in the simulation model to determine whether the MELD system could be refined. The similarities and differences of various MELD-based models are presented. Among the MELD-based models, the delta-MELD is discussed as a potential donor liver allocation criterion. In the literature review, I reveal an unresolved problem of delta-MELD being a viable criterion for donor liver allocation. Kalman filter is a statistical methodology that could provide MELD estimations in order to establish consistently measured delta-MELDs. The Kalman filter can provide MELD data based on existing MELD values in order to compute delta-MELD.

In the second section of the literature review, I evaluate the decision-making process of the donor allocation system that meets the objectives of urgency (sickest first), utility, and survivability. The DRI and the expanded criteria donor (ECD) are also evaluated. The survival outcomes following liver transplantation (SOFT) scoring system is reviewed which meets both urgency and utility objectives. The AHP model for multiple criteria and objectives was proposed for decision-making of the liver allocation system. In the literature review, I evaluate the usefulness and flexibility of AHP for analyzing the delta-MELD parameter as an additional criterion for donor liver allocation.

The literature search was based on reviewing liver transplant topics through the journals of *American Journal of Transplantation*, *Liver Transplantation*, *Journal of*

Hepatology, and *BMC Gastroenterology*. These peer-reviewed journals are either open access or are on-line journals accessible via the internet. Subsequently, the references used by these articles led to other peer-reviewed journals such as *Langenbecks Archive of Surgery*, *International Journal of Hepatology*, *Transplant International*, and *Hepatology International*. As the search developed, other medical, health, and management journal databases were referenced. Initially, the key words used to search these articles were *liver allocation*, *liver transplant*, *liver procurement*, *waitlist*, *MELD*, and *liver donation*.

In the literature review, I conclude by summarizing the major points of the literature review and describe the purpose of the research simulation. The simulation model includes the patients entering and exiting the waitlist, the progression of patients' illnesses, the arrival of available livers, and the selection of a compatible recipient reflected by theoretical frameworks of the Kalman filter and AHP.

The MELD Era

MELD Variable

In this first section of the literature review, I define the MELD, provide background and history of the MELD as it relates to the liver allocation system, identify and review the MELD-based models where the studies have been concluded, and identify and review the MELD-based models where the studies were inconclusive due to conflicting findings. In this literature review section, I also summarize how the MELD had positively impacted but also provided limitations to the liver allocation system, and review the effect of MELD in the transplant community of the United States, Europe, and Brazil. In the next section of the literature review, I reveal the trends of the MELD

system and discuss how the objective of attending to the *sickest first* policy (urgency) further led to the implementation of urgency, utility, and survivability objectives for decision-making on donor liver allocation. Bernardi et al. (2011); Teixeira de Freitas et al. (2010); Bahra and Neuhaus (2011); and Quante et al. (2012) acknowledged the advantages of the MELD system in the United States, Europe, and Brazil, particularly for shortening the waiting list and wait time for available donor livers.

In 2000, Malinchoc et al. (2000) from the Mayo Clinic in Rochester, Minnesota, devised a mathematical model to predict the probability of death within three months after the procedure. The model name was called MELD and it was successfully validated in patients with different liver disease severity, and from different geographical and temporal origin. Bernardi et al. (2011) explained that what made MELD an ideal prognostic model is its reliability in predicting the probability of survival.

Bernardi et al. (2011) further noted that MELD possessed predictive characteristics, the ability to provide a continuous ranking of disease severity, and the characteristic of being independent from the etiology of the liver disease. The impact of the MELD-based liver allocation policy had been impressive. New registrations on the waiting list suddenly dropped and the removal rate for death or disease progression also steadily declined.

Bernardi et al. (2011) explained that the adoption of the MELD to select and prioritize patients for liver transplantation represented a turning point in donor liver allocation. Prioritization of transplant recipients had switched from time accrued on the waiting list to the principle of "sickest first" (p. 1297). The simplicity of the MELD

score is in the incorporation of three laboratory parameters, serum creatinine, bilirubin, and INR for prothrombin time. Patients are also then stratified according to the disease severity in an objective and continuous ranking scale. Further advantages associated with implementing the MELD was a decrease in the median waiting time to transplant. Yet, the MELD limitations are related to the variability of the parameters and in the inability to predict mortality accurately in specific settings. Bernardi et al. discussed that these limitations of the MELD include not properly accounting for factors related to certain liver diseases where their progression are not weighted into the MELD scores such as with HCC.

While Bahra and Neuhaus (2011) also noted that MELD is limiting when scoring patients with HCC, Bahra and Neuhaus introduced the definitions of MELD-based allocations, labMELD, and a MELD-based concept called matchMELD which is a modification of the calculated MELD. The calculated MELD, or labMELD, was developed primarily for viral or ethyltoxic liver cirrhosis. In cases of HCC, the labMELD fails to indicate the urgency for liver transplantation. In this case, the MELD score will increase after implementing a defined criterion of standard exceptions.

Bahra and Neuhaus (2011) further noted that in the pre-MELD era, organ allocation was usually center-based, which meant that waiting list management was in the hands of the transplant center. The transplant center has the opportunity to decide which patient on the waiting list would receive the next available organ. Factors such as patient priority, clinical conditions of the recipient, donor organ quality, donor age, and other logistic aspects are included in the decision. The criteria of this allocation system

included patient urgency, blood group similarity, length of time a patient has been waiting for a transplant, size, other compatibility, and geographic location of the recipient hospital with regard to the donor hospital. In addition, a byproduct of the MELD allocation system was that good quality organs were usually preferentially allocated to patients with high morbidity (comparable to patients with MELD > 30). Bahra and Neuhaus questioned whether this allocation system was always fair. However, currently, there is a general consensus that organs with a high DRI should not be allocated to a high MELD scoring patient because of a significant increase in chance of posttransplant complication and death.

Bahra and Neuhaus (2011) argued that patients with HCC would usually achieve their matchMELD scores only through standard exceptions. Those patients usually have a labMELD score of less than 20, leading to a decreased rate of patients requiring intensive care after liver transplantation. Patients with matchMELD of 37 are not comparable to patients with labMELD of 37. Bahra and Neuhaus believed that this is a reflection of how the MELD system has some significant weaknesses compared to the diagnosis of a team of experienced physicians.

Teixeira de Freitas et al. (2010) likewise confirmed that end-stage liver disease is considered one of the major causes of death in the United States and its treatment is a major health dilemma. Teixeira de Freitas et al. explained that the MELD was introduced in Brazil for organ allocation in 2006. MELD score would help assess the severity of cirrhosis and predicted mortality. It would help provide priority to candidates waiting for liver transplants with more severe diseases. It also would help prioritize patients with

HCC. Before the MELD was introduced, organ allocation was based mainly on chronological waiting time.

Texiera de Freitas et al. (2010) explained that in one Brazilian center, after the introduction of MELD score as priority criterion for liver transplantation, there was an increase in the number transplants for patients with HCC. In the pre-MELD era 16.0% of receptors had HCC and in the MELD era 37.5%. There was no difference in the general MELD score of patients with HCC in the two eras. Excluding the cases of HCC, the transplants were performed in patients with more advanced cirrhosis. Furthermore, there were no increases in the indicators of worse prognosis or complications after the transplantations and there were no changes in the 3-month and 1-year posttransplant survival rate.

Teixeira de Freitas et al. (2010) also revealed that the MELD scores of patients without hepatocellular carcinoma was 18.2 ± 6 in the MELD era, which is similar to the study of Bahra and Neuhaus (2011), and this value was higher in the MELD era than the MELD score in the pre-MELD era which was 15.8 ± 4. Texiera de Freitas et al. explained that in Brazil, patients with cirrhosis and hepatocellular carcinoma were listed for liver transplantation with MELD exception points as according to the Milan criteria. This means that for one nodule of less than 5m in diameter or a maximum of 3 nodules is deemed safe, where each of the nodules is less than 3 cm in diameter. To avoid tumor growth beyond the Milan criteria while the patient is on the waiting list, extra points could be added to the MELD score. Therefore, some patients with HCC would be transplanted earlier in the evolution of cirrhosis. According to Brazilian legislation,

patients with HCC would initially receive 20 extra points to their MELD scores.

Quante et al. (2012) explained that the MELD was implemented on December 16th, 2006 as the basis for new allocation system in many countries within the Eurotransplant area. The MELD model provided a prediction of 3-month mortality without liver transplantation. Quante et al. similarly noted that there are many different possible underlying diseases that despite chronic liver decompensation, often have only a modest impact on laboratory results, and standard exceptions to the MELD system with adjustment of the score, have been defined. For example, patients suffering from HCC are given an adjustment in their MELD score because of the underlying malignancy and the consequent anticipated tumor growth during the waiting period. In addition to HCC, there are other risk factors that may not be reflected by laboratory results.

Quante et al. (2012) noted a few significant findings regarding waitlist mortality and MELD score at time of organ allocation and donor graft quality. In their center, there was a significant reduction in waiting-list mortality from 18% in the year before to 10% in the year after the MELD was introduced. Other single-center results within Europe have also confirmed a reduction in waitlist mortality since the introduction of MELD. Quante et al. also noted that after the MELD was implemented, there was a significant increase in the mean MELD score at the time of liver allocation, reflecting the intention to give priority to sicker patients on the waiting list. In their center, the mean MELD score increased from 16.3 points in the year before to 22.4 points in the year after MELD introduction. Since then, there has been a steady increase in mean MELD score within the Eurotransplant area, especially in Germany. In 2010, a mean MELD score of 34

points for standard liver allocation, without standard exceptions and without high-urgency status, was reported in Germany. In addition, Quante et al. detailed that this resulted in worse posttransplant outcomes in the group of high-MELD recipients, which in 2010 represented 43% of all liver graft recipients in Germany.

In addition to providing exception MELD points for patients with HCC, recent literature revealed how the MELD was compared against the Child-Turcotte-Pugh (CTP) model, which was previously applied in the United States, and how the MELD is considered a better prognostic model. Recent literature also revealed the outcomes of various MELD-based studies. Finally, a deeper look into the literature would identify and detail the various limitations, strengths, and benefits of MELD-based models.

MELD and MELD-based Models

Bernardi et al. (2011) explained that the impact of MELD scoring on the donor liver transplant allocation system had such an impact, that the period following the implementation of the MELD system is referred to as the "MELD era" (p. 1298). However, MELD has its weaknesses and many attempts had been underway to improve the applicability and reliability of the MELD formula with specific conditions. These attempts were based on the original MELD score as the original MELD is such that it can be employed in many settings. An analysis in the literature review of Huo et al. (2008) and Biselli et al. (2010) provided a comparison of multiple MELD-based models. This literature review details the benefits of the various MELD-based models.

Huo et al. (2008) and Biselli et al. (2010) both introduced, evaluated, and assessed various MELD-based models. Regarding the analysis of MELD-based models, Huo et al.

(2008) analyzed four MELD-based models by comparing and contrasting their risk of mortality prediction at 3 and 6 months. The four MELD-based models included MELD, MELD-Na, iMELD, and MESO index scores. Huo et al. explained that the MELD has been shown to be more accurate in predicting survival than the Child-Turcotte-Pugh (CTP) classification for patients with cirrhosis awaiting liver transplantation in the United States.

According to Huo et al. (2008), the MELD-Na, iMELD, and MESO index formulas are as follows (p. 838):

$$MELD(Na) = MELD + 1.59 * (135 - Na) \quad (2)$$

$$iMELD = MELD + (age * 0.3) - (0.7 * Na) + 100 \quad (3)$$

$$MESO\ index = (MELD/Na,\ mEq/L) * 100 \quad (4)$$

Huo et al. (2008) compared the short-term prognostic ability of the four models, MELD, MELD-Na, iMELD, and MESO index, to determine which MELD-based system have a better predictive accuracy in patients with cirrhosis. The criteria to select eligible patients included an initial Child-Pugh score of 6 or higher, with no coexisting hepatocellular carcinoma or human immunodeficiency virus infection, and a known initial MELD score at the time of evaluation and survival status at follow-up after 6 months.

Huo et al. (2008) explained that the iMELD tended to have the highest scores and that was followed by the MELD-Na, MELD, and MESO index at the time of initial evaluation. With the c-statistic and 3- and 6- month mortality as the endpoint, the estimated Area Under Curves (AUCs) for the four prognostic models in predicting

mortality were graphed. Of all patients, 83 patients or 10.1% of patients died at 3 months of follow-up, and 162 patients or19.6% of patients died at 6 months of follow-up. At 3 months, the iMELD had the highest AUC (0.807), and that was followed by the MELD-Na (0.801), MESO (0.784), and MELD (0.773). At 6 months, the iMELD still had the highest AUC (0.797), and was followed by the MELD-Na (0.778), MESO (0.747), and MELD (0.735).

Huo et al. (2008) concluded that to further improve the MELD-based liver allocation system, their studies have found that serum sodium (Na) is an important additional predictor of waitlist mortality. Hyponatremia is associated with severe complications of cirrhosis, including ascites, hepatorenal syndrome, and liver-related mortality. It has been suggested that Na should be incorporated into the MELD to further enhance the model's prognostic ability, and so a mathematical equation based on both MELD and Na known as the MELD-Na, has been developed to predict the 6-month mortality in patients with cirrhosis awaiting liver transplantation.

Huo et al. (2008) further concluded that the utilization of MELD has been demonstrated to have an equal or better ability in short-term or intermediate-term outcome prediction over the CTP system. In addition, the application of the MELD system has been shown to be a useful model in predicting the outcome of patients with cirrhosis undergoing surgical procedures for hepatocellular carcinoma and non-hepatocellular carcinoma conditions. Huo et al. noted that a potential limitation of their study is that the majority of the patients had chronic hepatitis B, were older with more males, and were taken from Taiwan’s Taipei Veterans General Hospital (p. 843). The

study results indicated that incorporation of Na into the MELD could enhance the prognostic accuracy of MELD for outcome prediction.

Similarly, Biselli et al. (2010) analyzed the results of six MELD-based score systems by comparing and contrasting the risk of mortality prediction. These six MELD-based parameters were MELD, UKELD, iMELD, MELD-Na, uMELD, and mCTP. At present, the MELD score is widely used for donor liver allocation, but it has shown some limitations. MELD is not directly influenced by other complications of cirrhosis associated with poor survival, such as persistent ascites and hyponatremia. For this reason, many recent studies have evaluated the effects of incorporating other variables into the model, such as serum sodium and age.

According to Biselli et al. (2010), the formulas of UKELD, iMELD, MELD-Na, and uMELD are as follows (p. 965).

$$UKELD = [(5.395 * \ln(INR)) + (1.485 * \ln(creatinen, \mu mol/L)) + (3.13 * ln(bilirubin, \mu mol/L)) = (81.565 * ln(Na, mmol/L)] + 435 \quad (5)$$

$$iMELD = MELD + (age * 0.3) - (0.7 * Na) + 100 \quad (6)$$

$$MELD(Na) = MELD - Na - [0.0225 * MELD * (140 - Na)] + 140 \quad (7)$$

$$uMELD = 1.266 * \ln(1 + creatinine, mg/dL) + 0.939 * \ln(1 + bilirubin, mg/dL) + 1.658 * \ln(1 + INR) \quad (8)$$

Biselli et al. (2010) explained that survival was calculated from the time of listing to drop-out, liver transplant, or end of the observation period. Drop-outs included

patients being removed from the list because of either death or worsening of their disease up to the point that they were too sick to undergo a liver transplant. Biselli et al. noted that at 6 months, the best calibrated score was iMELD. Furthermore, an iMELD cutoff of 39 identified listed patients with a worse prognosis more reliably than standard MELD of 15, whereas no significant difference was found with respect to standard MELD of 18. Similar to the study of Huo et al. (2008), at 3 months the iMELD had the highest AUC, showing an excellent diagnostic accuracy, followed by MELD-Na, but the comparison between AUCs showed that only MELD-Na had a better prognostic power than the standard MELD because of a very small standard error in the difference between the areas. At 6 months, the comparison between AUCs showed that only iMELD and MELD-Na had a better prognostic power than the standard MELD.

Biselli et al. (2010) explained that their study was a comparison of the short-term and intermediate-term prognostic ability of the standard MELD with respect to five alternative scoring models. The performance of these scoring models in relation to the varying severity of cirrhosis was specifically assessed by calibration analysis. The utilization of the MELD has been demonstrated to have an equal or better ability in short-term or intermediate-term outcome prediction in comparison with the CTP system. In addition, Biselli et al. similarly found that the application of the MELD system has been shown to be a useful model in predicting the outcome of patients with cirrhosis undergoing surgical procedures for hepatocellular carcinoma and non-hepatocellular carcinoma conditions. The studies of both Bisseli et al. (2008) and Huo et al. (2008) conclusively determined that the MELD-based models could be used to refine the

decision-making of liver allocation, particularly the iMELD and MELD-Na models.

In addition to providing exception MELD points for patients with HCC, Biselli et al. (2010) and Huo et al. (2008) revealed that when the MELD was compared against the CTP model, the MELD was considered a better prognostic model over CTP. Recent literature also included MELD-based studies that studied the delta-MELD. Bambha et al. (2004) defined the delta-MELD as the calculated difference between the *current-MELD* score from the lowest of all *serial-MELD* scores in the preceding 30 day window. Thus, delta-MELD is defined as the maximum change in MELD over the 30 day period. Delta-MELD was covered in several studies and was hypothesized to have prognostic predictive value (Cholongitas et al., 2006; Foxton et al., 2006; Gotthardt et al., 2009; and Young et al., 2006). However, these delta-MELD studies revealed that there are research problems regarding how delta-MELD was computed and interpreted.

Delta-MELD: Debate and Gap in Literature

The research of Bambha et al. (2004), Foxton et al. (2006), Young et al. (2006), Gotthardt et al. (2009), and Cholongitas et al. (2006) were reviewed on how delta-MELD can help refine the MELD system. These studies also focused on waitlist mortality while searching for a refinement for the MELD scoring system. A summary of their studies regarding the delta-MELD would reveal the conflict and inconsistency on the data collection for delta-MELD. Also, a methodology is presented and suggested regarding MELD estimation for the derivation of delta-MELD values.

The studies of Cholongitas et al. (2006), Bambha et al. (2004), Young et al. (2006), and Foxton et al. (2006) not only included MELD-based models, their studies

included the analysis of delta-MELD. Cholongitas et al. acknowledged that the MELD is now used for allocation of donor livers, and it has successfully replaced the Child-Turcotte-Pugh (CTP) model. However, there are still debates on whether the MELD is really superior to the CTP model in predicting mortality in patients with cirrhosis on the liver transplant waiting list and after liver transplant. Cholongitas et al. found from multiple studies, that only 4 of 11 showed MELD to be superior to CTP in predicting short-term mortality. In addition, two of three studies evaluating the changes in MELD score, delta-MELD, had shown that the delta-MELD had better prediction for mortality then the baseline MELD score. Finally, Cholongitas et al. also noted that several studies have shown that the predictive ability of MELD score increases by adding clinical variables, such as hepatic encephalopathy, ascites, and laboratory sodium parameters.

Cholangitas et al. (2006) evaluated the change in MELD, delta-MELD, in large cohorts of candidates on the liver transplant waiting list. In this evaluation, the delta-MELD score had better prediction for mortality than the baseline MELD score. For example, an increase of 5 points in delta-MELD during a 30-day period predicts a significantly increased risk of death. Cholangitas et al. also suggested that delta-MELD score be a tiebreaker for patients on the waiting list with identical MELD scores. When Cholangitas et al. explained how the delta-MELD, baseline MELD, and CTP were compared, the c-statistic showed that delta-MELD at 6 and 12 months was significantly better predictors compared to baseline MELD and CTP. Although it was found that MELD was significantly better than CTP in 4 of 11 studies, whereas 7 studies showed no statistical difference, there were no studies that showed MELD to be statistically inferior

to CTP scores.

Gotthardt et al. (2009) argued that an effort to improve the MELD system should involve analyzing the change in MELD scores over time while bearing in mind that this dynamic variable would reflect the dynamic of diseases in patients. When Gotthardt et al. analyzed a study of delta-MELD of 60 patients, the delta-MELD scores had better prediction for mortality than the baseline MELD score.

On the other hand, Bambha et al. (2004) explained that their research was focused on monitoring waiting list mortality and refining the MELD scoring system. Using an institutional liver transplantation database of serial MELD measurements for each patient, Bambha et al. found that the most recent MELD score for a patient awaiting liver transplantation was significantly associated with waitlist mortality. Bambha et al. also found that increasing MELD score, estimated by the slope of the line representing the changes of MELD scores over the 30-day period preceding the most recent MELD, conferred to an increased mortality risk on the waiting list, while decreasing MELD could be associated with a decrease in mortality.

However, Bambha et al. (2004) noted that an increasing MELD score may simply represent an intrinsic, irreversible component of the death process rather than being predictive of death in the future. For example, patients in the terminal phase of their disease may be expected to have increasing daily MELD scores during the last few days of life due to progressive organ failure. Bambha et al. further elaborated that when collecting laboratory data for calculation of delta-MELD scores, the potential for detection bias exists. For example, patients with acute liver illnesses, regardless of the

status of their liver disease, will undergo frequent laboratory tests producing multiple observations of MELD scores.

Foxton et al. (2006) explained that the MELD score is based on a methodology to predict poor survival in patients undergoing a liver transplantation. Foxton et al. also explained that MELD has been validated among multiple groups with liver disease and was shown to retain a high concordance with 3-month mortality. Foxton et al. noted that a change in MELD score, delta-MELD, while awaiting transplant has not only been suggested as a method of refining liver allocation, but delta-MELD should be examined for its impact on patient survival and intensive care stay. Foxton et al. found that using delta-MELD could subsequently help to improve overall outcomes. Foxton et al. found that delta-MELD over a period of 30 days was predictive of waiting list mortality and was significantly better than MELD score at the time of listing. Foxton et al. calculated delta-MELD by simply taking the MELD at transplant minus the initial listing MELD. Foxton et al. explained that various researchers may vary on how they calculate the delta-MELD value.

Young et al. (2006) also explained that the usefulness of MELD can be enhanced if it could also predict posttransplant outcomes. Predicting posttransplant outcome is important, as this would enable a more rational utilization of scarce resources to achieve their maximum benefit. The MELD score has been validated as an accurate tool for predicting 3-month mortality in different groups of patients with end-stage liver disease. Young et al. further explained that while MELD uses three readily measurable and objective parameters of bilirubin, creatinine, and INR in a logarithmic formula to produce

a score between 6 and 40, the wider range of MELD values when compared with CTP values can more easily allow the sickest patient to be prioritized. In addition, Young et al. similarly found delta-MELD and hyponatreamia to be significant parameters for predicting which patients would be placed on the waiting list and would not proceed to undergo a liver transplant. Young et al. noted that while a move to allocating donor liver solely by MELD is not justified for the U.K. allocation system, there is value in using MELD, delta-MELD, and hyponatreamia at predicting which patients should be placed on the waiting list and which would not proceed to transplant.

Young et al. (2006) confirmed in their study that hyponatreamia was highly significant in predicting which patients would not get a liver transplant. And delta-MELD has been studied before but until now has not been shown to be of significant value in determining allocation or predicting outcomes. Young et al. postulated that by identifying those patients who are hyponatreamic and who had large delta-MELD scores, it may be possible to prioritize them earlier and so have them liver transplanted before they become too sick to transplant. Hyponatreamia and delta-MELD, however, were not shown to be significant predictors of posttransplant outcome. By transplanting this group of patients sooner should not result in posttransplant outcomes that are worst off, but could even result in improved overall outcomes. Finally, data revealed that matching a poor quality organ with a sicker recipient will lead to much worse outcomes. Therefore, to allocate solely based on disease severity may sometimes discourage the use of marginal organs due to bad outcomes.

Young et al. (2006) concluded that one of the problems with delta-MELD is that

previous studies were biased due to the various collection method of the MELD data. Young et al. sought to minimize this bias by using the MELD scores at entry and exit from the waiting list. However this reasoning is flawed as this assumed a linear progression of MELD serial values while on the waiting list which is unlikely. Young et al. explained that in the study of Bambha et al. (2004), it was suggested that delta-MELD may be of limited value due to having too short a lead time to play a role in decision-making. On the contrary, Young et al. explained that it has been shown that despite relatively short waiting times, MELD can increase considerably prior to a liver transplant. A large prospective study designed to minimize collection bias is needed to fully clarify the role of delta-MELD in allocation. In conclusion, Young et al. noted that there is much data supporting MELD as a valuable tool in assessing potential liver transplant recipients in the U.K. Young et al. suggested that by using MELD and delta-MELD combined with a measure of hyponatreamia may improve the overall outcomes of allocating donor livers with more efficient and optimized timing to transplant while reducing waiting list deaths.

Even though Young et al. (2006) mentioned that a large prospective study should be designed to remove collection bias and to fully clarify the role of delta-MELD in liver allocation, there was no further research to either disregard or accept the delta-MELD as a valid criterion for determining a donor liver recipient for transplant. Furthermore, while Foxton et al. (2006) had calculated delta-MELD as the transplant MELD minus the listing MELD, Young et al. had calculated delta-MELD differently and as dxMELD, which is calculated by dividing delta-MELD by the time spent on the waiting list. While

the formula of dxMELD given by Young et al. seems to be more accurate, the time spent on the waiting list, the divisor, could be different among patients and the formula of dxMELD is different from the formula of delta-MELD given by Bambha et al. (2004).

To refine the concept of the *sickest first* policy based on a delta-MELD criterion, a methodology is needed for the consistent computation of delta-MELD prior to incorporating a multiple criteria decision-making model for analysis. A time series methodology can provide consistent estimation of data for delta-MELD computation. Huth et al. (2010) had suggested a methodology for tracking cell progress by the technique of the Kalman filter. The Kalman filter was used as a methodology for estimating data for the consistent computation of delta-MELD values.

Literature Review on Time Series Analysis Method

Kalman Filter

Not only was there a need for a methodology to estimate MELD values in a consistent manner based on patient MELD values, but a simulation was also needed to progressively track the MELD scores, stratify the groups of MELD values, and keep inventory of the patients on the waitlist. A simulation was needed to advance the study of delta-MELD in order to determine whether delta-MELD would be a useful criterion for donor liver allocation.

In this section, Kalman Filter, I examine the theoretical framework of the Kalman estimation and forecasting for the purpose of removing bias of delta-MELD computation. In the next section, *Kalman State Space, Prediction, and Estimation,* I examine the literature review of Asemoto (2010), Huth et al. (2010), Baker, Poskar, and Junker

(2011), and Zhou and Hu (2010) who demonstrated and highlighted the modeling of state space, estimation and prediction, and error management with Kalman filter.

Asemoto (2010) explained that the traditional time series analysis is primarily directed towards univariate data. The Kalman filter is a statistical algorithm that enables certain computations to be carried out for a model cast in state space form. Asemoto mentioned that even though the Kalman filter algorithm was proposed as far back as 1960, many statisticians are still unaware of the simplicity and succinctness of this methodology.

Asemota (2010) noted that state space models and the state-space representation of data are important tools for modeling time series data. State space models of random processes are based on the Markov property which implies the independence of the future of the process from its past, given the present state. In other words, the state of a Markov process summarizes all the information from the past that is necessary to predict its future. A state space model consists of two equations: the state equation, which is also called the transition or system equation, and the observation equation, which is also called the measurement equation. The measurement equation relates the observed variables or data, and the unobserved state variables, while the transition equation describes the evolution of the state variables (p. 7).

Asemoto (2010) explained that in a Kalman filter formulation, one can let Y_t, Y_{t-1}, ..., Y_1 be denoted as the observed values of an endogenous variable of interest at times t, $t-1$, ... 1 which depends on the unobservable quantity β_t and exogenous variable X_t, X_{t-1}, ... X_1 (which may be either scalars or vectors) through the following relationship:

$$Y_t = \beta_t X_t + \varepsilon_t \quad (9)$$

$$\beta_t = F_t \beta_{t-1} + V_t \quad (10)$$

where β_t is a vector of unobserved state variable, X_t is a vector of exogenous or predetermined observed variables.

The observation error ε_t and state error V_t are assumed to be Gaussian white noise sequences with zero mean and a covariance matrix. The covariance matrix is for the vector *X* of dimension *n x 1* is defined to be an *n x n* matrix that contains the ij^{th} elements that is the covariance between the i^{th} and j^{th} components of *X*. Equation (9) is known as the observation equation, and (10) is known as the state, system or transition equation. The system of equations (9) and (10) with their assumptions is called the state-space model. The essential difference between the state-space model and the conventional linear model representation is that in the state-space model, the dynamic nature of state is not assumed to be constant but may change with time. This dynamic feature is incorporated in the transition equation. The overall objective of the state space analysis is to study the development of the state over time using the observed values of the series (p. 7).

Zhou and Hu (2010) explained that a complementary Kalman filter differs from other similar work by adopting a refined noise model which could lead to an efficient computation of the Kalman filter. Zhou and Hu proposed a strategy consisting of four components, a prediction model, an error model, a standard Kalman filter, and a correction model. Zhou and Hu compared results by using a Kalman filter with those not using a Kalman filter, which only used a direct integration algorithm, and a kinematic

model to reconstruct the trajectories of joints from their studies. The results showed that a Kalman filter can significantly reduce errors in the orientation estimates when compared with the models in which no Kalman filter was used.

Huth et al. (2010) similarly explained that a method through the tracking of positions of individual cells over time, marked in consecutive images, through the use of the Kalman filter was compared to the migratory behavior of cells through the use of time-lapse microscopy. Huth et al. explained that the tracking of the positions of individual cells through the technique of Kalman filter was markedly improved and accurate over the technique of a tracking procedure that is commonly performed manually through a *point and click* imaging systems. In addition to being labor-intensive, a *point and click* method is highly susceptible to user-dependent errors regarding both the selection of subsets of cells for analysis as well as the manual determination of cell centroids serving as measuring points for cell positions.

Baker et al. (2011) explained that the focus of systems biology is to study the dynamic, complex, and interconnected functionality of living organisms. To have a systems-level understanding of these organisms, it is necessary to integrate experimental and computational techniques to form a dynamic model. Baker et al. elaborated that a Kalman filter designed for inference in a linear dynamic system can subsequently result in inaccurate results when applied to nonlinear systems. Instead, a number of extensions to the Kalman filter have been proposed to deal with nonlinear systems such as the extended Kalman filter (EKF) and the unscented Kalman filter (UKF).

In addition to establishing a state space model consisting of two equations that

contain both the state and the observation equations, the observed variables or data and the unobserved state variables need to be either collected or computed, while the transition equation describes the evolution of the state variables.

Kalman State Space, Prediction, and Estimation

Asemoto (2010), Huth et al. (2010), Baker et al. (2011), and Zhou and Hu (2010) applied the Kalman filter by setting their time series data into a state space model, performing prediction using the dynamic data of their systems, and using a correction model to reduce errors once their observed data are made available. These state space models were set up to achieve better cell tracking, positional estimation, and error reduction of time series data.

Asemoto (2010) also demonstrated how the Kalman filter recursive method can be applied to a model cast in state space form. The main advantage of the state space model is that it is based on a structural analysis of the problem that includes trend, seasonal, cycle, explanatory variables, and interventions that are put together into the state space model. The state space models are based on modeling the observed structure of the data.

Zhou and Hu (2010) explained that by including an error model in their study, a Kalman filter considers a state space representation and models the relation between the errors in the estimated orientation angle and the errors in an inclination predicted by the model. And by including a correction model, the error or noise can be reduced at a later stage.

Huth et al. (2010) demonstrated that systems represented as state space with the

Kalman filter can produce objective and highly reproducible measurements, outperforming manual tracking procedures. Huth et al. further explained that the precision of automated cell identification and centroid placement was very high, resulting in cell detection rates ranging from 96% to 99%. For the subsequent tracking of individual cell centroids through image sequences, Kalman filtering, commonly employed in multi-target tracking systems in military radar surveillance applications was utilized. Kalman filters are a set of mathematical equations allowing state ahead predictions of object positions as well as the estimation of optimized object states in noisy environments.

Baker et al. (2011) similarly noted for a nonlinear function of random variables, the use of the UKF is a technique that gives more accurate results than analytical linearization techniques, such as Taylor series linearization, as it considers the spread of the random variables. UKF is itself an extension of the unscented transform, a deterministic sampling technique which implements a native nonlinear transformation to derive the mean and covariance of the estimates. This transformed mean and covariance are then supplied to the Kalman filter equations to estimate the state variables.

These state space models were set up to achieve better cell tracking, positional estimation, and error reduction of time series data. This was made possible by knowing the observed data in order to refine the estimation of optimized object states in noisy environments. For the MELD data, the literature review provided insights on how to compute estimated MELD data when there are not enough MELD data available.

Asemoto (2010), Huth et al. (2010), Baker et al. (2011), and Zhou and Hu (2010)

discussed estimation and prediction techniques for their state space model. Estimation or prediction is only a part of their Kalman filter's iterative process. Their state space models were set up for the computation of not only predictions of the time series data, but for the estimation of errors and covariances.

Zhou and Hu (2010) explained that their prediction model included predictions of acceleration or gyroscope data that were generated based on previous estimates and sensor readings. The predicted estimates of the angular velocity at any time can be expressed as the summation of angular velocity estimate (world coordinate) and measurement errors which vary over time. Zhou and Hu then defined their estimated and their predicted estimate. Zhou and Hu explained that their correction model included error or noise reduction. Before proceeding to the correction model, Zhou and Hu needed to know the predicted estimates based on the prediction and error models. Zhou and Hu computed angular velocity error and acceleration error which were used as intermediate variables. Zhou and Hu explained that evaluation was performed during the iteration. The inclination difference was first minimized by the Newton method for the inclination difference. This was followed by optimizing using the proposed Kalman filter until the discrepancy is smaller than a fixed threshold.

Similarly, Huth et al. (2010) explained that the applied discrete Kalman filter algorithm consists of two alternating steps, which were repeated for each iteration and each new frame, prediction and correction. In the prediction step, the filter makes an assumption (prediction) about the future state of the observed object. In the correction step, an optimized state estimate was computed using a weighted difference between the

a priori state and an actual or noisy measurement. Huth et al. further explained that the weighting term was updated iteratively according to the quality of the previous prediction. If the prediction was good, the weighting term will suppress the influence of the measurement in the next iteration and show more weight in the state ahead prediction than in the measurement. If the prediction was poor, the weights are applied to the measurement more heavily in the next iteration while suppressing the influence of a predicted estimate.

Baker et al. (20111) argued that at the core of the UKF is the unscented transformation, which operates directly through a nonlinear transformation, instead of relying on analytical linearization of the system. Overall, the UKF has been found to be more robust and converges faster than the EKF due to increased time update accuracy and improved covariance accuracy. Baker et al. further noted that parameter estimation is highly dependent on the availability and quality of the measurement data. It could be difficult to obtain reliable estimates of unknown kinetic parameter values. Baker et al. found that in order to compare the parameter estimation methods, the nonidentifiable parameter was fixed to known values. In general, however, these parameters would not be known beforehand.

Kalman Error Management

Error management plays an important role in the processing of the Kalman filter. Asemoto (2010), Huth et al. (2010), Baker et al. (2011), and Zhou and Hu (2010) discussed their error reduction techniques for their state space models. Error reduction is a major part of the Kalman filter's iterative process.

