

DYNAMIC PREDICTION OF PRE-TRANSPLANT SURVIVAL IN A MULTI-CENTER STUDY OF PEDIATRIC ACUTE LIVER FAILURE

by

Drew Michael Strader Donnell

B.S., University of Michigan, 2010

Submitted to the Graduate Faculty of
the Graduate School of Public Health in partial fulfillment
of the requirements for the degree of
Master of Science

University of Pittsburgh

2014

UNIVERSITY OF PITTSBURGH
GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Drew Michael Strader Donnell

It was defended on

April 25, 2014

and approved by

Thesis Advisor:

Chung-Chou H. Chang, PhD
Associate Professor, Departments of Medicine, Biostatistics, and Clinical and Translational Science,
School of Medicine and Graduate School of Public Health
University of Pittsburgh

Committee Members:

Cindy L. Bryce, PhD
Associate Professor, Departments of Health Policy and Management,
Medicine, and Clinical and Translational Science
Graduate School of Public Health and School of Medicine
University of Pittsburgh

Jeanine L. Buchanich, PhD
Research Assistant Professor, Department of Biostatistics
Graduate School of Public Health
University of Pittsburgh

Mark S. Roberts, MD
Professor, Departments of Health Policy and Management,
Medicine, Industrial Engineering, and Clinical and Translational Science
Graduate School of Public Health, School of Medicine, and Swanson School of Engineering
University of Pittsburgh

Andrew J. Schaefer, PhD
W.K. Whiteford Professor, Departments of Industrial Engineering,
Bioengineering, Medicine, and Clinical and Translational Science
Swanson School of Engineering and School of Medicine
University of Pittsburgh

Copyright © by Drew Michael Strader Donnell

2014

DYNAMIC PREDICTION OF PRE-TRANSPLANT SURVIVAL IN A MULTI-CENTER STUDY OF PEDIATRIC ACUTE LIVER FAILURE

Drew Michael Strader Donnell, M.S.

University of Pittsburgh, 2014

ABSTRACT

Acute liver failure (ALF) is a clinical syndrome characterized by the rapid onset of illness and disruption of critical hepatic processes. The natural history and clinical recognition of ALF in children differs considerably from that observed in adults largely due to both heightened etiologic variation and delayed onset of clinical encephalopathy within the pediatric population. Despite efforts to implement multidisciplinary management strategies and understand optimal timing of orthotopic liver transplantation (OLT), current prognostic models are unreliable and fail to identify high-risk patients. We propose a dynamic prediction model of pre-transplant survival for pediatric patients with ALF, specifically to inform the sequential medical decision making process and consequently improve clinical outcomes. **Public Health Significance:** Dynamic prediction models are of great interest to clinicians and patients alike, enabling well-informed decisions in light of the unpredictable nature of clinical and pathophysiological systems. Extensions of our model may be utilized to facilitate proper allocation of scarce resources, such as donor organs.

Keywords: Acute liver failure; dynamic prediction; landmark analysis; liver transplantation; medical decision making; personalized medicine; survival analysis.

TABLE OF CONTENTS

PREFACE	X
1.0 INTRODUCTION	1
1.1 OVERVIEW	1
1.2 PREDICTION MODELS IN MEDICINE	5
1.3 PEDIATRIC ACUTE LIVER FAILURE (PALF)	6
1.4 PALF STUDY	7
1.5 REVIEW OF PALF MODELING LITERATURE	8
2.0 MATERIALS AND METHODS	11
2.1 STATISTICAL CONCEPTS	11
2.1.1 Estimation of Marginal Survival	11
2.1.2 Estimation of Conditional Survival	13
2.1.3 Crude Landmark Model	14
2.1.4 Crude Landmark Model with Time-Dependent Covariates	15
2.1.5 Stratified Landmark Supermodel with Time-Dependent Covariates	16
2.1.6 Proportional Baselines Landmark Supermodel with Time-Dependent Covariates	17
2.2 PALF STUDY DATA	18
2.2.1 Hospital Evaluation Form	19
2.2.2 Follow-up Form	19
2.2.3 Follow-up Vital Status Form	20
2.2.4 Hospital Flow Sheet Form	20

2.2.5	Summary Descriptive File.....	21
2.3	CONSTRUCTION OF ANALYSIS DATASET	21
2.3.1	Selection of Candidate Covariates	21
2.3.2	Creation of Super Prediction Dataset	22
2.4	ANALYSIS	25
2.4.1	Descriptive Statistics.....	25
2.4.2	Univariable Cox PH Models	26
2.4.3	Multivariable Cox PH Models.....	26
2.4.4	Proportional Baselines Landmark Supermodel	27
2.4.5	Prediction of Conditional Survival and Fixed Window Risk Estimates...	28
2.4.6	Evaluation of Model Performance	28
3.0	RESULTS	29
3.1	DESCRIPTIVE STATISTICS	29
3.2	UNIVARIABLE COX PH MODELS	33
3.3	MULTIVARIABLE COX PH MODELS	35
3.4	PROPORTIONAL BASELINES LANDMARK SUPERMODEL	37
3.5	PREDICTION OF CONDITIONAL SURVIVAL AND FIXED WINDOW RISK ESTIMATES	40
3.6	EVALUATION OF MODEL PERFORMANCE	46
4.0	DISCUSSION	51
	APPENDIX A : ALTERNATIVE LANDMARK SUPERMODELS	54
	APPENDIX B : R CODE	59
	BIBLIOGRAPHY	73

LIST OF TABLES

Table 1: Summarization of the published literature on predictions models of PALF.	10
Table 2: Patient demographics, clinical data, and laboratory values for the analysis cohort of PALF patients at the time of study enrollment.	31
Table 3: Regression parameter estimates from the univariable time-dependent Cox PH models of pre-transplant survival in PALF patients.	34
Table 4: Regression parameter estimates from the multivariable time-dependent Cox PH models of pre-transplant survival in PALF patients.	36
Table 5: Regression parameter estimates from the proportional baselines landmark supermodel of pre-transplant survival in PALF patients.	38
Table 6: Bootstrapped estimates of dynamic AUC for the proportional baselines landmark supermodel.	47

LIST OF FIGURES

Figure 1: Gross simplification of PALF natural history and the sequential decision making process.....	3
Figure 2: Simplified decision tree for each epoch in the sequential decision making process.....	4
Figure 3: Graphical representation of the data management process.	24
Figure 4: Kaplan-Meier estimate of survival in the analysis cohort of PALF patients.	30
Figure 5: Spaghetti plots of the four continuous predictors.....	32
Figure 6: Baseline hazard and landmark effects in the proportional baselines landmark supermodel.....	39
Figure 7: The trajectory of three-day conditional mortality probabilities for the “average patient”	41
Figure 8: Non-random selection of personalized conditional mortality trajectories for 4 censored PALF patients.	42
Figure 9: Non-random selection of personalized conditional mortality trajectories for 4 deceased PALF patients.	43
Figure 10: Distributions of conditional three-day mortality probability estimates and log mortality probably probability estimates aggregated over all landmark time points s	44
Figure 11: Distributions of conditional three-day mortality log probability estimates, stratified by landmark time point s	45
Figure 12: Bootstrapped estimate of AUC for the landmark supermodel, aggregated across all landmark time points s	47

Figure 13: Bootstrapped estimates of dynamic AUC for the landmark supermodel at distinct landmark time points s 49

Figure 14: Bootstrapped estimates of dynamic AUC for the landmark supermodel plotted as a function of landmark time point s 50

PREFACE

I would like to express my sincere gratitude to my ensemble of advisors and mentors, for it has truly been an honor to learn from every one of you. Looking back to the infancy of my experience in Pittsburgh, I simply cannot thank Mark Roberts and Cindy Bryce enough for making this wonderful experience possible in the first place. This thesis would not have been possible without the advising and insight I received from Joyce Chang, who also provided unwavering support from the beginning. I must thank Andrew Schaefer and Jeanine Buchanich for their willingness to join my thesis committee and provide reassurance along the way. I cannot forget to mention a quartet of faculty members who comforted and encouraged me throughout the process: Ken Smith, Lisa Weissfeld, Jonathan Yabes, and Ada Youk.

In addition to great faculty relationships, I am blessed to have compassionate relationships with my brilliant colleagues and classmates: Ark Fang, Heather Tomko, Kelsey Baker, and Jordan Zarone, to highlight the bunch. No letter of gratitude would be complete without mentioning my loving family. To my sisters Aubrey, Genevieve, and Teagan and my father Michael: thank you for the unique qualities and unconditional love each of you have provided to this point in my life. To my mother Cathy, my role model: I dedicate this work to you. My accomplishments are just as much yours as they are mine.

On lesser notes: I would like to thank Layla Parast for her contributions to the landmark methodology and kind words during a quick visit to Pittsburgh; I would also like to recognize Richard Hughes for sparking my initial interest in biomedical research when I was a senior in high school and providing honest mentorship thereafter. Finally, I must thank the entire Pediatric Acute

Liver Failure Study Group for the financial support and opportunity to pursue one of my lifelong interests, medical decision making.

1.0 INTRODUCTION

1.1 OVERVIEW

Acute liver failure (ALF) is a clinical syndrome characterized by the rapid onset of illness and disruption of critical hepatic processes. The natural history and clinical recognition of ALF in children differs considerably from that observed in adults largely due to both heightened etiologic variation and delayed onset of clinical encephalopathy within the pediatric population. Emergency orthotopic liver transplant (OLT) is the most effective intervention among limited therapies but is not without risk and long-term implications: reduced life expectancy and lifelong immunosuppression (McDiarmid, 2007; SPLIT Research Group, 2001; Feng, 2008). While children with ALF receive priority in the allocation according to policies established by the United Network for Organ Sharing (UNOS), access to quality donors is restricted by the overwhelming demand of patients with chronic illnesses. The medical decision making process is further complicated by the high rate of spontaneous recovery; an estimated 50% of patients with pediatric acute liver failure (PALF) return to normal levels of health within weeks (Lu, 2013). Despite extensive investigation of the underlying pathophysiological mechanisms and optimal therapeutic strategies, the short-term outcome for the population of PALF patients remains poor (Devictor, 2011; Ee, 2003; Squires, 2006). The status quo, specifically the absence of a standard method for evaluating illness severity or mortality, indicates an exciting opportunity to improve upon existing

methods of evidence-based patient evaluation and contribute to the growing body of acute pediatric care literature. The purpose of our study is to develop a dynamic prediction model of pre-transplant survival for pediatric patients with ALF, specifically to inform the sequential medical decision making process ([Figures 1, 2](#)).

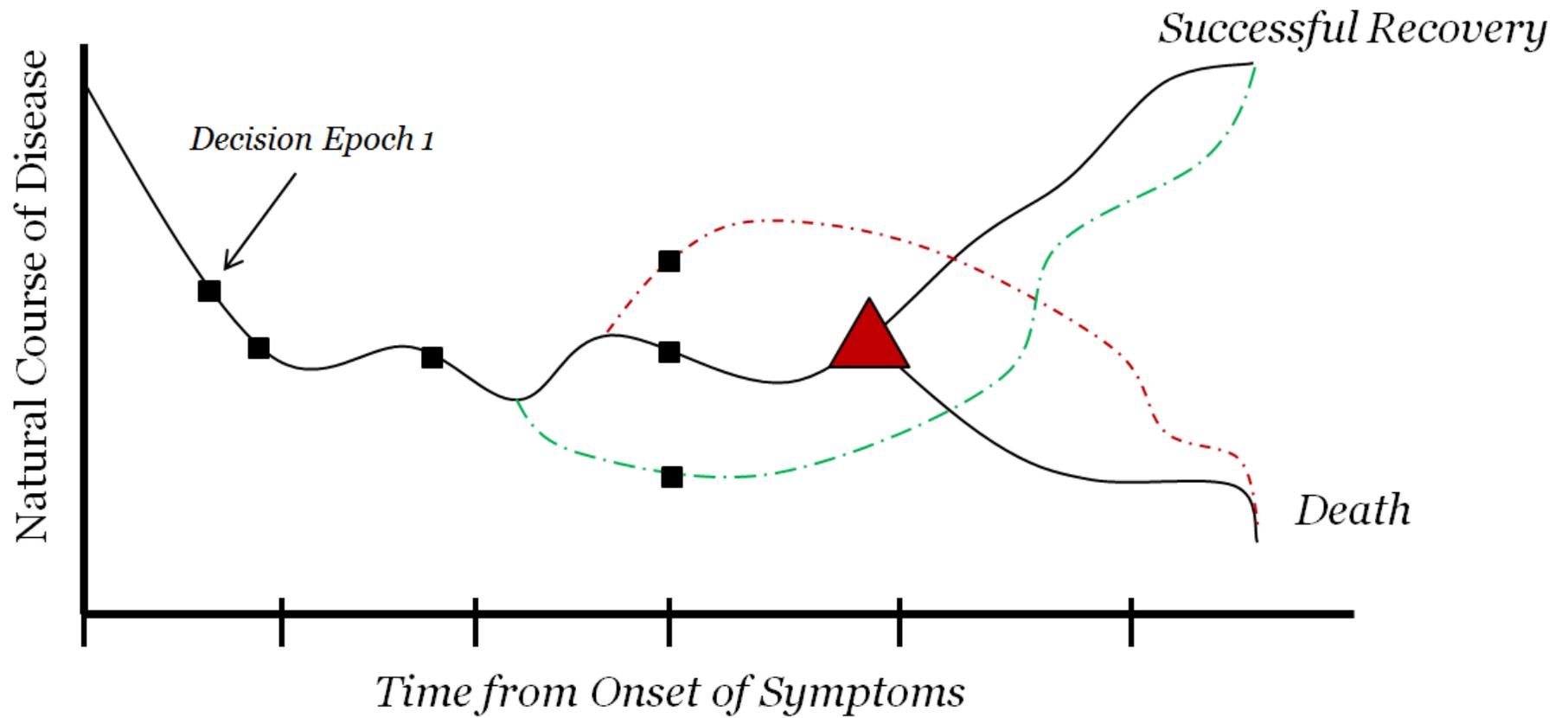


Figure 1: Gross simplification of PALF natural history and the sequential decision making process. Each square represents a single decision epoch.