Asemoto (2010) explained that the Kalman filter considers the mean square error (MSE) as the covariance matrix of the unobservable quantity minus a computed optimal estimator. Once current values become available, inference about the unobservable quantity on the basis of the observation, the MSE of the prediction can be computed. The prediction error in the state space model consists of two parts. These two parts include the prediction error due in making an inference about the unobservable quantity and the error in random shock of the observed values.

Zhou and Hu (2010) explained that their error model consisted of a state space representation and models the relation between the errors in the estimated orientation angles. The difference between the estimated and the corrected orientation angle and error was the difference between the gyroscope and accelerometer inclination estimates. Zhou and Hu compared the results of the Kalman filter with those not using a Kalman filter, a direct integration algorithm and a kinematic model, to reconstruct the trajectories. The result of Kalman filter significantly reduced errors in the orientation estimates.

Huth et al. (2010) managed the quality of the data set by taking an automatically generated track that was only regarded as valid if it followed one cell (and only one) through all frames in which the cell was visible. This stringent criterion was violated if a track failed to initialize, was prematurely terminated, or swapped between two cells. Huth et al. suggested that to simplify mitosis detection and track initialization / termination, backward tracking of the system was taken, meaning that cells were followed from the last to the first frame.

According to Baker et al. (2011), the UKF is more consistent throughout and in

estimating both larger and smaller values with a more consistent standard deviation. Baker et al. summarized that one of the benefits in integrating estimation and identifiability is the reuse of the variance generated by the UKF for the step size in the calculation and for the sensitivity coefficient for identifiability. The UKF is thus able to overcome one of the major bottlenecks in biological modeling, a lack of experimentally measured parameters.

MELD data can be set up into a state space model for reference, prediction, and error reduction. The methods reviewed are useful for estimating MELD data evenly and consistently for delta-MELD calculations. Kalman filter is useful in the estimation of MELD values into evenly time-spaced intervals that would be equivalent among all patients in order to compute consistently measured delta-MELD values. This would be an important aspect for removing bias that is described in the delta-MELD literature. Some of the techniques suggested in literature review of the Kalman filter included taking all the MELD values to compute estimated MELD data thresholds, backtracking data from the transplant time to the entry time of waitlist for better estimation of observed MELD, and error reduction by placing more weight to the observed MELD when estimating MELD values. These were all useful techniques for the simulation model.

Trends in the Liver Allocation System

MELD Era Objectives

The MELD era had brought on a major impact to the reduction of waitlist time and waitlist deaths while waiting for an available donor liver for liver transplant recipients without changing the overall outcomes of post liver transplants. The *sickest*

first policy has fulfilled its objective that was based on urgency (Bernardi et al., 2011). In order to simulate a decision scoring system to test whether delta-MELD would be a valid criterion, further literature review of the liver allocation system was reviewed. The following recent literature review would reveal a trend of not one objective but three objectives, urgency, utility, and survival, to refine the liver allocation system. In addition, the concepts of ECD and DRI were introduced. Additional trends were analyzed, such as increased age, MELD scores, and intensive care unit and hospital length of stays of transplant recipients. These literature reviews provided understanding on the objectives, criteria, and limitations of the liver allocation system that is informative for the simulation model.

Asrani and Kim (2010) acknowledged that the implementation of the MELD system has led to a reduction in waitlist registration and waitlist mortality. The MELD score had been useful in patient management, as well as providing an accurate gauge of liver disease severity. Asrani and Kim concluded that a future beyond MELD could be in updating the coefficients, adding terms that are better determinants of liver and renal functions, focusing on better donor-recipient matching, and updating the currently used urgency-based objective with the additional utility-based objective.

Weismuller et al. (2009) concluded that the prioritization of patients with higher labMELD scores for liver transplantation was followed by an increase in the mean MELD since the implementation of the Eurotransplant criteria in 2006. However, a decrease of post-liver transplantation survival was also observed. This led to the investigation of recipient and donor associated factors capable of determining outcome

after liver transplantation in the MELD-based allocation system, and thus provided insights to the variables influencing survival. Wiesmuller et al. also explained that in the United Kingdom, analyses indicated that delta-MELD and hyponatremia parameters were found to predict patients on the waitlist that did not reach transplantation. However, the prediction of posttransplant outcomes based on pretransplant parameters was much more difficult.

Wiesmuller et al. (2009) compared the graft quality between the two studied eras of transplant activity and recorded the parameters of donor age, cold ischemia time, split liver transplantations, gender matching, and ABO matching. In addition, the total time or duration of transplant surgery as a surrogate parameter for the technical complexity of the procedure was also reviewed. Wiesmuller et al. noted that the mean recipient age was found to be higher and the mean MELD rose from 14.3 years to 18.9 years in the MELD era. Wiesmuller et al. also noted that in addition to INR, bilirubin, and creatinine, there was an increase of blood urea nitrogen in the post-MELD group, which is an indicator of more severe renal or nutritional abnormalities. Mortality is also associated with the complexity of the surgery, along with cold ischemia time, age, and quality of donor graft. Wiesmuller et al. further noted that while the mean donor graft age did not differ between the groups, the mean cold ischemia time was significantly reduced in the MELD-era. However, the mean surgical procedure time was significantly longer. Further analysis showed that there was a significant correlation of the mean operation time with INR, and since INR is a component of MELD, also with MELD.

Wiesmuller et al. (2009) further noted that patients with high labMELD have the

highest probability of receiving a graft but they also exhibited the highest complexity and severity of disease. MELD encompassed two parameters, creatinine and INR, and in this analysis, these parameters were associated not only with the prediction of mortality before transplantation, but appear to be also predictive of 90-day survival. In a recent study, creatinine was also identified as an independent marker of posttransplantation survival together with cholinesterase and age. Wiesmuller et al. further noted that INR was associated with the length of operation time and this was found to also be an independent variable predicting 90-day mortality. Wiesmuller et al. explained that recent suggestions to modify the MELD score regarding INR and creatinine were confirmed. This indicated that the currently employed MELD used for prioritization could be modified to account for patients with complex morbidity to optimize overall posttransplantation survival.

Asrani and Kim (2010) similarly explained that the MELD-Na was associated with a higher risk of mortality independent of the MELD score in patients listed for liver transplantation. The effect was greater in patients with a lower MELD score. According to an analysis of 110 patients (23%), the difference between the MELD-Na and MELD scores was large enough to have affected allocation priority. Asrani and Kim argued that about 7% of the deaths on the waiting list could have been prevented if MELD-Na had been used rather than the MELD. Asrani and Kim also similarly explained, in the face of increasing use of ECD or high risk donors, identifying the right set of donor and recipient matching characteristics that would lend to a better outcome after liver transplant should be a significant objective.

Donor and recipient matching should occur at the time of organ procurement and transplantation with a substantial emphasis on selection going into accepting a donor liver. Given the importance of advanced donor age and graft quality, the arithmetic product of donor age and preoperative MELD (D-MELD) was recently evaluated as a predictive model (Halldorson et al., 2009).

Expanded Criteria Donor and Donor Risk Index

The MELD era brought along with it not only the conceptual framework of a *sickest first* policy for the reduction of time on waitlist to transplant but also the conceptual framework of optimizing *utilization* that would be based on how to best match donors and recipients. Blok et al. (2012) and Halldorson, Bakthavatsalam, Fix, Reyes, and Perkins (2009) reviewed, analyzed, and discussed the composition and issues of the DRI regarding donor-recipient compatibility for resource utilization in liver allocation. Blok et al. and Halldorson et al. also analyzed the risk, concerns, and modeling associated with DRI and posttransplant outcomes.

Blok et al. (2012) explained that a continuous scoring system for analyzing donor risk, DRI, has been developed within the OPTN. Blok et al. also validated the use of DRI in Eurotransplant. This was based on a database analysis of 5,939 liver transplants involving deceased donors and adult recipients from January 2003 to December 2007.

In addition, Halldorson et al. (2009) explained that recently, two developments have greatly impacted decision-making in liver transplantation. The first was the adoption of the MELD to prioritize the sickest patients for transplantation. The second was the increased use of higher risk donor livers to expand the donor pool and decrease

time to transplantation. Halldorson et al. argued that since posttransplant patient survival depends on both preoperative medical condition and donor liver quality, physicians are often faced with the difficult decision on whether to accept high risk donor liver offers for high risk patients. Halldorson et al. also elaborated that Feng et al. (2006) identified nine donor factors predicting graft failure after transplantation (donor age, donor height, donation after cardiac death donors, split liver donors, race, donor cause of death from cerebrovascular accident, regional sharing, local sharing, and cold ischemia time). Using these risk factors, a DRI was developed predicting the isolated and cumulative effects of these variables on graft survival. While highly informative, a DRI system is not easily translated into practical usage without making generalizations and extrapolations. In general, however, donor age is the predominant donor risk factor.

Blok et al. (2012) explained that when these data were analyzed, a significant correlation was shown between the DRI and outcomes. A multivariate analysis demonstrated that the DRI was the most significant factor influencing outcomes. Among all donor, transplant, and recipient variables, the DRI was the strongest predictor of outcomes. Blok et al. similarly described that with the increased need for liver allografts, the earlier and very strict criteria for liver donors have slowly become more liberal. However, the use of donors with additional risk factors may influence outcomes after liver transplantation. Currently, however, there is no unambiguous definition of what exactly these donor risk factors are and the extent of these risks, such as donor age, cause of death (COD), hypernatreamia, donation after cardiac death (DCD) status, and split liver status.

Halldorson et al. (2009) hypothesized that D-MELD, being the product of two continuous variables (donor age and calculated preoperative MELD), would result in an incremental gradient of risk for postoperative mortality and complications estimated in hospital LOS. Halldorson et al. also hypothesized that this gradient could then be used to identify a criterion where donor and recipient risks combined result in inferior outcome.

Blok et al. (2012) similarly noted the criteria used as risk factors for liver donation. These risk factors include donor age greater than 65 years, an ICU stay greater than 7 days, a high body mass index, steatosis, hypernatremia, high levels of aspartate aminotranferase (AST), alanine aminotransferase (ALT), and serum bilirubin. If any of these apply, a donor is considered marginal. However, Blok et al. argued that most of these donor criteria have never been validated, and parameters such as donation after cardiac death (DCD) status and split liver status were not included.

Blok et al. (2012) analyzed the set of factors contributing to DRI, which was developed by Feng et al. (2006) with OPTN data, into a continuous scoring system. These factors were based on only donor and transplant parameters found to significantly influence outcomes after liver transplantation in a multivariate analysis of a large cohort (20,023 transplants) from the Scientific Registry of Transplant Recipients (SRTR) database. Blok et al. noted these factors to be the donor's age, race, height, COD, split liver donation status, DCD status, type of allocation (local, regional, or national), and cold ischemia time.

Blok's et al. (2012) analysis showed that more than 48% of all transplants were from donors 47.6 ±16.5 years old. Fifty-three point eight percent of all transplants were

performed with livers recovered from male donors. Most donors died from cerebrovascular accident (CVA) (63%), only a little more than one-quarter died from a traumatic injury (27%). The DCD rate was 2.1% and the split liver donation rate was 4.4%. Among all donors, 0.9% was positive for hepatitis C core antibodies, and 5.8% were positive for hepatitis B core antibodies. Blok et al. also noted the differences between donor and transplant characteristics in OPTN/UNOS and Eurotransplant. The mean donor age was much higher in Eurotransplant versus UNOS (48 versus 39 years). The COD was more often CVA in Eurotransplant versus UNOS (63% versus 40.9%) and was less often trauma in Eurotransplant versus UNOS (26.7% versus 41.9%). The DCD and split liver donation rates were higher in Eurotransplant versus UNOS, and organs were more often allocated regionally than outside their regions in Eurotransplant versus UNOS. This resulted in a much higher mean DRI within Eurotransplant versus UNOS (1.71 versus 1.45). Similarly, Blok et al. (2012) noted that in Eurotransplant, 57.6% of all donors had a DRI > 1.5. This was the OPTN limit for twice as many discarded organs in comparison with donors with a DRI ≤ 1.1.

Halldorson et al. (2009) similarly noted that patients with MELD ≥ 30 and patients who received a donor liver aged ≥ 60 were analyzed as subgroups of the whole population and then studied as a population. Both MELD ≥ 30 and donor age ≥ 60 subgroups demonstrated worse survival when compared to the population as a whole. D-MELD was calculated as the simple product of donor liver age and laboratory-based MELD score capped at 40. The D-MELD scores were divided into groups of 400 and a D-MELD cutoff of 1,600 was found to best differentiate survival. Survival was

improved in both high risk groups if D-MELD was limited to less than 1,600. In the MELD ≥30 group, D-MELD < 1,600 demonstrated a 4-year survival of 71.3% versus 63.8% if D-MELD was 1,600 or greater. Halldorson et al. further noted that in all subgroups, with and without hepatitis C, survival was superior if D-MELD was limited to less than 1,600.

Blok's et al. (2012) analysis also showed that the outcome was strongly influenced by recipient factors. The recipient's age and the cause of liver disease were important factors influencing the outcome as well. The mean laboratory MELD score at transplantation was 20.3, and the mean age was 51.0 years. Halldorson et al. (2009) similarly explained that various donor and recipient risk factors influence patient and graft survival after liver transplantation. A major recipient risk factor is preoperative MELD score. The most influential donor risk factor is age. Halldorson et al. demonstrated that the product of these two factors, D-MELD, stratifies survival and LOS after liver transplantation. Halldorson et al. further noted that currently, liver allocation based on the MELD system is urgency which is based without regard for posttransplant survival. The merits of this current system lie in its simplicity, objectivity, and accuracy in predicting waitlist mortality. The downside was facing a dilemma in which centers in low donor to recipient ratio regions compete for organs by transplanting the sickest (highest MELD) patients or accepting the highest risk donor livers.

Blok et al. (2012) and Halldorson et al. (2009) did not only reviewed and identified the risk factors of donor livers, their studies analyzed the composition of DRI and established that the model D-MELD (Halldorson et al., 2009) can provide more

definitive measurements on the outcomes of recipient and donor matches based on donor age and recipient MELD score. The D-MELD measurement was the initial attempt to fulfill the objectives of *sickest first* and *resource utilization*.

Increased MELD Scores and Extended Intensive Care Unit Stays

Dutkowski et al. (2011) and Foxton et al. (2010) noted that the MELD era not only resulted in a higher mean MELD score, but also resulted in extended ICU stay, hospital LOS, and overall health care cost. Both Dutkowski et al. and Foxton et al. analyzed posttransplant data outside of the United States, although their results mirrored the trends of the United States.

Dutkowski et al. (2011) explained that there is currently an intense debate about whether liver grafts should be offered directly to a patient (the sickest one) or rather to a transplant center with the freedom to use an organ for the patient of their choice. In the United States, allocation of donor livers through the MELD system resulted in a substantial decrease in median time to transplant from 981 days in 2002 to 306 days in 2006 (p. 675). Dutkowski et al. described that despite this change leading to sicker patients at the time of transplantation, an initial analysis showed an excellent one-year survival after liver transplantation in the MELD era.

Foxton et al. (2010) similarly explained that liver allocation for transplantation worldwide has undergone dramatic change within the last 5 years, particularly with the introduction of the MELD. This policy change occurred because of increasing demand for liver transplant and increasing waitlist mortality. Foxton et al. explained that the MELD system removed the time variable on the waitlist as a discriminating factor in

allocation and mandated that organs are allocated to the sickest patients first. This has successfully resulted in a decrease in waitlist mortality with no corresponding worsening of mortality after liver transplantation.

Dutkowski et al. (2011) studied cost analysis of liver transplants that refers to all costs accumulating from the time of hospital admission prior to surgery until first posttransplant discharge. Dutkowski et al. further explained that as expected, introduction of the MELD policy increased the laboratory median MELD score of recipients from 13.5 to 20. One third of the transplanted patients, 32%, had a MELD score > 25 compared to only 14% in the pre-MELD era. Correspondingly, the preoperative incidence of hepatorenal syndrome increased in the MELD era from 14% to 35%. Significantly more patients in this group had to be hospitalized prior to liver transplant, 18% versus 35%.

Dutkowski et al. (2011) further noted that despite sicker transplant candidates in the MELD era, the proportion of patients with MELD ≥ 36 remained similar in both groups (4% versus 10%). Dutkowski et al. also noted in 6 months after liver transplant, the number of patients requiring renal replacement therapy was comparable and low in both groups. In addition, the median serum creatinine was not different in both groups after 6 months. Dutkowski et al. described that the median ICU and hospital LOS were 2 and 6 days longer during the MELD era, respectively. Also, the recipient MELD score correlated significantly with hospital LOS. Dutkowski et al. tallied the median cumulative cost per single case, which was from the time of admission to first discharge after liver transplant, and confirmed an increase in cost from U.S. $81,967 during the pre-

MELD era to U.S. $127,453 per case in the MELD era. Cost correlated strongly with the individual MELD score.

Foxton et al. (2010) similarly noted that the median cost associated with the ICU stay was U.S. $5,800 (IQR = U.S. $2,900 - $14,600). Forty-seven patients (11.7%) were admitted to ICU more than once following their liver transplant. Their median MELD score at transplant was 16. This was compared to a median MELD score of 14 for those who did not require ICU readmission. The need for renal replacement therapy (RRT) was associated with an ICU stay greater than 3 days. The median ICU cost of those receiving RRT was U.S. $52,812, whereas in those who did not require RRT post-liver transplant, the cost was U.S. $5,800. Foxton et al. explained that DRI was not associated with increased cost, whereas the MELD score was associated with increased cost. Also, when dividing DRI into groups, there was no correlation of any DRI group with increased health care cost or prolonged ICU stay.

Dutkowski et al. (2011) further compared the overall outcome of transplantation in a pre-MELD and MELD era. Previously, the number of patients with MELD > 25 at the time of listing was very low at 8%. Probably due to the fact that sick patients had no chance to receive a liver graft while waiting, and at that time, some end-staged candidates were not even placed on the waitlist. And it can be hypothesized that the true death rates were much higher in the pre-MELD era. Dutkowski et al. found that the number of patients requiring renal replacement therapy in the post group exceeded the number of comparable cases in the pre-group. However, 6 months after liver transplants, most kidneys recovered in both groups. Dutkowski et al. noted that countries with very low

donation rates are under higher pressure to use grafts from ECD. Dutkowski et al. found the price of MELD allocation to be an increase in postoperative morbidity, resulting in longer hospital stay, temporary renal complications, and higher health care cost.

Foxton et al. (2010) similarly found that their study demonstrated that ICU costs associated with liver transplantation increased with increasing MELD score particularly where MELD > 24. Recipient age, alcohol-related liver disease, and the severity of liver disease prior to transplantation, in the form of UNOS status or CTP score, have been shown to have significant impact on resource utilization according to the study of 711 patients who underwent liver transplantation in 3 U.S. transplant centers. Foxton et al. explained that in their study, they were not able to identify recipient age or alcoholic liver disease as factors that were associated with higher costs, although alcoholic liver disease was associated with a prolonged ICU stay.

Foxton et al. (2010) iterated Feng's et al. (2006*)* DRI formula as follows (p. 669).

$$
\begin{aligned}
DRI = \exp[&(0.154\ if\ 40 \le age < 50) + (0.274\ if\ 50 \le age < 60) \\
&+(0.424\ if\ 60 \le age < 70) + (0.501\ if\ 70 \le age) \\
&+(0.079\ if\ use\ of\ death = anoxia) \\
&+(0.145\ if\ cause\ of\ death = cerebrovascular\ accident) \\
&+(0.184\ if\ cause\ of\ death = other) \\
&+(0.176\ if\ race = Africa/Afro\ Caribbean) \\
&+(0.126\ if\ race = other\ ) \\
&+(0.411\ if\ donated\ after\ cardiac\ death) + (0.422\ if\ partial/split) \\
&+\{0.066\ [(170 - height)/10]\} + (0.105\ if\ regional\ share)
\end{aligned}
$$

$$+(0.244\ if\ national\ share) + (0.010 * cold\ ischemia\ time)] \quad (11)$$

Dutkowski et al. (2011) concluded that the MELD system addresses the goal of urgency and hence fairness the best. Despite the expected higher postoperative efforts, it still appears to be the most reliable tool for selecting liver transplant candidates. Foxton et al. (2010) likewise concluded that the group of patients with the highest MELD score, ≥24 points, also has the highest health care cost, reflecting significant increase in ICU costs and therefore overall transplant costs. However, patients with MELD score ≥24 represented only 8% of the cohort. Foxton et al. also showed that there is a significant increase in DRI over time which reflects the trend of transplant centers using more marginal grafts. This finding mirrored the responses to organ shortage that was experienced nationally and internationally. Foxton et al. noted that in practice, there was a clear attempt to match better organ quality with patients of higher MELD scores. This likely reflected appropriate matching of donor organs to recipients by experienced surgeons in an attempt to optimize outcomes and maximize donor organ utility. Foxton et al. also concluded that DRI failed as a predictor in determining total liver transplant, pretransplant, or posttransplant cost.

Although Foxton et al. (2010) and Dutkowski et al. (2011) did not establish a model fulfilling a liver allocation objective, per se, they showed the trends of the MELD era with higher median MELD scores and that higher MELD scores correlated with higher health care cost, longer ICU stays, and longer hospital LOS. Along with higher MELD scores, there was more liberal criteria of accepting ECD livers, and hence the need to stratify donor livers according to their DRI scores.

Decision-Making of Multiple Objectives: Urgency, Utility, and Survival

The literature review of Rana et al. (2008) and Rodrique, Hanto, and Curry (2011) did not only provide modeling and behavior perspectives of the ECD in liver allocation, but the perspective of patients' consent and feedback on accepting an ECD liver. This helped to shape the simulation model to meet the multiple objectives of urgency, utility, and survival with the possibility that an available liver is an ECD liver. The simulation model not only simulated the delta-MELD as being a criterion, it also simulated the liver allocation system realistically to the existing system, including the incorporation of ECD livers.

Rana et al. (2008) argued that because recipient factors alone were not predictive of survival following transplantation, a new model was required to accurately predict posttransplant survival. The lack of consideration of donor risk factors is one limitation of the existing standard, which is transplanting patients with a MELD greater than 15. Rana et al. explained that recently, the DRI has been proposed as a method to stratify outcomes associated with graft selection. However, the lack of recipient factors gave the DRI alone poor predictive value. In their analysis, Rana et al. combined both donor and recipient risk factors to construct the survival outcomes following liver transplantation (SOFT) score to accurately predict recipient posttransplant survival at 3 months. This score would allow clinicians to balance waitlist mortality at 3 months as predicted by the MELD score against 3-month mortality following liver transplantation as predicted by the SOFT score to determine which patients should undergo liver transplantation.

Rodrique et al. (2011) noted that despite notable efforts to increase rates of

deceased organ donation and living liver donation, the supply of livers has not kept pace with the growing demand for transplantation. Increased utilization of livers from higher risk deceased donors is one strategy to overcome the severe organ shortage, although there is no uniformly accepted definition for what constitutes an ECD liver. Rodrique et al. explained that a DRI has been developed and can be used to assess the relative risk for a potential graft and the relative risk for a specific recipient. Rodrique et al. further argued that the decision to utilize an ECD liver for transplantation is complex where patient's disease severity, comorbidities, and survival without transplantation are considered. During the pre-liver transplant period, patients are informed of ECD versus standard criteria donor (SCD) transplantations. Rodrique et al. explained that patients would need to provide explicit consent for ECD liver transplants. This may be a challenging decision for some patients, who must balance the risks and benefits of an earlier ECD liver transplant versus the risks and benefits of waiting for a SCD liver transplant at a time when their emotional, physical, and cognitive resources are likely compromised.

Rana et al. (2008) explained that their risk score was actually two different risk scores, the preallocation score to predict survival outcomes following liver transplantation (P-SOFT), and the score to predict survival outcomes following liver transplantation (SOFT). Rana et al. formulated two distinct scores. The P-SOFT is designed to evaluate patients on the waitlist and the SOFT includes both donor and recipient factors to evaluate transplant outcome at the time of transplantation. Rana et al. explained that since MELD has been proven to be an accurate predictor of 3-month

waitlist mortality, Rana et al. constructed the SOFT score to complement the MELD score by predicting 3-month posttransplant mortality. The SOFT score along with the MELD score would allow clinicians to make a more definitive decision on whether to accept a particular allograft. Rana et al. argued that the SOFT score can also be used to avoid wasteful transplants when predicted survival is below an acceptable standard. As the critical liver allograft shortage encourages more aggressive practices to utilize marginal donor allografts, the SOFT score can establish risk limits for particular liver transplant candidates.

Rana et al. (2008) concluded that candidates with a MELD score ranging from 17 to 19 points should only receive low-risk SOFT transplants. Candidates with a MELD score of 20-29 points should receive low or low-moderate risk SOFT transplants. Candidates with a MELD score of 30-39 points should receive low, low-moderate, or high-moderate risk SOFT transplants. And candidates with the highest waitlist mortality risk with a MELD score of greater than 40 should receive low, low-moderate, high-moderate, or high-risk SOFT transplants. These recommendations likely do not apply to patients with hepatic cancers since the benefit of early removal of tumor must also be considered in addition to the MELD and SOFT scores. Rana et al. emphasized that transplants in patients with a SOFT score of > 40 are likely futile since the predicted posttransplant mortality is greater than any waitlist mortality.

Rodrique et al. (2011), on the other hand, studied the willingness of patients accepting an ECD liver transplant. Transplant hepatologists or surgeons determine whether a patient is medically eligible to receive an ECD liver transplant and would

discuss this option with the patient. Eligibility includes the ability to read, speak, and understand English, and the ability to provide written informed consent. Expanded criteria donor liver transplant is discussed with patients during their initial visit with the transplant hepatologist and subsequently during their appointment with the transplant surgeon. Rodrique et al. noted that patients are also informed that they can pursue live donor liver transplant at another program or deceased donor liver transplant in another region, which may reduce the time they would otherwise have to wait for liver transplant in their current program. Finally, all patients need to attend a 90-minute liver transplant orientation class, which includes a discussion of ECD liver transplant, multiple listing, and live donor liver transplants. Rodrique et al. argued that while there is considerable discussion about the definition, breadth, and outcomes of ECD liver transplants, there were no studies examining patients' willingness to accept ECD liver transplant. Hence, the aim of their study was to assess patients' willingness to accept ECD liver transplant, identify the increase in mortality risk they are willing to assume relative to a SCD liver transplant, and examine the associations between sociodemographic variables and ECD liver transplant willingness.

Rodrique et al. (2011) found that patients were significantly less willing to accept ECD versus SCD liver transplants. Most patients were willing to accept a 1-year ECD liver transplant mortality risk that is higher than that expected for SCD liver transplant. Patients with high labMELD scores and patients of the white race were more willing to accept ECD livers and ECD livers with higher 1-year post-liver transplant mortality risk. Rodrique et al. found that in more than half of the study, patients reported a low

willingness to accept an ECD liver transplant, and one-third were unwilling to consider ECD liver transplant at all. Not surprisingly, patients with higher labMELD scores were more willing to accept ECD liver transplant and higher 1-year post-liver transplant mortality risk than those patients associated with SCD liver transplant. Rodrique's et al. analysis showed that the rate which patients are willing to accept an ECD liver with higher 1-year mortality risk was 25%.

Rodrique et al. (2011) hypothesized that since the MELD score is a reflection of the short-term survival probability without liver transplant, patients with higher MELD scores may feel a sense of urgency in trying to best balance the ECD liver transplant mortality risk with their risk of death while waiting for a higher quality SCD organ offer. One could reasonably hypothesize that patients with hepatocellular carcinoma (HCC), faced with potential malignancy and tumor progression, would be more willing to accept an ECD liver than patients without HCC. This study did not support this hypothesis, as patients with HCC did not differ from patients without HCC in their ECD liver transplant willingness. Furthermore, MELD score with exception points was not associated with ECD liver transplant willingness or mortality risk acceptability.

Regarding decision tools, Bernardi et al. (2011) described that an ideal decision tool should be able to achieve multiple objectives. It should quantify a patient's chances of survival in the short to medium-term for optimal allocation of patients waiting for liver transplants. It should classify patients according to their disease stage, while enabling doctors to determine whether it is too early, appropriate, or too late to perform a liver transplant. It should also be able to predict outcome regardless of the disease. Finally, it

should set aside subjective factors influencing the doctors' judgment, such as features of the transplant center, human resources, and physician's individual expertise.

Freeman, Jamieson, Schaubel, Porte, and Villamil (2009) added that the widening gap between the demand and supply of donor livers has prompted governments and medical policy-makers to develop strategies to optimize liver graft allocation. While realizing that the donor liver pool will never be sufficient to meet the demand, liver transplant practitioners have tried to expand the criteria that define graft quality acceptable for transplantation yet recognizing that expanding criteria often come with risks to transplant recipients. Some of the most difficult decisions have been focused on determining which patients with acute liver failure should receive transplants since most will die without transplantation while realizing that only a fraction of these patients will recover. Freeman et al. discussed that the allocation of liver graft should be based on patient-based models that consider urgency, utility, and survival benefits, while considering ECDs, HCC patients, and acute liver failure (ALF) patients.

The model formulas of DRI, D-MELD, P-SOFT, and SOFT were established to meet the utility and survivability objectives, while accommodating to the current trend of using ECD livers for liver transplant. Current literature revealed that ECD liver transplants would require consent from patients acknowledging the risks and awareness of accepting ECD livers. Hence, a decision-making model would take on multiple objectives, multiple criteria, and the dependency of patient consent of ECD livers for the research simulation.

In the next section, I review, evaluate, and assess the literature that is related to

modeling a multiple criteria and objectives decision-making tool appropriate for the liver donor allocation system, the analytic hierarchy process / analytic network process (AHP/ANP). Both AHP and ANP encompass the quantitative measurement of consolidating multiple objectives and criteria among many alternatives.

Literature on Multiple Criteria and Objective Decision-Making

Analytic Hierarchy Process / Analytic Network Process

The AHP's usage, purpose, and construction are the focus of the following literature review. The remaining literature review is based on the research by Parthiban and Goh (2011), Danner et al. (2011), Ishizaka, Balkenborg, and Kaplan (2011b), Sipahi and Timor (2010), and Ishizaka, Balkenborg, and Kaplan (2011a) on the use of AHP. This collection of literature highlights how AHP is flexible, consistent, simple in the development of pairwise comparisons, and straightforward in incorporating the decision-making requirements for the simulation model.

Danner et al. (2011) presented the AHP as a preference elicitation method in health technology assessment. Their AHP study included two AHP workshops where in these workshops, both patients and professionals rated their preferences with respect to the importance of different endpoints of antidepressant treatment by a pairwise comparison of individual endpoints. These comparisons were performed and evaluated by the AHP method and relative weights were generated for each endpoint.

Danner et al. (2011) explained that the six most important patient-relevant and professional-relevant outcome measures resulted in the same outcome for the two independent groups, and thus validating the consistency of the AHP. These six endpoints

covered 85% of the overall weights in the patient group and 89% in the professional group.

Parthiban et al. (2011) proposed that an integrated model consisting of performance measurements and quality factors measurements can be evaluated by using the AHP. Parthiban et al. provided a way to identify the current performance of an organization and a methodology for further improvement. An important contribution of the AHP model is that it combines both the qualitative and quantitative dimensions of manufacturing performance measurements. For Parthiban et al., both the objective and manufacturing quality factors have been converted into consistent dimensionless indices to measure system performance. Partiban et al. demonstrated that the applicability of the AHP model can support a manufacturing performance measure where AHP can be used to calculate the two different manufacturing units using time, cost, and service quality dimensions.

Ishizaka et al. (2011b) used experimental economics methods to test how well AHP fared as a choice support system in a real decision problem. Analytic hierarchy process provided a ranking that can statistically compare with three additional rankings, given by the subjects in the experiment, one at the beginning, one after providing AHP with the necessary pair-wise comparisons, and one after learning the ranking provided by AHP. While these rankings varied widely across subjects, it was observed that for each individual, all of the rankings were similar. Hence, AHP was able to replicate their rankings. Furthermore, the AHP ranking helped the decision-makers reformulate their choices by taking into account suggestions made by AHP.

Sipahi and Timor (2010) presented a detailed literature review of the recent applications of the AHP and analytic network process (ANP) group decision-making methodologies. The findings showed that during the years 2006-2009, the use of the AHP technique has continued to increase exponentially. Moreover, it is expected that ANP will gain more popularity in manufacturing, followed by the environmental management and agriculture field, power and energy industry, transportation industry, construction industry, and healthcare industry.

Ishizaka et al. (2011a) described a decision problem with an inherent trade-off between two criteria. For instance, a job may require two unrelated skills and workers tend not to be adept at both. Ishizaka et al. compared the additive AHP and its variant, the multiplicative AHP (MAHP), with the utility theory to evaluate the choice among three alternatives: two extremes and one compromise. The utility theory has a normative approach and AHP a descriptive or a practical orientation. In this study, Ishizaka et al. aimed to demonstrate the effects of the aggregation method of local priorities and the measurement scale of AHP on the selection of a compromise, and hence to the degree of agreement with the utility theory.

AHP Background and Applications

The experiments among the researchers were varied and the results of their studies would show that AHP is a valid decision-making tool, which can be integrated with other tools, and can be used to aid in decision-making processes. In addition, from the wide amount of research in literature, AHP and ANP are shown to be versatile, consistent in their technique, and are widely applied into decision problems across

multiple industrial sectors.

Danner et al. (2011) explained that AHP is an approach where a multi-attribute decision problem is first structured into a hierarchy of interrelated elements. This hierarchy is a tree-like structure that is used to decompose the decision problem, moving from main criteria to more specific sub-and sub-subcriteria. Pairwise comparisons of these criteria are separately performed at each level of a decision hierarchy from the lower-level to the upper-level criteria. Important methodological constraints within AHP regarding the decision hierarchy are the independence and comprehensiveness of criteria at each level. Danner et al. further explained the matrices of the pairwise comparisons, Saaty's mathematical algorithm as a key element within AHP allowing the calculations of an approximation vector representing preference-based weights for each of the decision criteria. While the preferences in AHP are recorded on a numbered but ordinal scale, calculation of preference weights is performed by transforming this scale into an approximation cardinal one. Danner et al. further explained that weights can be calculated for each endpoint and for each person, and the group geometric mean can be calculated for a group of individuals taking part in the AHP. In addition, because reciprocity and transitivity of preferences is required within AHP, AHP allows for calculation within a measure of consistency for each group of pairwise comparisons. This measure reflects how logical each pairwise comparison is with regard to the remainder of comparisons performed by the same individual. This consistency ratio, as a measure of performance within the AHP, has a threshold of 0.1 that should not be exceeded.

Parthiban et al. (2011) demonstrated the steps of AHP using criteria from performance measures that were classified into objectives and quality factors. A structured survey was conducted at two organizations, Unit A and Unit B using the same questions to elicit the performance measure classified into objective and quality factors. After the AHP was performed, the quality factor measure was then calculated from the results of AHP, followed by the service factor measure for both locations, A and B. This yielded a result that Unit A has a lower service factor measure value which meant Unit A needs improvement more than Unit B. Quality function deployment has been employed to facilitate this process. This was useful in establishing the priority of actions within the overall re-engineering strategy.

Sipahi and Timor (2010) reviewed recent literature that was comprised of a comprehensive literature review of recent applications of AHP and ANP as decision tools over the period of 2005-2009. Saaty (2001) developed the AHP technique, which constructed a decision-making problem in various hierarchies as goal, criteria, sub-criteria, and decision alternatives. Sipahi and Timor also explained that AHP provides decision-makers with a way to transform subjective judgments into objective measures. Due to its mathematical simplicity and flexibility, AHP has been a favorite decision tool for research in many fields, such as engineering, food, business, ecology, health, and government. Saaty (2001) also developed another technique, the ANP technique, as a generic form of AHP that allows for more complex interdependence in relationships, and feedback among elements in the hierarchy. Sipahi and Timor further explained that ANP has been used in several decision-making applications in the last decade, especially in the

study of risk and uncertainty.

Ishizaka et al. (2011b) noted that while the rankings vary widely from individual to individual, they found, by using a variety of non-parametric statistical tests, that for each individual the ranking generated by AHP is typically in reasonable agreement with the rankings provided by each participant. While the study did show that AHP detected a clear top and least priorities well, the study also found that the other rankings given by the subjects tend to be closer to each other than they are to the AHP ranking. Ishizaka et al. also noted that there is evidence that the subjects tend to follow the ranking provided by AHP and found the experiment showed that AHP is a useful decision tool and that AHP could be used as a decision aid.

Ishizaka et al. (2011a) elaborated that one of AHP's strengths is the possibility to evaluate quantitative and qualitative criteria and alternatives on the same preference scale. In Saaty's AHP, the verbal statements are converted into integers from 1 to 9. Theoretically there is no reason to be restricted to these numbers. Therefore, other scales have been proposed. With integers 1 to 9 being local weights, which could be unevenly dispersed, there could be a lack of sensitivity when comparing elements which are preferentially close to each other. Using a logarithmic scale could be smoother for these high values.

Ishizaka et al. (2011a) explained the decision technique of AHP and MAHP. This study described and discussed the hiring decision problem solved with AHP and MAHP. All the possible matrix combinations with an acceptable consistency were used with each preference scale. For the MAHP, four different weight normalizations were

applied. Then, the results of the AHP and MAHP were compared with the consumer choice theory. The final position of a compromise in a candidate was selected. Ishizaka et al. compared the result with the standard consumer theory where the consumer would prefer a compromise alternative B. The choice of a power or geometric scale excluded definitely (for AHP) or almost definitely (for MAHP) the compromise alternative. The MAHP captured the obvious case in which B should win.

AHP Decision-Making with delta-MELD for Simulation

The strength of the AHP method is in its method to reduce the cognitive burden of decision-making by decomposing a complex decision problem into a limited number of pairwise comparisons (Ishizaka et al., 2011b). The AHP with its applications of pairwise comparisons of criteria is shown to be in accordance with human behavior, especially if it is based on bounded rationality. Saaty's (1996) method of deriving priorities from pairwise comparisons based on matrix multiplication and the eigenvector calculation is not only mathematically sophisticated, but it is reflective of human decision-making. The decision-making of selecting a liver transplant recipient, using an integrated model of urgency (MELD-based), utility (DRI), and survivability (SOFT) can be set up, modeled, and constructed by AHP.

The literature review showed that AHP is a powerful decision tool for assessing decisions of many and various decision situations. Although AHP does not take into account dependencies and interrelationships among factors, real world problems usually consist of dependencies or feedback between elements (Sipahi & Timor, 2010). One such example is in the application of the simulation model where feedback is needed

from patients to agree on accepting ECD livers for liver transplant. For this reason, ANP was considered for taking into account patient consent.

To summarize the MELD-based aspect of the literature review, my analysis of Huo et al. (2008) and Biselli et al. (2010) provided a comparison and assessment of the MELD-based models of MELD-Na and iMELD with conclusive and known results for refining the liver allocation system. But the analysis of Foxton et al. (2006), Young et al. (2006), and Cholongitas et al. (2006) conflicted with Bambha et al. (2004) on the assessment of the MELD-based delta-MELD for refining the liver allocation system. These studies did not use AHP for decision-making analysis, but delta-MELD used as a major criterion in an AHP could help to resolve this conflict in literature regarding the delta-MELD utility for refining the liver allocation system.

Although Young et al. (2006) mentioned that a study should be designed to definitively disregard or accept the delta-MELD for refining the liver allocation system, there was still missing research on the assessment of delta-MELD as a valid criterion for liver allocation. Gotthardt et al. (2009) argued that an effort to improve the MELD system should involve analyzing the change in MELD scores, delta-MELD, over time as this dynamic variable would reflect the progression of disease in patients. AHP appeared to be an appropriate model to analyze whether the current liver allocation can be improved upon by taking into account the MELD as primary criterion and then comparing it to an AHP model with both the MELD and delta-MELDs as primary criteria.

Meanwhile, a recent trend of the liver allocation system has moved from the

objective of urgency to include utility and survivability. In addition, since the earlier and very strict criteria for liver donors have become more liberal, ECD livers have recently become more widely employed for liver transplant than in the past. This suggests that these factors needed to be considered for the study on the role of delta-MELD for liver allocation. A study needed to be conducted to deterministically conclude the utility of delta-MELD with consideration of the recent trends of the current liver allocation system. It was conceivable to consider delta-MELD along with the recent trends of the current liver allocation system in the construction of an AHP model for this study.

Conclusion

The literature review described the MELD and MELD-based development, and described the history of MELD in the evolution of the liver allocation system. The MELD-based variable, delta-MELD was defined, its background of its use in research was evaluated, and its gap regarding it being a viable criterion for liver allocation in literature was reviewed.

Young et al. (2006) sought to minimize the bias due to various collection method of the MELD data by using the MELD scores at entry and exit from the waitlist. The varying duration on the waitlist among patients suggested that a methodology should be considered to approximate the delta-MELD into consistently measured time intervals. Young et al. further explained that a study should be conducted to fully clarify the role of delta-MELD in liver allocation. Since decision-making in liver allocation system has evolved from the objective of urgency to include utility and survivability, decision-making should be multi-objective while integrating with the OPTN liver allocation.

Also, since ECD livers have recently been more widely employed for liver transplants than in the past, ECD livers were also considered in the study of delta-MELD for liver allocation.

In Chapter 3, I describe, outline, and define how the simulation was performed in two scenarios, with and without delta-MELD being used as criterion. Chapter 3 will explain how the average MELD scores and number of patients dropping off the waitlist from the two scenarios were compared. The decision-making technique behind the simulation was AHP.

Chapter 3: Research Method

In Chapter 3, I describe the research design, variables, parameters, and instrument (a simulation model). The simulation was run to determine whether delta-MELD should be used in addition to MELD as primary criteria in patient selection. As a patient selection criterion, the delta-MELD could improve the liver allocation system by reducing the number of pretransplant patients from dropping off the waitlist (Research Question and Hypothesis 1) and by lowering the average MELD score among pretransplant patients (Research Question and Hypothesis 2). The experiment utilized a simulation model of two scenarios which used secondary data from the OPTN/UNOS Standard Transplant Analysis and Research (STAR) database, additional estimated MELD values computed through a Kalman filter, and computed delta-MELD values. The additional MELD values through the Kalman filter supported consistent measurements of delta-MELD values among all patients on the waitlist. The theoretical frameworks of Kalman estimation and AHP for patient selection were applied in the simulation experiment. The simulation model, data collection by Kalman estimation, decision-making by the AHP technique, and the OPTN liver allocation policy are elaborated in the sections below.

In Chapter 3, I provide justifications on the validity of the research instrument and its methodologies. In addition, explanations of how the simulation outputs data for research summary, analysis, and conclusion are provided. Finally, in Chapter 3, I explain how the results of the experiment were designed to answer the research questions and hypotheses.

The Institutional Review Board (IRB) approval number for this study is 05-01-14-0175913.

The Research Design

Frankfort-Nachmias and Frankfort (2008) described quantitative research as a deductive research that deals directly with the operationalization, manipulation of variables, predictions, and testing. Hence, quantitative research places particular emphasis on the research methodology, procedures, and their validity. Consequently, quantitative research design should be arranged to show a clear progression from theory to operationalization of concepts, a correspondence from the choice of methodology to the procedures, and the association from statistical tests to findings and conclusions (p. 488). Furthermore, the findings and conclusions would relate and provide answers to the research hypotheses and questions.

Hillier and Lieberman (2010) outlined the steps that a major simulation study should contain. These steps include identifying the research problem, collecting the data, formulating the simulation model, constructing the computational program, planning the experiments to be performed, conducting analysis of the experiments, and summarizing and concluding the study. In addition, the research design details how the data are gathered, processed, and measured, and how the simulation is constructed and used to influence its outcome.

The description of the research questions and hypotheses, process and steps, variables, data collection through secondary data, data organization for the simulation, and simulation model are described below.

Research Questions and Hypotheses

The research questions and hypotheses below are repeated from Chapter 1.

1. Does a simulation model using the additional parameter of delta-MELD as a patient selection criterion reduce the number of pretransplant patients who dropped off the waiting list?

The research hypotheses for the first research question are as follows.

H_o: There is no difference in the number of patients who dropped off of the waiting list (*Total_Patients_Removed*) between simulation models with and without delta-MELD along with the MELD score as primary criteria for patient selection in donor liver allocation.

H_a: There is a difference in the number of patients who dropped off of the waiting list (*Total_Patients_Removed*) between simulation models with and without delta-MELD along with the MELD score as primary criteria for patient selection in donor liver allocation.

2. Does a simulation model using the additional parameter of delta-MELD as a patient selection criterion lower the average MELD score among pretransplant patients?

The research hypotheses for the second research question are as follows.

H_o: There is no difference in the average MELD score ($MELD_{mean}$) among pretransplant patients between simulation models with and without delta-MELD along with the MELD score as primary criteria for patient selection in donor liver allocation.

H_a: There is a difference in the average MELD score ($MELD_{mean}$) among pretransplant patients between simulation models with and without delta-MELD along with the MELD score as primary criteria for patient selection in donor liver allocation.

Research Process and Steps

The research process and steps in Table 1 comprise the research design.

Table 1

Research Process and Steps

	Research Process and Steps
Step 1	Identify/describe the problem and the plan of study. Describe research purpose, questions, and hypotheses for simulation.
Step 2	Describe the control factors and response variables.
Step 3	Formulate the simulation model. Describe key pieces of data, processes, parameters, and events in the model.
Step 4	Ensure the validity of the simulation model. Describe how the theoretical frameworks are incorporated into the simulation and their construct validity for the simulation.
Step 5	Test (verify) the simulation model.
Step 6	Plan the simulation experiment to be performed. Organize combinations of the control factors (inputs to the simulation/experiment) in some form of an experiment.
Step 7	Conduct the experimental simulation runs and analyze the results. Perform quantitative analysis on the outputs of the simulation runs by conducting statistical *t* tests of two independent populations in order to test the hypotheses.
Step 8	Provide analysis, conclusion, and summary of the research questions, based on the results of the hypothesis tests. Provide explanation of limitations and recommendations for future studies.

Simulation Overview

The simulation and its computations were implemented by using Microsoft Excel and C++ programming language with NetBeans' integrated development environment (IDE) for the Windows operating system. The simulation generated output data (the response variables) in two scenarios: (a) using only MELD as primary criterion, and (b) using delta-MELD and MELD as primary criteria in the simulation scenarios. Hence, the control factor for the experiment was the presence or absence of delta-MELD as a simulation parameter.

There were two response variables for this research: the number of pretransplant patients who dropped off the waitlist (*Total_Patients_Removed* for Research Question and Hypothesis 1) and the average MELD score among pretransplant patients ($MELD_{mean}$ for Research Question and Hypothesis 2). Detailed descriptions of the simulation variables, parameters, data organization, and validity are explained.

Variables and Parameters in this Research

The MELD and delta-MELD were the primary simulation parameters of interest. The MELD values were used to compute the delta-MELD parameters, where

$$deltaMELD = (MELD_t - MELD_{t-1}) / (time_{t-(t-1)}) \quad (12)$$

(Young et al., 2006). Regarding the use of delta-MELD, the control factor for the simulation experiment was indicated by *DM* for indicating whether or not to reference the simulation parameter delta-MELD into the AHP decision-making. Therefore, *DM* was a categorical variable which assumed one of two values (with delta-MELD, without delta-MELD). The simulation ran two scenarios: one with and one without delta-MELD.

The simulation model ensured that all patients' delta-MELD scores can be uniformly compared from an evenly distributed time series. In order to obtain an evenly distributed set of MELD values for consistent and unbiased computation of delta-MELD values, the Kalman estimation was performed. Patient MELD scores were referenced to support the Kalman estimation for additional MELD values. The MELD score is a parameter supplied by UNOS that is calculated using serum creatinine, serum total bilirubin, and the INR according to the following formula as is currently used by the UNOS organization.

$$MELD = 9.57 * log_e\ creantinine\ (mg/dL) + 3.78 * log_e\ bilirubin\ (mg/dL) + 11.20 * log_e\ INR\ (mg/dL)\ + 6.43 \quad (13)$$

The MELD data through Kalman estimation for delta-MELD values were derived data from the UNOS STAR database. This database was requested from the OPTN organization. The UNOS STAR database, provided by UNOS, contained waitlist, patient, donor liver, and posttransplantation information on all recipients undergoing liver transplantation in the United States since 1987 (Northup & Berg, 2004).

The delta-MELD values were computed in order to be used as a primary patient selection criterion for liver allocation when the indicator *DM* indicated that the scenario with delta-MELD is to be run. Patient selection was performed in the simulation by the AHP processing.

Other simulation parameters of this research included DRI and SOFT parameters. In addition to the two important and primary parameters used for donor liver allocation,

MELD and delta-MELD that were used to fulfill the medical urgency objective, the parameters DRI and SOFT were used to meet the utility and survivability objectives. These simulation parameters, DRI and SOFT, were referenced in the patient selection decision-making. Within the two scenarios with and without delta-MELD, the same donor liver data were referenced from the UNOS STAR database. The DRI and SOFT values were not manipulated or varied between the scenarios with and without delta-MELD. In addition, when patient data were retrieved from the STAR database and were referenced within a 180-day simulation interval, they were the same patient data used in both scenarios with and without delta-MELD. The computation of the DRI parameter is expressed using equation (11). The computation of the SOFT parameter is as follows.

$$SOFT = MELD + Donor\_Risk\_Factors + Recipient\_Risk\_Factors \quad (14)$$

Donor_Risk_Factors and *Recipient_Risk_Factors* are listed in Table 2 and Table 3.

Table 2

Donor Risk Factors

Donor Risk Factors	Risk Points
Age > 70	3
COD (anoxia, trauma)	2
Creatinine > 1.5	2
National Procurement	2

Table 3

Recipient Risk Factors

Recipient Risk Factors	Risk Points
Age > 70	4
BMI > 35	2
Albumin < 2.0	2
Previous Abdominal Surgery	2
Dialysis pretransplant	3
ICU pretransplant	6
Hospitalized pretransplant	3
MELD 30-39	4
MELD ≥ 40	4
Life support pretransplant	9
Encephalopathy at transplant	2
Portal vein thrombosis at transplant	5
Portal bleed within 48 hours pretransplant	6
Ascities pretransplant	3

In each 180-day interval, even with different replications, the simulation used the same pool of patients against the same pool of livers. With each replication, the same scenario in each interval was used in both scenarios (two decision-making criteria). However, even though both scenarios involved identical patients, the model

was stochastic because in each replication, the livers and their arrival times were randomly generated. In addition, not only were DRI and SOFT simulation parameters referenced as decision criteria for patient selection, but the DRI and SOFT parameters were stochastic since there was no way to know in advance the liver type, liver data, and the timing that a donor liver would become available for transplant. But the DRI and SOFT values were from the same liver data in both the scenarios with and without delta-MELD. The MELD and delta-MELD parameters, as well as DRI and SOFT parameters were referenced for AHP scoring in the simulation for patient selection.

Although the parameters, DRI and SOFT, were referenced in each of the ten 180-day simulation scenarios for patient selection, they were both considered to be much lower in importance compared to the MELD and delta-MELD scores. Subsequently they were rated (weighted) consistently much lower than MELD and delta-MELD in both scenarios whether using delta-MELD or not. The trend in literature, as discussed in Chapter 2, suggested that DRI and SOFT parameters were realistic factors in patient selection although the MELD and delta-MELD were the primary factors fulfilling the sickest first objective.

In addition to the MELD, delta-MELD, DRI, and SOFT parameters for decision-making within the simulation model, the response variables of *Average_MELD* and *Patients_Dropped_From_Waitlist* were both computed and output by the simulation model. They were compiled at the end of each simulation week and would subsequently be summed or averaged within a 180-day interval to compute the response variables that determined whether the research hypotheses would be accepted or rejected.

The *Patients_Dropped_From_Waitlist* parameter was processed from the simulation's Disease Progression and Waitlist Patient Management processes. *Patients_Dropped_From_Waitlist* was computed by adding all the patients who had dropped off from the waitlist at the end of each week based on patient data from the UNOS STAR database. The removal of patients from the waitlist was based on patients' MELD score, hazard ratios based on MELD scores, and patients who were deemed too sick to transplant.

The average MELD from the simulation model (dependent variable) was computed over ten 180-day intervals, producing ten MELD_{mean} values. For each interval,

$$\text{MELD}_{mean} = (1/26) * \sum_{n=1\ to\ 26\ weeks} Average\_MELD(n) \quad (15)$$

as there are 26 weeks in one 180-days interval. The total number of patients dropping off of the waitlist from the simulation model (dependent variable) was also computed over ten 180-day intervals, producing ten *Total_Patients_Removed* values. For each interval,

$$Total\_Patients\_Removed = \sum_{n=1\ to\ 26\ weeks} Patients\_Dropped\_From\_Waitlist(n) \quad (16)$$

For both the response variables, MELD_{mean} and *Total_Patients_Removed*, the control factor, *DM,* was postulated to be influential. This was because in addition to the parameters of MELD, DRI, and SOFT, the simulation model with delta-MELD as a decision criterion could affect the final decisions of winning patients differently from the scenario without delta-MELD. The response variable MELD_{mean} relied on *Average_MELD(n=1…26)* compiled over weeks on the waitlist, and like the response variable, *Total_Patient_Removed*, it could be affected by the choices of winning patients

which were based on MELD, delta-MELD, DRI, and SOFT. The processing of these variables and parameters are further elaborated in the sections, *Simulation Outputs*, *Experiment and Sample Size,* and *Hypothesis Testing*.

Data Collection through Secondary Data

The main source of data collection was through secondary archived data, which was received upon request from the OPTN and the UNOS organizations. I filled out an agreement form indicating that this data was not used to pursue contact of any individual patient. There was a programming fee of $200 for a university researcher not associated with a liver transplant hospital to use the UNOS STAR database requested from OPTN.

OPTN/UNOS STAR database was received by postal mail which consisted of patient data and donor liver data from the OPTN/UNOS organizations. Patient data included age group, gender, race, primary cause of disease, transplant history, blood type, MELD scores, date of MELD scores, time on wait list, and status. Donor liver data included donor age, donor height, donation after cardiac death donors, split liver donors, race, donor's cause of death from cerebrovascular accident, regional sharing, local sharing, and cold ischemia time.

In a study by Halldorson et al. (2009), the UNOS STAR national transplant database was referenced to analyze survival for first-time liver transplant recipients with chronic liver failure. In this study, following approval by the University of Washington Institutional Review Board, Halldorson et al. extracted all records of recipients transplanted for chronic liver disease from UNOS Standard Transplant Analysis and Research (STAR) files from January 1, 2003 through December 31, 2006. The data were

referenced to compute the D-MELD parameter. This required the referenced donor liver data in order to compile the DRI parameter. The DRI parameter consisted of a computation based on nine risk factors and it was referenced to compute D-MELD for the study of Halldorson et al. Similar to Halldorson et al., the simulation model referenced patient data and donor liver data from the UNOS STAR database.

I utilized OPTN data in the four processes of the research simulation. Donor liver data contained the DRI composition which, according to Feng et al. (2006), included nine parameters, were needed to compute the DRI parameter. Patient data were needed to determine compatibility and urgency considerations. Data from the five years of 2008-2012 were processed by the simulation model. The DRI parameter is expressed using equation (11).

Rana et al. (2008) developed the SOFT score by combining patients' MELD and risks scores. The risks scores come from both donor and recipient risk factors and their risks points are summarized in Table 2 and Table 3. Finally, the MELD, delta-MELD, DRI, and SOFT scores were normalized prior to being referenced in the AHP algorithm.

Data Organization for the Simulation

Foxton et al. (2010) explained that liver allocation for transplantation worldwide has undergone dramatic changes within the last 5 years, particularly with the introduction of the MELD. Liver allocation policy changed because of increasing demand for liver transplants and increasing waitlist mortality. Hence, sample data of years 2008-2012 were used as these years are recent. There are 11 Organ Procurement Organization (OPO) regions in the United States. The data requested were from the one of the most

populous yet confined OPO regions, Region 9, of the United States. Region 9 includes New York state and western Vermont.

Burr and Shah (2010) explained that the United States has 58 OPOs and 11 UNOS regions. Within the regions are multiple OPOs. Furthermore, the OPO centralizes the patients from the waiting lists of all centers within its coverage area and assigns them priority based on the MELD score, so that available organs will first be allocated to patients in descending MELD order within each specific OPO. Organ allocation is prioritized as local (within the OPO), then regional (within a UNOS region) and finally, national (p. 134).

Burr and Shah (2010) further explained that the purpose of this allocation is to reduce cold ischemia time by shipping the organs within a confined region. Reducing cold ischemia time improves the quality of the transplanted organ. Hence, the simulation model limited the study to one region, Region 9, and limited the scope outside the need to concern with cold ischemia time and varying MELD averages of additional regions, by focusing solely on one region, Region 9, which is confined to the area of New York state and western Vermont.

The patient and donor liver data from Region 9 were used for simulation input. However, the patient data were taken from the UNOS STAR database to enter in the waiting list according to the time interval according to their entry timeframe, but the donor liver was randomly selected from the UNOS STAR database when simulating the arrival of a donor liver. The timing of the arrival of donor liver was determined by a Poisson process. This way, the content of a patient and a donor liver were from real

patient and donor liver data, while the timing and order of available livers were not based on any actual occurrences. But the data of patients and the data of donor livers were from actual data.

MELD projections based on timed intervals. The patient data and donor liver data received from OPTN/UNOS were organized to accommodate an event driven simulation. The patient data were organized into one of the 10 simulation interval based on UNOS' date of entry. The arrival of an available liver was randomly selected from the pool of donor livers at the UNOS-based average rate determined by a Poisson process. This meant that the timing of the actual matching of liver to patient was not replicated according to the archived data, but by the arrival of livers processed by the simulation model. Donor livers were randomly selected from the list of donor livers which were from UNOS STAR of the 2008-2012 timeframe. The intent was to simulate the same sequence of events in both scenarios, without and with delta-MELD, with the purpose of comparing their outcomes.

Kalman estimation of MELD values. An additional method of gathering data was by the computation of additional and estimated MELD values through the Kalman algorithm. It was through this form of data collection that the simulation was able to process sufficient MELD values for a steady flow of available and consistently computed delta-MELD values. The delta-MELD values were used for decision-making in the simulation model.

The Kalman estimation was used to estimate a steady supply of time series MELD values for the consistent computation of delta-MELD values. The OPTN

stratification of MELD scores were applied in the simulation model as described in the *OPTN liver allocation policy* section. Steps were taken to use the Kalman estimation for additional MELD data that were computed from patient MELD scores.

Appendix A illustrates the steps and a basic example of an Extended Kalman Filter (EKF). The progression of actual disease progression is likely to be non-linear and hence, in the simulation model, the EKF was used for capturing the series of sequences that are likely to be non-linear approximations. After the time series of MELD values were computed, the time series of MELD values were referenced to compute the delta-MELD values. Then, patients were stratified according to their MELD scores in the following OPTN liver allocation policy levels for patient selection. This was done in the AHP processing of the simulation. Appendix B illustrated an example of an AHP processing.

OPTN liver allocation policy. In each OPO, the purpose of allocating livers is to enable physicians to apply their consensus medical judgment for the benefit of liver transplant candidates as a group. Each candidate is assigned a status or probability of candidate death that has been derived from their MELD score reflecting the degree of their medical urgency. The MELD score is the patients' mortality risk scores determined by prognostic factors. Candidates are then stratified by the MELD score and by blood type similarity (Organ Procurement and Transplantation Network, 2014).

Regarding the hierarchy of decision-making, donor livers are offered to candidates first with an assigned status of 1A and 1B (highest priority) in descending point sequence with the candidate having the highest number of points receiving the

highest priority before being offered for candidates of lower probability rankings. At each hierarchical level, adult livers are allocated in descending sequence order most urgent to least urgent in the OPO hierarchy groups as follows.

1. Status 1A / Status 1B candidates in descending point order (local and regional).
2. Candidates with MELD scores ≥35 in descending order of mortality risk MELD scores with local candidates ranked above regional candidates at each level of MELD score (local and regional).
3. Candidates with MELD scores 29-34 in descending order of mortality risk MELD scores with local candidates ranked above regional candidates at each level of MELD score (local).
4. Liver-intestine candidates in descending order of Status and mortality risk MELD scores (national).
5. Candidates with MELD Scores 15-28 in descending order of mortality risk MELD scores (local).
6. Candidates with MELD Scores 15-34 in descending order of mortality risk MELD scores (regional).
7. Candidates with MELD Scores < 15 in descending order of mortality risk MELD scores (local first, then regional).

These categorical levels were reflected in the AHP scoring algorithm which ensured that the OPTN categories were adhered to. In other words, the AHP algorithm ensured that a candidate in category 3 (with a MELD score of 29-34) would never supersede a candidate of the same blood type in category 2 (with a MELD score ≥ 35).

Figure 1 is an illustration of hierarchy which the AHP scoring would adhere to. Within a hierarchical level, the AHP scoring is ordered according to medical urgency (sickest first) and blood type, with the consideration of the donor liver quality and survivability factors.

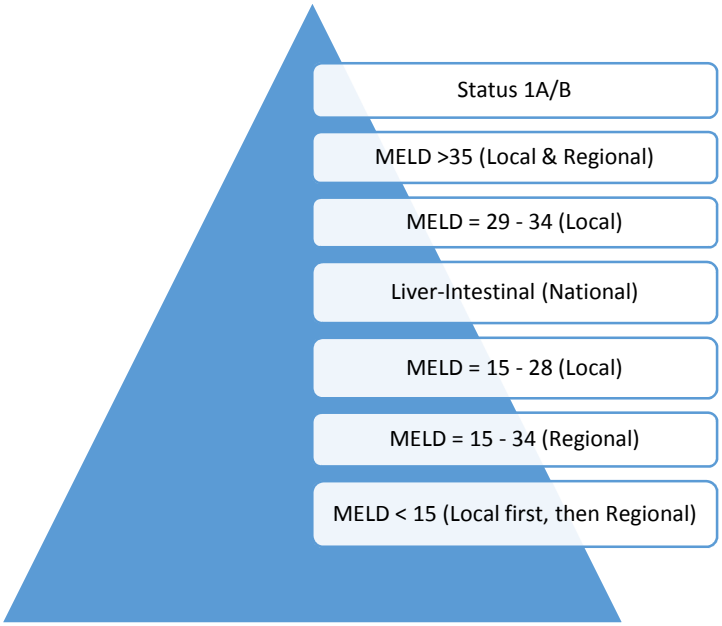


Figure 1. Hierarchy of MELD levels for liver allocation.

The priority of the decision-making is to consider patients at the highest MELD level first. When there are no candidates with a valid match at an existing priority level, then patients of the next lower priority level are considered.

Internal and External Validity of the Simulation Components

The internal validity of the simulation method lies in using secondary data for the simulation model to ensure sampling validity, and in the systems perspective of the simulation to simulate the actual liver allocation system. In addition, the simulation reflected the objectives and processes of the actual donor liver allocation system.

According to Frankfort-Nachmias and Nachmias (2008), one kind of validity that is primarily related to the instrument is the sampling validity. In the *Simulation validity* section, the sampling validity is described. The construct validity of the simulation is described in the sections, *Analytic hierarchy process validity* and *Kalman estimation validity* as it is the premise of these theoretical frameworks that the simulation was built upon.

Simulation validity. Frankfort-Nachmias and Nachmias (2008) defined research method validity as three basic types of validity, content validity, empirical validity, and construct validity. These validities relate to a specific type of evidence and conditions (p. 149). The simulation validity was established by relating the measuring instrument of the data collected which really was the actual data to the general theoretical framework. Secondary data were requested and OPTN allocation levels were implemented to reflect the actual MELD scoring and decision-making of the existing U.S. liver allocation system. Secondary data were also used to extend the MELD time series with additional MELD values.

Regarding sampling reliability, OPTN data are open for researchers to use for research and the OPTN organization is a reliable resource for the research of patient, donor liver, and waitlist history related to organ transplant. According to Northup and Berg (2004), the data requested from the OPTN/UNOS organizations is from the most comprehensive liver transplant reference database presently in existence in the United States (p. 1648).

The data pertaining to this research were based on the recent 5 years of 2008-2012. The simulation time duration was 180 days for each of the 10 intervals where there are approximately two 180 day intervals per year. Requesting for multiple years of data ensured that there was enough data to carry out the analysis for a statistical conclusion and ensured that the sampling sizes were adequate.

Regarding content validity of the simulation program, not only were patient and donor liver data based on actual data, but the OPTN liver allocation levels were integrated into the AHP decision-making. Also, the percentage of ECD acceptance was taken from peer-reviewed literature and it was accounted for in the simulation model. The objectives of urgency, utility, and survivability, by the measurements of MELD, DRI, and SOFT normalized scores, were integrated into the AHP algorithm.

Analytic hierarchy processing validity. Regarding the use of AHP as a valid tool for multi-criteria decision-making, Saaty (1996) explained that it is of high importance to recognize measurements of various kinds of scales and in particular, the ratio scales. An ordinal scale is a set of numbers that is invariant under monotone transformations. In other words, ordinal numbers can neither be multiplied nor added meaningfully. An interval scale is a set of numbers that is invariant under linear transformations, specifically, of the form, $ax+b$, where $a>0$, $b\neq0$. Different interval scales cannot be multiplied. However, numbers from the same scale can be added. A ratio scale, on the other hand, is a set of positive numbers that is invariant under a positive similarity transformation of the form, ax, where $a>0$. Different ratios scales can be multiplied and divided and still give rise to a ratio scale because the invariance of their

products and quotients is derived from the invariance of each one of these scales. Numbers from the same scale can also be added. Ratio scales enable us to relate alternatives of tangible action to criterion and values that are intangible. The ratio scale of the AHP decision-making scores enabled us to see that our preferences were measured and compared among different measuring units among all the patients that were subjected to the same criteria. Finally, the consistency ratio, ensured the transitivity property and numerical proportions of the AHP decisions are consistent by ensuring it does not exceed 0.1.

The simulation selected the patient with the highest AHP score of matching blood type. If the donor liver was an ECD or ECD 1-year liver, and the patient has not given consent and does not wish to proceed to transplant unless the donor liver is a SCD liver, then the next highest AHP scoring patient was selected. The structure of AHP decision-making is shown in Figure 2.

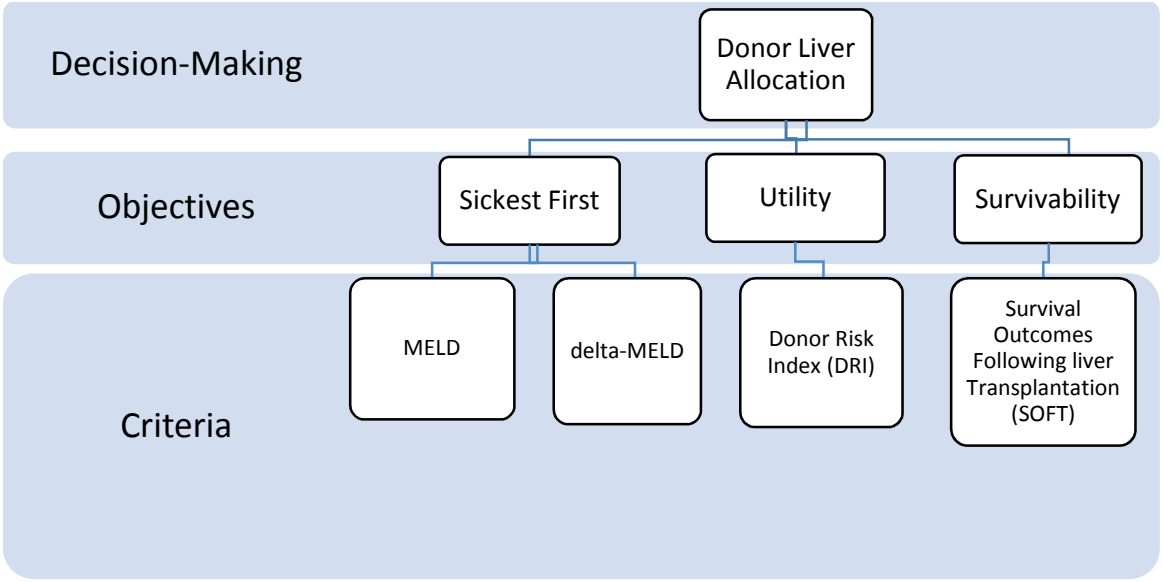


Figure 2. Analytic hierarchy process structure of objectives and criteria.

The liver allocation decision-making can be translated in its limited range into meaningful numbers reflecting criteria ranking and patient selection. The liver allocation selection process via AHP included criteria that were uniformly used among all pretransplant patients.

Kalman estimation validity. The Kalman filter is the theoretical framework to model liver disease progression of pretransplant patients for the simulation. Steps were taken to translate the Kalman filter theory into the simulation model by using patient data from the OPTN organization and by estimating additional MELD data consistent with actual MELD data. The main task of Kalman filter was to develop the system model where the goal is to determine the matrices reflecting the systems dynamics of disease progression, covariance values, and observation matrices of the observed disease progression.

The Kalman filter algorithm for the MELD and delta-MELD parameters consisted of two alternating steps, which were repeated for each iteration and each new frame, prediction, and correction. In the prediction step, the filter made an assumption about the future state. In the correction step, an optimized state estimate was computed using a weighted difference between the prediction state and an actual or averaged measurement. Hence, this provided a mechanism to ensure optimal MELD estimations. In the simulation model, estimates of MELD scores were based on actual MELD values and these values were used to propagate MELD data upon arrival of available donor livers. An error technique was setup in order to check that Kalman estimations did not exceed beyond certain thresholds.

The Simulation Model

This section describes the simulation model and its processing. The model simulated the liver allocation system by the way it handled available donor livers, the process of placing patients into the waitlist, and patient disease progression while waiting for available livers as reflected by the U.S. liver allocation system.

The model simulated four processes and outputs relevant values for analysis. The four processes included a process for the arrival of donor livers, transplant patient entry into the waitlist, liver disease progression, and waitlist patient management. These processes referred to patient and donor liver data.

A stochastic component of the simulation experiments was the interarrival timing of available donor livers. The available donor livers were randomly selected from the UNOS STAR database, and hence it was not in accordance to the timing as specified by the database. The interarrival timing of the liver was set accordingly to a Poisson process and based on the mean interarrival time from actual data. Also, the random selection of available liver presented an uncertainty regarding the type of available liver, whether it is SCD, ECD-1 year, or ECD donor liver.