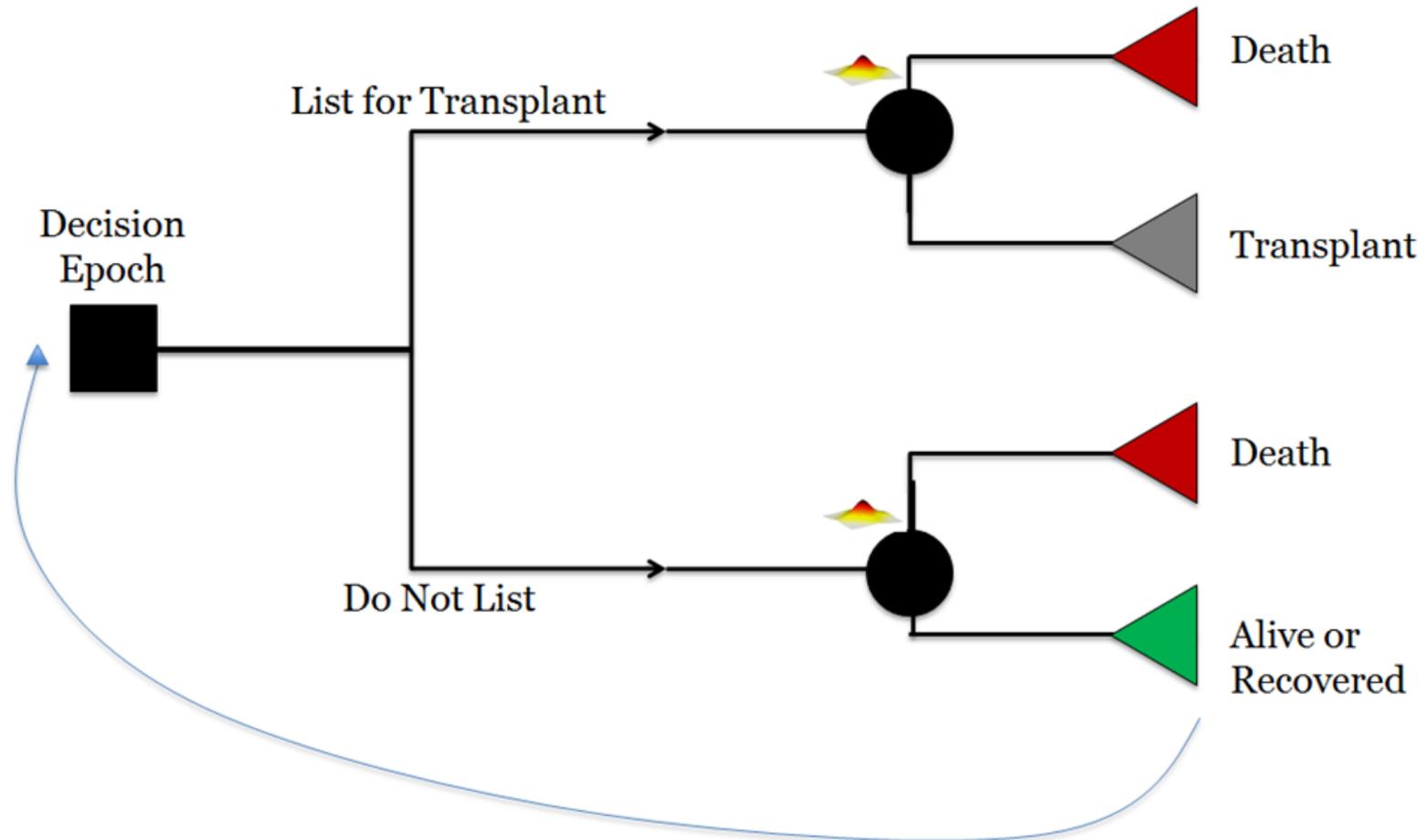


Figure 2: Simplified decision tree for each epoch in the sequential decision making process. Each circle represents a chance node. *Death* and *transplant* are absorbing states; realization of the non-absorbing state *alive or recovered* results in progression to the preceding epoch.

1.2 PREDICTION MODELS IN MEDICINE

The notion of prognosis, the probable course of a condition or disease, is central to clinical practice, medical research, public health, and countless disciplines of the biomedical sciences (Eagle, 2004; Friedman, 2002; Visser, 2002). Historical evidence of prognostication dates back to the early works of Hippocrates in 5th century BC (Garrison, 1966), a period seemingly dissimilar to the evidence-based era of Western medicine. Rather, the observed exponential growth of publications on prognostic modeling between 1970 and 2005 suggests the transition from subjective to objective evaluation of prognosis is a recent phenomenon (Steyerberg, 2009).

This application of the scientific method to prognostic modeling and strategies devised to improve prognosis on a population level (such as screening procedures, diagnostic tests, and therapies) reinforces emerging evidence-based standards and enhances shared medical decision making. The addition of a quantitative analyst to the joint patient-physician decision construct facilitates the communication of technical concepts and mitigates biased reporting of risks. Furthermore, the integration of quantitative analysts into clinical analysis provides a basis for development of the increasingly complex models required to address the issues of modern clinical medicine.

The development of personalized risk prediction models represents an ideal opportunity for such interdisciplinary efforts. Patients of similar identifiable clinical characteristics do not necessarily experience similar clinical outcomes. Except for random uncertainty, one of the reasons for this heterogeneity is contributed by the time-dependent characteristics, which include nonlinear prognosis over the course of time, and time-updated information on clinical events and

test results. Standard clinical prediction models are not suitable for such complicated designs. Within the context of survival, a set of markedly similarly patients may display statistically significant differences in time-to-event profiles (Dent 2007, Goldhirsch 2010). To reduce heterogeneity in survival estimation among patients with similar baseline characteristics, a method was recently developed to incorporate short-term events into the prediction of long-term survival estimation (Parast, 2013).

Dynamic prediction is an estimation procedure that accounts for time-dependent characteristics. van Houwelingen and colleagues proposed dynamic prediction models based on a landmarking method. With these models one not only can incorporate time-dependent information into risk prediction but also can efficiently make prediction at a series of predetermined time points (Nicolaie, 2013; van Houwelingen, 2007, 2008; van Houwelingen and Putter, 2012). Moreover, the landmarking method van Houwelingen introduced is robust against misspecification of the proportional hazards assumption. In this thesis, we applied these landmark dynamic prediction models to estimate the probability of survival at a set of future time points based on the collected time-dependent information for children with ALF.

1.3 PEDIATRIC ACUTE LIVER FAILURE (PALF)

Acute liver failure (ALF) is a clinical syndrome characterized by the rapid onset of illness and disruption of critical hepatic processes. Often health deteriorates within a period of days, resulting in the immediate need for intensive medical treatment. The accurate estimation of incidence is difficult, as with many rare conditions, yet its impact on the transplant community is clear: adults and children with ALF account for a disproportionate number of deceased-donor

liver transplants after adjusting for the number of listings ([ANZLTR, 2010](#); [Rajanayagam, 2013](#)). Causes of pediatric acute liver failure include accidental acetaminophen overdose, viral hepatitis, metabolic errors, ischemia, and indeterminate in roughly 50% of the population ([Squires, 2008](#)). The delayed onset of hepatic encephalopathy (HE) until terminal stages of natural history opposes the classic criteria for the diagnosis of ALF. Thus the management of PALF requires a pediatric-specific definition and an advanced framework of etiologic assessment. Marginal prognosis is better in children than adults but conditional prognosis indicates significant disparities among select clinical features ([Squires, 2008](#)). Accordingly, the volume of PALF research is growing with improved understanding of the pathophysiology, therapeutic interventions, influential clinical predictors, and outcome profiles. Unfortunately, short-term outcomes remain bleak in spite of efforts to implement multidisciplinary management strategies and understand optimal timing of orthotopic liver transplantation: PALF results in mortality or resource-transplant in up to 45% of patients ([Lu, 2013](#)).

1.4 PALF STUDY

The primary objective of the PALF Study is to collect, maintain, analyze, and report clinical, epidemiological, and outcome data in children with ALF, including information derived from biospecimens. A secondary objective is to develop data-driven methods to predict the likelihood of a child spontaneously recovering without requiring OLT. The PALF Study Group began collecting prospective patient data in December 1999, following the operationalization of the clinical syndrome and approval of institutional review boards at 20 unique international clinical sites (n=958), 17 within the borders of the United States (n=756). Now in its second phase of

funding from the National Institute of Diabetes and Digestive and Kidney Diseases, the study currently supports 12 sites in North America. The PALF working group, consisting of 21 investigators, defines the entry criteria for children under the age of 18 years: (1) no known evidence of chronic liver disease; (2) biochemical evidence of acute liver injury; and (3) hepatic-based coagulopathy defined as a prothrombin time (PT) ≥ 15 seconds or international normalized ratio (INR) ≥ 1.5 not corrected with vitamin K in the presence of clinical HE or a PT ≥ 20 seconds or INR ≥ 2.0 regardless of the presence or absence of clinical HE (Squires, 2006). Collection of demographic, clinical, and laboratory information for up to 7 consecutive days begins following enrollment. Upon informed consent from a parent or legal guardian, the clinical diagnosis is assessed and classified in accordance to one of the numerous etiologies such as acetaminophen-overdose, indeterminate, and viral-induced. The following primary outcomes are assessed at 3 weeks after entry into the cohort: successful hospital discharge, liver transplantation, or death. Confirmation of current health status occurs at 30 days post-enrollment in addition to follow-up at 6 months and 12 months, if appropriate.

1.5 REVIEW OF PALF MODELING LITERATURE

The existing literature of prognostic modeling for adult patients with ALF is appropriately large. Examples include the King's College Hospital criteria, Clichy criteria, serum group-specific component protein levels, liver volume on CT scanning, blood lactate levels, hyperphosphataemia, Acute Physiology and Chronic Evaluation II score, serum alfa-fetoprotein levels, and the Model for End Stage Liver Disease (MELD) (O'Grady, 1989; Yantorno, 2007; Schiodyt, 2007; Yamagishi, 2009; Bernal, 2002; Chung, 2003; Antoniades, 2007; Murray-Lyon, 1976; Zaman, 2006). The

corresponding literature of prognostic modeling for pediatric patients with ALF is underdeveloped in comparison (Table 1). The prior studies are simple in nature and rely upon standard logistic regression modeling procedures and validation methods (Sanchez, 2012; Liu, 2005; Lu, 2013; Rajanayagam, 2013; Squires, 2006). Rajanayagam et al. (2013) and Azhar et al. (2013) are two notable exceptions. This pair of manuscripts applies machine-learning methods to small datasets, with sample sizes of 54 and 49, respectively. Findings are consistent across the studies; however the latter analyses consider a significantly larger set of candidate variables, effectively increasing likelihood of false positives and improved fit by chance.

A common limitation throughout the literature is the misrepresentation of the stochastic nature of the multisystem disorder. PALF is a function of dynamic, interrelated physiological processes. Patients display a dynamic state of being; rarely is a static state observed. Given the time-dependent nature of ALF, the standard set of baseline covariate measures is insufficient. Dynamic conditions require dynamic approaches to modeling. The prediction of clinical outcomes for patients with such an unpredictable, volatile condition is arduous yet worthwhile of rigorous scientific investigation. Prognostic models of PALF can therefore be improved with early and exact evaluation of condition severity, incorporation of novel time-varying biological markers, and implementation of dynamic modeling principles. Therein the purpose of our study is to develop a dynamic prediction model of pre-transplant survival for patients with PALF, specifically to facilitate timely registration for deceased-donor OLT and guide utility-driven allocation decisions with organ procurement organizations (OPO).

Table 1: Summarization of the published literature on predictions models of PALF.