Each simulation run was conducted over ten 180-day intervals of simulation time. There were two scenarios for the simulation which defined the values for the control variables, one not using delta-MELD and one using delta-MELD as a primary criterion for patient selection. The sequence of the donor liver arrival was the same for both scenarios, with or without using delta-MELD.

Waitlist Entry

The management of waitlist entry included processing new patients waiting for an available liver for transplantation. These patients have undergone a liver medical assessment and have an initial MELD score with a start date into the waitlist. The Waitlist Entry process referenced the patient data from OPTN/UNOS. Once a patient had entered into the waitlist, the Disease Progression process would approximate future patients' MELD scores in regular time intervals by Kalman estimation. The Waitlist Entry's data included patient's waitlist start date, patient initial MELD score, and the number of patients entering into the waitlist.

Burr and Shah (2010) explained that the way the current allocation system worked is that patients are prioritized on the waitlist according to blood type by descending MELD order. This would mean that organs are offered to the waitlisted patient with the highest MELD score and blood type identical to the patient. To avoid an inequitable distribution of organs, blood type O livers are only assigned to blood type O patients. The system allows patients with special situations such as very small size adult patients or AB-type patients to be listed for more than one match of blood types.

Donor Liver Arrival

The events of donor liver arrival were simulated by a Poisson process and the data were randomly selected to be one of the donor livers from the UNOS STAR database. The Donor Liver Arrival process selected matching patients from the waitlist according to patient medical urgency, donor liver's DRI, and patient-donor SOFT scores. Hence, the Donor Liver Arrival process computed the AHP scores for all the patients on the

waitlist. The winning patient was selected as the best scoring patient for the available donor liver. However, if the available donor liver was an ECD liver, consent needs to be retrieved from the patient. The chances of the patient accepting an ECD liver was determined in this process. The Donor Liver Arrival's output data included the ECD/SCD status, DRI score, whether a patient consented to accept an ECD liver, patient MELD and AHP scores, the number of SCD and ECD livers, and the number of patients transplanted. In Figure 3, processing is shown for handling ECD and SCD livers.

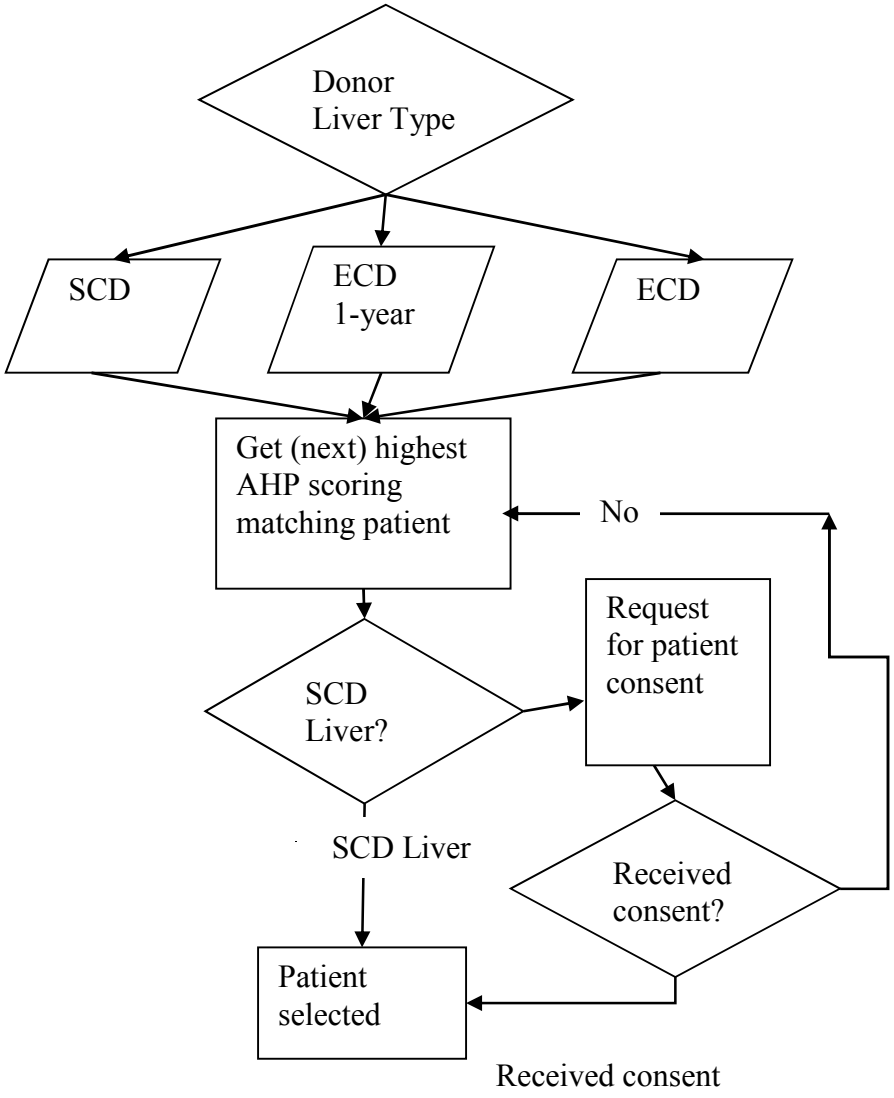


Figure 3. Donor liver processing based on liver quality.

Upon arrival of an ECD or ECD 1-year liver, the model determined whether a patient accepted an ECD or ECD 1-year liver, by a random function, with a chance of 25% or 15%, respectively (Rodrique et al., 2011).

Disease Progression

The Disease Progression process estimated patients' disease progression by updating MELD and delta-MELD parameters. The patient's waitlist status was updated indicating whether the patient was still on the waitlist awaiting for an available donor liver, or has dropped off from the waitlist due to being too sick, death, or for other reasons. The Disease Progression's output data included patient MELD, delta-MELD, and waitlist status. In addition, the output data provided the average MELD score of patients on the waitlist and the count of patients dropping off from the waitlist at the end of each week.

Unlike the Waitlist Entry, Donor Liver Arrival, and Waitlist Patient Management processes, which were event-driven, the Disease Progression process was a timer process that runs continuously at a one second rate. Every instance when the Disease Progression process runs simulated one day on the waitlist. The timer process ran for 180-days to fulfill one of ten intervals in a simulation scenario.

In order to simulate and track the occurrence of patients being removed from the waitlist because the patient was deemed too sick to transplant, the *Patients_Dropped_From_Waitlist* parameter was updated at the end of each week. The survival table based on MELD score and patient risk factors according to the study of Rana et al. (2008) was referenced as a guide when the weekly average MELD exceeded the actual Region 9 average MELD of 21. The risks based on patient risk factors are tallied up and patients with the highest risks are considered for removal based on MELD

scores and survival rates. The survival rates were based on the hazard ratio according to MELD scores as follows (Sharma, Schaubel, Gong, Guidinger, & Merion, 2012).

Table 4

Hazard Ratios based on MELD

MELD	15-17	18-20	21-23	24-26	27-29	30-32	33-35	36-37	38-39	40
HR	0.03	0.04	0.08	0.12	0.22	0.39	0.5	0.82	0.98	1
P-value	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.1	0.89	0.87

Waitlist Patient Management

The Waitlist Patient Management process managed patient disease updates due to actual MELD scores and removed patients off of the waitlist reported as being too sick to receive transplant or died while waiting for available liver. These updates were based on actual patient reports provided by the OPTN data. The waitlist patient output data included patients' MELD scores, statuses of removal from the waitlist, as well as the count of patients dropping off of the waitlist.

Simulation Inputs and Processing

Input Data

The simulation model input data included a control factor, assumptions, and limitations. From the UNOS STAR database, and similar to the study by Northup and Berg (2004), recipients listed for liver transplantation with MELD exclusions such as hepatocellular carcinoma patients, all status 1 (acute hepatic failure) recipients, and patients with incomplete laboratory or survival data were excluded from the simulation.

Also, the data set was queried for MELD score on the day of transplant. Patients with only a single MELD score reported to UNOS or with no MELD scores were excluded from the analysis. The parameters of DRI and SOFT scores were computed upon the arrival of a donor liver.

The uncertainties of the process, simulated by the model, included the arrival of available livers and the type of quality of these donor livers, whether they were SCDs, 1-year ECDs, or ECDs. The interarrival times of liver donor were varied according to a Poisson distribution. The liver and patient data and their arrival times were the same in both scenarios with and without delta-MELD. The objective was to measure and compare the scenarios with and without delta-MELD to the same livers, pool of patients, and their arrival times. Finally, since there were two 180-day intervals in a year, and data were based on 2008-2012 timeframe, there were 10 different model runs. The entire experiment ran 7 replications of these model runs, to meet the minimum required sample size, which was 70, for a *t* test of two independent populations of means. There was an equal number of replications for each of the two groups (one group for each level of the independent variable).

Input Data for Simulation without delta-MELD

In the simulation scenario without delta-MELD, the AHP integrated the OPTN allocation levels of urgency and played a major part in patient selection. The following AHP decision tables were initialized for decision-making where the delta-MELD was not used as a criterion.

The patient AHP scores were based on the AHP initial values, donor liver DRI, patient-donor SOFT, and MELD normalized scores. The AHP score was computed as follows, where the coefficients and multiplier are according to Table 5.

$$AHP_{(DM=0)} = Multiplier * (a_1 * MELD_{norm} + a_2 * DRI_{norm} + a_3 * SOFT_{norm}) \quad (17)$$

Table 5

AHP Weights and Ranking without delta-MELD

Allocation Level	MELD a_1	DRI a_2	SOFT a_3	Multiplier
Status 1A/B	0.570305	0.214847	0.062941	0.320454
≥35 (Loc & Reg)	0.457376	0.271311	0.214847	0.216472
29-34 (Local)	0.347082	0.347082	0.326458	0.149638
Liver-Int (National)	0.257338	0.257338	0.371330	0.109904
15-28 (Local)	0.177135	0.177135	0.411432	0.081728
15-34 (Regional)	0.122618	0.122618	0.438690	0.065517
<15 (Local, Reg)	0.882129	0.088212	0.455893	0.056284

Input Data for Simulation with delta-MELD

In the simulation scenario with delta-MELD, the AHP integrated the OPTN liver allocation levels of urgency that also played a major part in patient selection while using the delta-MELD parameter as a criterion for liver allocation. The following AHP Table 6 contains the weights of the criteria (coefficients), MELD, delta-MELD, DRI, and SOFT, and the ranking (multiplier) of priority groups for liver allocation used in AHP scoring.

The patient AHP scores were based on the AHP initial values, delta-MELD, donor liver DRI, patient-donor SOFT, and MELD normalized scores. The AHP score was computed as follows, where the coefficients and multiplier are according to Table 6.

$$deltaMELD = (MELD_t - MELD_{t-1}) \,/\, (time_{t-(t-1)}) \quad (18)$$

$$AHP_{(DM=1)} = Multiplier * (b_1 * MELD_{norm} + b_2 * deltaMELD_{norm} + b_3 * DRI_{norm} + b_4 * SOFT_{norm}) \quad (19)$$

Table 6

AHP Weights and Ranking with delta-MELD

Allocation Level	MELD b_1	delta-MELD b_2	DRI b_3	SOFT b_4	Multiplier
Status 1A/B	0.363181	0.363181	0.138186	0.136818	0.352108
≥35 (Loc & Reg)	0.313835	0.313835	0.186164	0.186164	0.229594
29-34 (Local)	0.257655	0.257655	0.242344	0.242344	0.150847
Liv-Int (National)	0.204669	0.204669	0.295330	0.295330	0.104031
15-28 (Local)	0.150480	0.150480	0.349519	0.349519	0.070833
15-34 (Regional)	0.109225	0.109225	0.390774	0.390774	0.051732
<15 (Local, Reg)	0.081062	0.081062	0.418937	0.418937	0.040853

The AHP technique for the selection of the most suited and sickest patient was appropriate for the liver allocation decision-making because there were multiple objectives and criteria to weigh into the consideration of many patients awaiting a donor liver. Winston (2004) described that a multi-criteria decision-making process could be complex because when multiple objectives are important to a decision-maker, it may be difficult to choose among the many alternatives (patients). The AHP was a tool used for

the selection of matching patient as it integrated the objectives with the OPTN liver allocation levels into the decision-making. AHP scores were outputted for analysis from the donor liver process of the simulation. In addition to AHP, number of patients dropping out of the waitlist, and average MELD scores were outputted for analysis. In the Figure 4, an overview of the simulation's data processing is presented.

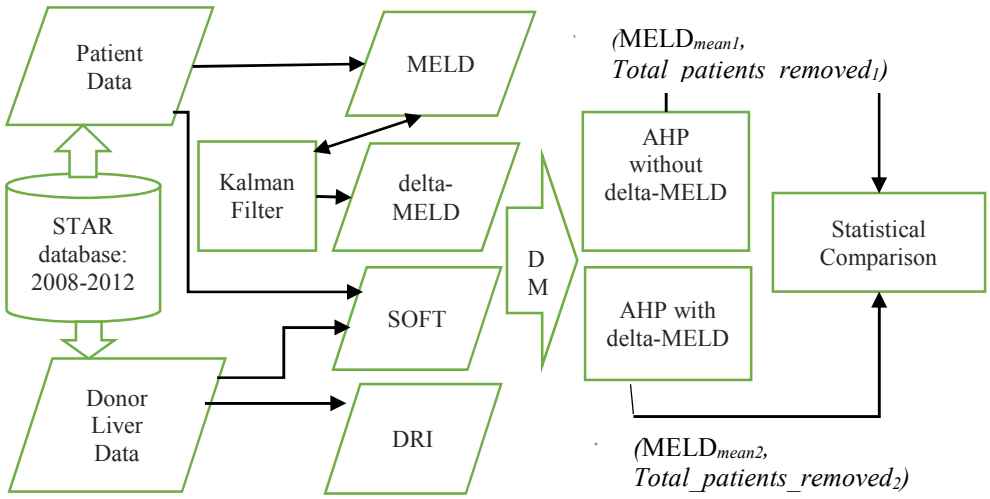


Figure 4. Overview of simulation data processing.

Table 7 lists the actual number of liver transplants performed from deceased donors in the years 2008-2012 according to OPTN (2014). This list indicates that when considering the limitation of actual patient data used in the simulation due to non-HCC disease, non-status 1 patients, etc., the simulation processing a unique pool of 100 patient records and 130 donor liver records within any 180-day simulation interval is feasible.

Table 7

Number of Region 9 Liver Transplants from Deceased Donors

Year	Number of Region 9 Liver Transplant Patients from Deceased Donors
2012	313
2011	343
2010	322
2009	403
2008	402

Table 7 shows that the patient sampling within the simulation of 100 patients in each of the ten 180-day intervals is a feasible sampling size, where the actual patients from the time spanning 2008-2012 from the UNOS Region 9 database was *1,783.* The simulation sampled 100 patients for each 180-day interval, where there were two 180-day intervals per year, producing 1,000 patients in ten intervals. The simulation patient sampling of 1,000 patients compared to the actual population of 1,783 UNOS Region 9 patients represent more than 50% of the actual patient population. This showed that in each 180-day interval, it was feasible to set up a unique pool of 100 patients and at least 130 donor livers, reflective of actual data from UNOS, for each interval to produce a set of response variables. Each simulation of 26 weekly reports per 180-day interval was processed after removing patients with HCC disease, status 1 patients, and patient data with only one MELD score.

Simulation Outputs

To summarize the simulation output data, each simulation run generated output data for analysis for each of the two response variables: $MELD_{mean}$ and *Total_Patients_Removed*. But most importantly, the disease progression output data included patient's MELD, delta-MELD, as well as the average MELD score of patients on the waitlist. The simulation output data included the number of patients dropping off of the waitlist at the end of each week. The average MELD and number of patients removed from the waitlist from the scenario without delta-MELD was compared against the average MELD and number of patients removed from the waitlist from the scenario with delta-MELD.

The response variables from model runs were used to compute $MELD_{mean}$, equation (15), and *Total_Patients_Removed*, equation (16), for each 180-day interval. Both *Average_MELD(n)* and *Patients_Dropped_From_Waitlist(n)* parameters were output from both scenarios with and without delta-MELD. The only control factor of the simulation experiment was the *DM* that indicated the use or non-use of the delta-MELD as a decision criterion in the simulation.

Experiment and Sample Size

The experiment covered 5 years of data (spanning 2008-2012), measured in 180-day intervals for a total of 10 timeframes. For each of these timeframes, the simulation model ran twice, once for each level of the control factor (once with and once without delta-MELD). Thus, there were 20 simulation scenario runs (10 timeframes times 2 runs, with and without delta-MELD). However, the simulation model was stochastic, so I

conducted the experiment to run the simulation model multiple times (7 replications) for each timespan and control factor level. This ensured I met the minimum sample size for *t* tests of two independent populations' means to meet the desired power, effect size, and confidence.

I computed the minimum experimental sample size using the following methodology. In determining the minimum experimental sample size, Aczel and Sounderpandian (2008) provided the following formula for this purpose (p. 243):

Minimum experimental sample size for a *t* test of two independent and equal populations is as follows:

$$n = 2 * \left(Z_{\alpha/2} + Z_{\beta}\right)^2 * \sigma^2/E^2 \quad (20)$$

$Z_{\alpha/2} =$ the normal distribution critical value for a probability of $\alpha/2$ in each tail,

$Z_{\beta} =$ the normal distribution critical value for a probability of β,

$\sigma^2 =$ population variance, and

$E = d\sigma$, the standard error.

I used a 95% confidence interval, hence, $Z_{\alpha/2}$ is 1.96 (Aczel & Sounderpandian, 2008), and 80% power, hence, Z_{β} is 0.84 (Gelman & Hill, 2007, p. 441). I ensured a power of at least 0.8, which is $1 - \beta$, where β is the percentage of Type II error, and Type II error is the error of not rejecting the null hypothesis when it is false.

For the dependent variables, $MELD_{mean}$ and *Total_Patients_Removed*, I used Cohen's *d* effect size, where a large effect size is 0.80, a medium effect size is 0.50, and a small effect size is 0.20 (Cohen, 1992). The standard error, *E* is rewritten as $d\sigma$, as $E = d\sigma$, where I chose a medium effect size, $d = 0.5$. Cohen (1992) explained that the effect

size is that measure which can determine whether the null hypothesis may likely be wrong. For $MELD_{mean}$ and *Total_Patients_Removed,* I chose a medium effect size since a MELD score that is off by 1 is a reasonable and noticeable effect size. $E = d\sigma$, where $\sigma = 2$ (Figure 11), $d = 0.5$ times ($\sigma = 2$) is 1, a medium effect size. Similarly, for *Total_Patients_Removed*, one person removed from the waiting list is a reasonable and noticeable effect size.

Using an effect size, *0.5*, the formula for sample size is reduced to the following:

$$n = 2 * \left(Z_{\alpha/2} + Z_{\beta}\right)^2 * \sigma^2/E^2 \quad (21)$$

$$= 2 * \left(Z_{\alpha/2} + Z_{\beta}\right)^2 * \sigma^2/d^2\sigma^2 \quad (22)$$

$$= 2 * \left(Z_{\alpha/2} + Z_{\beta}\right)^2/d^2 \quad (23)$$

With $d = 0.5$, $\alpha = 0.05$, $\beta = 0.80$, the sample size is

$$n = 2 * \left(Z_{\alpha/2} + Z_{\beta}\right)^2 * \sigma^2/E^2 = 2 * (1.96 + 0.84)^2/(0.5)^2 = 62.72 \quad (24)$$

n is rounded up to 70. The sample size of 70 required 7 simulation runs as there are ten 180-days intervals (sample units) per simulation (replication). The normality, homogeneity of variances, and independence of the *t* tests of two independent populations were verified (see Appendix D).

Hypothesis Testing

I conducted statistical tests based on output data from the experiment. *Total_Patients_Removed* and $MELD_{mean}$ values were hypothesized to vary with changes in the control factor: without and with delta-MELD. I utilized *t* tests for two independent populations, and the experiment was based on the random sampling of arriving livers

with an UNOS-based average random interarrival time of five days. The experiment involved two scenarios, one scenario with decision-making not including delta-MELD, and another scenario with decision-making including the delta-MELD (hence, generating samples of two different and independent populations). Upon completion of the simulation experiment, two *t* tests of two independent populations were performed (one for each hypothesis) to evaluate the difference in means between the case where delta-MELD is not used as a criterion and the case where delta-MELD is used as a criterion.

Aczel and Sounderpandian (2008, p. 313) provided the formula for the *t* test of two independent populations for the case where σ_1 and σ_2 may be unknown and may be unequal:

$$t = \frac{(x_1 - x_2) - (\mu_1 - \mu_2)}{\sqrt{S_1^2/N_1 + S_2^2/N_2}} \tag{25}$$

Since the null hypotheses for both dependent variables stated that there is no difference in their values for both scenarios without and with delta-MELD as criterion, $(\mu_1 - \mu_2)$ is equal to 0. Hence,

$$t = \frac{(x_1 - x_2)}{\sqrt{S_1^2/N_1 + S_2^2/N_2}} \tag{26}$$

(Aczel & Sounderpandian, 2008, p. 314). The formula for the *t* test of two independent populations for $MELD_{mean}$ of the scenarios with and without delta-MELD is as follows.

$$t_{(MELDmean)} = \frac{(x_1 - x_2)}{\sqrt{S_1^2/N_1 + S_2^2/N_2}} \tag{27}$$

$$MELD_{mean} = (1/70) * \sum_{n=1\ to\ 70} [MELD_{mean}(n)], \quad (28)$$

where x_1 is the $MELD_{mean}$ average from scenarios without delta-MELD,

x_2 is the $MELD_{mean}$ average from scenarios with delta-MELD,

S_1 is the standard deviation of $MELD_{mean}$ without delta-MELD,

S_2 is the standard deviation of $MELD_{mean}$ with delta-MELD,

N_1 is the number of $MELD_{mean}$ in experiment without delta-MELD,

N_2 is the number of $MELD_{mean}$ in experiment with delta-MELD.

The *t*-statistic is positive when the $MELD_{mean}$ in the scenario without delta-MELD is larger than the $MELD_{mean}$ in the scenario with delta-MELD, and negative when the $MELD_{mean}$ in the scenario with delta-MELD is larger than the $MELD_{mean}$ in the scenario without delta-MELD. A similar *t* test of two independent populations was performed for *Total_Patients_Removed* of the scenarios without delta-MELD and the scenarios with delta-MELD.

Similarly, the *t* test of two independent populations for the number of patients removed for scenarios without delta-MELD and with delta-MELD is as follows.

$$t_{(PatientsRemoved)} = \frac{(x_1 - x_2)}{\sqrt{S_1^2/N_1 + S_2^2/N_2}} \quad (29)$$

$$Total\_Patients\_Removed_{avg} = (1/70) * \sum_{i=1\ to\ 70} [Total\_Patients\_Removed(i)], \quad (30)$$

where x_1 is the *Total_Patients_Removed* mean for scenarios without delta-MELD,

x_2 is the *Total_Patients_Removed* mean for scenarios with delta-MELD,

S_1 is the standard deviation of *Total_Patients_Removed* without delta-MELD,

S_2 is the standard deviation of *Total_Patients_Removed* with delta-MELD,

N_1 is the number of *Total_Patients_Removed* for scenario without delta-MELD,

N_2 is the number of *Total_Patients_Removed* for scenario with delta-MELD.

Aczel and Sounderpandian (2008) explained that the degrees of freedom for this *t* test is computed by

$$df = \frac{[(S_1^2/N_1)+(S_2^2/N_2)]^2}{[(S_1^2/N_1)^2/(N_1-1)] + [(S_2^2/N_2)^2/(N_2-1)]} \quad (31)$$

df is then rounded down to the nearest integer. Here N_1 and N_2 are equal to 70. *df* is computed to be 69.

I utilized a two-tailed *t* test of two independent populations. The scenarios were treated as two independent samples with different means for the dependent variables that were compared using a *t* test. The *t*-statistic was compared to the critical value of *t.* The alpha level was set to 0.05. This means that five times out of a hundred, I will reject a null hypothesis when I should have failed to reject it (a false positive result, or Type I error); that is, a difference between the means is in truth due to random variability in the stochastic process (simulation model) even if no difference exists in reality (Green & Salkind, 2011). Confidence is the inverse of alpha (1 - α), indicating the confidence I have that I will avoid incorrectly seeing an effect that is not present in the population.

The power of the test was set to 0.80. The power of a test is 1 - β, where β is the probability of a false negative (Type II error)—failing to reject a null hypothesis that should have been rejected (Aczel & Sounderpandian, 2008). Power, therefore, is the

probability of properly detecting an effect, such as a true difference in population means.

At the end of running the simulation 7 times, there were 70 sets of response variables, $MELD_{mean}$ and *Total_Patients_Removed,* for the *t* tests of two independent populations. The *t*-statistic was compared to the critical value of *t* for 69 degrees of freedom, 95% confidence, and a two-tailed test. When the experimental *t*-statistic is greater than the critical value of *t* or less than the negative value of the critical *t*, the null hypothesis is rejected. The results of statistical tests provided answers to both the research questions and hypotheses.

Pilot Testing and Scenario Runs

Hillier and Lieberman (2010) explained that after the computer simulation program has been constructed and debugged, the next key step is to test whether the simulation would provide valid results for the system it is representing (p. 961). Hillier and Lieberman suggested some ways to test the simulation model which may include observing animations and logs of simulation runs as a useful way to check the validity of the simulation model. Another suggestion provided is to construct and verify a prototype simulation which is a smaller version of the simulation (p. 962).

The simulation was designed to generate logs from each of the simulation's four processes. An important purpose for this was to provide simulation verification based on a pilot dataset. A pilot dataset was used and generated to run each of the simulation processes and to review the logs for simulation verification. This verification process included review of data initialization and data setup of patient, liver, and disease progression data, as well as the processing verification of Kalman estimation of patients'

disease progression, AHP scoring for patient selection, removal of patients from the waitlist, and *t* test of two independent populations results. More specifically, verifications of the Waitlist Entry process included checking whether the number of patients and their entry sequences matched that of actual data. Verifications of the Donor Liver Arrival process included checking donor liver mean interarrival times, percentages of SCD, ECD and ECD 1-year liver types accepted by patients, and AHP selection of patients against actual data and intended processing. Verifications of the Disease Progression process included checking whether patients were dropping off of the waitlist according to actual data and intended processing.

Both the pilot and scenario datasets went through the same simulation processing and similar log review. The log review of pilot runs provided description, understanding, and verification of simulation steps and output data. The dataset for the pilot simulation consisted of the first 180-day interval of 2008 in the UNOS STAR dataset with the same patient and liver sample data in the scenario runs as in the experimental scenario runs. The log review of scenario runs provided data analysis and interpretation of output results that was based on the UNOS STAR's dataset. Descriptions for generating logs for data analysis and simulation verification are described in Appendix C, *Simulation Programming Notes*.

Summary

The research design was formulated to ensure the research questions and hypotheses can be answered. The research instrument was the simulation, where the design of the research included the simulation parameters MELD and delta-MELD as the

primary criteria. The simulation parameters MELD, delta-MELD, DRI, and SOFT were used to compute a recipient selection by AHP. This operationalization of criteria included the stratification of MELD scores into OPTN's categories of MELD levels. The design of the research also included secondary patient data and donor liver data from the UNOS STAR database of Region 9 from 2008-2012. However, in the simulation, the donor liver data were not sequenced according to its archived occurrence, but randomly, where the donor livers were randomly selected from the database of the same year. The research experiment included distributions of SCD, ECD-1 year, and ECD donor livers with varying interarrival times of available donor livers. The random selection of liver data and interarrival times were the same for both scenarios with and without delta-MELD.

The simulation used the same patient data for the two scenarios (without delta-MELD and with delta-MELD), in multiple replications, and are treated as having two independently separate sets of patients. With each experimental run, the purpose of the simulation was to generate the average MELD scores and number of patients removed from the waitlist over a 180-day interval, for cases with and without using the delta-MELD parameter as a criterion. The experiment's objective was to determine whether the delta-MELD parameter would be a viable criterion to refine the current liver allocation system. I ran the simulation 7 times for each of the 10 intervals in 2008-2012 with delta-MELD and 7 times without delta-MELD. This resulted in a total of 70 model runs for each scenario (with and without using delta-MELD as a criterion).

The research methodology described in this chapter was made possible by the theoretical frameworks of the Kalman filter, AHP, the OPTN liver allocation policy, and the patient, liver and waitlist history data provided by the OPTN/UNOS organizations. In this chapter, I explained the theoretical frameworks, policy, and use of the OPTN/UNOS data regarding how they were applied to the simulation. This information provided the conceptual framework for the *t* tests of two independent populations of response variables, *Total_Patients_Removed* and $MELD_{mean}$, where these *t* tests helped to address the research questions and hypotheses.

Chapter 4: Simulation Results and Analysis

The purpose of this research was to investigate whether using the delta-MELD criterion can improve the liver transplant patient selection process by reducing the number of patients dropping off the waitlist and lowering the average MELD score. Pidd (2004) explained that simulation model-building should ideally include four main steps. These steps include conceptual model-building, computer implementation, validation, and experimentation (p. 35). Pidd further explained that the conceptual model-building is an activity in which the analyst tries to capture the essential features of the system that is being modelled. In the research model, I emphasized the allocation aspect of the donor liver system to study the effects of having an additional criterion, delta-MELD, for patient selection. The description of the conceptual model-building and computer implementation for this simulation was described in Chapter 3. In Chapter 4, I describe the validation and experimentation results.

Chapter 4 includes four sections, *Research Questions and Hypotheses*, *Pilot Testing and Verification of the Simulation Model*, *Experimental Outcome*, *Results*, and *Summary*. I utilized *t* tests of two independent populations to determine if there was a difference in outcomes of the two scenarios, with and without using delta-MELD for decision-making.

Research Questions and Hypotheses

The research questions and their respective hypotheses are repeated here from Chapter 1:

1. Does a simulation model using the additional parameter of delta-MELD as a patient selection criterion reduce the number of pretransplant patients who dropped off the waiting list?

H_o: There is no difference in the number of patients who dropped off of the waiting list (*Total_Patients_Removed*) between simulation models with and without delta-MELD along with the MELD score as primary criteria for patient selection in donor liver allocation.

H_a: There is a difference in the number of patients who dropped off of the waiting list (*Total_Patients_Removed*) between simulation models with and without delta-MELD along with the MELD score as primary criteria for patient selection in donor liver allocation.

2. Does a simulation model using the additional parameter of delta-MELD as a patient selection criterion lower the average MELD score among pretransplant patients?

H_o: There is no difference in the average MELD score ($MELD_{mean}$) among pretransplant patients between simulation models with and without delta-MELD along with the MELD score as primary criteria for patient selection in donor liver allocation.

H_a: There is a difference in the average MELD score ($MELD_{mean}$) among pretransplant patients between simulation models with and without delta-MELD along with the MELD score as primary criteria for patient selection in donor liver allocation.

Pilot Testing and Verification of the Simulation Model

As introduced and explained in Chapter 3, the model simulated processes for Waitlist Entry, Waitlist Patient Management, Disease Progression, and Donor Liver Arrival. The Waitlist Entry process handled the days which patients enter into the waitlist within the simulation interval. The Waitlist Patient Management process ensured that MELD updates were processed for all patients according to the UNOS data. The Disease Progression process handled the Kalman estimation for all patients' MELD scores according to existing UNOS data on days where there were no patient updates. The Donor Liver Arrival handled the arrival of donor livers according to computer randomly generated days by a Poisson process. In addition, pilot testing and verification of the simulation model included weekly reports of the 180-day interval and statistics from one simulation run which were output for review.

Table 8 highlights the data derived by the simulation as well as data provided by the UNOS STAR database for simulation processing. Table 8 identifies whether the data were derived, stochastically generated, or retrieved from the UNOS STAR database. The simulation data, which were initially described in Chapter 3, are summarized in Table 8.

Table 8

Simulation Variables and Parameters

Variable / Parameter	Description
MELD	Model for end-stage liver disease retrieved from UNOS STAR database and derived for patient on waitlist.
Delta-MELD	This derived parameter is the calculated difference between a current MELD and the previous MELD score divided by the days between the two MELD scores.
DM	This control parameter is a categorical variable where 1 = scenario with delta-MELD as criterion, and 0 = scenario without delta-MELD as criterion.
DRI	The *donor risk index* is a derived parameter that is a measurement of liver quality based on nine factors.
SOFT	The *survival outcomes following liver transplantation* is a parameter and measurement of survivability based on risk factors.
AHP	*Analytic hierarchy process* score is a derived parameter and measurement of importance or preference among the alternatives in decision-making.
Patients_Dropped_From_Waitlist	This parameter contains the number of patients dropped from the waitlist weekly.

$MELD_{mean}$	This is the mean of weekly *Average_MELDs* for the simulation scenarios with or without delta-MELD and is expressed using equation (15).
Total_Patients_Removed	This parameter is expressed using equation (16) and is computed within each simulation interval.
Average_MELD	*Average MELD* is compiled weekly from patient waitlist.

I performed a verification process on the simulation's four processes, Waitlist Entry, Waitlist Patient Management, Disease Progression, and Donor Liver Arrival. The purpose of pilot testing and verification was to verify that the features of the simulation were working as intended and that the simulation input data, which were provided by UNOS STAR database, were interpreted correctly. I verified these four processes by running the simulation graphical user interface (GUI), reviewing the outputs against the UNOS data, and reviewing the GUI panels and logs generated by the simulation. The data used for pilot testing were the first 180-day interval of 2008. In addition, I verified weekly reports of one 180-day interval and statistics from one simulation run.

Waitlist Entry and Waitlist Patient Management Processes

Table 9 lists the verification steps and corresponding variables or parameters being observed for verification of the Waitlist Entry process.

Table 9

Waitlist Entry and Waitlist Patient Management Processes Verification

Step	Waitlist Entry and Waitlist Patient Management Processes Steps
1	Initialize patient data with valid entry day #, delta-MELD, disease group, and status.
2	Ensure the interval duration goes from 1 through 180 days.
3	Ensure that both scenarios, one without using delta-MELD and one with using delta-MELD as decision-making, are processed.
4	Ensure that the delta-MELD derived field is computed properly for every patient.

The UNOS STAR data contained all patients who were on the waitlist from 1987 to March 2014. The data included all patients and waitlist history of patients from all regions of the United States, Regions 1 through 11. The UNOS tabbed delimited data files from the UNOS STAR files were processed in Excel spreadsheets for filtering patient waitlist data from Region 9, from 2008 to 2012. The data were further filtered to exclude patients who were status 1 (emergency patients), HCC disease type, and pediatric end-stage liver disease (PELD) patient type. The Waitlist data included the patient MELD updates for the Patient ID of the Patient data. The data, Patient and Waitlist data, were merged into the Patient data for the simulation, joined by Patient ID. More specifically, the Waitlist Entry process handled the initial patient entry into the waitlist while the Waitlist Patient Management process handled the updates of patient MELD and patient statuses through the course of a simulation interval.

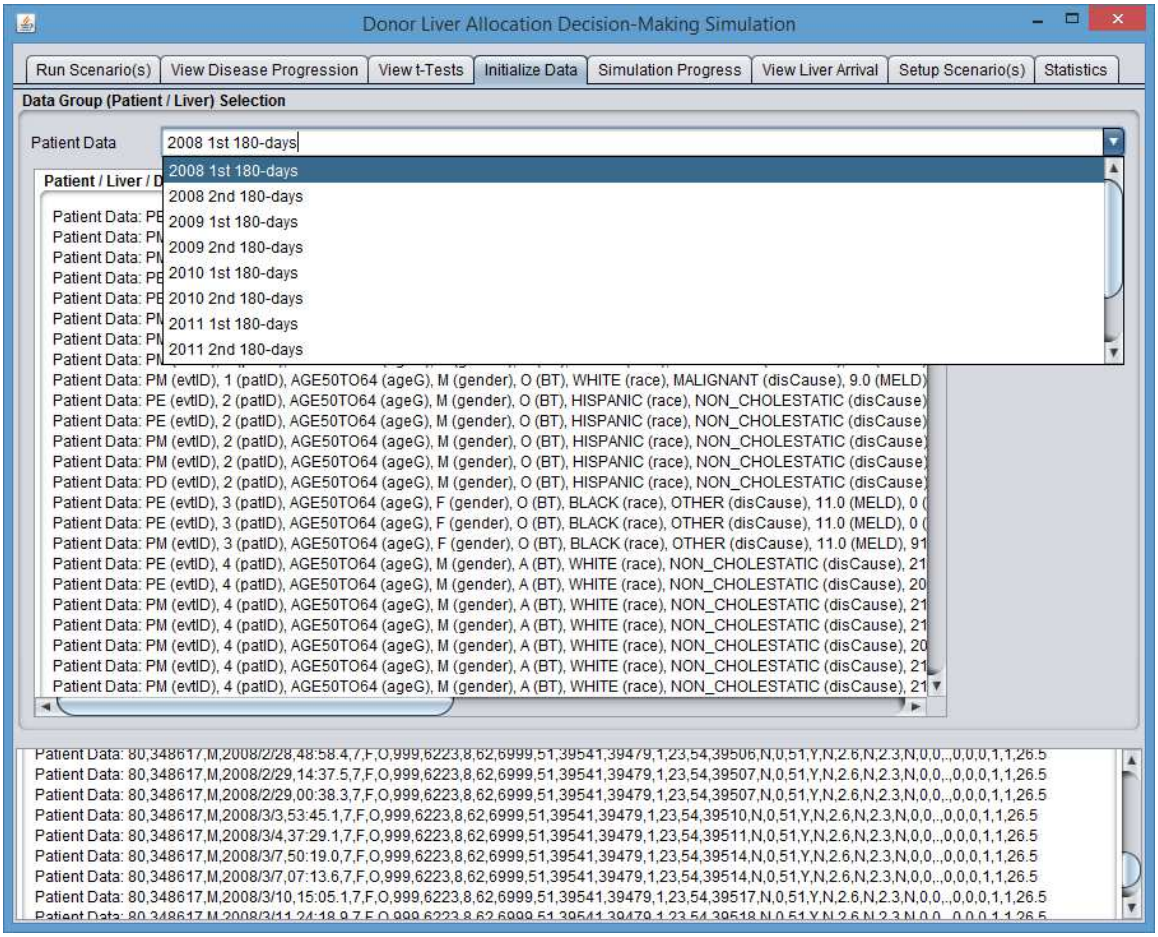


Figure 5. Waitlist Entry and Waitlist Patient Management processes.

Figure 5 shows the patient data processed into the simulation's patient file in the top panel from the raw UNOS STAR patient data. The data on the top panel were correctly processed by observation against the UNOS liver and patient data on the bottom panel. The bottom panel contains the UNOS liver and patient data. I selected the first 180-day interval of 2008 for pilot testing but any of the simulation intervals could have been selected for pilot testing. In actual experimental runs, all 10 intervals are processed in the simulation.

Disease Progression Process

The Disease Progression process ran from day 1 through 180 within the simulation interval. This included processing the Kalman estimation for MELD score progression, taking weekly reports of MELD averages and the number of patients who were dropped from the waitlist, and updating any patient records for each day.

Table 10

Disease Progression Process Steps for Verification

Step	Disease Progression Process Steps
1	Ensure that the processing is performed once for the scenario where AHP would use delta-MELD and once for where AHP would not use delta-MELD.
2	Ensure the processing performs Kalman estimation of disease progression by propagating the MELD scores properly when computed by Kalman estimation.
3	Ensure that patients are removed from the waitlist when their statuses indicate they are too sick or they have died.
4	Ensure that this process proceeds from day 1 through 180.

As shown in Figure 6, verification steps 2 and 3 of Table 10 confirmed that the Kalman estimation was propagating MELD scores and only patient status of "WAITING" was processed. In Figure 7, verification steps 1 and 4 of Table 10 were accomplished by showing that the scenarios, without and with delta-MELD, were performed with the simulation progressing from day 1 through 180.

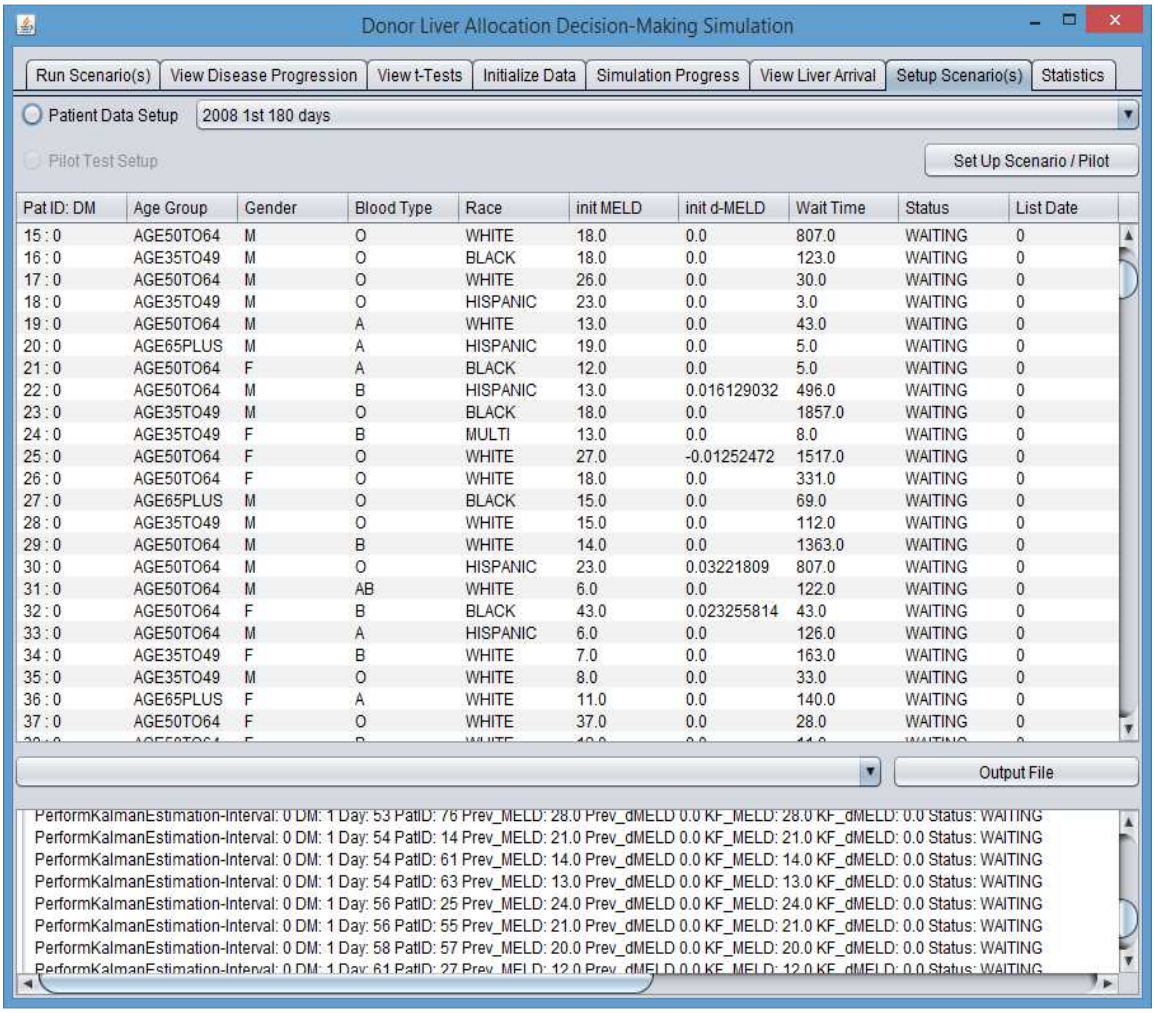


Figure 6. Disease Progression process-1.

Figure 6 shows not only that the initial delta-MELD were computed for patients as shown on the top panel, but the bottom panel shows the internal daily Kalman estimated MELD and delta-MELD computed for patients who do not have a waitlist update record for that day as the scenario progressed from day 1 through 180 of the scenario. Patients with the status of "WAITING" were filtered for the initial setup and the internal Kalman estimation were performed for patients with "WAITING" statuses only.

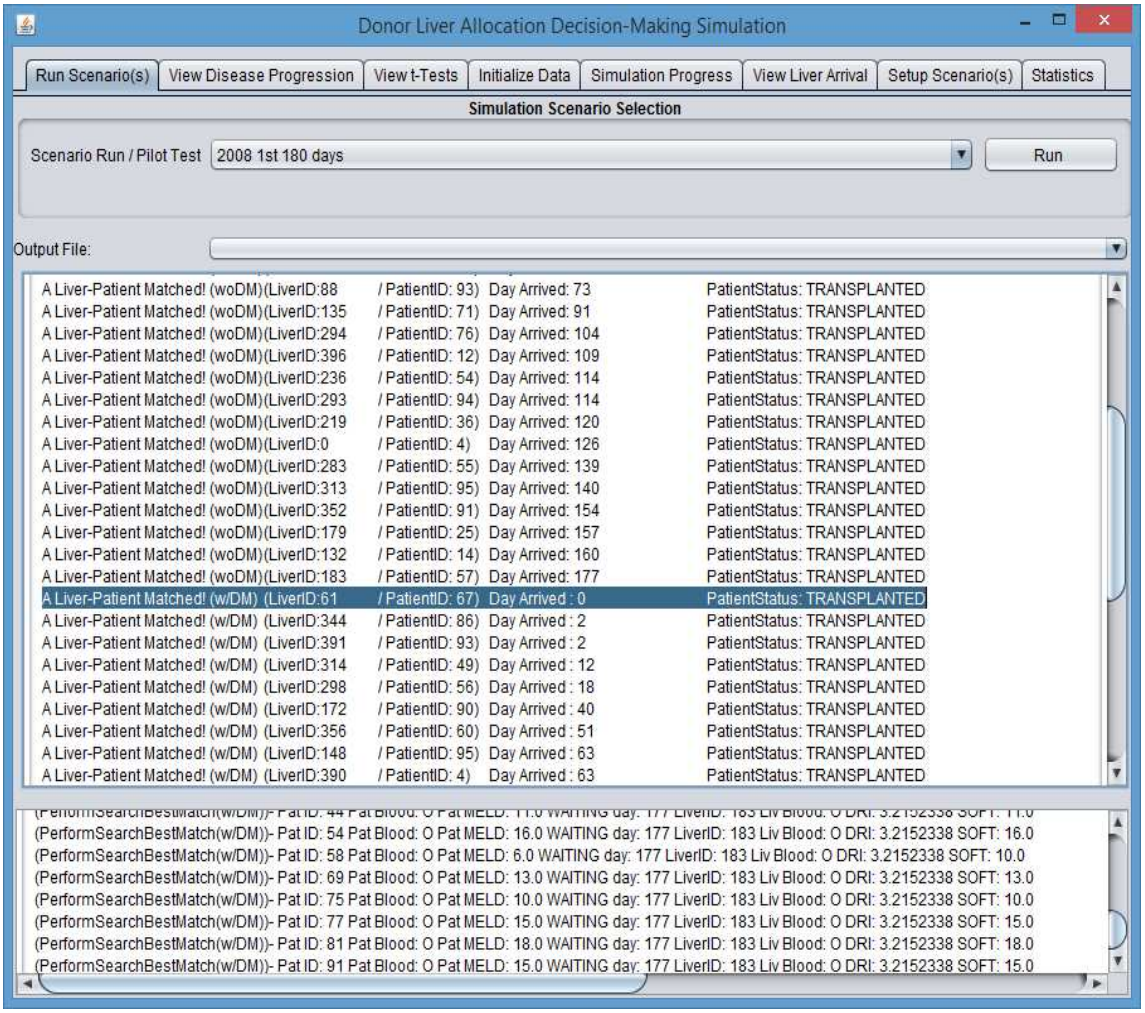


Figure 7. Disease Progression process-2.

Figure 7 shows the highlighted line on the top panel indicating the beginning of the Liver Patient match processing using delta-MELD as criterion. In the panel's previous lines, the processing was performed without delta-MELD. The liver arrivals were determined by a Poisson process with a computed UNOS-based average interarrival time of 5 days. Once a patient match was determined, the patient status was changed from "WAITING" to "TRANSPLANTED". The parameters of DRI and SOFT logged on the bottom panel were confirmed against data retrieved from patient and liver data.

Donor Liver Arrival Process

The Donor Liver Arrival process verification method includes steps shown in Table 11.

Table 11

Donor Liver Arrival Process Steps for Verification

Step	Donor Liver Arrival Processing Steps
1	Ensure the Donor Liver Arrival process is performed once with delta-MELD and once without delta-MELD.
2	Ensure that the day of liver arrival is based on the current day and the number of days computed by the Poisson process.
3	Ensure the correct computation of DRI and SOFT derived data.
4	Ensure patient statuses are translated correctly when processing the UNOS patient data.
5	Ensure the correct processing of winning patient based on AHP score, blood type, and DM indicator is performed.
6	Ensure that upon patient selection, the patient is not a candidate for later liver donor arrivals.

As shown in Figure 8, I performed verification steps 1 and 2 of Table 11 by showing that the scenarios, without and with delta-MELD, were performed and the livers randomly arrived (by the C++ random function) between the days 1-180. In Figure 9, I performed verification steps 3 and 4 of Table 11 by showing the computation of parameters DRI and SOFT were correct by tracing back to their patient and liver data. Also in Figure 9, I performed verification steps 5 and 6 of Table 11 by showing that AHP scores, blood types, and the *DM* indicator were used to filter the selection of patients,

determining best patient match, and ensuring that patients were taken off of the waitlist after transplant.

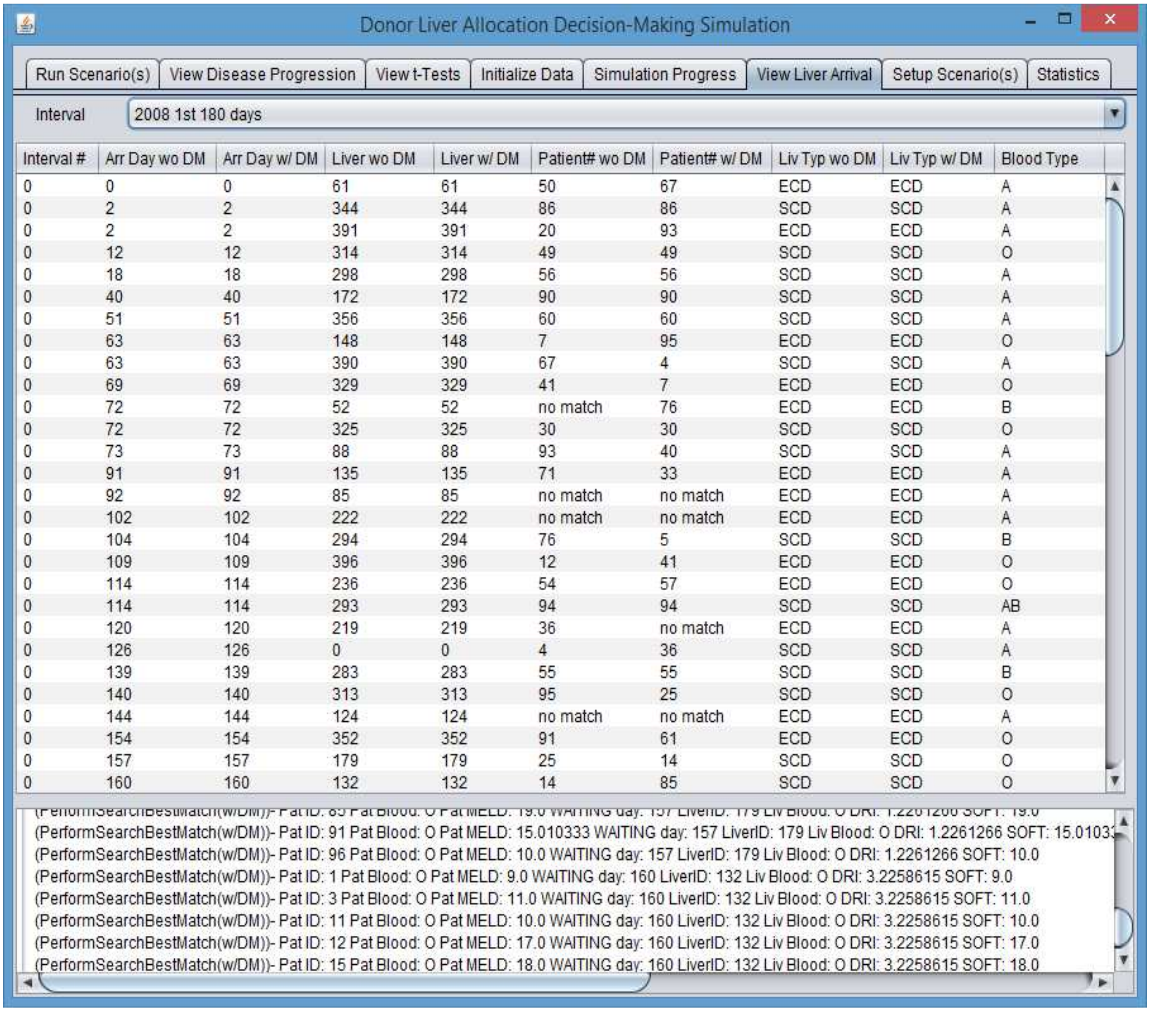


Figure 8. Donor Liver Arrival process-1.

Figure 8 shows that both scenarios, without and with delta-MELD, always shared the same liver arrival day, liver ID, and initial pool of patients. Patient selection in each scenario may have varied for the same liver arrival, but may sometimes have resulted in the same patient selection. When there was no patient selected for an arriving liver, it

was due to “no match” of patient blood type to donor liver blood type. Often, these donor livers would be transferred to another region.

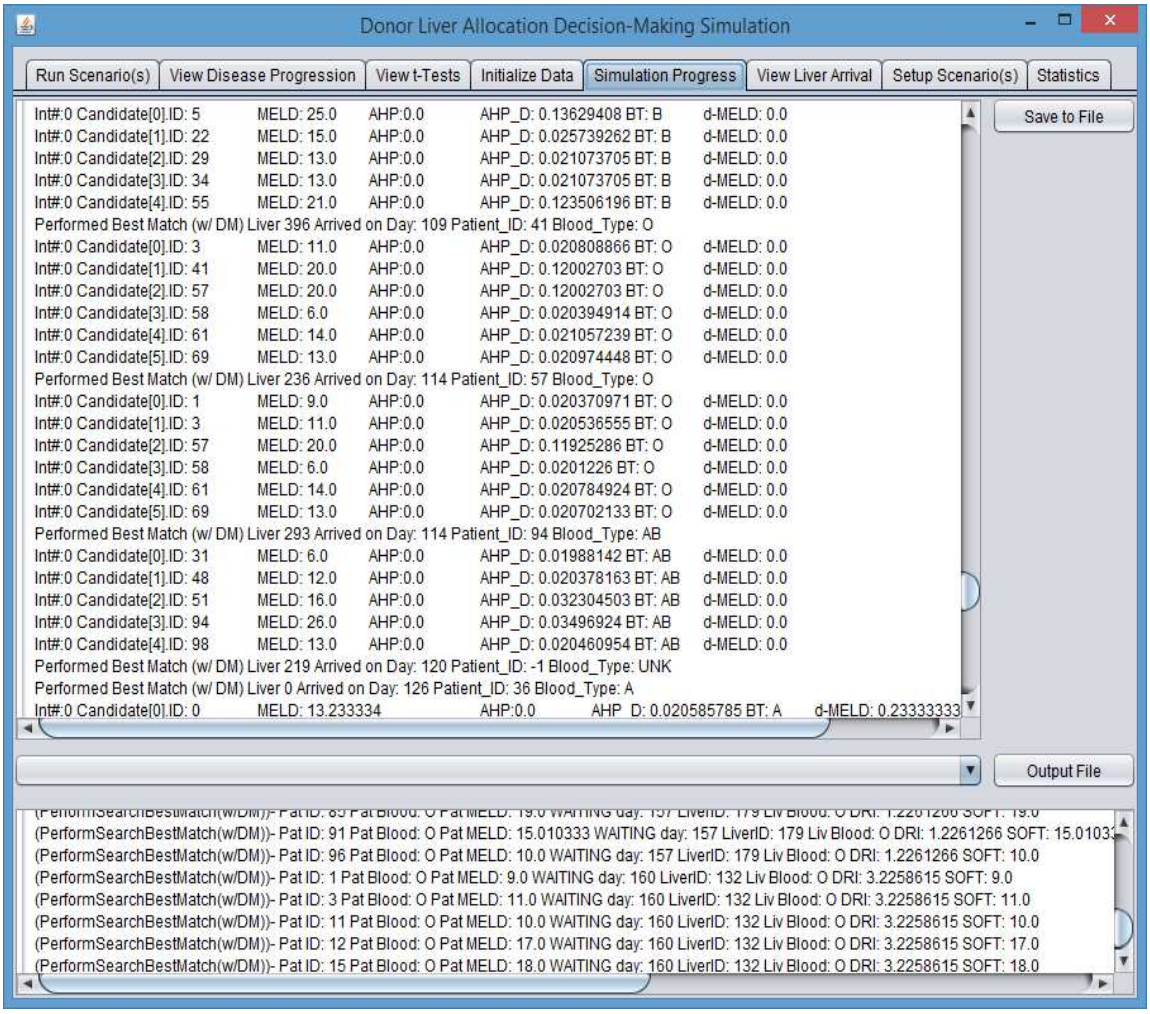


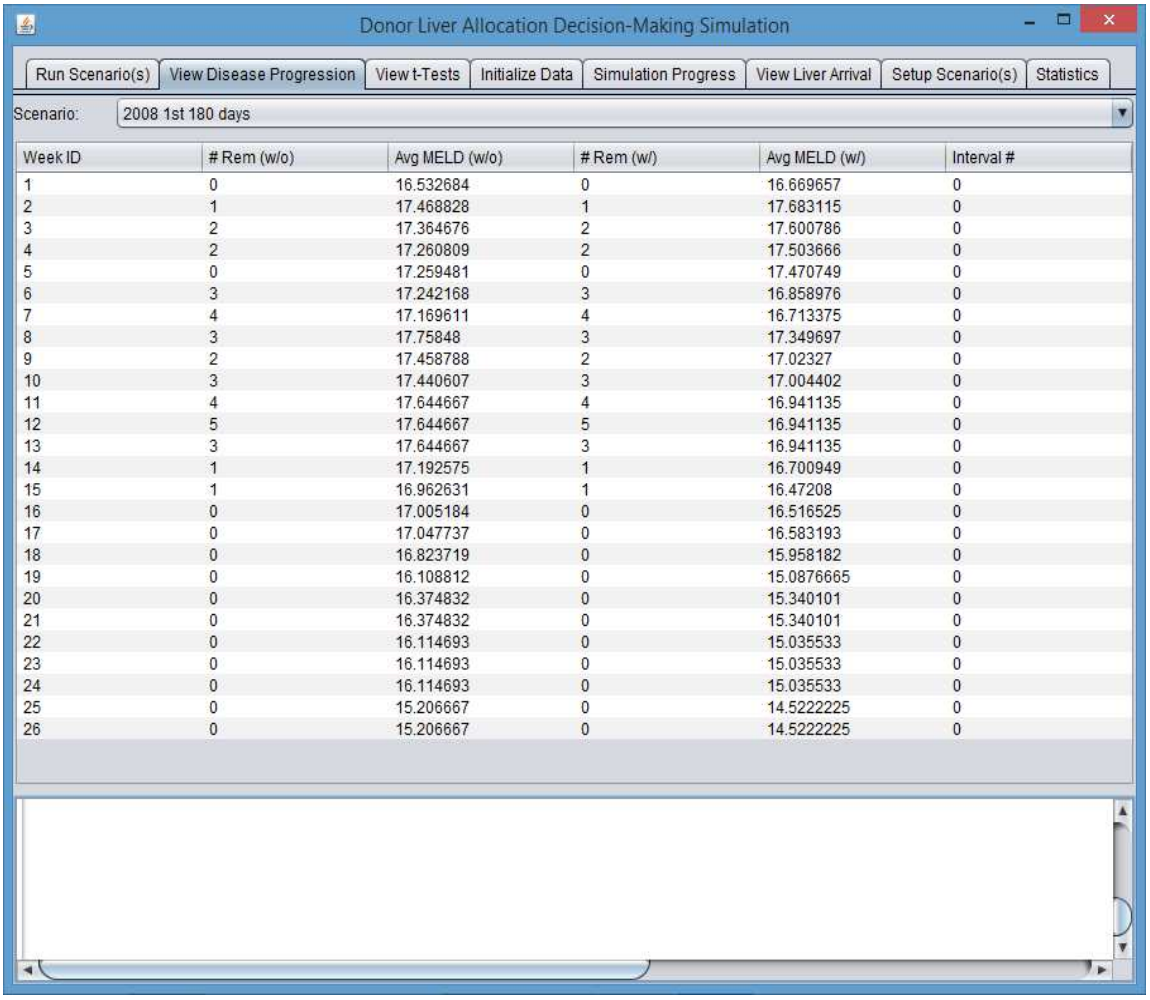
Figure 9. Donor Liver Arrival process-2.

Figure 9 shows that on Day 109, Patient ID 41 was matched with Liver 396. The candidate and donor liver blood types were both “O”. On Day 114, Patient ID 57 was matched with Liver 236. Patient ID 41 was not listed again for the same blood type as Patient 41 had received a liver transplanted already on Day 109. On both days, 109 and

114, the patients with the highest AHP score were matched for the arriving liver. Also, ECD acceptance may have affected whether a patient was listed on the candidate list.

Weekly Reports

The response variables from the simulation were verified to ensure that computations were performed from weekly statuses. The $MELD_{mean}$ and *Total_Patients_Removed* were captured in two scenarios, one with and one without delta-MELD.



Week ID	# Rem (w/o)	Avg MELD (w/o)	# Rem (w/)	Avg MELD (w/)	Interval #
1	0	16.532684	0	16.669657	0
2	1	17.468828	1	17.683115	0
3	2	17.364676	2	17.600786	0
4	2	17.260809	2	17.503666	0
5	0	17.259481	0	17.470749	0
6	3	17.242168	3	16.858976	0
7	4	17.169611	4	16.713375	0
8	3	17.75848	3	17.349697	0
9	2	17.458788	2	17.02327	0
10	3	17.440607	3	17.004402	0
11	4	17.644667	4	16.941135	0
12	5	17.644667	5	16.941135	0
13	3	17.644667	3	16.941135	0
14	1	17.192575	1	16.700949	0
15	1	16.962631	1	16.47208	0
16	0	17.005184	0	16.516525	0
17	0	17.047737	0	16.583193	0
18	0	16.823719	0	15.958182	0
19	0	16.108812	0	15.0876665	0
20	0	16.374832	0	15.340101	0
21	0	16.374832	0	15.340101	0
22	0	16.114693	0	15.035533	0
23	0	16.114693	0	15.035533	0
24	0	16.114693	0	15.035533	0
25	0	15.206667	0	14.5222225	0
26	0	15.206667	0	14.5222225	0

Figure 10. Weekly reports: Disease progression.

Figure 10 shows the weekly *Average_MELD* and *Patients_Dropped_From_Waitlist*. Patients who received liver transplants were removed from the waitlist and were not included in the weekly count. The simulation interval for viewing weekly *Average_MELD* and *Patients_Dropped_From_Waitlist* values is user-selectable.

Statistics from One Simulation Run

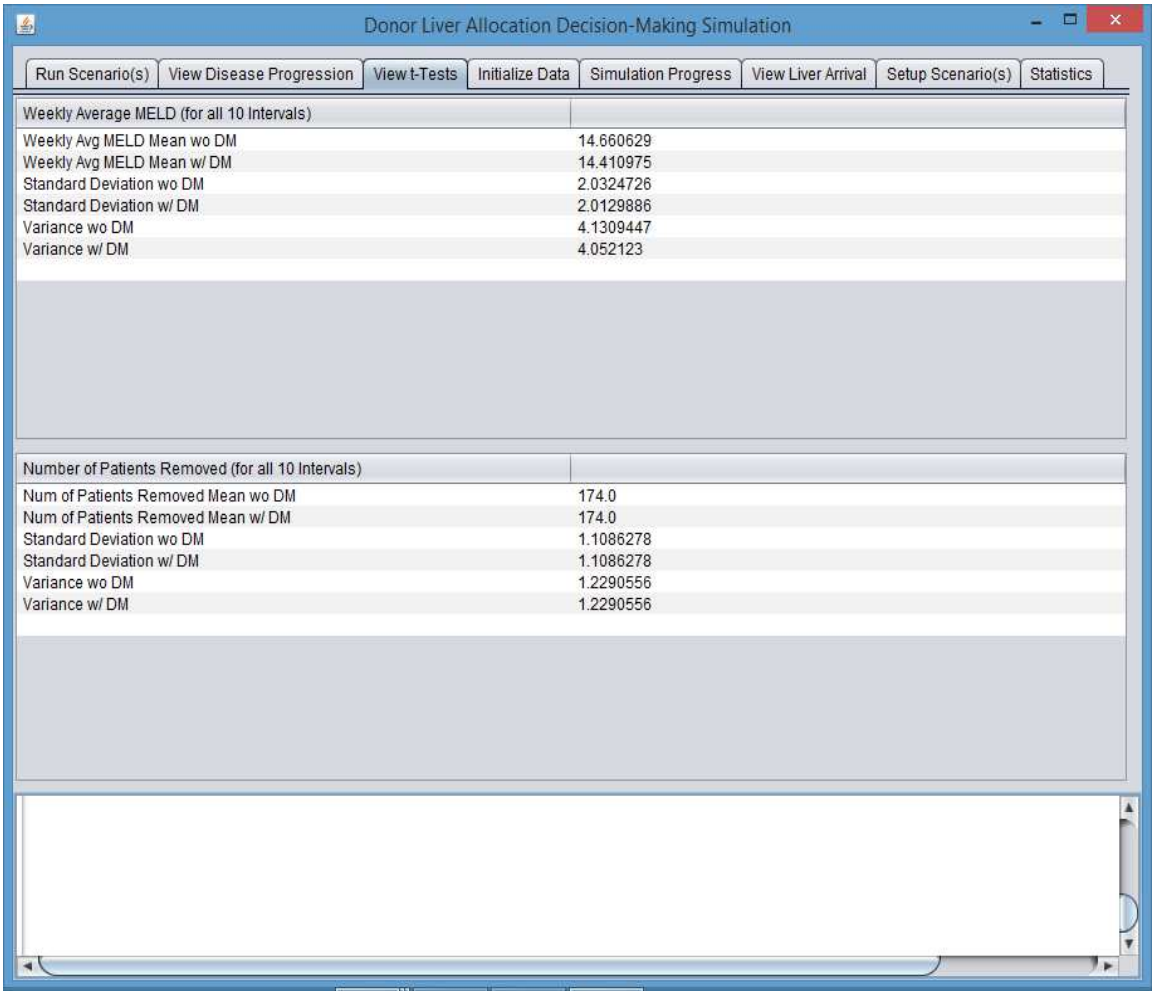


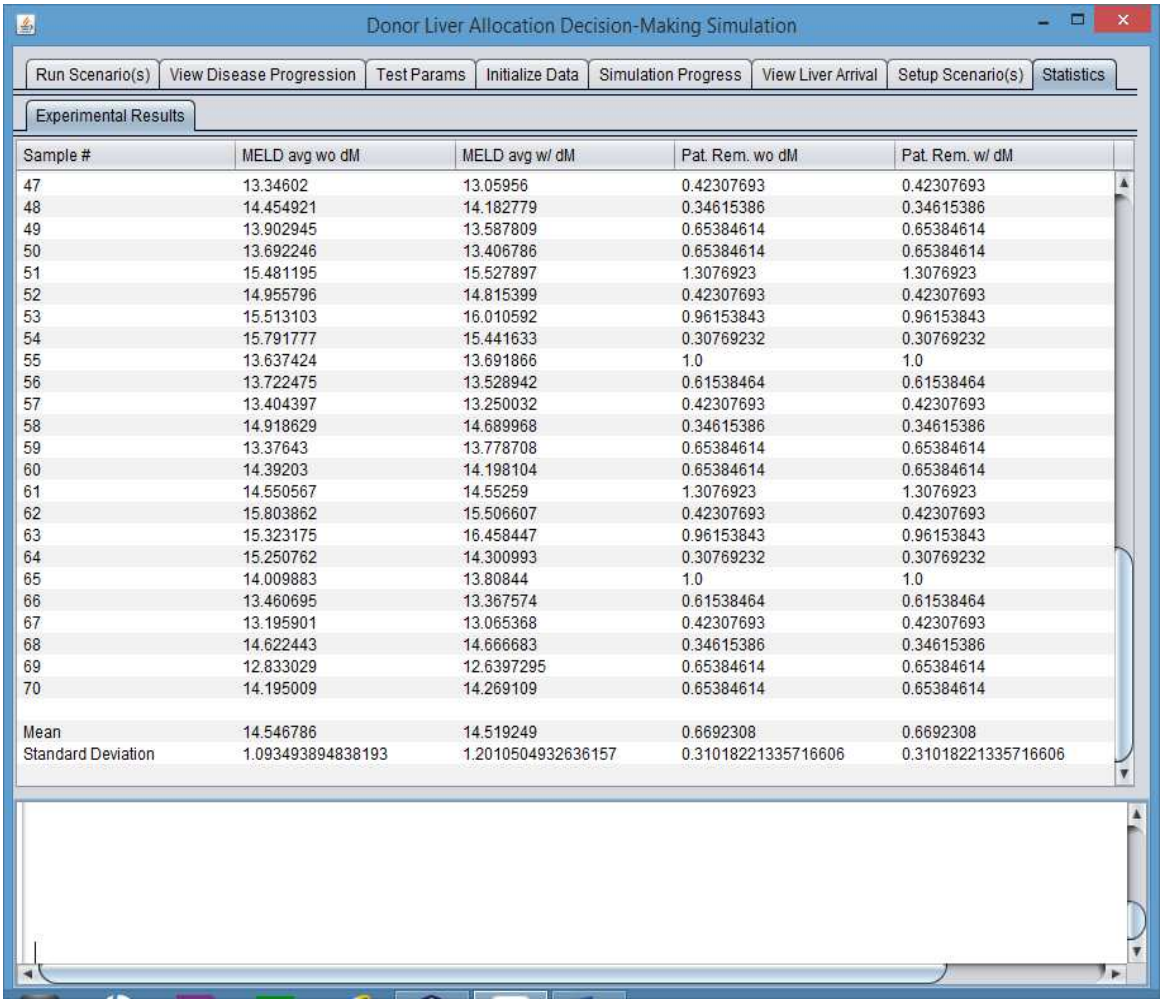
Figure 11. Statistics from one simulation run.