Primary Author	Year	Country	n	Method	Outcome(s)	Time Horizon	AUC	Inputs
Rajanayagam J	2013	Australia	54	artificial neural network	OLT/ death without OLT (composite)	6 months	0.96	etiology, ALT, AST, GGT, ALP, ammonia, albumin, LDH, lactate PT, Factor V, fibrinogen, WCC, platelets, INR (admission, diagnosis, day 7, peak), serum bilirubin (diagnosis and peak)
Rajanayagam J	2013	Australia	53	logistic regression, ROC	OLT/ death without OLT (composite)	3 months	0.86	PELD scores (peak)
				logistic regression, ROC	OLT/ death without OLT (composite)	3 months	0.71	PELD scores (admission)
Azhar N	2013	USA	49	dynamic Bayesian network	alive, OLT, or death without OLT (3 events)	3 weeks	-	array of serum inflammatory measures (up to 7 days)
Sanchez M	2012	USA	40	logistic regression, ROC	OLT/ death without OLT (composite)	length of hospitalization	0.89	PELD scores (time of diagnosis)
Lu B	2013	USA	578	logistic regression, ROC	OLT/ death without OLT (composite)	3 weeks	0.76	bilirubin, INR, ammonia (admission)
			546	logistic regression, ROC	OLT/ death without OLT (composite)	3 weeks	0.76	bilirubin, PT, ammonia (admission)
			461	logistic regression, ROC	OLT/ death without OLT (composite)	3 weeks	0.81	bilirubin, INR, ammonia (peak)
			454	logistic regression, ROC	OLT/ death without OLT (composite)	3 weeks	0.84	bilirubin, PT, ammonia (peak)
			399	logistic regression, ROC	OLT (death without OLTs removed from risk set)	3 weeks	0.84	bilirubin, INR, ammonia (peak)
			294	logistic regression, ROC	death without OLT (OLTs removed from risk set)	3 weeks	0.76	bilirubin, INR, ammonia (peak)
Liu E	2006	USA	81	logistic regression, ROC	OLT/ death without OLT (composite)	4 weeks	0.91	bilirubin, INR, ammonia (peak)
			81	logistic regression, ROC	OLT/ death without OLT (composite)	4 weeks	0.89	bilirubin, PT, ammonia (peak)

2.0 MATERIALS AND METHODS

2.1 STATISTICAL CONCEPTS

2.1.1 Estimation of Marginal Survival

Let T_i^* denote the event time for the i -th subject ($i = 1, \dots, n$) and C_i be the corresponding censoring time. Suppose $\{T_i, \delta_i, \mathbf{X}_i(\cdot)\}$ indicates the observed ordinary right-censored survival data for n individuals from our target population, where $T_i = \min(T_i^*, C_i)$ is the observed follow-up time, $\delta_i = I(T_i^* \leq C_i)$ the event indicator, $I(\cdot)$ an indicator function that takes value 1 only when $T_i^* \leq C_i$, and $\mathbf{X}_i(\cdot)$ a vector of p baseline covariates. Given the assumptions of conditional independence of event and censoring times and independence among individuals, we define the marginal cumulative distribution and marginal survival functions at time t as:

$$F(t) = P(T \leq t),$$

$$S(t) = 1 - F(t) = P(T > t).$$

The Kaplan-Meier (KM) estimator ([Kaplan and Meier, 1958](#)), also referred to as the product-limit estimator, is a non-parametric maximum likelihood estimator (MLE) of the marginal survival probability. The KM estimator for a given dataset can be expressed as follows:

$$\hat{S}(t) = \prod_{t_i \leq t} \left(1 - \frac{d_i}{n_i}\right),$$

where d_i denotes the number of events at time t_i and n_i the number of at-risk individuals just prior to time t_i .

If we are interested in estimating marginal survival adjusting for covariates \mathbf{X} , a Cox proportional hazards (PH) regression model (Cox, 1972) is commonly used for estimation. The Cox PH model, subject to the proportionality assumption, is defined as:

$$h(t|\mathbf{X}) = h_0(t) \exp(\mathbf{X}'\boldsymbol{\beta}),$$

where $h_0(t)$ denotes the hazard of an unspecified form for individuals with baseline covariates, \mathbf{X} a vector of p baseline covariates, and $\boldsymbol{\beta}$ a vector of p -dimensional regression coefficients. The marginal survival function $S(t)$ of the Cox PH model takes the form:

$$\begin{aligned} S(t|\mathbf{X}) &= \exp \left\{ - \int_0^t h_0(u) \exp(\mathbf{X}'\boldsymbol{\beta}) du \right\} \\ &= \exp\{-H_0(t)\} \exp(\mathbf{X}'\boldsymbol{\beta}). \end{aligned}$$

The estimated marginal survival function $\hat{S}(t|\mathbf{X})$ can be obtained by substituting unknown parameters $\boldsymbol{\beta}$ with the partial likelihood estimator $\hat{\boldsymbol{\beta}}$ and substituting the unspecified cumulative baseline hazard function $H_0(t)$ with the Breslow estimator $\hat{H}_0(t)$. When the PH assumption is violated, other regression models can be used to estimate marginal survival. These models usually incorporate a more complex procedure to estimate time-varying covariate effects. For example, Gray's time-varying coefficients model (Gray, 1992) is a commonly used alternative for the Cox PH model if the PH assumption is violated. Gray's model expresses time-varying effects via a linear combination of B-spline basis functions. The unknown regression parameters are estimated by the penalized partial likelihood estimators.

2.1.2 Estimation of Conditional Survival

For dynamic prediction, we are interested in the conditional survival function, that is the survival function at a specified time s conditional on surviving event free until time $t - (s > t)$. Accounting for the covariates \mathbf{X} , we express the conditional survival function as a ratio of two marginal survival functions with the form:

$$S(s|t, \mathbf{X}) = P(T > s | T \geq t, \mathbf{X}) = \frac{S(s|\mathbf{X})}{S(t|\mathbf{X})}$$

for $s \geq t$. If we rewrite the future prediction time t as a function of current time s plus a prediction window of length w , the probability of the event occurring within the fixed window w conditional on surviving event free until time $t -$, can be defined as:

$$F_w(t|\mathbf{X}) = P(T \leq t + w | T \geq t, \mathbf{X}) = 1 - S(t + w|t, \mathbf{X}).$$

This conditional failure probability is also called the fixed width failure function. In practice, the selection of the fixed width w depends on the length of the follow-up, number of events, and the overall survival rate. For small values of w the fixed width failure function approximates the instantaneous risk of the event for the individuals still at risk, the basis of the flexible class of regression models.

It is worth noting that based on the definition of the conditional survival, the process of estimating fixed width failure probability is a two-step sequential process by separately estimating the marginal survival at times t and $t + w$ separately. Computationally this two-step estimation process is not efficient. In addition, theoretical derivation of the estimated standard errors for the resulting estimated failure probability is difficult.

2.1.3 Crude Landmark Model

van Houwelingen (2007) proposed a Cox-based model that can be used to accurately estimate the fixed width failure probability in one step. Moreover, the proposed model is not sensitive to the violation of the PH assumption. He called this extension of methodology the crude landmark model.

Suppose we are interested in predicting failure probability at time $t_{LM} + w$ conditional on surviving up to a pre-specified landmark time t_{LM} . The crude landmark model first creates a dataset including all individuals at risk at $t = t_{LM}$ and ignores all events occurring beyond the fixed window width w by means of adding an administrative censoring at $t_{LM} + w$. Therefore, the hazard function of the crude landmark model can be expressed as follows:

$$h(t|t_{LM}, \mathbf{X}) = h_0(t|t_{LM}) \exp(\mathbf{X}'\boldsymbol{\beta}_{LM})$$

for $t_{LM} \leq t \leq t_{LM} + w$. We can then compute the corresponding conditional failure probability as:

$$\begin{aligned} \hat{F}_w(t_{LM}|\mathbf{X}) &= 1 - \exp\left\{-\int_{t_{LM}}^{t_{LM}+w} \hat{h}_0(u|t_{LM}) \exp(\mathbf{X}'\widehat{\boldsymbol{\beta}}_{LM}) du\right\} \\ &= 1 - \exp\left[-\exp(\mathbf{X}'\widehat{\boldsymbol{\beta}}_{LM})\{\widehat{H}(t_{LM} + w) - \widehat{H}(t_{LM} -)\}\right], \end{aligned}$$

where $\widehat{\boldsymbol{\beta}}_{LM}$ is the partial likelihood estimator of regression parameters $\boldsymbol{\beta}_{LM}$ and $\widehat{H}(\cdot)$ is the Breslow estimator of the baseline cumulative hazard $H(\cdot)$. Upon violation of the proportional hazards assumption, $\widehat{\boldsymbol{\beta}}_{LM}$ is a weighted average of the true time-varying effects $\boldsymbol{\beta}(t)$ on the interval $[t_{LM}, t_{LM} + w]$ and $\hat{F}_w(t_{LM}|\mathbf{X})$ well approximates the conditional failure, assuming: follow-up is not too long, the hazard ratios do not vary too much, and the covariate effects are not excessively large (van Houwelingen and Putter, 2012).

2.1.4 Crude Landmark Model with Time-Dependent Covariates

Prior to this point we restricted prognostic information to baseline covariates, or information available at the beginning of the follow-up period. In the presence of p time-dependent endogenous covariates we denote the dynamic patient information as $\mathbf{X}(t)$. A common approach to predicting conditional survival given time-dependent covariates requires the specification of a joint model for $\{T, \mathbf{X}(t)\}$ (Wulfsohn and Tsiatis, 1997; Henderson, 2000; Hashemi, 2003; Tsiatis and Davidian, 2004). Such joint models are computationally complex and often inappropriate for prediction modeling given that the distribution of $\mathbf{X}(t)$ is unknown.

When dynamic prediction is the primary concern, the landmark approach represents a viable alternative to joint modeling. We assume the time-dependent covariates to be Markovian in nature. That is, the future values of $\mathbf{X}(t)$ are solely dependent upon the present value of the covariate process, $\mathbf{X}(t_{LM})$. The hazard function for the crude landmark model with time-dependent covariates can be expressed as follows:

$$h\{t|t_{LM}, \mathbf{X}(t_{LM})\} = h_0(t|t_{LM}) \exp\{\mathbf{X}(t_{LM})' \boldsymbol{\beta}_{LM}\}$$

for $t_{LM} \leq t \leq t_{LM} + w$. We can then compute the corresponding conditional failure probability as:

$$\begin{aligned} \hat{F}_w(t_{LM} | \mathbf{X}(t_{LM})) &= 1 - \exp \left[- \int_{t_{LM}}^{t_{LM}+w} \hat{h}_0(u|t_{LM}) \exp\{\mathbf{X}(t_{LM})' \widehat{\boldsymbol{\beta}}_{LM}\} du \right] \\ &= 1 - \exp \left[- \exp\{(\mathbf{X}(t_{LM})' \widehat{\boldsymbol{\beta}}_{LM})\} \{ \widehat{H}(t_{LM} + w) - \widehat{H}(t_{LM} -) \} \right], \end{aligned}$$

where $\widehat{\boldsymbol{\beta}}_{LM}$ is the partial likelihood estimator of regression parameters $\boldsymbol{\beta}_{LM}$ and $\widehat{H}(\cdot)$ is the Breslow estimator of the baseline cumulative hazard $H(\cdot)$, and $t_{LM} -$ indicates the instantaneous moment prior to t_{LM} . As before, upon violation of the proportional hazards assumption, $\widehat{\boldsymbol{\beta}}_{LM}$ is a weighted

average of the true time-varying effects $\boldsymbol{\beta}(t)$ on the interval $[t_{LM}, t_{LM} + w]$ and $\hat{F}_w(t_{LM} | \mathbf{X}(t_{LM}))$ well approximates the conditional failure, given the set of assumptions previously stated in section 2.1.3.

2.1.5 Stratified Landmark Supermodel with Time-Dependent Covariates

When there is a set of landmark time points of interest in dynamic prediction, van Houwelingen (2007) proposed a stratified landmark supermodel to model the regression parameters $\boldsymbol{\beta}_{LM}$ as a function of the landmark time t_{LM} . The hazard function of the model follows the form:

$$h(t | s = t_{LM}, \mathbf{X}(s = t_{LM})) = h_0(t | s = t_{LM}) \exp\{\mathbf{X}(s = t_{LM})' \boldsymbol{\beta}_{LM}(s = t_{LM})\}$$

for $s \leq t \leq s + w$, where $h_0(t | s = t_{LM})$ is the unspecified baseline hazard and $\boldsymbol{\beta}_{LM}(s = t_{LM})$ is an arbitrarily-defined smooth function (e.g. polynomial, spline) of the landmark time $s = t_{LM}$. In practice we posit a linear model of $\boldsymbol{\beta}_{LM}(s)$ on s :

$$\boldsymbol{\beta}_{LM}(s) = \sum_{j=1}^{m_b} \theta_j f_j(s)$$

with a set of m_b basis functions $\{f_1(s), f_2(s), \dots, f_{m_b}(s)\}$ and a vector of θ parameters. The consistent estimators of unknown parameters can be obtained by maximizing the integrated partial log-likelihood (IPL) function:

$$IPL(\boldsymbol{\beta}_{LM}) = \sum_{i=1}^n d_i \left[\int_0^{t_i} \mathbf{X}_i(u)' \boldsymbol{\beta}_{LM}(u) \psi(u) du - \int_0^{t_i} \ln \left\{ \sum_{t_j \geq t_i}^n \int_0^{t_i} \exp(\mathbf{X}_j(u)' \boldsymbol{\beta}_{LM}(u)) \right\} \psi(u) du \right]$$

where $\psi(\cdot)$ is an indicator function that takes on value 1 in the specified window $[s, s + w]$ and 0 otherwise.

The stratified supermodel conveniently estimates smooth landmark-dependent covariate effects $\boldsymbol{\beta}_{LM}(s)$, however it provides separate estimated baseline hazards at the event time t_i for each landmark stratum under the following expression:

$$\widehat{h}_0(t_i|s) = \left[\sum_{t_i \leq t_j}^n \exp\{\mathbf{X}_j(s)' \widehat{\boldsymbol{\beta}}_{LM}(s)\} \right]^{-1},$$

limiting the generalizability of predictions between landmark time points. Note that $\widehat{h}_0(t_i|s)$ does not depend on s if $\mathbf{X}(s)$ and $\boldsymbol{\beta}_{LM}(s)$ are constant.

2.1.6 Proportional Baselines Landmark Supermodel with Time-Dependent Covariates

To address the issue of separate baseline hazards for each landmark stratum, [van Houwelingen \(2007\)](#) proposed another model called the proportional baselines landmark supermodel. The premise is to model a common baseline hazard through a multiplicative dependence on two components: the set of landmark-specific baseline hazards and a smooth function of the landmark time $s = t_{LM}$. Formally, $h_0(t|s) \equiv h_0(t) \exp\{\boldsymbol{\gamma}(s)\}$. Therefore, the hazard function with time-dependent covariates follows the form:

$$h(t|s, \mathbf{X}(s)) = h_0(t) \exp\{\mathbf{X}(s)' \boldsymbol{\beta}_{LM}(s) + \boldsymbol{\gamma}(s)\},$$

for $s \leq t \leq s + w$. In practice, we fit the gamma function via a linear model:

$$\boldsymbol{\gamma}(s) = \sum_{j=1}^{m_h} \eta_j g_j(s)$$

with a set of m_h basis functions $\{g_1(s), g_2(s), \dots, g_{m_h}(s)\}$ and a vector of η parameters. The unknown regression parameters (θ, η) will be estimated by maximizing a generalization of the integrated partial log-likelihood (IPL*):

$$IPL^*(\boldsymbol{\beta}_{LM}, \boldsymbol{\gamma}) = \sum_{i=1}^n d_i \left[\int_0^{t_i} \mathbf{X}_i(s)' \boldsymbol{\beta}_{LM}(s) \psi(s) ds + \int_0^{t_i} \boldsymbol{\gamma}(s) \psi(s) ds \right. \\ \left. - \int_0^{t_i} \psi(s) ds \ln \left\{ \sum_{t_j \geq t_i}^n \int_0^{t_i} \exp(\mathbf{X}_j(s)' \boldsymbol{\beta}_{LM}(s) + \boldsymbol{\gamma}(s)) \right\} \psi(s) ds \right],$$

where $\psi(s)$ is an indicator function that takes on value 1 in the specified window $[s, s + w]$ and 0 otherwise. The corresponding common baseline hazards will be estimated via the following formula:

$$\widehat{h}_0(t_i) = \frac{\int_0^{t_i} \psi(s) ds}{\sum_{t_j \geq t_i} \int_0^{t_i} \exp\{\mathbf{X}_j(s)' \widehat{\boldsymbol{\beta}}_{LM}(s) + \widehat{\boldsymbol{\gamma}}(s)\} \psi(s) ds}.$$

Note that the estimated hazard $\widehat{h}_0(t_i)$ no longer depends on s . For both the stratified and proportional baselines supermodels, we implement the sandwich estimators of [Lin and Wei \(1989\)](#) to correctly estimate standard errors by accounting for correlation due to the clustering of observations within an individual.

2.2 PALF STUDY DATA

The PALF Registry contains center- and patient-specific data from 20 sites throughout Canada, the United States, and the United Kingdom. For the purpose of our study, we restrict our attention to the sample of US patients only (corresponding to the exclusion of site codes 39, 40,

and 41), spanning admissions from December 25, 1999 to January 7th, 2011. Clinical information including medical history, demographics, and laboratory test results are collected and recorded on de-identified forms. Conditional on consent, blood samples and additional biospecimens (including but not limited to urine, liver tissue, bile, and skin) are gathered over the course of the following seven days only when clinical measures require such sampling. Additional data are collected at 6 and 12 months following enrollment, if applicable, given the current status of the patient. The University of Pittsburgh Epidemiology Data Center (EDC) receives all registry data and remotely stores the potential sources of analysis data in one of the five data structures described below. (Note: descriptions of data structures are adaptations of original PALF-EDC documentation.)

2.2.1 Hospital Evaluation Form

The hospitalization evaluation (HE) form database contains 756 unique entries, accounting for all US patients. The research team distributes the HE form immediately following patient enrollment in the PALF registry. The form captures patient demographics, admission, family, and medication histories, and in-hospital information through the first of the following outcomes: successful hospital discharge, liver transplantation, or death. The majority of baseline information regarding characteristics of patients at the time of enrollment is found within this file.

2.2.2 Follow-up Form

The follow-up (FF) form database contains 352 entries, representing 194 unique patients. (Note: 36 patients with one entry; 158 patients with two entries.) The research team follows up

with each registry patient at 6 and 12 months following enrollment. The FF form records the current status of patients alive with and without a transplant at the last time of registry assessment, either at the original hospital discharge or prior follow-up evaluation. Specifically, the form captures changes in patient location, list status, vital status, and final diagnosis during the follow-up interval.

2.2.3 Follow-up Vital Status Form

The follow-up vital status (FV) form database contains 170 entries, representing 93 unique patients. (Note: 16 patients with one entry; 77 patients with two entries.) The research team administers the FV form to patients who underwent liver transplantation, either during the initial hospitalization or during the interval following discharge. Follow-up vital status form is completed for patients who underwent liver transplantation during the initial hospitalization or during the previous follow-up interval. The form captures the same changes as indicated by the description of the FF form.

2.2.4 Hospital Flow Sheet Form

The hospital flow sheet (HF) form contains 6257 entries, representing 756 unique patients. The median number of entries is 9, the mode 10, the interquartile range 4, and the range 1 to 14. The HF form is essential to any dynamic analysis, as it contains the time-varying covariate information. The HF records daily in-hospital laboratory, procedural, treatment, and event information in the PALF registry from the time of enrollment up through 7 days post-enrollment in the PALF registry or N-acetylcysteine Trial (until the first of successful hospital discharge,

transplantation, or death). The flow sheet is setup as an Excel spreadsheet and is easily accessible on the project website. In combination with the HE form, the HF form comprises the majority of the patient information necessary to build dynamic prediction models such as van Houwelingen's landmark supermodel.

2.2.5 Summary Descriptive File

The summary descriptive (SUMDS) file contains 756 entries, one for each unique patient in the registry. The summary file is an unofficial aggregation of select variables that data managers from the EDC believe to be particularly useful in describing the study population and outcomes. The file includes useful variables such as date of hospitalization, date of enrollment, time of death, time of transplant, and multiple assessments of diagnosis, but fails to capture information pertaining to the patient listing with UNOS for transplantation.

2.3 CONSTRUCTION OF ANALYSIS DATASET

2.3.1 Selection of Candidate Covariates

Extensive exploratory analysis precedes the development of dynamic prediction models. With expert clinical guidance we considered the proceeding variables from the HE and HF form databases: discretized age, albumin (serum), albumin (total), ammonia (arterial), ammonia (venous), alanine aminotransferase (ALT), ascites, aspartate aminotransferase (AST), blood pH, blood pressure (diastolic), blood pressure (systolic), blood urea nitrogen, creatinine, primary

diagnosis, encephalopathy grade, gamma-glutamyl transpeptidase (GGT), gender, hematocrit, international normalized ratio (INR), prothrombin time (PT), race, sodium, UNOS Status 1, treatment in the intensive care unit, ventilator support, and white blood cell count. We excluded potential covariates for a myriad of reasons including substantial proportions of missing values, multi-collinearity, and homogeneity within subgroups. High proportions of missing values prevented the selection of many desirable predictors: GGT (33.1% missing), Factor VIIa (53.1% missing), sodium (56.9% missing), and venous ammonia (55.3% missing), to name a few. Following the identification of the candidate variables we generated a repeated measures survival dataset utilizing information concerning outcomes from all four data collection forms, baseline characteristics from the HE form, and time-varying information of laboratory values from the HF form. The final repeated measures dataset contained 3349 observations from a sample of 658 unique PALF patients.

2.3.2 Creation of Super Prediction Dataset

Landmark analysis necessitates intricate data management routines ([Figure 3](#)). The method requires construction of a particular form of data frame, called a super prediction dataset, through an involved procedural algorithm. The process is summarized as follows:

- i. Fix the prediction window w based on clinical knowledge. For the analysis we elected a prediction window of 3 days, an empirical estimate of the median elapsed between listing and emergency OLT.

- ii. Select a set of uniformly spaced landmark prediction time points and define the set $s_L = \{s_1, \dots, s_l\}$ based on clinical knowledge. For the analysis we elected to use the set $s_L = \{0, \dots, 18\}$ such that the analysis spans the full 21 days prescribed in the PALF study design.
- iii. Create a prediction dataset for each landmark time point s by left truncation and right administrative censoring at end of the prediction window ($s + w$); collectively we refer to each of these datasets as a stratified data frame.
- iv. Stack all 19 stratified data frames (one for each landmark time point s within the set s_L) vertically into a single super prediction dataset. Note that passing from one stratum to the next one corresponds to sliding the window over the range of time points.
- v. Transform the stacked super prediction dataset by the time-dependent covariates. In practice, this equates to censoring and resetting the clock for all individuals within a landmark risk set for each observed event.

The repeated process of generating overlapping stratified data frames, re-arranging the components as a super stacked dataset, and transforming the data by the time-varying covariates results in the rapid inflation of the sample size. The final stacked super prediction dataset contains 25,666 observations.

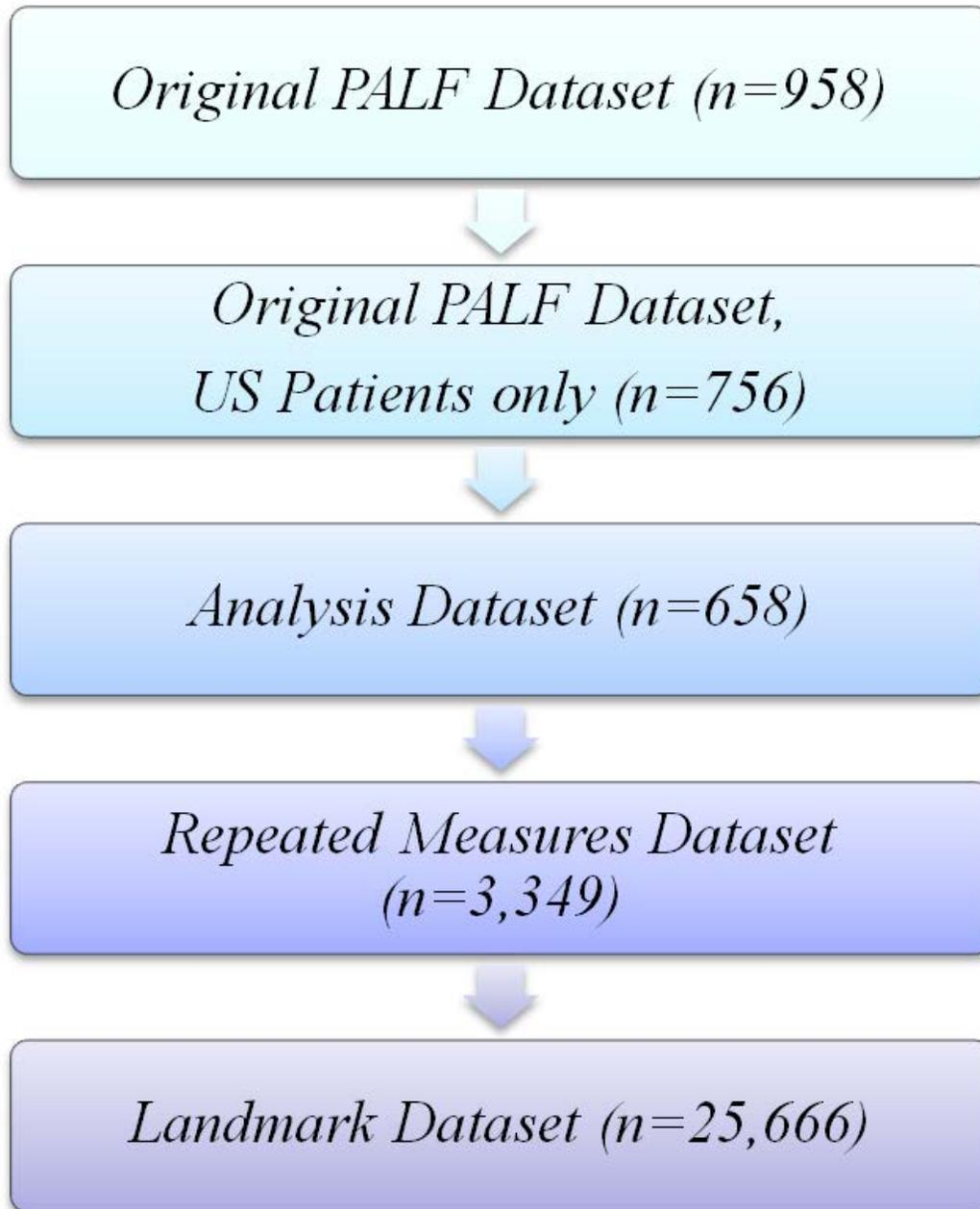


Figure 3: Graphical representation of the data management process.

2.4 ANALYSIS

We managed data and performed analyses in the statistical packages SAS version 9.3 (SAS Institute, Cary, North Carolina) and R version 3.0.3 (CRAN, <http://cran.r-project.org/>), respectively. The analysis called for a sequential approach to construct a dynamic prediction model of three-day pre-transplant death risk for patients with PALF. Upon performing descriptive statistics, we developed the traditional survival regression models and progressed to the landmarking method of van Houwelingen, and then examined its predictive ability. Detailed specifications of this process are found below.

2.4.1 Descriptive Statistics

We reported summary measures of candidate covariates as two-way proportions and sequences of mean-median-standard deviation for categorical and continuous variables, respectively. We implemented Fisher's Exact Tests to assess differences among distributions of categorical variables between outcomes of (1) alive or OLT and (2) dead without OLT. Analogously, we implemented the Mann-Whitney U Test to assess differences in the distributions of continuous laboratory values between the aforementioned outcomes.

2.4.2 Univariable Cox PH Models

We modeled the outcome of regression models as pre-transplant survival, measured from the time of enrollment in the PALF Study. We applied right censoring to the patients who left the study for extraneous reasons, underwent emergency OLT, or did not experience an event within the short-term period of interest defined as 21 days. All thirteen of the candidate covariates entered univariable regression models; extended Cox regression models with time-dependent covariates assessed the continuous variables with repeated measures.