Figure 11 shows the statistics from only one simulation run, with and without delta-MELD, which included 10 intervals covering the years 2008-2012. This simulation run was conducted to verify that the simulation would properly provide statistics of the $MELD_{mean}$ and *Total_Patients_Removed* response variables and that a simulation can run correctly to completion. The number of intervals for the experiment was 10 x 7 = 70. The experiment was conducted from 7 replications of simulation runs.

Experimental Outcome

Figure 11 shows the simulation results from running through all 10 intervals, of only one simulation run for review of the *Total_Patients_Removed* and $MELD_{mean}$ statistics. Figure 11 also shows that there was no difference in *Total_Patients_Removed* between the scenarios using and not using the delta-MELD parameter as criterion for patient selection. This is because the average MELD scores from the weekly report did not vary by much between the two scenarios with and without delta-MELD; and hence, the chances of death for the patients did not change between the two scenarios. Therefore, the simulation removed patients off of the waitlist with the same probability based on Table 4 within the two scenarios, yielding identical numbers of patients removed.

For the simulation experiment, Figure 12 captured the *Total_Patients_Removed* and $MELD_{mean}$ values from 7 simulation runs, 70 intervals runs, which the experiment covers from the beginning to the end of 2008-2012 time span, seven times.



Sample #	MELD avg wo dM	MELD avg w/ dM	Pat. Rem. wo dM	Pat. Rem. w/ dM
47	13.34602	13.05956	0.42307693	0.42307693
48	14.454921	14.182779	0.34615386	0.34615386
49	13.902945	13.587809	0.65384614	0.65384614
50	13.692246	13.406786	0.65384614	0.65384614
51	15.481195	15.527897	1.3076923	1.3076923
52	14.955796	14.815399	0.42307693	0.42307693
53	15.513103	16.010592	0.96153843	0.96153843
54	15.791777	15.441633	0.30769232	0.30769232
55	13.637424	13.691866	1.0	1.0
56	13.722475	13.528942	0.61538464	0.61538464
57	13.404397	13.250032	0.42307693	0.42307693
58	14.918629	14.689968	0.34615386	0.34615386
59	13.37643	13.778708	0.65384614	0.65384614
60	14.39203	14.198104	0.65384614	0.65384614
61	14.550567	14.55259	1.3076923	1.3076923
62	15.803862	15.506607	0.42307693	0.42307693
63	15.323175	16.458447	0.96153843	0.96153843
64	15.250762	14.300993	0.30769232	0.30769232
65	14.009883	13.80844	1.0	1.0
66	13.460695	13.367574	0.61538464	0.61538464
67	13.195901	13.065368	0.42307693	0.42307693
68	14.622443	14.666683	0.34615386	0.34615386
69	12.833029	12.6397295	0.65384614	0.65384614
70	14.195009	14.269109	0.65384614	0.65384614
Mean	14.546786	14.519249	0.6692308	0.6692308
Standard Deviation	1.093493894838193	1.2010504932636157	0.31018221335716606	0.31018221335716606

Figure 12. Experimental outcome.

Figure 12 shows the response variables of two independent populations from running 70 (180-day) intervals (Figure 12 shows an excerpt of interval runs 47-70). Table 12 and Table 13 contain the $MELD_{mean}$ without delta-MELD and with delta-MELD as criteria. There was no difference in patients removed between the two scenarios because the $MELD_{mean}$ in the two scenarios did not differ significantly, and probabilities referenced from Table 4 resulted in the same number of patients who dropped off of the waitlist.

Table 12

$MELD_{mean}$ without delta-MELD

15.8438	16.1996	15.9228	15.1181	13.6869	13.5392	13.7071
14.2018	13.6028	14.2833	15.5608	16.3485	16.6443	16.7870
14.1635	13.2521	12.7992	14.3478	13.2947	13.7007	15.8071
15.1035	16.8286	15.7062	13.9185	13.2357	13.2084	14.4254
14.0263	13.7512	15.3257	15.9791	16.7212	16.3308	13.9717
13.4543	13.2394	14.5611	13.3201	14.0769	15.1372	15.6286
15.5249	14.8552	13.8493	13.5476	13.3461	14.4549	13.9029
13.6922	15.4811	14.9557	15.5131	15.7917	13.6374	13.7224
13.4043	14.9186	13.3764	14.3921	14.5505	15.8038	15.3231
15.2507	14.0098	13.4606	13.1959	14.6224	12.8331	14.1951

Table 12 contains response values of $MELD_{mean}$ from 70 sample interval scenarios not using delta-MELD as criterion. Its standard deviation was 1.093 and the mean of the $MELD_{mean}$ values was 14.5467.

Table 13

$MELD_{mean}$ with delta-MELD

15.5884	15.6231	16.4457	15.3546	13.6629	13.5124	13.7133
14.1556	13.6045	14.1657	15.7969	16.6854	17.6333	17.3875
14.0139	13.1232	12.7618	14.2003	13.2199	13.6131	15.8623
15.1584	16.6119	15.5911	14.0714	13.4139	13.2076	14.3658
13.9046	13.6221	15.5583	15.6571	16.9211	16.0705	13.8258
13.2952	13.1861	14.3737	13.4701	13.9819	15.1675	15.7436
15.7462	14.5906	13.7309	13.1515	13.0595	14.1815	13.5878
13.4068	15.5278	14.8153	16.0105	15.4416	13.6918	13.5289
13.2501	14.6899	13.7787	14.1981	14.5525	15.5066	16.4584
14.3009	13.8084	13.3675	13.0653	14.6666	12.6397	14.2691

Table 13 contains response values of $MELD_{mean}$ from 70 sample interval scenarios

that used delta-MELD as criterion. Its standard deviation was 1.2010, and the mean of the $MELD_{mean}$ values was 14.5192.

The values for the *t*-statistic are as follows.

$$N_1 = 70 \quad (32)$$

$$N_2 = 70 \quad (33)$$

$$x_1 = 14.5467 \quad (34)$$

$$x_2 = 14.5192 \quad (35)$$

$$S_1 = 1.0930 \quad (36)$$

$$S_2 = 1.2010 \quad (37)$$

These values were applied to the formula for the *t*-statistic described in Chapter 3.

$$t = (14.5467 - 14.5192)/\sqrt{(1.0930^2/70) + (1.2010^2/70)}) \quad (38)$$

$$= 0.02750/0.1940 \quad (39)$$

$$= 0.1417 \quad (40)$$

The critical *t* value for 69 degrees of freedom is 1.995. The 95% confidence interval was computed according to the formula,

$$(X_1 - X_2) \pm t_{\alpha/2} * \sqrt{({\sigma_1}^2/N_1) + ({\sigma_2}^2/N_2)} \quad (41)$$

(Aczel & Sounderpandian, 2008). Applying the appropriate values,

$$X_1 = 14.5467, \quad (42)$$

$$X_2 = 14.5192, \quad (43)$$

$$t_{\alpha/2} = 1.995, \quad (44)$$

$$\sigma_1^2 = (1.0930)^2, \quad (45)$$

$$\sigma_2^2 = (1.2010)^2, \quad (46)$$

$$N_1 = 70, \quad (47)$$

$$N_2 = 70. \quad (48)$$

95% confidence interval is

$$= 0.0275 \pm 1.995 * \sqrt{1.1946/70 + 1.4424/70} \quad (49)$$

$$= 0.0275 \pm 1.995 * (0.1940) \quad (50)$$

$$= 0.0275 \pm 0.3872 \quad (51)$$

$$= (-0.3597, 0.4147). \quad (52)$$

Since $t = 0.1417$ is not greater than 1.995, or less than -1.995, the null hypothesis for $MELD_{mean}$ is not rejected. Thus there is insufficient evidence to conclude that there is a difference in the average MELD score ($MELD_{mean}$) among pretransplant patients between simulation models with and without delta-MELD where MELD score is the primary criteria for patient selection in donor liver allocation.

Results

I can conclude that from the data gathered, there was not enough evidence to say there was a difference in $MELD_{mean}$ between simulation scenarios with and without using delta-MELD as decision-making criterion for liver transplant patient selection. Based on the experimental outcome, I can answer the two research questions.

Research Question One

Does a simulation model using the additional parameter of delta-MELD as a patient selection criterion reduce the number of pretransplant patients who dropped off the waiting list?

There is not enough evidence to say there is a difference in *Total_Patients_Removed* between the simulation scenarios with and without delta-MELD since that difference is 0. Therefore, to answer the research question of whether a simulation model using the additional parameter of delta-MELD as a patient selection criterion would reduce the number of pretransplant patients who dropped off of the waiting list, the answer is no.

Research Question Two

Does a simulation model using the additional parameter of delta-MELD as a patient selection criterion lower the average MELD score among pretransplant patients?

To answer the research question of whether a simulation model using the additional parameter of delta-MELD as a patient selection criterion would lower the average MELD score among pretransplant patients, the answer is also no.

Summary

Based on the simulation results, we conclude that there is no difference in outcomes whether or not I used delta-MELD as a decision criterion for the liver allocation system. In Chapter 5, I interpret my experimental results, provide detailed insights from the results, and provide new ideas for refining the donor liver allocation

system. These suggestions were based on the observations and limitations of the simulation data.

Chapter 5: Research Conclusion

In Chapter 5, I offer my interpretation of the simulation results which leads to some recommended future studies that would complement this research. These recommended future studies were based on observations of the OPTN/UNOS data and limitations of the OPTN/UNOS data for the simulation model. Chapter 5 includes four sections: *A Summary of the Research*, *Explanation of Simulation Results*, *Recommendations for Future Studies,* and *Conclusion*.

A Summary of the Research

I investigated whether the U.S. donor liver allocation system could be improved upon by including the delta-MELD parameter into the decision-making process for patient selection. More specifically, I evaluated the influence of the delta-MELD parameter on the $MELD_{mean}$ and *Total_Patients_Removed* response variables in a simulation model. The model was based on Kalman estimation of missing MELD scores for the computation of delta-MELD values, and the AHP algorithm for decision-making. The main objective was to compare the outcome of the decision-making using delta-MELD against the decision-making not using delta-MELD as a criterion for patient selection. I used *t* tests for two independent populations to determine whether the postulated improvement in decision-making was significant enough to justify adding delta-MELD into the decision-making process for patient selection to refine the liver allocation system.

As detailed in Chapter 4, *Pilot Testing and Verification of the Simulation Model*, I verified the simulation model's four processes through pilot testing and verified that the

simulation processed the UNOS STAR data correctly and as intended. From the simulation, I learned that the UNOS STAR data used in the simulation, such as the MELD average, the limitation of data to only OPO Region 9, and the exclusion of HCC and status 1 patients can make a difference in the simulation outcome. As a result of a rigorous verification and validation process, I believe the simulation model can be used in future research, particularly with data from other OPO regions that have higher MELD averages than OPO Region 9.

The simulation results showed that given the limitation of data from OPO Region 9, exclusion of HCC and status 1 patients, and sample intervals of 180-days, the results were not enough to be statistically significant given my sample size, level of significance, and hypothesis tests. From the simulation results based on the UNOS STAR data that I used, there was insufficient statistical evidence to conclude that including the delta-MELD parameter for decision-making could improve the liver allocation system. The simulation results showed that there was a slight improvement in the liver allocation system, a small drop in average MELD, and this may be operationally significant. Even a modest improvement to the donor liver allocation system, like a 1% reduction of patients removed from waitlist or lowering of the average MELD, could mean saving additional lives when refining the allocation system. Freeman (2009) explained that the delta-MELD parameter may have more variability near the end-stage of patients' liver disease. This suggests that an evaluation of the delta-MELD of patients who dropped off the waitlist should be analyzed against patients who remained on the waitlist, while extending the simulation sample interval from 180 to 360 days in future research. The

duration of the simulation sample intervals may not have been long enough when evaluating the influence of delta-MELD up to 180 days, as the median wait time for liver transplant is 11 months (Gift of Life Donor Program, 2014). Hence, future research should be performed by updating the simulation sample intervals from 180 to 360 days to ensure that the duration of simulation intervals would be sufficient for all expected outcomes to occur in its due time.

Explanation of Simulation Results

The simulation results failed to provide sufficient statistical evidence to reject the null hypothesis and to then conclude that the use of delta-MELD values were influential on the simulation's response variables. Looking more closely at the UNOS STAR data and the simulation results to explain this outcome, I observed that for most patients, their delta-MELD scores did not change by much within a short timeframe such as within a month; and when their delta-MELD scores did change, their MELD scores did not always increase. Many times the MELD scores decreased. There were not many patient cases where the delta-MELD increased significantly along with a high MELD score. This may be because when such cases occur, the statuses of these patients are changed to status 1, where these patients are then removed from the waitlist. Status 1 patients are normally in ICU with less than 7 days to live (Cherkassky, 2011). Status 1 and HCC patients undergo additional decision-making based on physicians' knowledge and experience, such as exception MELD point assignment or deciding whether a status 1 patient should undergo transplant after all, that are outside of the decision-making of the OPO's hierarchy of priority levels for liver allocation. The additional decision-making for these patients

should be researched and implemented in an updated simulation model, and these patients should be included in future study.

In my research, I replicated the decision-making process of the OPO's hierarchy of priority levels into the simulation, used actual MELD data of patients and donor livers from OPO Region 9 of the United States from 2008-2012, and used Kalman estimation to get uniformed and unbiased delta-MELD values. While the simulation reflected OPO's hierarchy of priority levels of the liver allocation system and I used actual donor liver and patient data, I excluded status 1 and HCC patients, as these subgroups go through additional decision-making for patient selection. However, when Young et al. (2006) concluded that there was value in using both MELD and delta-MELD in decision-making regarding donor liver allocation, all patients listed for liver transplant between July 1998 and June 2003 from St. James University Hospital in the U.K. of their study were included. Young et al. explained that their data included all patients who were removed from or died on the waiting list during this period. Data collected included demographic, clinical, survival, and donor data that had been prospectively recorded in the U.K. transplant database (p. 332). This means that patients who were the equivalent of status 1 and HCC patients of the United States were included. Freeman (2009) explained that delta-MELD has been associated with increased waiting list mortality and the most significant changes in delta-MELD tend to occur late in the course of the disease. This leads me to conclude that the delta-MELD should be studied among status 1 and HCC patients with simulation sample intervals extending to 360 days, as well as analyzing the utility and survivability aspects of liver transplants of these patients.

On the other hand, in my study, the MELD average from OPO Region 9 is 20, which is much lower compared to the MELD averages from other studies in the literature review. Quante et al (2012) explained that in Germany, there has been a steady increase in the MELD average. In 2010, the MELD average for standard liver allocation was 34 points, without standard exceptions and without high-urgency status. This is 14 MELD points higher than the MELD average in OPO Region 9 of the United States for a similar time frame of 2008-2012. A higher MELD score reflects a more urgent need of a liver transplant and a condition closer to the end-stage of liver failure, which is accompanied by more variability in the delta-MELD (Freeman, 2009). Hence, other OPO regions and other countries that use the MELD system that have higher MELD averages than OPO Region 9 are also suitable for future studies regarding the use of the delta-MELD parameter.

Another observation to note regarding the simulation results was that there were instances where donor livers could not be matched to patients on the waitlist. The simulation model did not handle the occurrence of the case where an available donor liver needs to be transferred to another OPO. This suggests that there could be benefits for patients to enroll into another OPO in order to increase their chances of getting an earlier transplant. A study should be conducted to evaluate the decision-making of patients deciding on whether to stay in the current waitlist or enrolling into another or nearby OPO that has a shorter wait time.

Recommendations for Future Studies

Myers et al. (2013) explained that multiple studies used the UNOS STAR database for refining the liver allocation system. Similarly, in this study, the UNOS STAR database was referenced to identify patients registered on the liver transplant waiting list in the United States. In fact, the data inclusion and exclusion criteria of this study were similar to the study of Myers et al. These criteria excluded patients listed for multiple organs and live donor liver recipients, and status 1, temporarily inactive, and HCC patients. Patients with missing laboratory data necessary for calculation of MELD were excluded as well (p. 2). A more complete model would have included a separate study for some of these subgroups as these subgroups were excluded from this delta-MELD study.

Additionally, applying a modeling perspective into a research topic such as this one matters because it provided a basis on how to extend this research. Expanding this research could mean going beyond the limitations of pretransplant data such as to include posttransplant data for the study of survivability. Young et al. (2006) explained that the usefulness of MELD can be enhanced if it could also predict posttransplant outcomes in some way. Predicting posttransplant outcome could enable a more rational utilization of scarce resources to achieve the maximum benefit. Also, a modeling perspective could provide a basis for the analysis of donor liver allocation beyond the geographical location of UNOS Region 9 which covers New York and western Vermont to a geographical expanded region covering all of continental United States.

Pidd (2010) explained that some models are intended for routine use on a frequent

basis, with little or no human intervention. Other models provide assistance to the human decision process. Model validation and data requirements can add value or provide insights to possible pitfalls that can lead to a theory. But most importantly, modeling is grouped into four categories. These four categories of decision modeling are decision automation, routine decision support, investigation and improvement, and generation of insights for debate (p. 14). The models of AHP and Kalman estimation models can be extended into new and future research for the study of posttransplant survivors, for patients that were in the exception subgroups, and for an expanded geographical UNOS region, for the purpose of decision support, and investigation and improvement to the allocation system, as well and generation of insights for discussion.

Schaubel et al. (2009) explained that currently, patients awaiting deceased-donor liver transplantation were primarily prioritized by medical urgency. More specifically, waitlist chronic liver failure patients are sequenced in decreasing order of MELD scores. In order to maximize lifetime gain through liver transplantation, posttransplant survival should also be considered in the prioritization of liver allocation for patients on the waiting list. Schaubel et al. evaluated that a survival benefit-based system should be applied for allocation of deceased-donor livers to chronic liver failure patients. Under this proposed system, the transplant survival benefit score would be computed for each patient active on the waiting list (p. 1). Schaubel et al. also explained that this proposed score should be based on the difference in 5-year mean lifetime (with vs. without a liver transplant) and should account for patient and donor characteristics. There is an overlap in the distribution of benefit score across MELD categories, since waiting list mortality is

significantly affected by several factors (p. 1). Schaubel et al. further argued that their simulation study results indicated that over 2,000 life-years could be saved if benefit-based allocation was implemented. Schaubel et al. explained that while the shortage of donor livers increases, the need to maximize the life-saving capacity of procured livers has become more pressing. Allocation of deceased-donor livers to chronic liver failure patient efficiency could make the liver allocation system more effective by also prioritizing patients based on transplant survival benefit (p. 1).

Schaubel et al. (2009) further explained that one can envision an extreme case where medical urgency-based allocation does not result in fewer deaths, but merely shifts mortality from the pretransplant to the posttransplant side. Conversely, a utility-based allocation system would ensure that transplanted organs are received by patients with lowered posttransplant mortality. However, patients with the best posttransplant outcomes may also have the best waiting list outcomes. In an extreme case, an ordering that is based on utility could also result in transplantation having no effect on the mortality experience of the patient population, since the low death rate faced by the low-risk patients is merely traded for a low posttransplant death rate. In both cases, the lifetime experienced by the patient population is equal to that in the absence of access to transplantation (p. 2).

Pidd (2010) also explained that operations research/ management science modeling is an external and explicit representation of a part of reality that is seen by people who use models to understand, change, manage, and control that part of reality (p. 10). It is not really so important whether a model is based on a sophisticated

mathematical formulation or whether it is just a simple flow diagram showing how entities are believed to relate to one another. It is more important to recognize that models are approximations, built with intended use in mind and that models are the product of human thought and ingenuity (p. 14). Therefore, specific and additional models can be built for that part of the allocation model that deals specifically on exception subgroups or status 1 patients where different sets of decision-making criteria are usually implemented.

Regarding exception patient groups, Bernal et al. (2010) explained that acute liver failure which is one of the major exception groups within status 1 patients is the clinical manifestation of sudden and severe hepatic injury which can arise from many causes. After abrupt loss of hepatic metabolic and immunological function, it leads to hepatic encephalopathy, coagulopathy, and in many cases progressive multi-organ failure. Although uncommon, this illness occurs mostly in young adults and is associated with high mortality and resource cost. In many countries, it is the most frequent indication for emergency liver transplantation. In the past 10 years, there have been major changes in the understanding of the cause and pathogenesis of this disease.

Bernal et al. (2010) further explained that the main causal agents for the hepatic injury that triggers the onset of liver failure show wide geographical variation, and is normally dependent on the prevalent hepatotoxic virus infections and patterns of drug use. In the developing world, viral causes predominate, with infection by hepatitis A, B, and E viruses accounting for most cases. By contrast, acute viral infection is an uncommon cause in the United States and much of Western Europe, where drug induced

liver injury instead, predominates (p. 190). Bernal et al. further explained that drug-induced injury is the second main cause of acute liver failure and predominates in much of the developed world. In the United States and northern Europe, non-prescription paracetamol (acetaminophen) is the analgesic that is most commonly consumed in overdose, either inadvertently or with intent for deliberate self-harm. Paracetamol-induced hepatoxicity is characteristically hyper-acute from of acute liver failure (p. 3). Perhaps this subgroup could be a candidate group for future research simulation of delta-MELD.

For these cases of acute liver failure, survival has been transformed by the introduction of emergency transplantation, which is now part of routine care in many countries for those patients with acute liver failure who meet criteria indicative of a poor prognosis. However, emergency transplantation outcomes are consistently lower than those of elective surgery, with high early posttransplant mortality. Surgical outcomes have shown progressive and substantial improvement, where 1 year survival exceeds 80% (Bernal et al, 2010). However, the ideal means for identification and selection of patients who are likely to benefit from emergency liver transplantation remains controversial. Inaccurate selection will have serious effects where a patient who would otherwise have survived with medical management and who has incorrectly received a transplant will be subjected to an unnecessary surgical procedure and lifelong immunosuppression, which is associated with major resource cost and increased risk of death. But more significantly, a graft that could have been used for a more suitable candidate would have been lost. The result of failure to identify a patient with acute liver

failure who would have survived only with emergency liver transplantation is of equal magnitude, because of a potentially preventable death (p. 196). This suggests that the delta-MELD parameter could be an area of focus for these subgroups in future research.

Finally, Pidd (2010) argued that the view of complexity can be the property of a real-world system that has manifested from the inability to apply any one formal method as being adequate to capture all its properties. A single approach may not be sufficient to capture the rich behavior of real-world systems. Hence, a single model may not be sufficient to fully represent its behavior. This is not based on the view that different interpretations are due to cognitive limitations, but that different interpretations may be necessary to provide clarity and understanding to the problem (p. 15).

Pidd (2010) argued that in many simulation studies, complexity is a function of the number of elements in a system and of the number of interactions between the elements. This definition treats complexity as an issue of scale and is better regarded as being concerned with complicated, rather than complex, systems (p. 15). In the context of complexity, Feglar and Levy (2005) explained that methods of AHP and ANP are powerful for combatting complex system requiring decision-making with tradeoff considerations. At the same time, design of an appropriate hierarchical structure (AHP) can help in setting up the control structure (ANP). An additional application of these methods could be significantly simplified when integrated with decision models such as benefits, opportunities, cost, and risk (BOCR) models, to a simulation framework. Feglar and Levy suggested using AHP and ANP methods while synthesizing BOCR models. Future studies of the donor liver allocation system can be applied in a nation-wide scope

by covering the continental United States while integrating the BOCR model into multiple AHP regional decision-making models. This would enable various decision-making considerations to be taken into account, such as meeting survivability, travelling time that add risks to the donor liver's cold ischemia time, and the number of patients on the waitlist.

Finally, the social change implication of this research is that donor liver allocation systems in the United States, Europe, and Brazil can continuously be made more efficient to save more lives through liver transplant. The MELD and MELD-based modeling has been studied world-wide, and the MELD and MELD-based modeling has continued to attract researchers from the areas of medicine, health science, and decision-science management disciplines. As new research and additional findings help to refine the liver allocation system, the opportunity to effect social change of the liver transplant community through the discipline of operation research techniques and modeling continues to be a crucial role.

Conclusion

In this study, I investigated whether the donor liver allocation system could be improved by including the delta-MELD parameter into decision-making for selecting a matching recipient. The variables for analysis in the simulation were the average MELD scores and the number of patients who dropped off the waitlist. In this study, I created a simulation that mimics the actual U.S. donor liver allocation system. The simulation model was based on Kalman estimation of MELD score progression and the AHP, an operations research technique for decision-making. The main objective was to compare

the outcome of decision-making using delta-MELD against decision-making not using delta-MELD as a criterion for patient selection. Statistical *t* tests were used for statistical analysis and comparison. This simulation did not result in an improvement in patient average MELD scores or patient wait time after reviewing the data and final results of the simulation. Although a gap in literature in determining the usefulness of the delta-MELD parameter towards a significant improvement in the donor liver simulation system is closed, there are still unanswered cases of whether, due to the research scope, delta-MELD could have been an effective criterion for excluded subgroups.

While this research did not show a significant impact from using delta-MELD in decision-making in a simulation, this research led to the recommendations for future research to study decision-making using delta-MELD for status 1 patients with acute liver failure, analyze the survivability objective with posttransplant data, and study the different dynamics of integrating regional OPOs into a nation-wide study.

References

Aczel, A., & Sounderpandian, J. (2008). *Complete business statistics, 7th Edition.* New York, NY: McGraw-Hill.

Asemota, O. (2010). Understanding the Kalman filter: A classical approach with an application. *European Journal of Scientific Research*, *45*(1), 6-15.

Asrani, S., & Kim, R., (2010). Organ allocation for chronic liver disease: Model for end-stage liver disease and beyond. *Current Opinions Gastorenterology*, *26*(3), 209-213.

Bahra, M., & Neuhaus, P. (2011). Liver transplantation in the high MELD era: A fair change for everyone? *Langenbecks Archives of Surgery*, *396*, 461-465.

Baker, S., Poskar, C., & Junker, B. (2011). Unscented Kalman filter with parameter identifiability analysis for the estimation of multiple parameters in kinetic models. *EURASIP Journal on Bioinformatics and Systems Biology*, *2011*, 1-8.

Bambha, K., Kim, W., Kremers, W., Therneau, T., Kamath, P., Wiesner, R., & Rosen, C. (2004). Predicting survival among patients listed for liver transplantation: An assessment of serial MELD measurements. *American Journal of Transplantation, 4*, 1798-1804.

Bernal, W., Auzinger, G., Dhawan, A., & Wendon, J. (2010). Acute liver failure. *Institute of Liver Studies, 376*(1), 190-201.

Bernardi, M., Gitto, S., & Biselli, M. (2011). The MELD score in patients awaiting liver transplant: Strengths and weaknesses. *Journal of Hepatology, 54*, 1297-1306.

Biselli, M., Gitto, S., Gramenzi, A., Di Donato, R., Brodosi, L., Ravaioli, M., & Grazi, G. (2010). Six score systems to evaluate candidates with advanced cirrhosis for orthotopic liver transplant: Which is the winner? *American Journal of Transplantation*, *16*, 964-973.

Blok, J., Braat, A., Adam, R., Burroughs, A., Putter, H., Kooreman, N., & Rahmel, A. (2012). Validation of the donor risk index in orthotopic liver transplantation within the Eurotransplant region. *Liver Transplantation*, *18*, 113-120.

Box, G., Jenkins, G., & Reinsel, G. (2008). *Time series analysis: Forecasting and control*. Hoboken, NJ: John Wiley and Sons.

Brockwell, P., & Davis, R. (2006). *Time series: Theory and methods*. New York, NY: Springer.

Burr, A., & Shah, S. (2010). Disparities in organ allocation and access to liver transplantation in the USA. *Expert Review of Gastroenterology & Hepatology*, *4*(2), 133-140.

Cherkassky, L. (2011). Does the United States do it better? A comparative analysis of liver allocation protocols in the United Kingdom and the United States. *Cambridge Quarterly of Healthcare Ethics*, *20*(3), 418-433.

Cholangitas, E., & Burroughs, A. (2012). The evolution in the prioritization for liver transplantation. *Annals of Gastroenterology*, *25*(1), 6-13.

Cholongitas, E., Marelli, L., Shusang, V., Senzolo, M., Rolles, K., Patch, D., & Burroughs, A. (2006). A systematic review of the performance of the model for

end-stage liver disease (MELD) in the setting of liver transplantation. *Liver Transplantation*, *12*, 1049-1061.

Cohen, J. (1992). Statistical power analysis: Current directions in psychological science. *Association for Psychological Analysis*, *1*(3), 98-101.

Danner, M., Hummel, J., Volz, F., Van Manen, J., Weigard, B., Dintsios, C., & Bastian, H. (2011). Integrating patients' views into health technology assessment: Analytic hierarchy process (AHP) as a method to elicit patient preferences. *International Journal of Technology Assessment in Health Care*, *27*(4), 369-375.

Dutkowski, P., Oberkofler, C., Bechir, M., Mullhaupt, B., Geier, A., Raptis, D., & Clavien, P. (2011). The model for end-stage liver disease allocation system for liver transplantation saves lives, but increases morbidity and cost: A prospective outcome analysis. *Liver Transplantation*, *17*, 674-684.

Feglar, T., & Levy, J. (2005). Dynamic analytic network process: Improving decision support for information and communication technology. In J. Levy (Chair), *Proceedings of the 8th International Symposium on Analytic Hierarchy Process: Multi-criteria Decision Making*. Honolulu, HI: University of Hawaii, Information and Computer Science Department.

Feng, S., Goodrich, N., Bragg-Gresham, J., Dykstra, D., Punch, J., DebRoy, M., & Greenstein, S. (2006). Characteristics associated with liver grafts failure: The concept of a donor risk index. *American Journal of Transplantation*, *6*, 783-790.

Foxton, M., Al-Freah, M., Portal, A., Sizer, E., Bernal, W., Auzinger, G., & Rela, M. (2010). Increased model for end-stage liver disease score at the time of liver

transplant results in prolonged hospitalization and overall intensive care unit costs. *Liver Transplantation*, *16*, 668-677.

Foxton, M., Kendrick, S., Sizer, E., Muiesan, P., Rela, M., Wendon, J., & Heaton, N. (2006). Change in model for end-stage liver disease score on the transplant waiting list predicts survival in patients undergoing liver transplantation. *European Society for Organ Transplantation*, *19*, 988-994.

Frankfort-Nachmias, C., & Nachmias, D. (2008). *Research methods in the social sciences*. New York, NY: Worth Publishers.

Freeman, R. (2009). Impact and benefits of the MELD scoring system for liver allocation. *Trends in Transplantation, 3*, 70-76.

Freeman, R., Jamieson, N., Schaubel, D., Porte, R., & Villamil, F. (2009). Who should get a liver graft? *Journal of Hepatology, 50*, 664-673.

Gelman, A., Hill, J. (2007). *Data analysis using regression and multilevel / hierarchical models*. New York, NY: Cambridge University Press.

Gift of Life Donor Program (2014). *Understanding the organ transplant waiting list*. Retrieved on June 04, 2015 from http://www.donors1.org/patient/waitinglist/

Green, S., & Salkind, N. (2011). *Using SPSS for Windows and Macintosh: Analyzing and understanding data*. Upper Saddle River, NJ: Pearson Education, Inc.

Gheorghe, L., Iacob, S., Iacob, R., Gheorghe, C., & Popescu, I. (2007). Variation of the MELD score as a predictor of death on the waiting list for liver transplantation. *Journal of Gastrointestinal Liver Disease, 16*, 267-272.

Gotthardt, D., Weiss, K., Baumgartner, M., Zahn, A., Stremmel, W., Schmidt, J., & Bruckner, T. (2009). Limitations of the MELD score in predicting mortality or need for removal from waiting list in patients awaiting liver transplantation. *BioMed Central Gastrolenterology, 9*(72), 1-7.

Halldorson, J., Bakthavatsalam, R., Fix, O., Reyes, J., & Perkins, J. (2009). D-MELD, a simple predictor of post liver transplant mortality for optimization of donor/recipient matching. *American Journal of Transplantation*, *9*, 318-326.

Hillier, F., & Lieberman, G. (2010). *Introduction to operations research.* New York, NY: McGraw Hill.

Hoover, W. (1984). *Algorithms for confidence circles and ellipses (NOAA Technical Report NOS 107 C&GS 3)*. Rockville, MD: National Oceanic and Atmospheric Administration.

Huo, T., Lin, H, Huo, S., Lee, P., Wu, J., Lee, F., & Hou, M. (2008). Comparison of four model for end-state liver disease-based prognostic systems for cirrhosis. *Liver Transplantation*, *14*, 837-844.

Huth. J., Buchholz, M., Kraus, J., Schmucker, M., Von Wichert, G., Krndija, D., & Seufferlein, T. (2010). Significantly improved precision of cell migration analysis in time-lapse video microscopy through use of a fully automated tracking system. *BioMed Central Cell Biology*, *11*(24), 1-12.

Ishizaka, A., Balkenborg, D., & Kaplan, T. (2011a). Influence of aggregation and measurement scale on ranking a compromise alternative in AHP. *Journal of Operational Research Society*, *2011*(62), 700-710.

Ishizaka, A., Balkenborg, D., & Kaplan, T. (2011b). Does AHP help us make a choice? An experimental evaluation. *Journal of Operational Research Society*, *2011*(62), 1801-1812.

Kanwal, F., Dulai, G., Spiegel, B., Yee, H., & Gralnek, M. (2005). A comparison of liver transplantation outcomes in the pre- and post-MELD eras. *Aliment, Pharmacology, and Therapy, 21*, 169-177.

Malinchoc, M., Kamath, P., Gordon, F., Peine, C., Rank, J., & ter Borg, P. (2000). A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. *Hepatology*, *31*, 864-871.

Massie, A., Caffo, B., Gentry, S., Hall, E., Axelrod, D., Lentine, K., & Schnitzler, M. (2011). MELD exceptions and rates of waiting list outcomes. *American Journal of Transplantation, 11,* 2362-2371.

Myers, R., Shaheen, A., Faris, P., Aspinall, A., & Burak, K. (2013). Revision of MELD to include serum albumin improves prediction of mortality on the liver transplant waiting list. *PLoS ONE*, *8*(1), 1-10.

Northup, P., & Berg, C., (2004). Preoperative delta-MELD score does not independently predict mortality after liver transplantation. *American Journal of Transplantation, 4*, 1643-1649.

O'Grady, J. (2012). Liver transplantation for acute liver failure. *Best Practice & Research Clinical Gastroenterology, 26*, 27-33.

Organ Procurement and Transplantation Network. (2014). *Resources*. Retrieved on January 17, 2014 from http://optn.transplant.hrsa.gov/.

Parthiban, P., & Goh, M. (2011). An integrated model for performance management of manufacturing units. *Benchmarking*, *18*(2), 261-281.

Pidd, M. (2004). *Computer Simulation in Management Science, 5th Edition*. West Sussex, England: Wiley.