2.4.3 Multivariable Cox PH Models

Variables deemed statistically significant at a conservative cutoff of $P < 0.20$ from the univariable Cox models, along with all possible covariate-by-covariate interactions, then entered the multivariable time-dependent Cox regression model. We employed an Akaike Information Criterion-based (AIC) forward selection procedure to produce a parsimonious multivariable model. To validate the selection of variables in the multivariable model, we entered all significant variables from the univariable models, along with all possible interactions, into a multivariable time-dependent Cox regression model and performed manual backwards selection with a slightly less conservative cutoff of $P < 0.15$. Specifically, we assessed the significance of all predictors using the Wald Test, to identify a parsimonious model. Categorical variables with multiple levels remained in the model only if the simultaneous testing of the associated vector of parameters proved significant.

2.4.4 Proportional Baselines Landmark Supermodel

The construction of the landmark prediction model requires the R package *dynpred* (Putter, 2011). We carried forward the final model from the multivariable Cox regression and developed the landmark supermodel in a fashion similar to that described in section 2.4.3. Generally speaking, the process of fitting a parsimonious landmark supermodel is cumbersome in the presence of multiple time-varying covariates.

Here we simplify the process for the purpose of illustration and translatability. Beginning with the final predictors from before, we clustered the model on patient to account for multiple observations per patient, inserted two necessary proportional baseline parameters ($\gamma(s)$), and added a landmark time s -by-covariate interaction ($\beta(s)$) for each of the main effects. (The models quickly become large and difficult to test.) As before, we employed an AIC forward selection procedure to admit entry into the multivariable model and produce less biased estimates of the parameters in a parsimonious model. For the purpose of validating the selection of candidate variables, we entered all variables mentioned previously in 2.4.3 into a separate multivariable time-dependent Cox regression model and performed manual backwards selection. Just as before, we assessed the significance of all predictors using the Wald Test. In the presence of repeated measure survival outcomes, the likelihood ratio test (LRT) assumes independence among clustered observations whereas the Wald Test correctly accounts for their dependence. Thus the use of the Wald Test is pertinent to testing parameters within all landmark regression models with repeated measures.

2.4.5 Prediction of Conditional Survival and Fixed Window Risk Estimates

To compute estimates of conditional survival and the related estimate of fixed window risk probabilities, we invoked a user-modified version of the *Fwpredict* function in the *dynpred* package (Putter, 2011). (The function is practical but unable to handle models more complex than those found in the van Houwelingen and Putter’s textbook.) We plotted the trajectory of the average patient’s three-day death probabilities, in addition to the trajectories of all 658 patients from the PALF Registry with complete data and aggregated the data in various forms.

2.4.6 Evaluation of Model Performance

To examine the performance of the model in the absence of a validation cohort, we calculated a marginal estimate of the area under the receiver operating characteristic curve (AUC) for all landmark times s (Green and Swets, 1966; Heagerty and Zheng, 2005). As an extension, we calculated the dynamic AUC at each of the original landmark time points $s_L = \{0, \dots, 18\}$ and plotted the performance as a function of time.

3.0 RESULTS

3.1 DESCRIPTIVE STATISTICS

Among the 756 pediatric patients enrolled in the PALF Study within the United States, we identified 658 with complete information (87.0%) based upon expert clinical guidance. The Kaplan-Meier function estimated a 3-week survival of 82.6% (95% CI: 79.0 – 86.4) ([Figure 4](#)). Overall, 78 patients succumbed to ALF (11.9%) during the 3 week period of study with a median time-to-death of 5 days (range 0 - 20). The remaining 580 patients spontaneously recovered, underwent orthotopic liver transplant, or left the study (88.1%) with a median time-to-censoring of 11 days (range 0 - 21).

Patient demographics, clinical data, and baseline laboratory values for the PALF patients meeting the inclusion study for the analysis are outlined in [Table 2](#). The following patient demographics, clinical factors, and laboratory values indicated statistically significant differences by outcome with respect to the Fisher Exact and Mann-Whitney U Tests for categorical and continuous variables, respectively: age group, ascites, diagnosis, encephalopathy, gender, race, UNOS Status 1, ventilator support, ICU, AST, ALT, PT, and total bilirubin. Spaghetti plots of the four continuous laboratory values are displayed in [Figure 5](#).

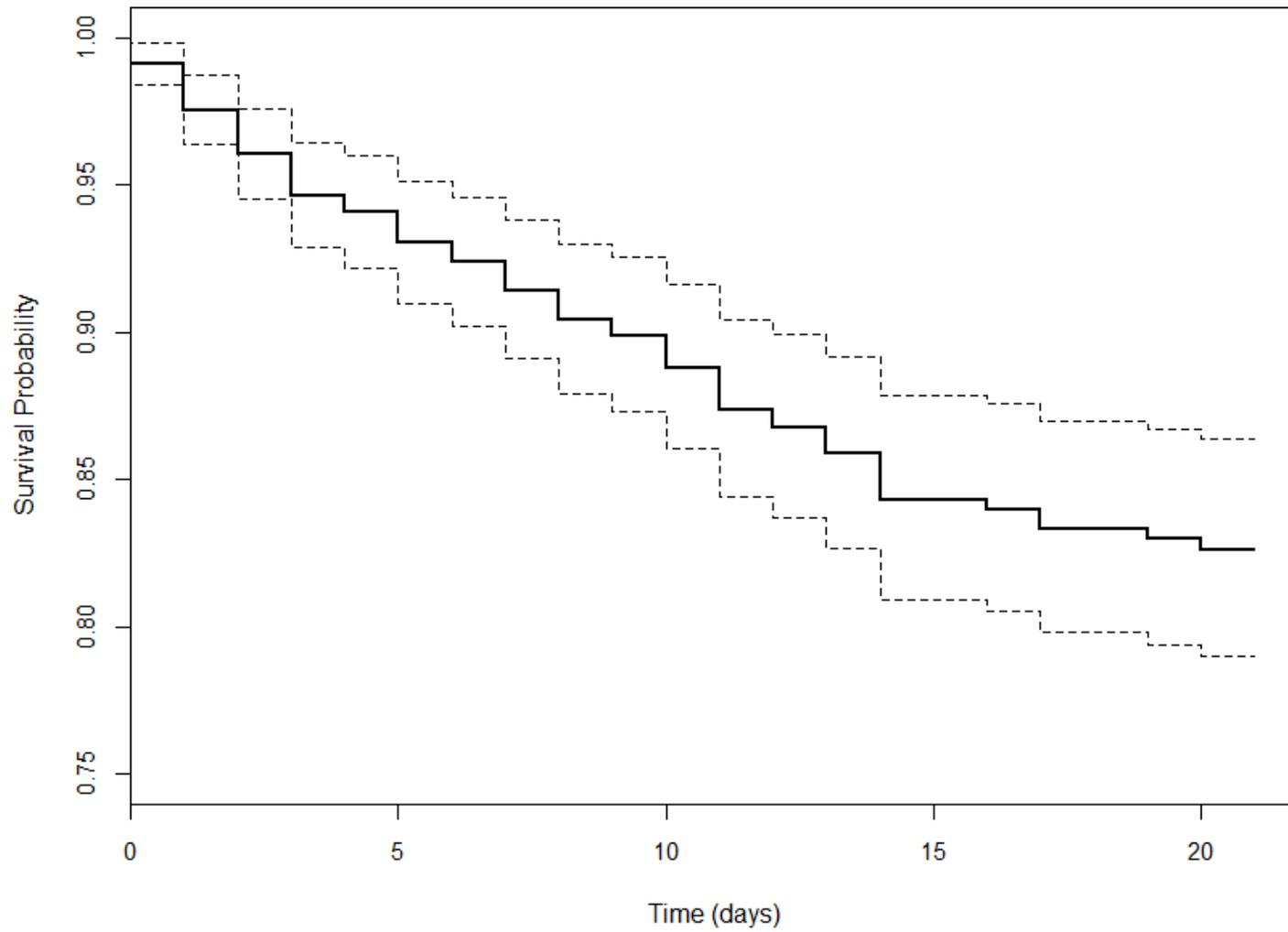


Figure 4: Kaplan-Meier estimate of survival in the analysis cohort of PALF patients.

Table 2: Patient demographics, clinical data, and laboratory values for the analysis cohort of PALF patients at the time of study enrollment.

Covariate	Total (N = 658)	Censored (N = 580; 88.1%)	Died without LT (N = 78; 11.9%)	P-Value
Age Group				
0-<6 mo	117 (17.8%)	91 (15.7%)	26 (33.3%)	0.0008 †**
6 mo-<3 y	130 (19.8%)	114 (19.7%)	16 (20.5%)	
>=3 y	411 (62.5%)	375 (64.7%)	36 (46.2%)	
Gender				
Male	333 (50.6%)	290 (50.0%)	43 (55.1%)	0.4018 †
Female	325 (49.4%)	290 (50.0%)	35 (44.9%)	
Race				
Black	94 (14.3%)	82 (14.1%)	12 (15.4%)	0.9560 †
White	332 (50.5%)	293 (50.5%)	39 (50.0%)	
Other	232 (35.3%)	205 (35.3%)	27 (34.6%)	
Ascites				
No	530 (80.5%)	483 (83.3%)	47 (60.3%)	<0.0001 **
Yes	128 (19.5%)	97 (16.7%)	31 (39.7%)	
Diagnosis				
Acetaminophen	92 (14.0%)	89 (15.3%)	3 (3.8%)	0.0004 †
Indeterminate	313 (47.6%)	282 (48.6%)	31 (39.7%)	
Other	253 (38.4%)	209 (36.0%)	44 (56.4%)	
Ventilator Support				
No	485 (73.7%)	466 (80.3%)	19 (24.4%)	<0.0001 **
Yes	173 (26.3%)	114 (19.7%)	59 (75.6%)	
Intensive Care Unit				
No	188 (28.6%)	184 (31.7%)	4 (5.1%)	<0.0001 **
Yes	470 (71.4%)	396 (68.3%)	74 (94.9%)	
UNOS Status 1				
No	583 (88.6%)	510 (87.9%)	73 (93.6%)	0.1829
Yes	75 (11.4%)	70 (12.1%)	5 (6.4%)	
Encephalopathy				
Grade 0	300 (45.6%)	274 (47.2%)	26 (33.3%)	<0.0001 †**
Grade I-II	265 (40.3%)	239 (41.2%)	26 (33.3%)	
Grade III-IV	93 (14.1%)	67 (11.6%)	26 (33.3%)	
Lab Values, median, mean (sd)				
Log Aspartate Aminotransferase	7.32, 7.07 (1.48)	7.32, 7.05 (1.44)	7.47, 7.15 (1.72)	0.2394
Log Alanine Aminotransferase	7.33, 6.96 (1.61)	7.37, 7.02 (1.59)	6.96, 6.51 (1.69)	0.0121 **
Log Prothrombin Time	3.25, 3.30 (0.37)	3.24, 3.28 (0.36)	3.40, 3.48 (0.43)	<0.0001 **
Log Total Bilirubin	2.22, 1.91 (1.18)	2.22, 1.88 (1.21)	2.33, 2.15 (0.94)	0.0250 **

† indicates testing of multiple factors within a given variable

* statistical significance at $\alpha = .10$

** statistical significance at $\alpha = .05$

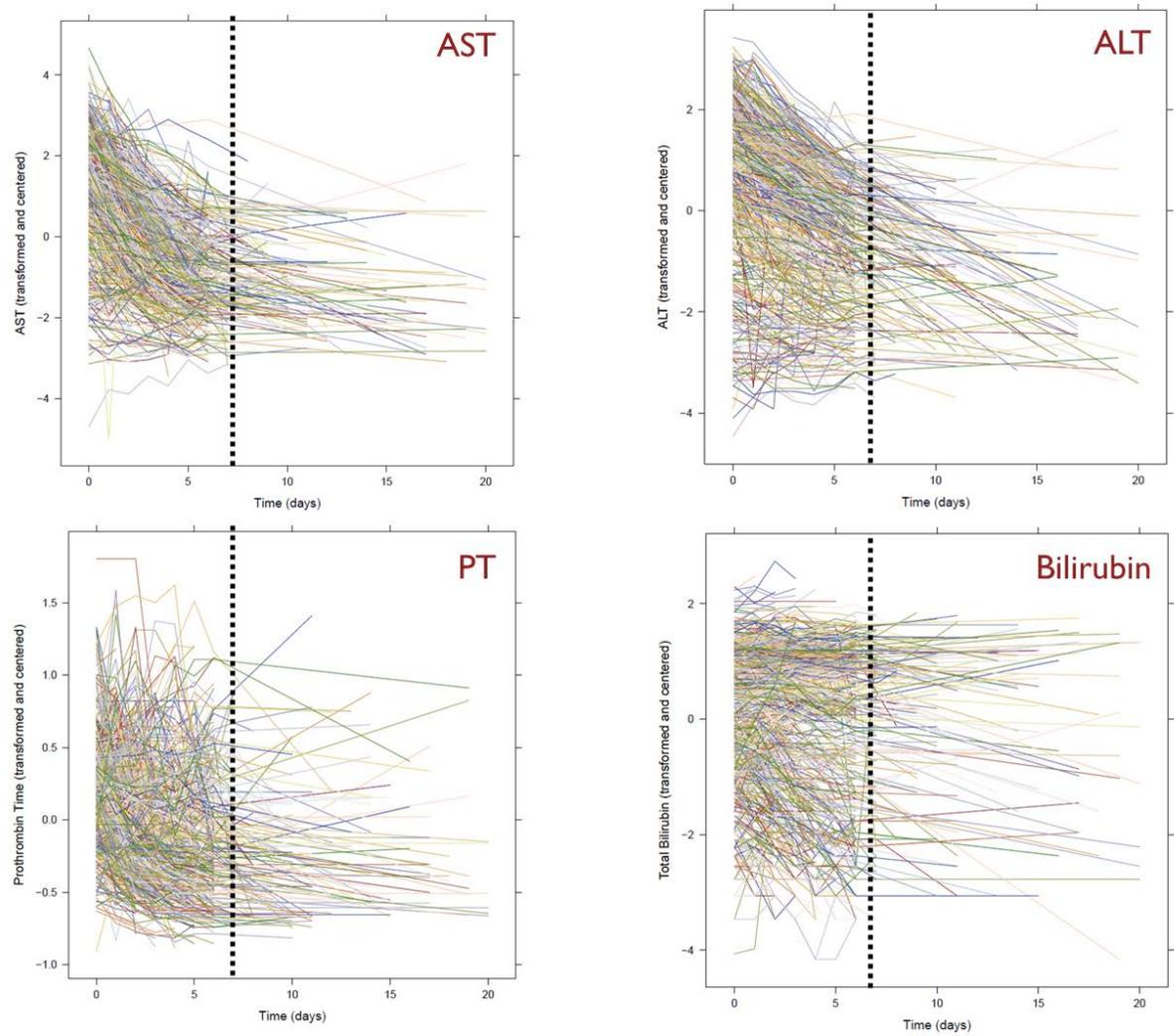


Figure 5: Spaghetti plots of the four continuous predictors AST, ALT, PT, and Bilirubin. Observe the sudden drop in the density of measurements beyond day 7 of the PALF study.

3.2 UNIVARIABLE COX PH MODELS

Univariable Cox regression models indicated that six time-invariant predictors influenced the survival of PALF patients: age group ($P = .0348$), ascites (risk factor, $P < .0001$), diagnosis ($P = .0145$), ventilator support (risk factor, $P < .0001$), ICU (risk factor, $P < .0001$), and encephalopathy ($P < .0001$) (Table 3). Within the specific levels of the predictors, we observed significant relationships: age of at least 3 years positively impacted survival ($P = .0081$) compared to age less than 6 months, a diagnosis of etiology other than acetaminophen-related or indeterminate negatively impacted survival ($P = .0091$) compared to an acetaminophen-related diagnosis, and an encephalopathy grade of III or IV negatively impacted survival ($P < .0001$), compared to a grade of 0. Among the time-varying lab values, the extended Cox regression models indicated that elevated levels of three predictors influenced the survival of PALF patients: alanine aminotransferase ($P < .0001$), prothrombin time ($P < .0001$), and total bilirubin ($P < .0001$).

Table 3: Regression parameter estimates from the univariable time-dependent Cox PH models of pre-transplant survival in PALF patients.

Covariate	Log Hazard Ratio (95% CI)	P-Value
<i>β</i> -parameters		
Age Group		0.0348 †**
6 mo-<3 y	-0.521 (-1.144 , 0.102)	0.1001
≥3 y	-0.685 (-1.193 , -0.177)	0.0081 **
Gender	-0.157 (-0.603 , 0.289)	0.4890
Race		0.8507 †
Black	0.177 (-0.470 , 0.825)	0.5910
Other	-0.003 (-0.494 , 0.488)	0.9910
Ascites	1.008 (0.554 , 1.461)	<0.0001 **
Diagnosis		0.0145 †**
Indeterminate	1.159 (-0.026 , 2.344)	0.0553 *
Other	1.558 (0.388 , 2.727)	0.0091 **
Ventilator Support	2.312 (1.794 , 2.829)	<0.0001 **
Intensive Care Unit	2.183 (1.176 , 3.189)	<0.0001 **
UNOS Status 1	-0.176 (-1.086 , 0.734)	0.7050
Encephalopathy		<0.0001 †**
Grade I-II	0.265 (-5.180 , 5.710)	0.3400
Grade III-IV	1.457 (0.912 , 2.001)	<0.0001 **
Lab Values		
Log Aspartate Aminotransferase	0.133 (-0.054 , 0.320)	0.1965
Log Alanine Aminotransferase	-0.369 (-0.514 , -0.224)	<0.0001 **
Log Prothrombin Time	1.310 (0.858 , 1.762)	<0.0001 **
Log Total Bilirubin	0.728 (0.481 , 0.975)	<0.0001 **

Log hazard ratios are reported with 95% confidence intervals, calculated with robust standard errors.

3.3 MULTIVARIABLE COX PH MODELS

Per the forward selection procedure, the multivariable time-dependent Cox regression model indicated nine influential predictors of survival in PALF patients: ventilator support (risk factor, $P < .0001$), ICU (risk factor, $P = .0580$), aspartate aminotransferase ($P < .0001$), alanine aminotransferase ($P = .0088$), prothrombin time ($P = .0003$), total bilirubin ($P = .0006$), aspartate aminotransferase-bilirubin interaction ($P = .1200$), aspartate aminotransferase-prothrombin time interaction ($P = .0840$), and bilirubin-prothrombin time interaction ($P = .0680$) (Table 4). Manual backwards selection confirmed the identification of pertinent predictors.

Table 4: Regression parameter estimates from the multivariable time-dependent Cox PH models of pre-transplant survival in PALF patients.

Covariate	Log Hazard Ratio (95% CI)	P-Value
<i>β</i> -parameters		
Ventilator Support	2.071 (1.515 , 2.627)	<0.0001 **
Intensive Care Unit	1.095 (-0.038 , 2.228)	0.0580 *
Log Aspartate Aminotransferase	0.701 (0.413 , 0.989)	<0.0001 **
Log Alanine Aminotransferase	-0.361 (-0.631 , -0.091)	0.0088 **
Log Prothrombin Time	1.185 (0.551 , 1.819)	0.0003 **
Log Total Bilirubin	0.645 (0.275 , 1.015)	0.0006 **
Interactions		
AST-Bilirubin	-0.143 (-0.323 , 0.037)	0.1200
AST-PT	-0.318 (-0.679 , 0.043)	0.0840 *
PT-Bilirubin	-0.692 (-1.435 , 0.051)	0.0680 *

Log hazard ratios are reported with 95% confidence intervals, calculated with robust standard errors.

3.4 PROPORTIONAL BASELINES LANDMARK SUPERMODEL

The final multivariable landmark supermodel included a set of predictors similar to that of the multivariable Cox model: ventilator support (risk factor, $P < .0001$), ICU (risk factor, $P = .0400$), aspartate aminotransferase ($P < .0001$), alanine aminotransferase ($P = .0082$), prothrombin time ($P < .0001$), total bilirubin ($P = .3900$), alanine aminotransferase-bilirubin interaction ($P = .0930$), aspartate aminotransferase-prothrombin time interaction ($P = .0790$), and bilirubin-prothrombin time interaction ($P = .0037$) (Table 5). Furthermore the model indicated the following supermodel-related covariate dependencies: aspartate aminotransferase as function of landmark time s ($P = 0.1300$), prothrombin time as a function of landmark time s ($P = 0.0083$), hazard as a function of landmark time s ($P = .0140$), and hazard as a function of landmark time s^2 ($P = .0004$) (Figure 6).

Table 5: Regression parameter estimates from the proportional baselines landmark supermodel of pre-transplant survival in PALF patients.

Covariate	Log Hazard Ratio (95% CI)		P-Value
<i>β</i> -parameters			
Ventilator Support	2.1500	(1.586 , 2.714)	<0.0001 **
Intensive Care Unit	1.1810	(0.052 , 2.310)	0.0400 **
Log Aspartate Aminotransferase			
<i>constant</i>	0.6310	(0.333 , 0.929)	<0.0001 **
<i>s</i>	-0.4950	(-1.134 , 0.144)	0.1300
Log Alanine Aminotransferase	-0.3720	(-0.648 , -0.096)	0.0082 **
Log Prothrombin Time			
<i>constant</i>	1.4400	(0.893 , 1.987)	<0.0001 **
<i>s</i>	2.6260	(0.678 , 4.574)	0.0083 **
Log Total Bilirubin	0.2160	(-0.280 , 0.712)	0.3900
Interactions			
ALT-Bilirubin	0.2060	(-0.035 , 0.447)	0.0930 *
AST-PT	-0.2590	(-0.547 , 0.029)	0.0790 *
PT-Bilirubin	-0.7940	(-1.329 , -0.259)	0.0037 **
<i>γ</i> -parameters			
Proportional Baseline			
<i>s</i>	3.0810	(0.629 , 5.533)	0.0140 **
<i>s2</i>	-5.3600	(-8.333 , -2.387)	0.0004 **

Log hazard ratios are reported with 95% confidence intervals, calculated with robust standard errors.

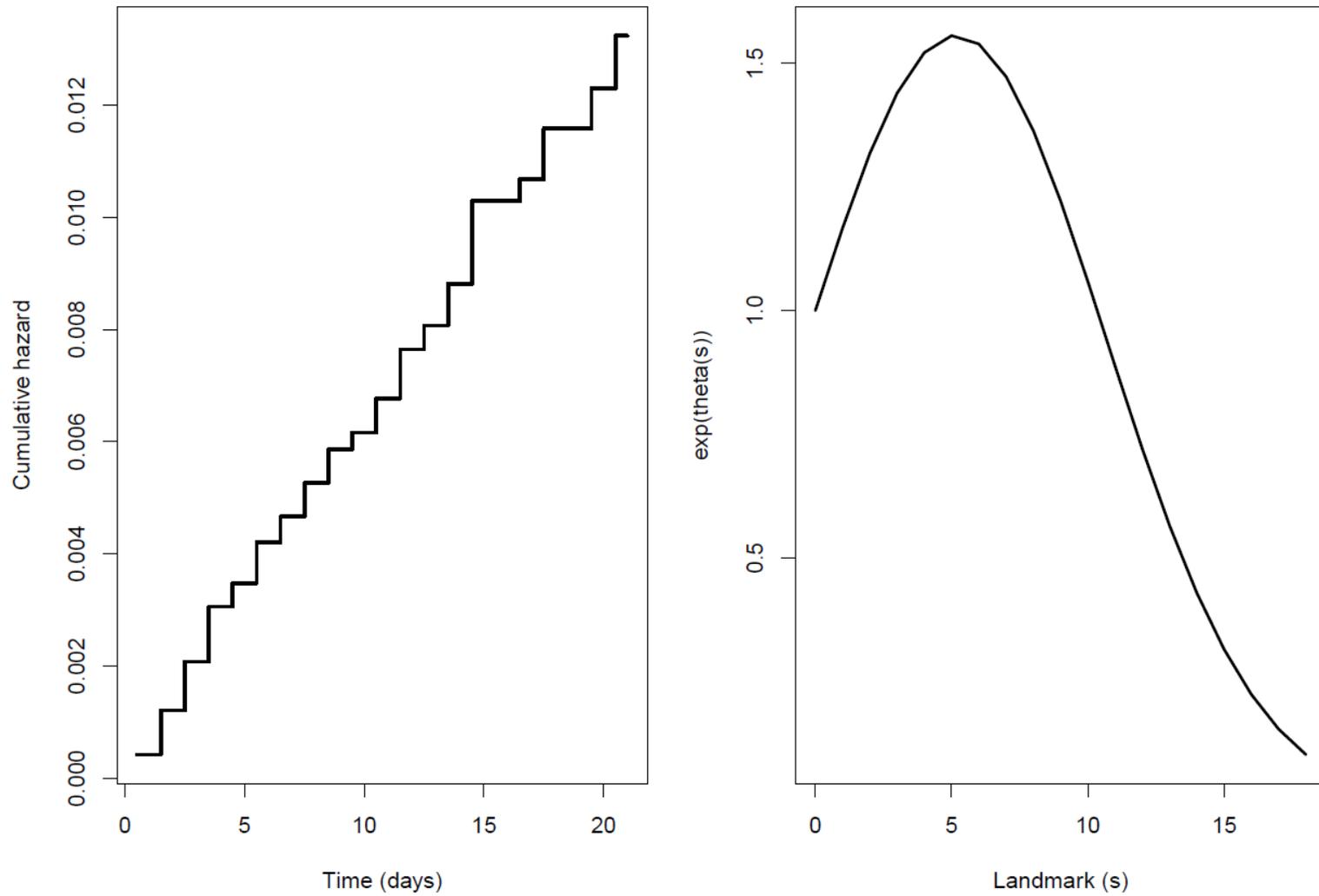


Figure 6: Baseline hazard and landmark effects in the proportional baselines landmark supermodel.

3.5 PREDICTION OF CONDITIONAL SURVIVAL AND FIXED WINDOW RISK ESTIMATES

The conditional three-day mortality probability estimates of the average patient ranged from 0.75% at the time of registry enrollment to 0.06% at the eighteenth day of follow-up. The maximum observed probability of 1.1% for the average patient occurred at the second day of the study (Figure 7). The personalized dynamic predictions of risk are noticeably different than that of the “average patient”, characterized by a substantial degree of variability (Figures 8, 9). Personalized predictions for all 658 patients of the analysis cohort are found in the supplementary file (separate document). The distribution of conditional three-day mortality probabilities aggregated across all landmark points in time and the distributions of the log probabilities stratified by landmark time point s provide a graphical representation of the variation in estimates probabilities and the potential to discriminate between events and non-events, or rather deaths and non-deaths (Figures 10, 11).