Pidd, M. (2010). Why modelling and model use matters. *Journal of Operational Research Society*, *2010*(61), 14-24.

Quante, M., Benckert, C., Thelen, A., & Jonas, S. (2012). Experience since MELD implementation: How does the new system deliver? *International Journal of Hepatology, 2012*, 1-5. doi:10.1155/2012/264015.

Rana, A., Hardy, M., Halazun, K., Woodland, D., Ratner, L., Samstein, B., & Guarrera, R. (2008). Survival outcomes following liver transplantation (SOFT) score: A novel method to predict patient survival following liver transplantation. *American Journal of Transplantation*, *8*, 2537-2546.

Rahman, T., & Hodgson, H. (2001). Clinical management of acute hepatic failure. *Intensive Care Medicine*, *27*, 467-476. doi 10.1007/s001340100873.

Rodrique, J., Hanto, D., & Curry, M. (2011). Patients' willingness to accept expanded criteria donor liver transplantation. *American Journal of Transplantation*, *11*, 1705-1711.

Saaty, T. (1996). *Decision making with dependence and feedback: The analytic network process*. Pittburgh, PA: RWS Publications.

Saaty, T. (2001). *The analytic network process*. Pittsburgh, PA: RWS Publications.

Schaubel, D., Guidinger, M., Biggins, S., Kalbfleisch, J., Pomfret, E., Sharma, P., & Merion, R. (2009). Survival benefit-based deceased-donor liver allocation. *American Journal of Transplantation*, *9*(4), 970-981.

Schiano, T. (2012). MELD-good for many, not as good for others…at least for now. *Hepatology International, 2012,* 1-3. doi:10.1007/s12072-012-9393-0.

Scientific Registry for Transplant Recipients. (2012). *Supporting the transplant community with statistical analyses to improve patient outcomes*. Retrieved on October 24, 2012 from http://www.srtr.org/annual_reports/2010/

Sharma, P., Schaubel, D., Gong, Q., Guidinger, M., & Merion, R. (2012). End-stage liver disease candidates at the highest model for end-stage liver disease scores have higher wait-list mortality than Status-1A candidates. *Hepatology*, *55*(1), 192-198.

Siciliano, M., Parlati, L., Malderelli, F., Rossi, M., & Corradini, S. (2012). Liver Transplantation in adults: Choosing the appropriate timing. *World Journal of Gastrointestinal Pharmacology and Therapeutics, 3*(4), 49-61.

Sipahi, S., & Timor, M. (2010). The analytic hierarchy process and analytic network process: An overview of applications. *Management Decision*, *48*(5), 775-808.

Teixeira de Freitas, A., Itikawa, W., Kurogi, A., Stadnik, L., Parolin, M., & Coelho, J. (2010). The impact of the model for end-stage liver disease (MELD) on liver transplantation in one center in Brazil. *Arquivos de Gastroenterologia*, *47*(3), 233-237.

United Network for Organ Sharing Organization. (2014). *Technology for transplantation*. Retrieved on January 17, 2014 from http://www.unos.org/donation/index.php?topic=data_collection

Weismuller, T., Negm, A., Becker, T., Barg-Hock, H., Klempnauer, J., Manns, M., & Strassburg, C. (2009). The introduction of MELD-based organ allocation impacts 3-month survival after liver transplantation by influencing pretransplant patient characteristics. *Transplant International*, *22*, 970-978.

Winston, W. (2004). *Operations research applications and algorithms.* Belmont, CA: Brooks/Cole.

Young, A., Rajagenashan, R., Asthana, S., Peters, C., Toogood, G., Davies, M., & Lodge (2006). The value of MELD and sodium in assessing potential liver transplant recipients in the United Kingdom. *Transplant International*, *20*, 331-337.

Zhou, H., & Hu, H. (2010). Reducing drifts in the inertial measurements of wrist and elbow positions. *IEEE Transactions on Instrumentation and Measurement*, *59*(3), 575-585.

Appendix A: Kalman Filter and Error Ellipse

A state space model consists of two equations: the state equation and the observation equation. The essential difference between the state-space model and the conventional linear model representation is that the state space nature is not assumed to be constant but can change with time. This dynamic feature is incorporated and reflected in the transition equation.

In the example of Kalman filter, the state vector is (MELD*, deltaMELD*). The state space nature is not assumed to be constant but can change with time. This dynamic feature is reflected in the transition equation. The transition equation would project the current MELD and the delta-MELD values, patient's progression of liver disease to a specific time. In a state space system, the state vector can be propagated to the specific time when a donor liver is made available and upon a change in a patients' health status. The covariance matrix can also be propagated to the time when an available liver is made available or upon a patient's health status is updated. The eignenvalues of the corresponding covariance matrix are computed and use these eigenvalues are used to derive an error ellipse. The state vector is propagated to an instance in time as follows where *x* is MELD and *y* is *deltaMELD*.

$$\Delta X = \Delta t * Xdot, \tag{A1}$$

$$\Delta Y = \Delta t * Ydot, \tag{A2}$$

$$where\ \Delta t = T_{update} - T_{val} \tag{A3}$$

T_{val} is the last valid time of state vector, and T_{update} is the time of current update.

$$X = MELD, \tag{A4}$$

$$Y = deltaMELD, \tag{A5}$$

Xdot is the rate of *X*, and *Ydot* is the rate of *Y*. T_{update} is the time of current update, and *Xdot* and *Ydot* are the *X* and *Y* rates of change of the MELD score progress.

The covariance matrix is computed to propagate the covariance matrix to time t_{update}, where T_{MELD} is the current MELD score and $T_{deltaMELD}$ is the current *deltaMELD*. The 4 x 4 covariance matrix is computed from the patients' last health update.

The covariance matrix use the matrix equation,

$$P_{k+1} = \psi P_k \psi^T \tag{A6}$$

to propagate the 4 x 4 covariance matrix of MELD score update.

P_{k+1} is the covariance matrix valid at time $t + 1$,

ψ is the transition matrix,

P_k is the covariance matrix from the matrix valid at time t, and

ψ^T is the transpose of the transition matrix.

Using matrix equation, $P_{k+1} = \psi P_k \psi^T$, we calculate this into its final algebraic form.

$$\psi\ P_k\ \psi^T = \begin{pmatrix} 1 & \Delta & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} A & B & C & D \\ B & E & F & G \\ C & F & H & I \\ D & G & I & J \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ \Delta & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \Delta & 1 \end{pmatrix} \tag{A7}$$

where P_{k+1} is the covariance matrix valid at time $t_{update.}$ ψ is the transition matrix P_k is the covariance matrix from the MELD score update matrix valid blood work at time $T_{val.}$ ψ^T is the transpose of the transition matrix.

The algebra for the covariance propagation is computed as follows.

$$\begin{pmatrix} [(A+\Delta B)+\Delta(B+\Delta E)] & (B+\Delta E) & [(C+\Delta F)+\Delta(D+\Delta G)] & (D+\Delta G) \\ (B+\Delta E) & E & (F+\Delta G) & G \\ [(C+\Delta D)+\Delta(F+\Delta G)] & (F+\Delta G) & [(H+\Delta I)+\Delta(I+\Delta J)] & (I+\Delta J) \\ (D+\Delta G) & G & (I+\Delta J) & J \end{pmatrix} \quad \text{(A8)}$$

$where\ \Delta = \Delta t$ (A9)

$$A = \sigma_X^2 \quad \text{(A10)}$$

$$B = \sigma_{X(Xdot)} \quad \text{(A11)}$$

$$C = \sigma_{XY} \quad \text{(A12)}$$

$$D = \sigma_{X(Ydot)} \quad \text{(A13)}$$

$$E = \sigma_{Xdot}^2 \quad \text{(A14)}$$

$$F = \sigma_{Xdot(Y)} \quad \text{(A15)}$$

$$G = \sigma_{Xdot(Ydot)} \quad \text{(A16)}$$

$$H = \sigma_Y^2 \quad \text{(A17)}$$

$$I = \sigma_{X(Ydot)} \quad \text{(A18)}$$

$$J = \sigma_{Ydot}^2 \quad \text{(A19)}$$

$for\ X = MELD\ and\ Y = deltaMELD.$

The elements of the 2 x 2 covariance matrix needed for the error ellipse calculation are as follows.

$$(\sigma x)^2 = \sigma_x^2 \quad \text{(A20)}$$

$$\sigma xy = \sigma_{xy}\ , \quad \text{(A21)}$$

$$(\sigma y)^2 = \sigma_y^2. \quad \text{(A22)}$$

To calculate the eigenvalues of the 2 x 2 covariance matrix, the eigenvalues and the roots of the characteristic equation of a matrix are calculated by determining the following.

$$DET(A - \lambda I) = 0 \quad \text{(A23)}$$

$for\ x = MELD, and\ y = deltaMELD.$

In the case of our 2 x 2 covariance matrix, we get the following.

$$(\sigma_x^2 - \lambda)(\sigma_y^2 - \lambda) - \sigma_{xy}^2 = 0 \quad \text{(A24)}$$

$$a = \sigma_x^2 \quad \text{(A25)}$$

$$b = \sigma_y^2 \quad \text{(A26)}$$

$$c = \sigma_{xy}^2 \quad \text{(A27)}$$

$$where\ we\ have\ (a - \lambda)(b - \lambda) - c = 0 \quad \text{(A28)}$$

The following steps calculate out the following.

$$(a - \lambda)(b - \lambda) - c = 0 \quad \text{(A29)}$$

When in its equation into its quadratic form, we have,

$$\lambda^2 - (a + b)\lambda - c + ab = 0 \quad \text{(A30)}$$

$$where\ A = -(a + b)\ and\ B = -c + ab \quad \text{(A31)}$$

$$z_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad \text{(A32)}$$

$$z_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \quad \text{(A33)}$$

Then λ_1 and λ_2 are as follows, and they would be used to formulate our *error ellipse*.

$$\lambda_1 = (-A + \sqrt{A^2 + 4B})/2 \quad \text{(A34)}$$

$$\lambda_2 = (-A - \sqrt{A^2 + 4B})/2 \quad \text{(A35)}$$

$$where\ A = -(a\ +\ b)\ and\ B = -c\ +\ ab, and\ where$$

$$a\ =\ \sigma^2_{MELD}$$

$$b\ = \sigma^2_{deltaMELD}\ \ and$$

$$c = \sigma^2_{MELD*deltaMELD}$$

The elements of a covariance matrix include σ^2_x, σ^2_y, and σ^2_{xy}, where the eigenvalues of λ_1 and λ_1 are derived to provide the axes for the *Error Ellipse*. The aspect ratio of the ellipse,

$$(\Delta y/y)\ /\ (\Delta x/x) \quad \text{(A36)}$$

is computed which is the counter-clockwise rotation *angle* of the ellipse as follows.

$$Angle = ½ * tan^{-1}\ ((1/aspect\ ratio) * 2\sigma_{xy}/(\sigma_x^2 - \sigma_y^2)) \quad \text{(A37)}$$

Since σ_x and σ_y represent the standard deviations of stochastically independent random variables, additional theorem for the chi-square distribution can be used to show that the probability associated with a confidence ellipse is given by $p = 1 - e^{-(1/2)k^2}$. Conversely, the semimajor ($k * \sigma_x$) and semiminor ($k * \sigma_y$) axes of a confidence ellipse having specified probability *p* can be calculated from (σ_x, σ_y).

$$k\ =\ \sqrt{(-2 * ln(1 - p))} \quad \text{(A38)}$$

Hence, the error ellipse is a confidence ellipse with elliptical scale factor *k = 1* and probability approximately p = 0.3935. The 50% and 95% confidence ellipses have

elliptical scale factors approximately 1.1774 and 2.4477, respectively (Hoover, 1984).

If $\sigma_x > \sigma_{y,}$ then the semi-axis length parallel to the x-axis is equal to $\sqrt{\max(\text{eigenvalues})} * \text{scalefactor}$ is computed. Then the semi-axis length parallel to the y-axis is equal to $\sqrt{\min(\text{eigenvalues})} * \text{scalefactor}$.

If $\sigma_y > \sigma_{x,}$ then the semi-axis length parallel to the x-axis is equal to $\sqrt{\min(\text{eigenvalues})} * \text{scalefactor}$ is computed. The semi-axis length parallel to the y axis is equal to $\sqrt{\max(\text{eigenvalues})} * \text{scalefactor}$. The ellipse can now be rotated counter clock-wise from this orientation (angle).

Appendix B: Analytic Hierarchy Process (AHP)

Saaty (1996) developed the analytic hierarchy process (AHP) and stated that AHP is a general measurement that derives ratio scaled values from both discrete and continuous paired comparisons of multilevel hierarchy structures. These comparisons can be taken from actual measurements or from a fundamental scale that reflects the relative strength of preferences and feelings. AHP is widely used for multiple criteria decision-making in planning, resource allocation, and conflict resolution. In using the AHP to model a problem, one would need a hierarchy or a network structure to represent the problem. Pairwise comparisons are used to establish relations within the hierarchy or network structure.

Saaty (1996) explained that the AHP is a heuristic algorithm for scoring multiple criteria and alternatives in decision-making. AHP is best demonstrated by reviewing its step with an example. The details of the three major AHP algorithm steps are as follows.

1. Develop the weights for the criteria.
    a. Develop a single pair-wise comparison matrix for the criteria.
    b. Multiply the values in each row together and calculate the n^{th} root of the product.
    c. Normalize the n^{th} root of products to get the appropriate weights.
    d. Calculate and check the Consistency Ratio.
2. Develop the ratings for each decision alternative for each criterion.

a. Develop a pair-wise comparison matrix for each criterion, with each matrix containing the pair-wise comparisons of the performance of decision alternatives on each criterion.
b. Multiply the values in each row together and calculate the n^{th} root of the product.
c. Normalize the n^{th} root of product to gate the corresponding ratings.
d. Calculate and check the Consistency Ratio.

3. Calculate the weighted average rating for each alternative. Then choose the one with the highest score.

The following demonstrates the above steps with an example of decision-making for selecting a liver transplant recipient. The criteria of the example are MELD score, blood type, and body structure. Furthermore, the decision alternatives are three possible alternatives (patients). Pair-wise comparison is used to establish the relative priority of each criterion against every other criterion. Then the relative priority of the alternatives is pair-wise compared against every other alternative for each criterion.

The main element to the AHP technique is the use of pair-comparisons. The pair-wise comparisons use a scale that ranges from equally preferred to extremely preferred. The following illustrates the values associated with the level of preference used to scale the results of pair-comparisons.

Table B1

Paired Comparison: Value-Description

Value	Description
1	Equally preferred
2	Equally to moderately preferred
3	Moderately preferred
4	Moderately to strongly preferred
5	Strongly preferred
6	Strongly preferred to very strongly preferred
7	Very strongly preferred
8	Very strong
9	Extremely preferred

In the example below, the value "9" is used to denote that the preference of MELD score is "extremely preferred" over body structure. Also, as indicated in the example the MELD score is "moderately preferred" over blood type. When comparing blood type to body structure, the paired comparison shows that blood type is "strongly preferred" (number 5) over body structure.

Table B2

AHP Table of Weights of the Criteria

	MELD	Blood Type	Body Structure	3rd root of product	Priority Vector
MELD	1	3	9	3	0.67162545
Blood Type	0.33333333	1	5	1.1856311	0.26543334
Body Structure	0.11111111	0.2	1	0.2811442	0.06294121
Sum	1.44444444	4.2	15	4.4667753	1
Sum*PV	0.97012565	1.11482004	0.94411808	3.0290637	NA
LambaMax	3.02906377	NA	NA	NA	NA
CI	0.02906377	NA	NA	NA	NA
CR	0.05019004	NA	NA	NA	NA

Additional notes on Table B2 are as follows.

1. The “Sum” row is the value of the sum of the criteria column (Ex: MELD Column = 1 + 0.33333333 + 0.11111111).

2. The “3rd Product Root” is the product of the row of criteria (Ex: MELD row = 1 * 3 * 9) taken to the 3rd root divided by the sum of all 3rd Product Root.

3. The “Sum * PV” row is the value of the sum of the previous row times the corresponding “Priority Vector”.

4. The “LambdaMax” value is the sum of all “Sum*PV” values.

5. The “Consistency Index” is the value of “LambdaMax” minus 3. 3 is the number of criteria.

6. The “Consistency Ratio” is computed by taking the “Consistency Index” and dividing by 0.58. 0.58 is the value to divide for three criteria.

Now suppose that there are three alternatives (three types of patients). And these three types of patient are the AHP alternatives that have the highest MELD score within a Transplant Center (TC), an alternative whose MELD score is the next highest within the same Transplant Center, and an alternative of highest MELD score outside of the Transplant Center where the donor is. In this example, the hierarchy of priority is such that the highest priority is the alternative with the highest MELD score within the same Transplant Center. The next in priority is the next highest MELD scoring alternative outside of the Transplant Center but within the same OPO. If there are no suitable recipients within the OPO, then the highest MELD scoring alternative outside of the OPO

is selected. We now develop the ratings for the each alternative with respect to each criterion (MELD, blood type, and body structure).

Table B3

AHP Table of Weights of Alternatives according to MELD

	Highest MELD in TC	Next Highest MELD in TC	Highest MELD out of TC but within OPO	Outside of OPO	4rd root of product	Priority Vector
Highest MELD in same TC	1	5	5	7	3.63713	0.62248
Next Highest MELD in TC	0.2	1	3	5	1.44225	0.246835
Highest MELD out of TC but within OPO	0.2	0.2	1	5	0.58480	0.100086
Outside of OPO	0.1428571	0.2	0.2	1	0.17878	0.030597
Sum	1.5428571	0.2	0.2	1	1.78781	0.030597
Sum*PV	0.9603987	1.579744	0.92079	0.55076	4.01169	NA
LambaMax	4.011670	NA	NA	NA	NA	NA
CI	0.0116969	NA	NA	NA	NA	NA
CR	0.0129966	NA	NA	NA	NA	NA

Additional notes on Table B3 are as follows.

1. The "Sum" row is the value of the sum of the criteria column (Ex: Column = 1 + 0.2 + 0.2 + 0.1428571).

2. The "4th Product Root" is the product of the row of criteria (Ex: Row = 1* 5 * 5 * 7) taken to the 4th root divided by the sum of all 4th Product Root.

3. The "Sum * PV" row is the value of the sum of the previous row times the corresponding "Priority Vector".

4. The “LambdaMax” value is the sum of all “Sum*PV” values.

5. The Consistency Index (CI) above is the value of “LambdaMax” minus 4. 4 is the number of criteria.

6. The Consistency Ratio above is computed by taking the Consistency Index (CI) divided by 0.90. 0.90 is the value to divide by for four criteria.

Table B4

AHP Table of Weights of Alternatives according to Blood Type

	Highest MELD in same TC	Next Highest MELD in TC	Highest MELD out of TC but within OPO	Outside of OPO	4rd root of product	Priority Vector
Highest MELD in same TC	1	1	1	1	1	0.25
Next Highest MELD in TC	1	1	1	1	1	0.25
Highest MELD out of TC but within OPO	1	1	1	1	1	0.25
Outside of OPO	1	1	1	1	1	0.25
Sum	4	4	4	4	4	1
Sum*PV	1	1	1	1	4	NA
LambaMax	4	NA	NA	NA	NA	NA
CI	0	NA	NA	NA	NA	NA
CR	0	NA	NA	NA	NA	NA

The values on Table B5 are similarly derived as the values from Table B4. The paired comparisons of the alternatives are all identically preferred over each other. Hence the Priority Vectors are 0.25 each for all four alternatives.

Table B5

AHP Table of Weights of Alternatives according to Body Structure

	Highest MELD in TC	Next Highest MELD in TC	Highest MELD score within OPO	Outside of OPO	4rd root of product	Priority Vector
Highest MELD in same TC	1	1	1	1	1	0.25
Next Highest MELD in TC	1	1	1	1	1	0.25
Highest MELD out of TC but within OPO	1	1	1	1	1	0.25
Outside of OPO	1	1	1	1	1	0.25
Sum	4	4	4	4	4	1
Sum*PV	1	1	1	1	4	NA
LambaMax	4	NA	NA	NA	NA	NA
CI	0	NA	NA	NA	NA	NA
CR	0	NA	NA	NA	NA	NA

Similar to Table B4, the values in Table B5 are similarly derived by paired comparisons. The paired comparisons of the alternatives are all identically preferred over each other. Hence the Priority Vectors are 0.25 each for all four alternatives. Finally, Table B6 shows the calculations of the weighted average rating for each decision alternative. The highest weighted average rating is selected as the "Winner".

Table B6

AHP Table Ranking of Alternatives

Criteria and Alternatives	MELD	Blood Type	Body Structure	Sum	Multi-plier	
Highest MELD in same TC	0.671624	0.265433	0.062941	1.12248	0.50016	Winner
Next Highest MELD in TC	0.2468350	0.25	0.25	0.74683	0.24787	
Highest MELD out of TC but within OPO	0.1000867	0.25	0.25	0.60008	0.14931	
Outside of OPO	0.030597	0.25	0.25	0.53059	0.10264	

Appendix C: Simulation Programming Notes

The simulation Graphical User Interface (GUI) panels allow the simulation experiments to be broken into logical steps. The simulation allows the user to choose the interval data or pilot test to run, and provides simulation GUI panels and output file for simulation verification and data analysis of scenario and pilot runs. Below are the simulation GUI panels and their corresponding programming notes.

Initialize Data Panel

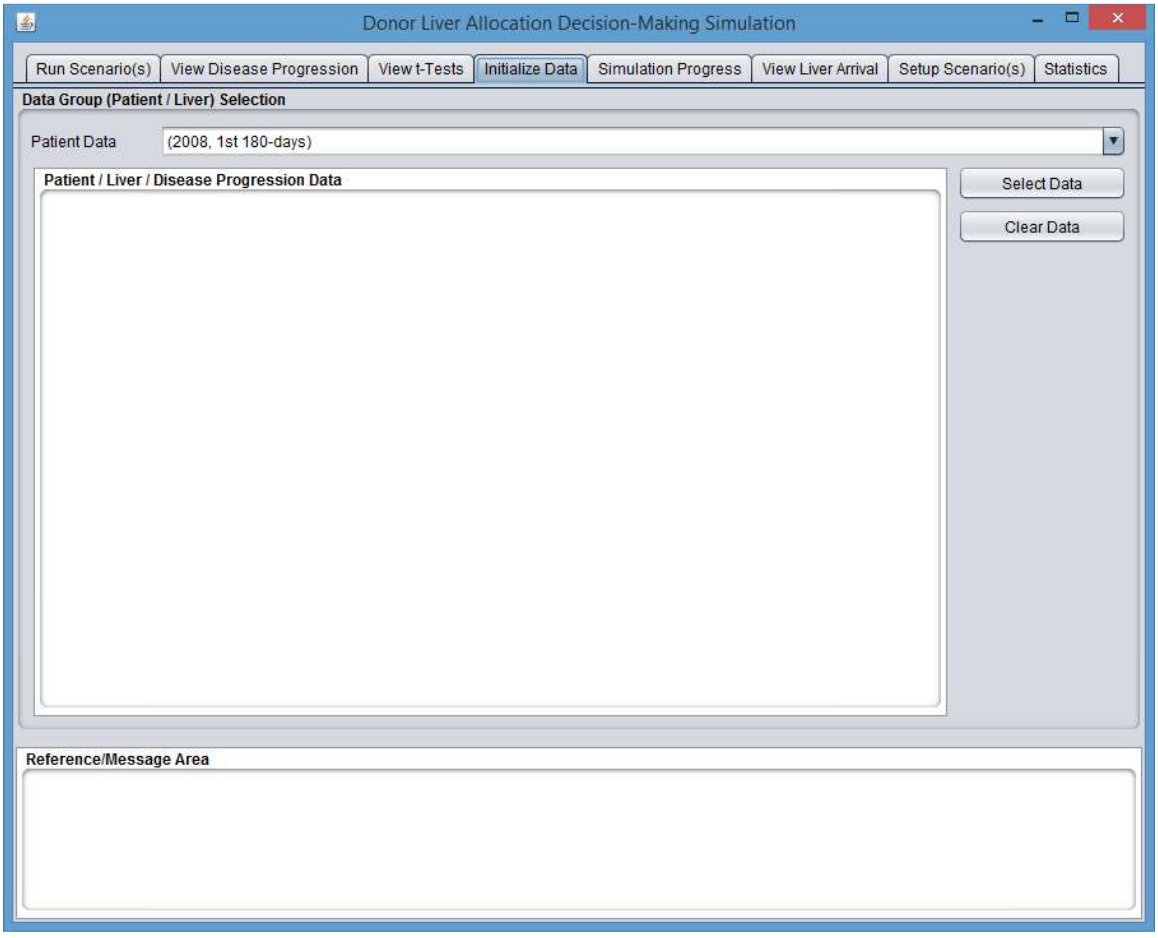


Figure C1. Initialize data panel.

Programming Notes:

1. Upon selection of *Select Data* push button, Patient data that include age group, gender, race, primary cause of disease, transplant history, blood type, MELD scores, date of MELD scores, time on wait list, and status are initialized. This Patient data are translated from the UNOS STAR database depending on the interval selection or pilot dataset.
2. Donor Liver data that include donor age, donor height, donation after cardiac death donors, split liver donors, race, donor's cause of death by cerebrovascular accident, regional sharing, local sharing, and cold ischemia time are initialized. This data are translated from the UNOS STAR database depending on the interval selection or pilot dataset.
3. Patient data also include the additional initialized fields of patient ID, interval ID, and Day_# fields.
4. Donor Liver data also include the additional initialized fields of Donor ID, interval ID, ECD/ECD-1 year/SCD status fields. Day_# is cleared here and set by the *Setup Scenario(s) panel*.
5. Patient data processing is prescreened for Waitlist Entry processing of new patients by ensuring new patients have multiple MELD scores and that their waitlist start dates exist.
6. Patient data processing is prescreened for patients with non-HCC disease, non-status 1 patients, and having completed laboratory or survival data. Hence, there are 100 patient records per user defined interval and 130 donor liver records annually.
7. Patient data processing include prescreening for MELD scores on the day of transplant as well as at least 30 days prior to medical transplant.
8. The AHP decision table parameters are initialized for the simulation of decision-making where the delta-MELD is not used as a criterion according to *AHP Weights and Ranking without delta-MELD* table.

9. The AHP decision table parameters are initialized for the simulation of decision-making where the delta-MELD is used as a criterion according to *AHP Weights and Ranking with delta-MELD*.
10. Initialize the Recipient and Donor Risk factor tables.
11. Patient data, Donor Liver data, and Disease Progression data updated into the *Initialize Data panel* are also updated onto the *Simulation Progress panel* for data analysis of scenario runs and simulation verification of pilot runs.

Setup Scenario(s) Panel

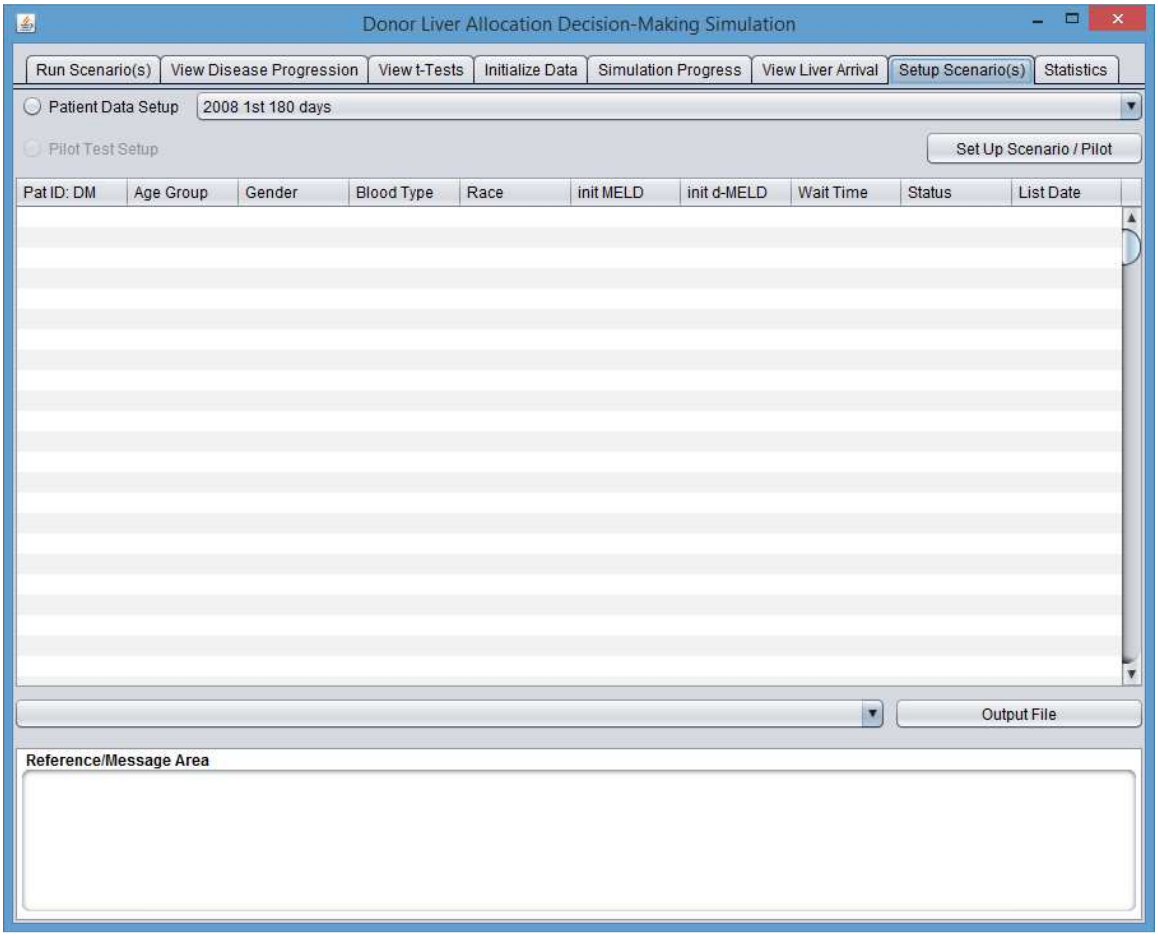


Figure C2. Setup scenario(s) panel.

Programming Notes:

1. Upon selection of (*Pilot Test Setup* radio button or *Patient Data Setup* radio button), and *Set up Scenario / Pilot* push button, perform the following steps.
2. Upon selection of the *Patient Data Setup* radio button and *Set up Scenario / Pilot* push button, the program ensures that the scenarios for including and not including delta-MELD are setup consistently in patient data regarding the patients' arrival times (day_#s) into the simulation event queue.
3. Upon selection of the *Patient Data Setup* radio button and *Set up Scenario / Pilot* push button, the program ensures that the corresponding Donor Liver data are set up

consistently regarding the interarrival times and donor liver parameters for both scenarios including and not including delta-MELD into the event queue. The day_# is processed randomly by a Poisson function.

4. Once Patient and Donor Liver data for a specified time interval are setup, that specified scenario can be run. A new setup would override an existing setup.
5. Upon selection of *Patient Data Setup* radio button and selection of a *180-day interval*, the liver quality of type SCD, ECD 1-year, or ECD is determined.
6. The programming of both Patient and Donor Liver data are setup up consistently, such that patient and donor liver arrival times (day_#s) for both without and with delta-MELD scenarios are the same.
7. The programming of donor liver types are setup randomly but consistently in both scenarios with and without delta-MELD, and the liver types are based on the proportion of OPTN data.
8. The Patient Entry, Donor Arrival Liver, and Waitlist Patient Management simulation process data are setup as event messages queued for the *Run Scenario(s) panel* to be processed.
9. Compute the derived parameter DRI based on the Donor Livers' donor risk factors.
10. The Patient Entry, Donor Arrival Liver, and Waitlist Patient Management event messages are updated onto the *Simulation Progress panel* for data analysis in scenario runs and simulation verification in pilot runs.

Run Scenario(s) Panel

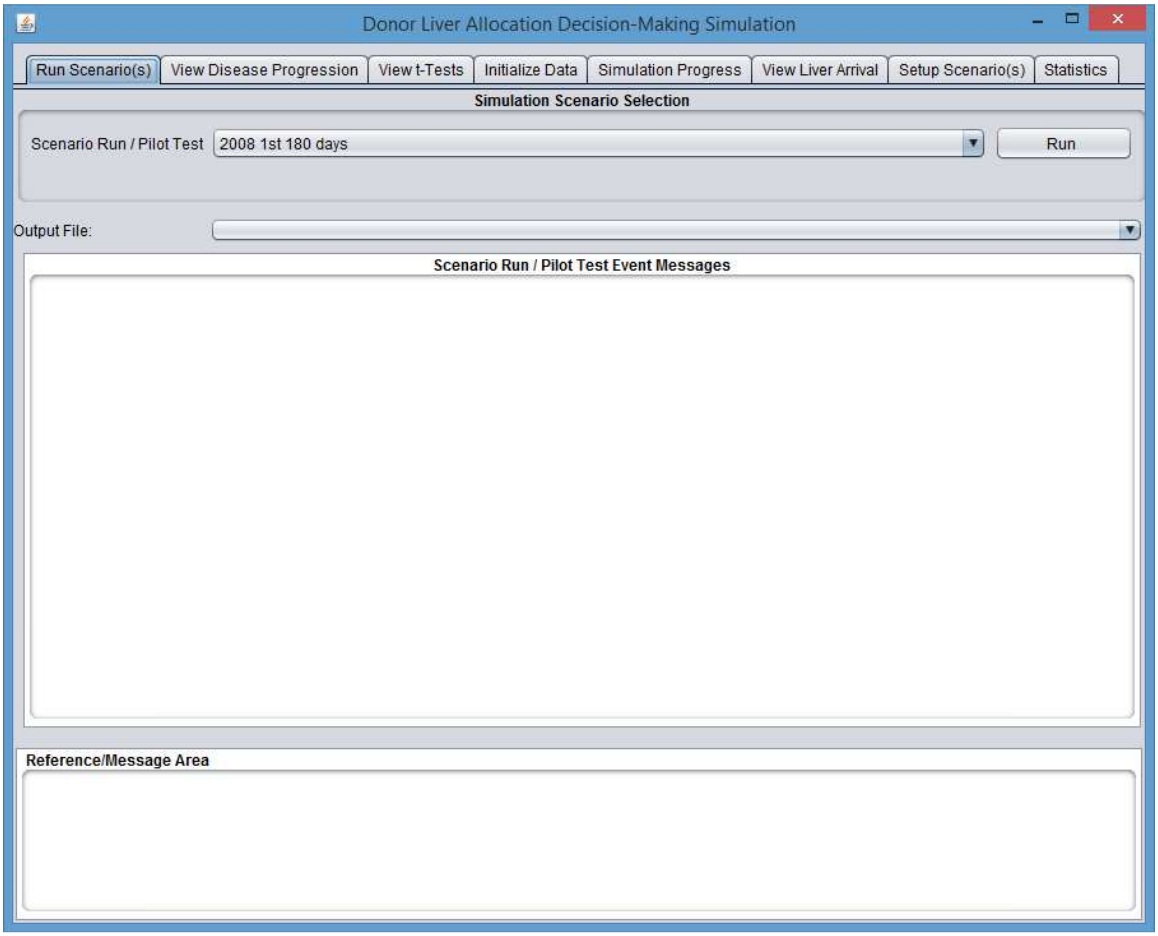


Figure C3. Run scenario(s) panel.