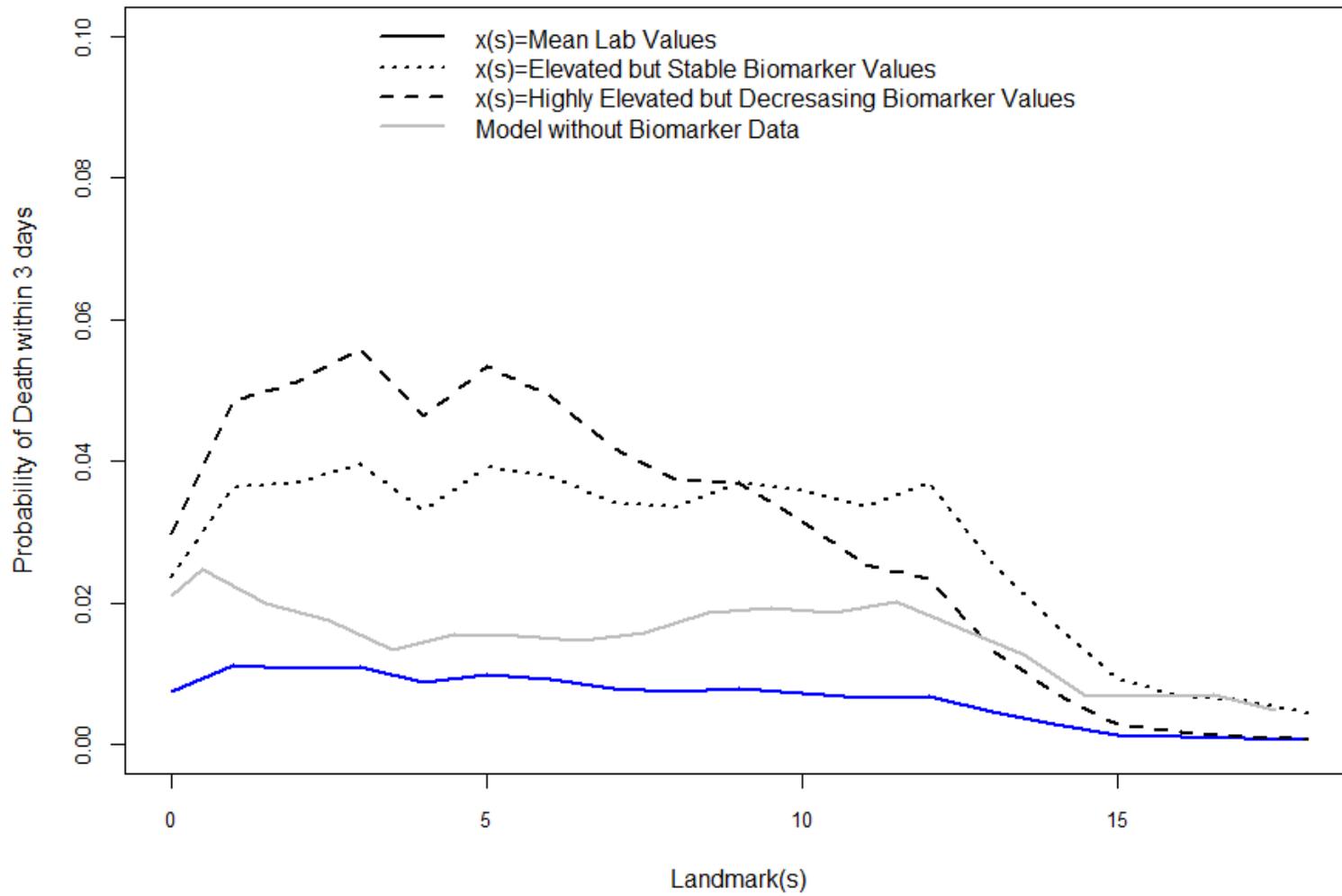


Figure 7: The trajectory of three-day conditional mortality probabilities for the “average patient,” or an individual with mean lab values, indicated in blue.

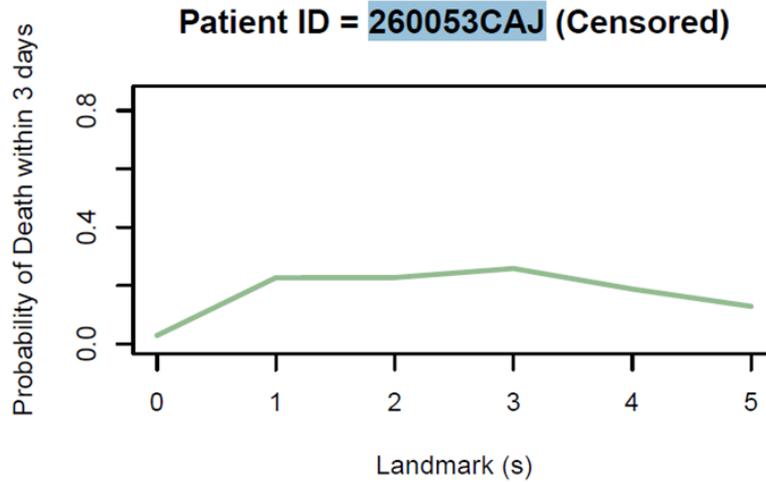
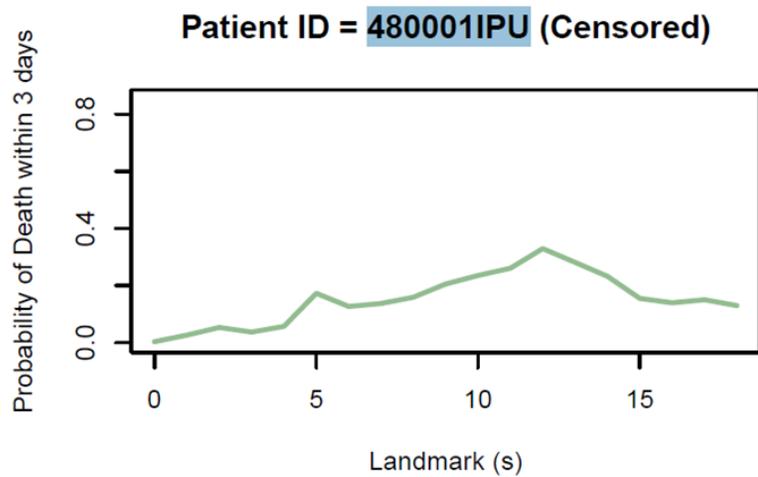
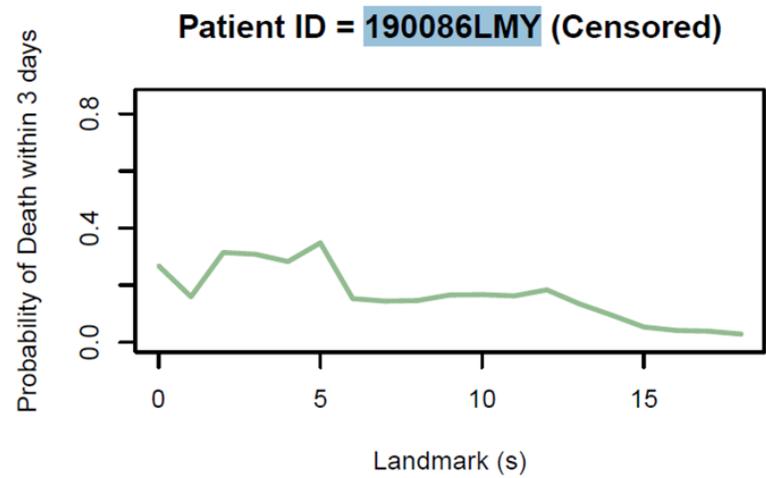
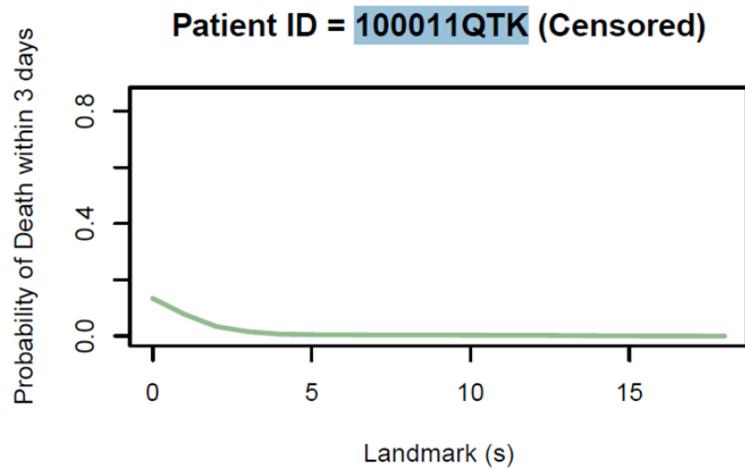


Figure 8: Non-random selection of personalized conditional mortality trajectories for 4 censored (successfully recovered or transplanted) PALF patients.

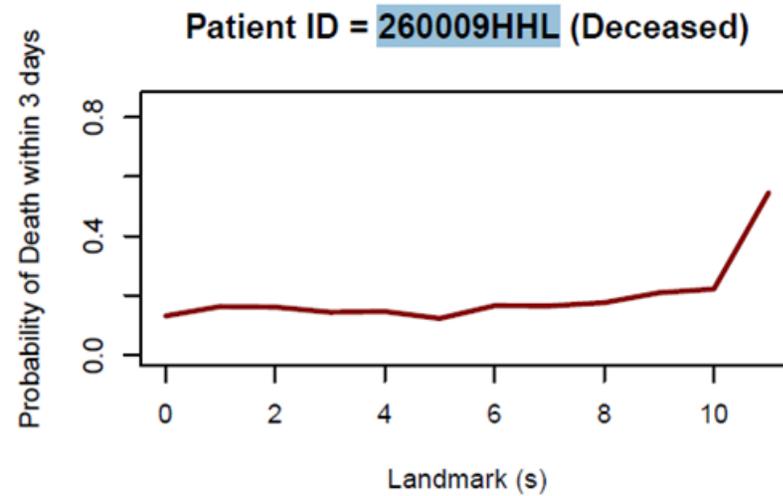
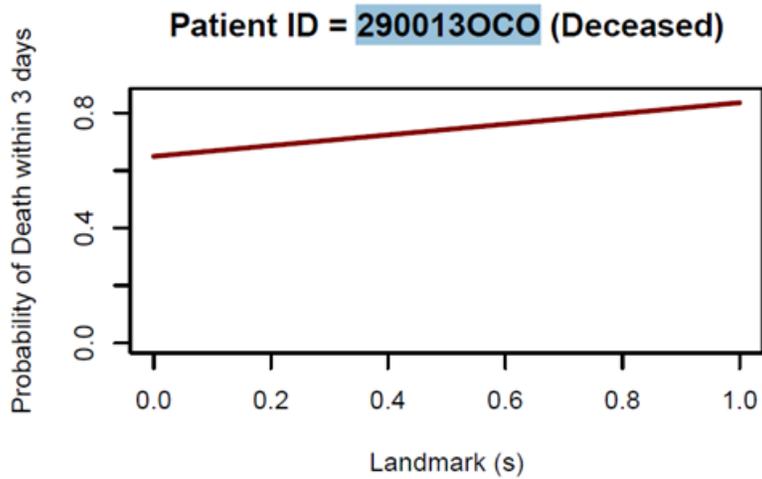
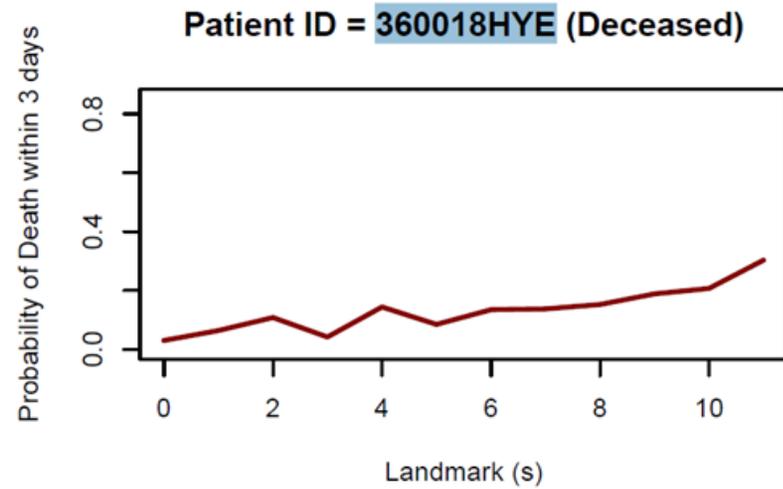
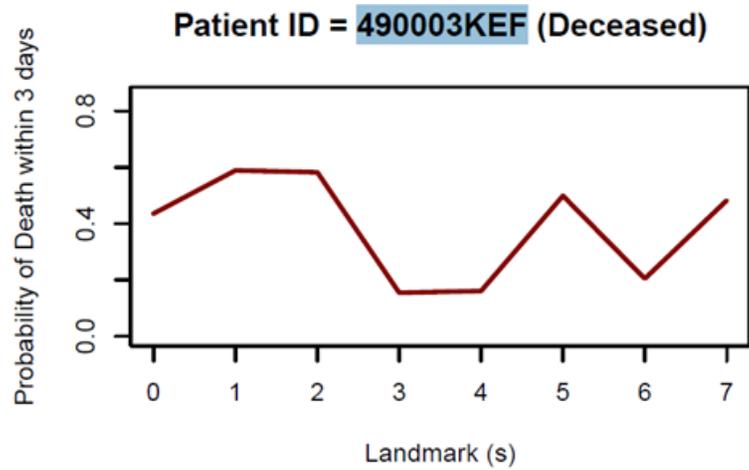


Figure 9: Non-random selection of personalized conditional mortality trajectories for 4 deceased PALF patients.