Programming Notes:

1. The *Run Scenario(s) panel* takes scenario event messages (Patient, Liver, or Waitlist Patient Mgmt. messages) and their day_# setup by *Setup Scenario(s),* and run the simulation processes of Waitlist Entry, Donor Liver Arrival, and Waitlist Patient Management processes.
2. Each message processed are updated in the *Scenario Runs / Pilot Test Event Message(s)* text area and *Simulation Progress* text area.
3. The *Run Scenario(s) panel* also processes the simulation Disease Progression process for 180 days (iterations).

Programming Notes (Waitlist Entry):

1. A patient is entered into the waitlist by the day_# (there are 180 days in each scenario and pilot run).
2. Upon patient's waitlist entry, the program ensures that the number of patients entering into the waitlist is being tallied at the end of each week.
3. Patient's waitlist entry along with initial MELD and MELD dates are updated into the *Simulation Progress* text area for data analysis of scenario runs and simulation verification of pilot runs. The *Simulation Progress panel* would tag this data as from the Waitlist Entry process.

Programming Notes (Donor Liver Arrival):

1. The Donor Liver Arrival's output data includes the ECD/SCD status, DRI score, whether a patient has accepted an ECD liver if donor liver is an ECD, patient MELD, AHP scores, and the number of SCD and ECD livers.
2. Upon arrival of an ECD or ECD 1-year, the simulation would determine whether a patient would accept an ECD 1-year or ECD by a random function with a chance of 25% or 15%.
3. The Donor Liver Arrival process computes the AHP scores for all the patients on the waitlist.
4. The parameters of DRI and SOFT scores are computed upon the arrival of a donor liver according to the formula that is based on donor risks and recipient risks. Compute the AHP scores for decision-making where the delta-MELD is not used as a criterion.

$$AHP_{(DM=0)} = Multiplier * (a_1 * MELD_{norm} + a_2 * DRI_{norm} + a_3 * SOFT_{norm}) \quad \text{(C1)}$$

5. Compute the AHP scores for decision-making where the delta-MELD is used as a criterion and where delta-MELD and AHP scores are calculated as follows.

$$deltaMELD = (MELD_t - MELD_{t-1}) \,/\, (time_{t-(t-1)}) \quad \text{(C2)}$$

$$AHP_{(DM=1)} = Multiplier * (b_1 * MELD_{norm} + b_2 * deltaMELD_{norm} + b_3 * DRI_{norm} + b_4 * SOFT_{norm}) \quad \text{(C3)}$$

6. Patients are prioritized on the waitlist according to blood type by descending AHP scores. The livers are offered to the waitlist patient with the highest AHP score, patient consent to accept ECD liver if liver is ECD or ECD 1-year, and blood type identical to the patient. To avoid an inequitable distribution of organs, blood type O livers are only assigned to blood type O patients. If there is a note indicating the patient is a very small size adult patients or AB-type, that patient would be listed for more than one blood type.
7. The *Simulation Progress panel* would be updated of successful patient selection along with patient ID, liver ID, blood-type, liver type, AHP score, MELD, and delta-MELD score for data analysis of scenario runs and simulation verification of pilot runs. This update would be tagged as from the Waitlist Entry process.

Programming Notes (Disease Progression):

1. This Disease Progression process computes the Kalman estimation of MELD and delta-MELD parameters based on the patients' MELD parameters, and patients' covariance matrices.
2. This process references Disease Progression data updated from the *Initialization Data panel* and performs a Kalman propagation for every patient.
3. Every patient's MELD and delta-MELD (Kalman estimation of MELD and delta-MELD) are output into the *Simulation Progress* text area for data analysis of scenario runs and simulation verification of pilot runs. The *Simulation Progress panel* would tag this data as coming from the Disease Progression process along with the day_#.
4. The Disease Progression outputs include patient MELDs, delta-MELDs, and waitlist statuses. Additional outputs include the average MELD scores of patients

on waitlist and the counts of patients removed from the waitlist, *Patients_Dropped_From_Waitlist*, updated at the end of each week for 26 weeks.

5. The risks based on patient risk factors are tallied up and patients with the highest risks are considered removal based on MELD scores and survival rates. The survival rates are based on the *Hazard Ratios based on MELD Table*.
6. The patient waitlist status would be updated indicating whether the patient is still on the waitlist awaiting for an available donor liver, or who has dropped off from the waitlist due to being too sick, death, or other reasons.
7. The response variables are computed at the end of 26 weeks per interval:

$$\text{MELD}_{mean} = (1/26) * \sum_{n=1\ to\ 26\ weeks} Average\_MELD(n) \quad \text{(C4)}$$

$$Total\_Patients\_Removed = \sum_{n=1\ to\ 26\ weeks} Patients\_Dropped\_From\_Waitlist(n) \quad \text{(C5)}$$

8. The *Average_MELD(n)* and *Patients_Dropped_From_Waitlist(n)* are updated into the *Simulation Progress* text area for data analysis of scenario runs and simulation verification of pilot runs for week *n = 1 through 26*. The *Simulation Progress panel* would tag this data as coming from the Disease Progression process along with the week_#s and patient IDs of patients removed.

Programming Notes (Waitlist Patient Management):

1. The Waitlist Patient Management processes the messages indicating patient statuses have been updated which may require these patients to be removed from the waitlist depending on the reason description.
2. The patient IDs and the patient statuses are updated into the *Simulation Progress panel* for data analysis of scenario runs and simulation verification of pilot runs.

View Disease Progression Panel

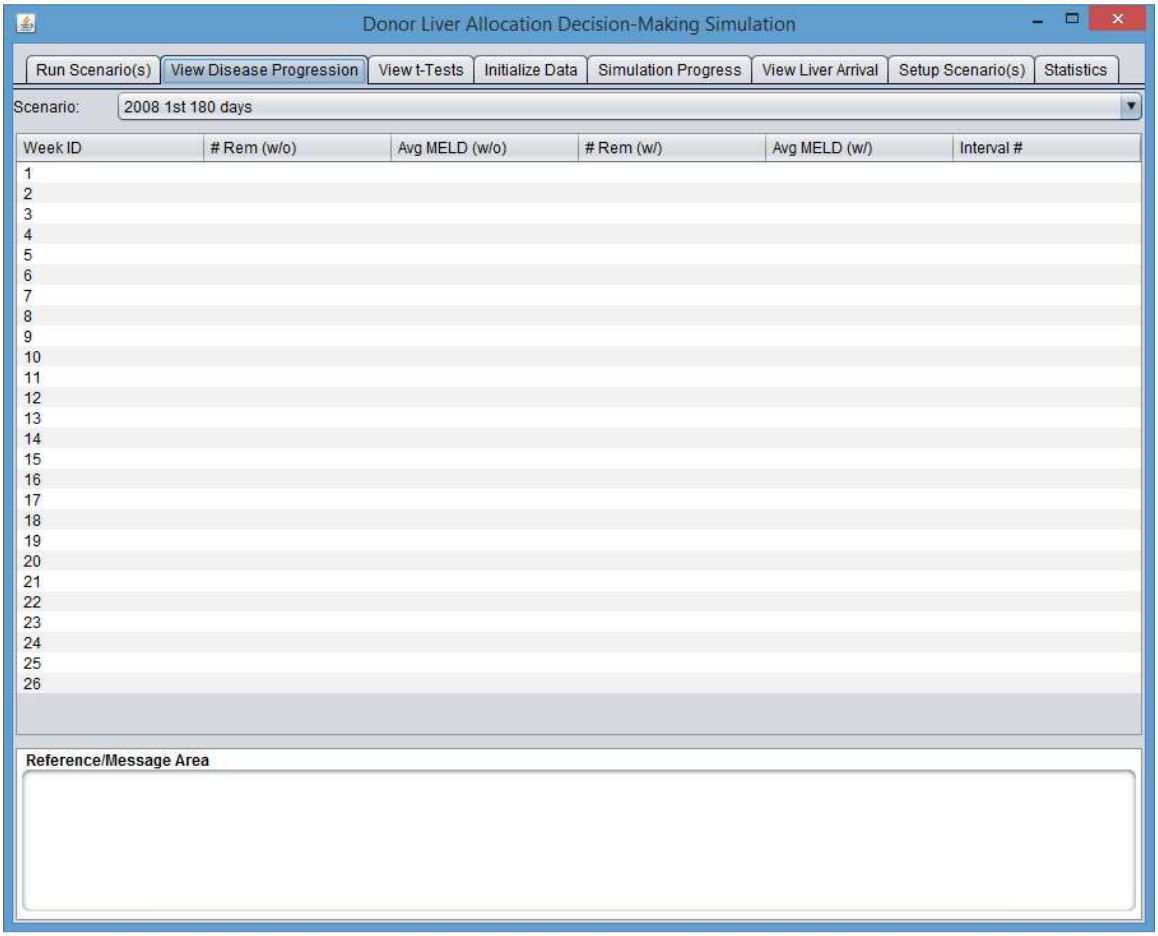


Figure C4. View disease progression panel.

Programming Notes:

1. Disease Progression output data include interval_#, week ID, patients who dropped from the waitlist, and average MELD. These fields are tagged "(wo)" for the scenario without delta-MELD, and "(w/)" for the scenario with delta-MELD.
2. The two dependent variables: *Average_MELD* for "Avg MELD" and *Patients_Dropped_From_Waitlist* for "# Rem" are updated into the *View Disease Progression panel* at the end of each week from the Disease Progression process.
3. The dependent variables from scenario or pilot runs are used to compute response variables $MELD_{mean}$ and *Total_Patients_Removed,* and are updated as the last entry of the *View Disease Progression panel*.
4. Both "Avg_MELD" and "# Rem" dependent variables updated into the *View Disease Progression panel* are also updated into *the Simulation Progress panel* for data analysis of scenario runs and simulation verification of pilot runs.

View Liver Arrival Panel

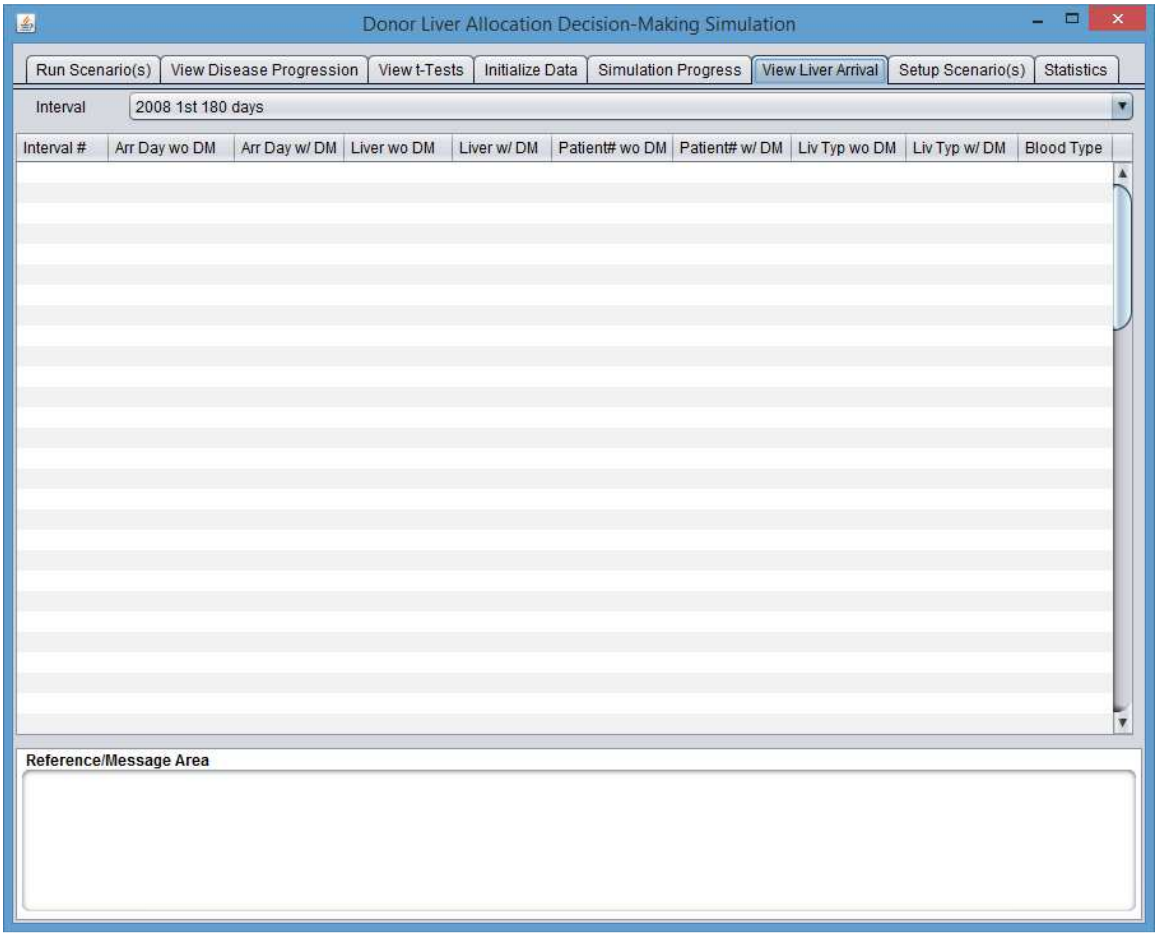


Figure C5. View liver arrival panel.

Programming Notes:

1. To differentiate the data between the scenarios without delta-MELD and with delta-MELD, the Liver ID field is tagged either with "(wo)" or "(w/)". "(wo)" indicates the output is from the scenario without delta-MELD. "(w/)" indicates the output is from the scenario with delta-MELD.
2. The data fields for scenarios with and without delta-MELD include Liver ID, Patient ID, blood type, interval #, and Arrival Day and it is updated upon successful patient selections for transplant from the Donor Liver Arrival process.

3. The data updated into the *View Liver Arrival panel* are also updated into the *Simulation Progress panel's* text area where the data can be updated into a log file for data analysis of scenario runs and simulation verification of pilot runs.

View *t* tests of two independent populations Parameter Panel

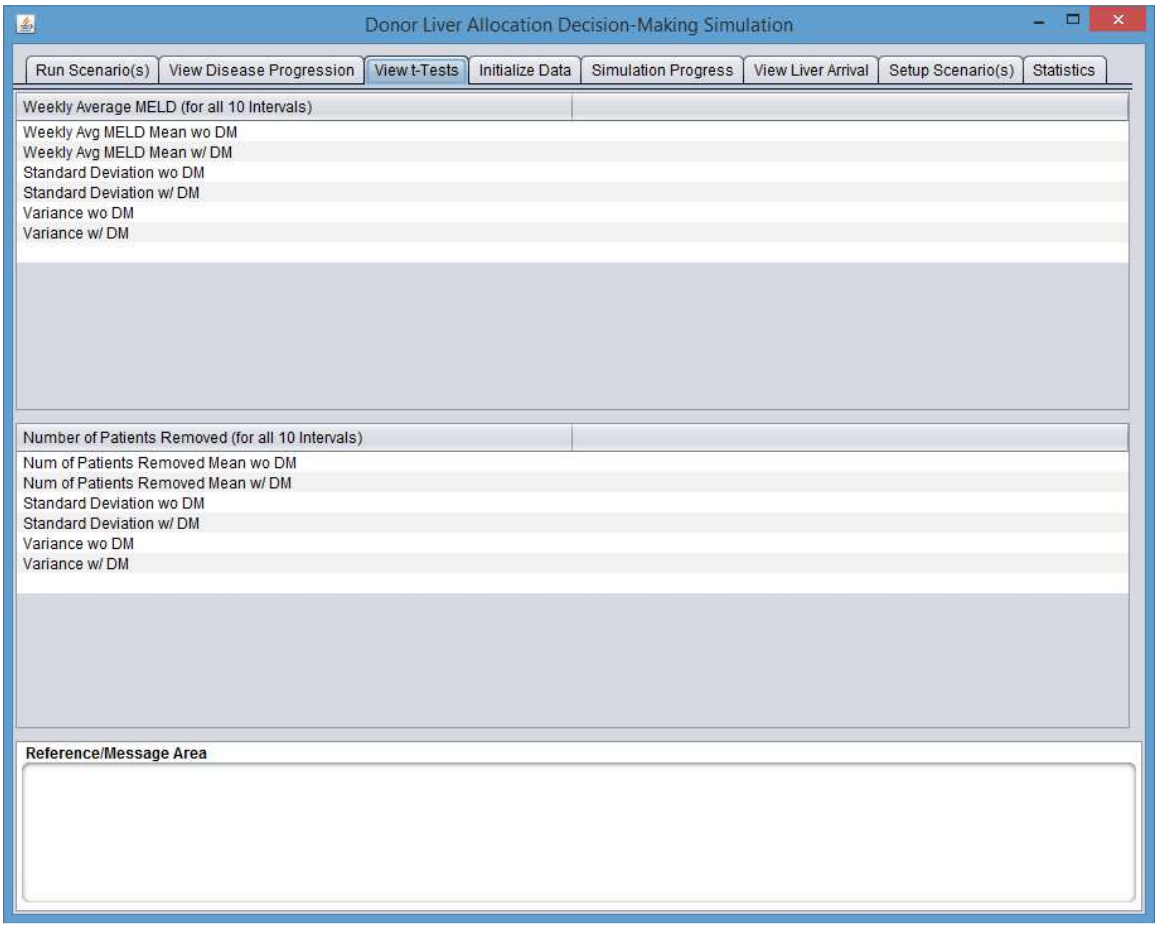


Figure C6. View *t* test of two independent populations' parameters panel.

Programming Notes:

1. Statistical tests are computed upon completion of scenario run.
2. The *t* test of two independent populations parameters for average MELD scores are computed;

$$t_{(MELDmean)} = \frac{(x_1 - x_2)}{\sqrt{S_1^2/N_1 + S_2^2/N_2}} \quad \text{(C6)}$$

$$MELD_{mean} = (1/10) * \sum_{n=1\ to\ 10} [MELD_{mean}(n)]\ (pilot), \quad \text{(C7)}$$

$$MELD_{mean} = (1/70) * \sum_{n=1\ to\ 70} [MELD_{mean}(n)]\ (exp), \quad \text{(C8)}$$

n = 1 to 10 for pilot test, *n = 1 to 70* for experimental run.

x_1 is the $MELD_{mean}$ in scenario without delta-MELD,

x_2 is the $MELD_{mean}$ in scenario with delta-MELD,

S_1 is the standard deviation of $MELD_{mean}$ in scenario without delta-MELD,

S_2 is the standard deviation of $MELD_{mean}$ in scenario with delta-MELD,

S_1^2 is the variance of $MELD_{mean}$ in scenario without delta-MELD,

S_2^2 is the variance of $MELD_{mean}$ in scenario with delta-MELD, and

$$std\ dev_{pilot} = \sqrt{(1/9) \sum_{n=1\ to\ 10} (MELD_{mean}(n) - MELD_{mean})^2} \quad \text{(C9)}$$

n = 1 to 10 in pilot test supporting Figure 11,

$$std\ dev_{exp} = \sqrt{(1/69) \sum_{n=1\ to\ 70} (MELD_{mean}(n) - MELD_{mean})^2} \quad \text{(C10)}$$

n = 1 to 70 in experimental run supporting Figure 12.

The *t* test of two independent populations parameters are computed for the *Total_Patients_Removed* response variable;

$$t_{(PatientsRemoved)} = \frac{(x_1 - x_2)}{\sqrt{S_1^2/N_1 + S_2^2/N_2}} \quad \text{(C11)}$$

$$Total\_Patients\_Removed_{avg} = (1/10) * \sum_{i=1\ to\ 10} (Total\_Patients\_Removed(i))\ (pilot), \quad \text{(C12)}$$

$$= (1/70) * \sum_{i=1\ to\ 70} (Total\_Patients\_Removed(i))(exp), \quad \text{(C13)}$$

i = 1 to 10 for pilot test, *i = 1 to 70* for experimental run.

x_1 is the *Total_Patients_Removed* in scenario without delta-MELD,

x_2 is the *Total_Patients_Removed* in scenario with delta-MELD,

S_1 is the standard deviation of *Total_Patients_Removed* without delta- MELD,

S_2 is the standard deviation of *Total_Patients_Removed* with delta-MELD, and

$$T_{avg} = Total\_Patients\_Removed_{avg} \quad \text{(C14)}$$

$$T(n) = Total\_Patients\_Removed(n), \text{ } n \text{ is the interval index,} \quad \text{(C15)}$$

$$std\ dev_{pilot} = \sqrt{(1/9)\sum_{n=1\ to\ 10}[T(n) - T_{avg}]^2} \quad \text{(C16)}$$

n = 1 to 10 in pilot test supporting Figure 11,

$$std\ dev_{exp} = \sqrt{(1/69)\sum_{n=1\ to\ 70}[T(n) - T_{avg}]^2} \quad \text{(C17)}$$

n = 1 to 70 in experimental run supporting Figure 12.

S_1^2 is the variance of *Total_Patients_Removed* without delta-MELD,

S_2^2 is the variance of *Total_Patients_Removed* with delta-MELD,

3. The data updated onto the *View t test of two independent populations panel* are also updated into the *Simulation Progress* text area for data analysis of scenario runs and simulation verification of pilot runs.

Simulation Progress Panel

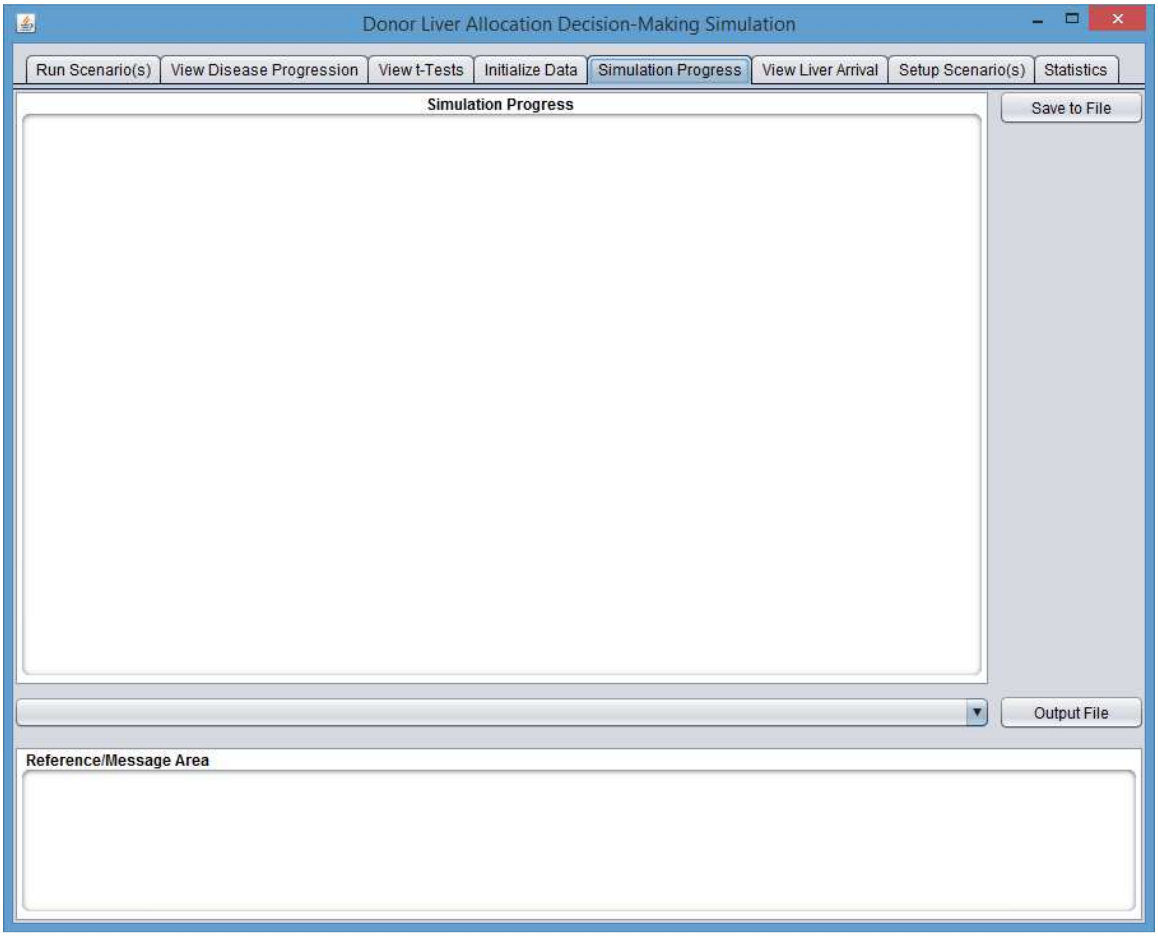


Figure C7. Simulation progress panel.

Programming Notes:

1. Each messages updated into the *Simulation Progress panel* have the tag of *Initialize Data*, *Setup Scenarios(s)*, *Run Scenario(s)*, *View Liver Arrival*, *View Disease Progression* depending on the source of panel output.
2. When messages are generated as a result of the *Run Scenario(s)*, the simulation process names are also tagged (Waitlist Entry, Disease Progression, Donor Liver Arrival, and Waitlist Patient Management).
3. The selection of *Save to File* push button logs all messages from startup of the simulation GUI into the designated output file.
4. The *Message Area* text area allows the user to view completion statuses of user requests and of any simulation's informational, warning, or error messages.

Appendix D: Statistics Notes

The *F*-ratio test was used for testing homogeneity of variances. From an example of running the pilot test, the standard deviations were computed from Figure D1. These standard deviations were squared to yield their variances and these variances were set into an *F*-ratio, where the smaller of the two variances is the denominator, and the larger of the two variances is the numerator. The critical value for 25 degrees of freedom for both variances is 1.35 according to the F distribution critical values table provided by Aczel and Sounderpandian (2008). From Figure D1, the variances were computed from the 2008 1st 180 days interval into an *F-ratio*. The *F*-ratio is $(1.5363)^2 / (1.3861)^2 =$ 1.228 and is below the critical value of 1.35. Since the *F-ratio* of 1.228 is not greater than the critical value of 1.35, this indicates the variances of the average MELD scores are homogeneous.

Aczel and Sounderpandian (2008, p. 311) provided the formula for the test statistic Z, for the comparison of two populations, where the hypothesized value for the difference in the two population means is $(\mu_1 - \mu_2)$.

$$t = \frac{(x_1 - x_2) - (\mu_1 - \mu_2)}{\sqrt{S_1^2/N_1 + S_2^2/N_2}} \qquad \text{(D1)}$$

The research experiment was about a comparison of MELD scores based on the same set of arriving donor livers that were applied to the same pool of 100 patients, one scenario without delta-MELD, one scenario with delta-MELD. The MELD scores were normally distributed according to a similar study of Kanwal, Dulai, Speigel, Yee, and Gralnek. (2005). Aczel and Sounderpandian (2008) stated that in the case where X_1 and

X_2 each follows a normal distribution, (X_1 - X_2) would also follow a normal distribution. It can be verified whether (X_1 - X_2) assumed a normal distribution (p. 311).

Regarding whether the difference in average MELD scores between the scenarios with and without delta-MELD is normally distributed, the chi-square test for normality was utilized. Aczel and Sounderpandian (2008) explained that a chi-square goodness of fit method requires hypothesizing about the sample set with null and alternative hypotheses, computing frequencies of where the null hypothesis is expected, providing the expected counts of data points into different chi-square bins, and computing the difference between the observed and expected data leading to the chi-square statistics (p. 662). Aczel and Sounderpandian (2008) further explained that a goodness of fit test is a statistical test that tells whether data would support an assumption relating to a distribution or random variable. The simulation computed the difference in MELD averages against its mean, the mean and standard deviation of these differences, and the z-values that would provide the bin values for the goodness of fit chi-square bins. The following formula is the chi-square statistics.

$$\chi^2 = \sum(O_i - E_i)^2 / E_i, \quad \text{(D2)}$$

for i= 1 to k, where k is the number of fit cells.

The chi-square hypotheses are as follows:

H_o: The differences of average MELD scores are normally distributed.

H_a: The differences of average MELD scores are not normally distributed.

The differences of average MELD means, standard deviations, and variances are computed. With 26 weeks of average MELD scores, 5 bins are set up, where each bin

would have 1/5 = 0.20 probability. The bin boundaries of 0.20 probability are translated to their corresponding z-values which then provide the z value intervals of < -0.84, (-0.84, -0.255), (-0.255, 0.255), (0.255, 0.84), and >0.84. From the formula of

$$z = (x - xbar) / s \quad \text{(D3)}$$

Where *s* is the standard deviation, and *x-bar* is the mean, the formula is rewritten in terms of *x*,

$$x = s * z + xbar \quad \text{(D4)}$$

Table D1 supports the computation of chi-square value based on the differences of average MELD scores falling into the appropriate bins.

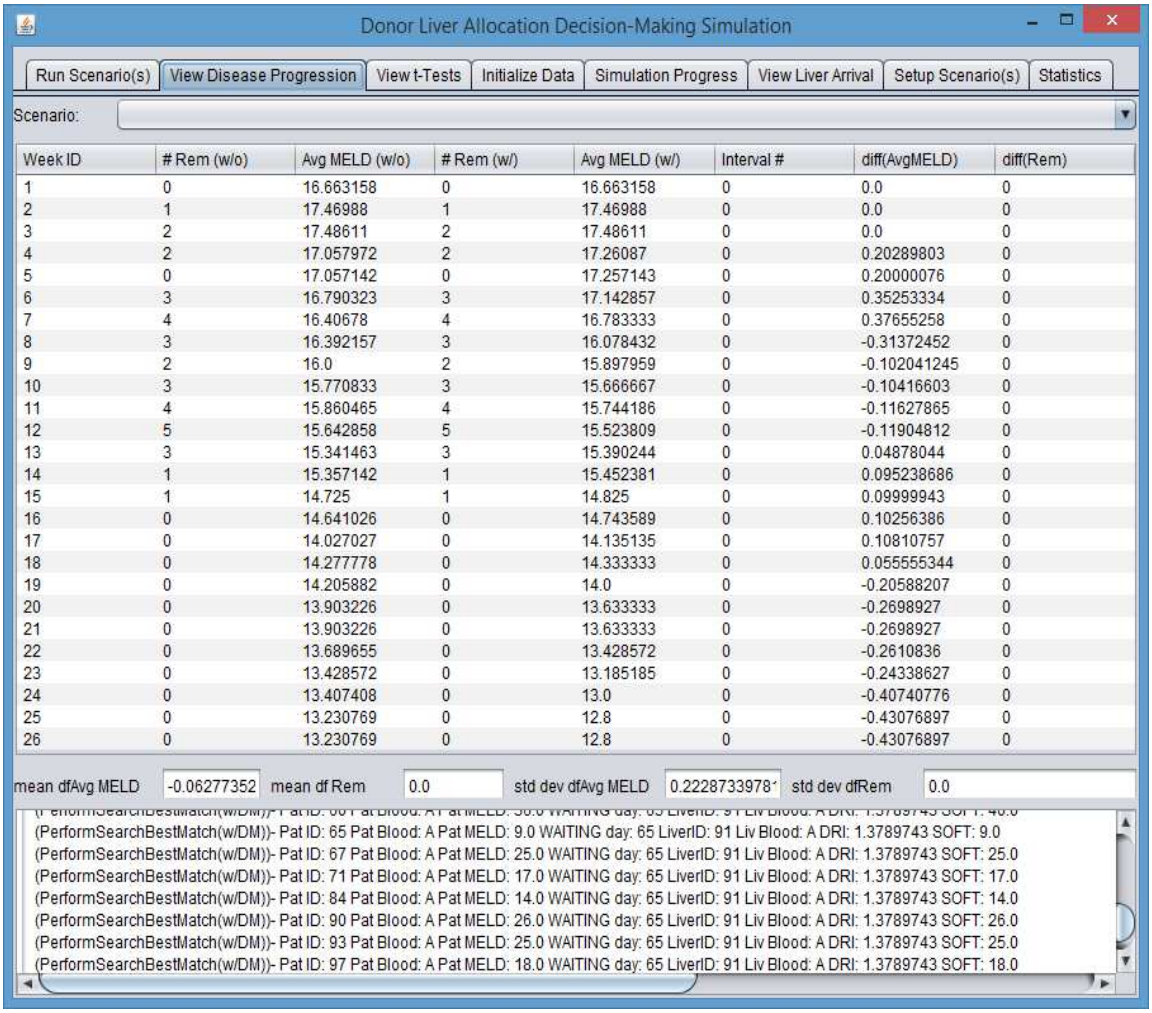


Figure D1. Goodness of fit-test for normality.

In Figure D1 example, the differences in MELD averages were computed and they were counted and allocated into Table D1. Table D1 was set up to determine the chi-square goodness of fit for the normal distribution. The differences of average MELDs were computed, along with their mean and standard deviation. These values were utilized to support a *Goodness of Fitness* computation by setting up bin boundary values for each of the five bins. Once these boundary values were computed for each bin, a count of the differences of average MELDs can be performed and be grouped into their

appropriate bins. Table D1 identifies the computational steps necessary to carry out the chi-square *Goodness of Fit* test for normality.

Table D1

Chi-square Goodness of Fit test for Normality

Bin #	Interval	Observed Frequency (f)	Expected Frequency (e)	$(f - e)$	$(f - e)^2$	$(f - e)^2/e$
1	<-0.249998	7	5	2	4	4/5
2	(-0.249998, -0.119605)	6	5	1	1	1/5
3	(-0.119605, -0.005940)	1	5	-4	16	16/5
4	(-0.005940, 0.124440)	7	5	2	4	4/5
5	>0.124440	4	5	-1	1	1/5

The intervals were computed, with

$$xbar = -0.062773, \tag{D5}$$

$$s = 0.222873, and \tag{D6}$$

$$z = -0.84, -0.255, -0.255, 0.84. \tag{D7}$$

$$x_i = s * z + xbar, \tag{D8}$$

where *i* is the interval boundary index,

$$x_1 = (0.222873)(-0.84) + (-0.062773) = -0.249998 \tag{D9}$$

$$x_2 = (0.222873)(-0.255) + (-0.062773) = -0.119605 \tag{D10}$$

$$x_3 = (0.222873)(0.255) + (-0.062773) = -0.005940 \tag{D11}$$

$$x_4 = (0.222873)(0.84) + (-0.062773) = 0.124440 \tag{D12}$$

$$\chi^2 = ((7-5)^2 + (6-5)^2 + (1-5)^2\,(7-5)^2 + (4-5)^2)/\,5$$

$$= (2^2 + 1^2 + 4^2 + 2^2 + 1)\,/\,5$$

$$= (4 + 1 + 16 + 4 + 1)\,/\,5$$

$$= (21 + 5)\,/\,5 = 26\,/\,5$$

$$= 5.2 \qquad \text{(D13)}$$

For this example, with 4 degrees of freedom, the chi-square value of 5.2 lies between the 0.05 and .95 confidence level of the chi-square region, where the corresponding chi-square value is between 1.06336 and 7.779. The goodness of fit test did not suggest rejecting the chi-square null hypothesis. Hence, one can conclude that there is not enough evidence to claim that the MELD scores are not normally distributed.