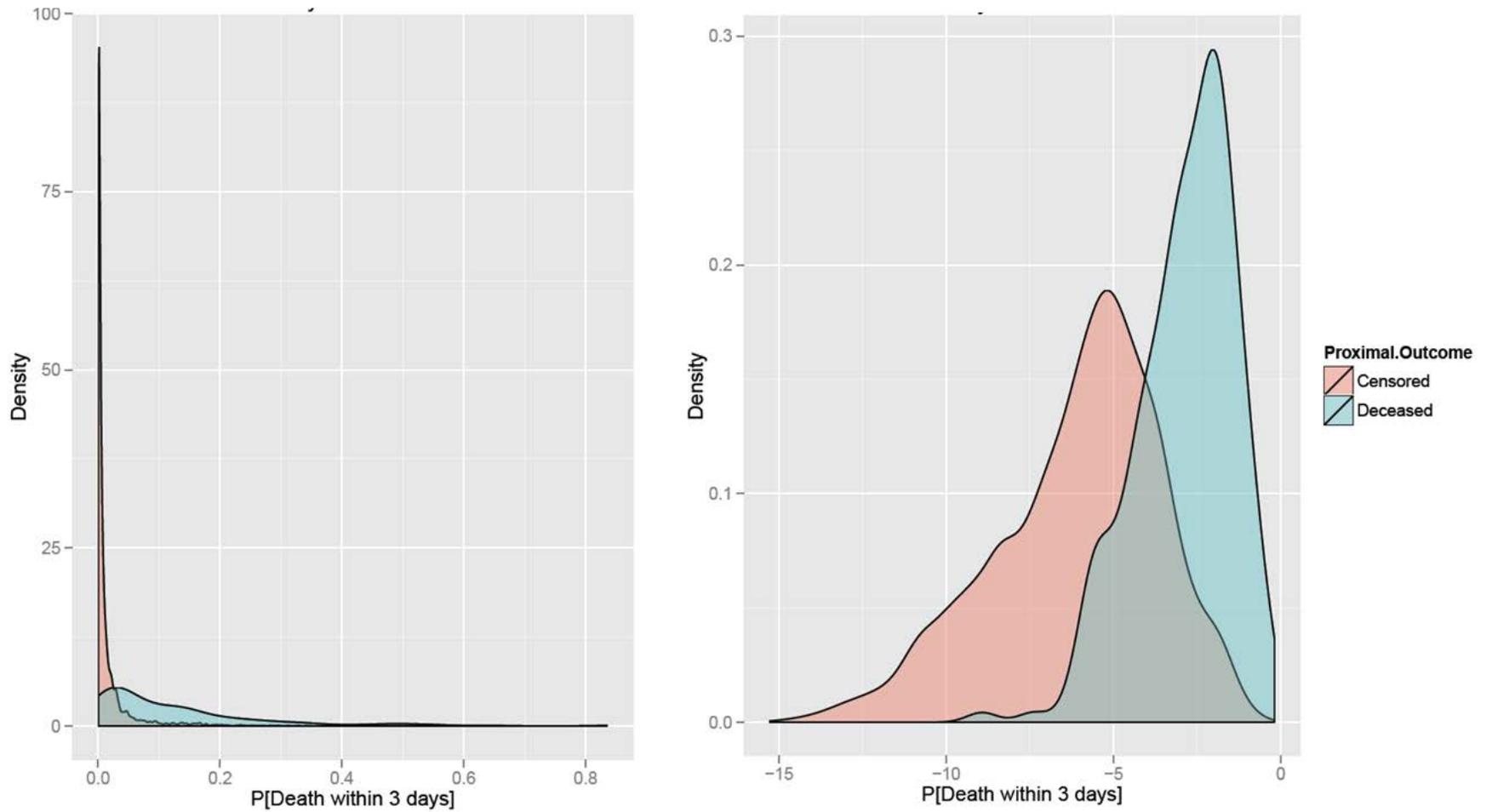


Figure 10: Distributions of conditional three-day mortality probability estimates and log mortality probably probability estimates, respectively, aggregated over all landmark time points s .

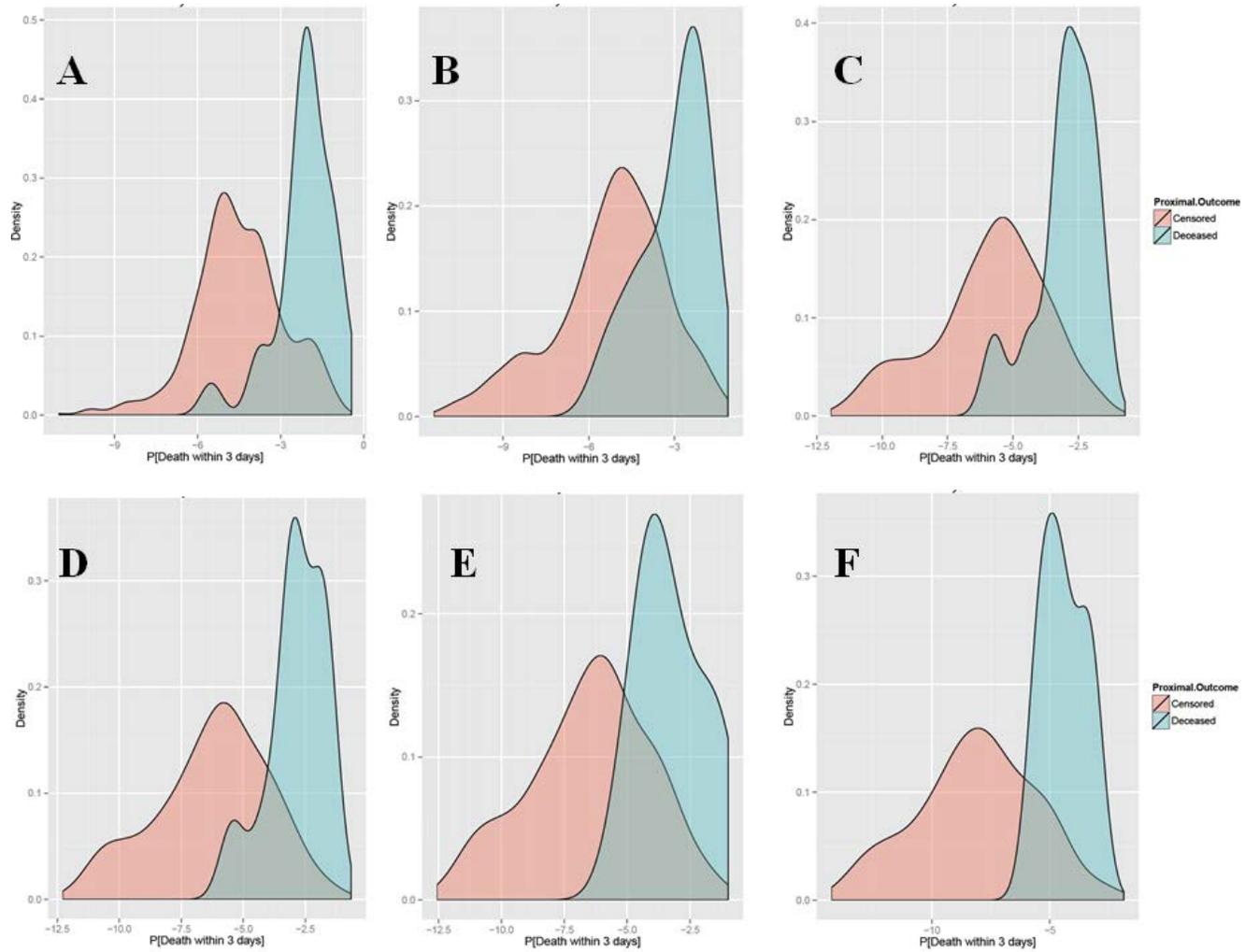


Figure 11: Distributions of conditional three-day mortality log probability estimates, stratified by landmark time point s . Probabilities are log transformed to enhance visual separation. (A) Day 0 (B) Day 3 (C) Day 6 (D) Day 9 (E) Day 12 (F) Day 15

3.6 EVALUATION OF MODEL PERFORMANCE

The bootstrapped estimates of the supermodel AUC aggregated across all landmark time points s indicated a moderate level of discriminative ability, with a marginal AUC of 73.3 (95% CI: 71.0 – 78.3) (Figure 12). Although such an evaluation of predictive value is technically improper due to the conditional nature of the estimates, the estimate provides a preliminary assessment of discriminative ability. The aggregated AUC estimate falls within the confidence interval of an oft-cited study in the literature (Lu, 2013). The estimates of dynamic AUC changed considerably as a function of landmark time, highlighted by a range of 66.8 – 92.1 (Table 6, Figures 13, 14). The lowest and highest predictive value of the landmark supermodel occurred on the third and sixteenth day following enrollment with estimates of 66.8 (95% CI: 63.9 – 85.1) and 92.1 (95% CI: 90.3 – 93.8), respectively.

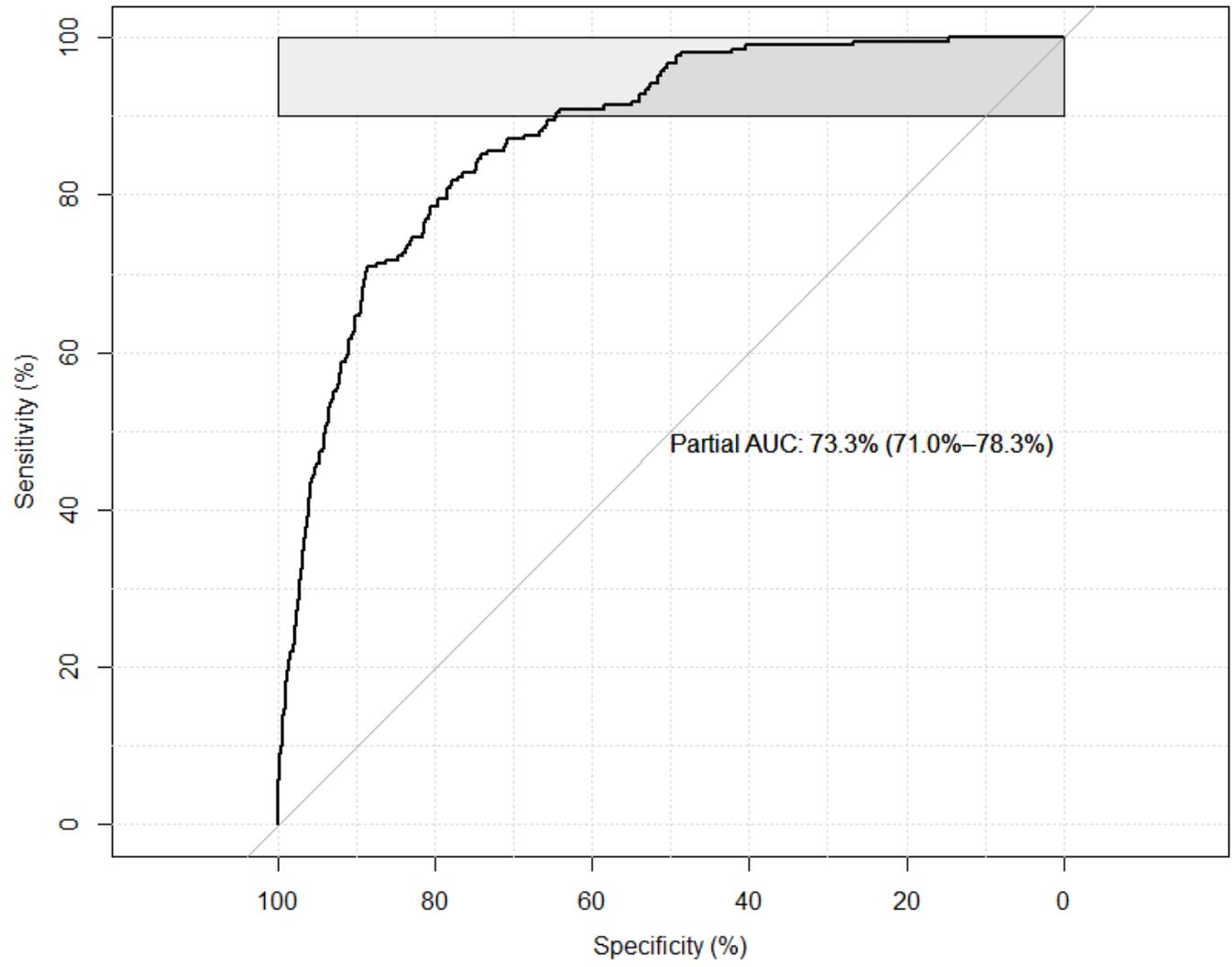


Figure 12: Bootstrapped estimate of AUC for the landmark supermodel (73.3%), aggregated across all landmark time points s .

Table 6: Bootstrapped estimates of dynamic AUC for the proportional baselines landmark supermodel.

Landmark Time Point s	AUC(s) (95% CI)
Day 0	83.23 (69.62 , 90.08)
Day 1	77.76 (69.64 , 86.21)
Day 2	72.46 (65.90 , 81.86)
Day 3	69.35 (62.28 , 87.14)
Day 4	66.82 (63.94 , 85.08)
Day 5	70.56 (66.48 , 84.33)
Day 6	76.35 (72.38 , 88.86)
Day 7	84.71 (77.91 , 91.96)
Day 8	85.69 (79.25 , 94.26)
Day 9	82.58 (79.07 , 93.10)
Day 10	82.24 (79.33 , 91.23)
Day 11	84.20 (80.22 , 90.88)
Day 12	84.40 (80.78 , 89.54)
Day 13	75.03 (69.41 , 90.55)
Day 14	76.90 (74.57 , 93.28)
Day 15	91.01 (89.52 , 93.21)
Day 16	92.12 (90.34 , 93.76)
Day 17	73.19 (70.23 , 94.21)
Day 18	74.93 (70.74 , 99.10)
Aggregate/ Marginal	73.30 (71.04 , 78.31)

Estimates of AUC are reported with 95% confidence intervals, calculated for 100 bootstrap iterations.

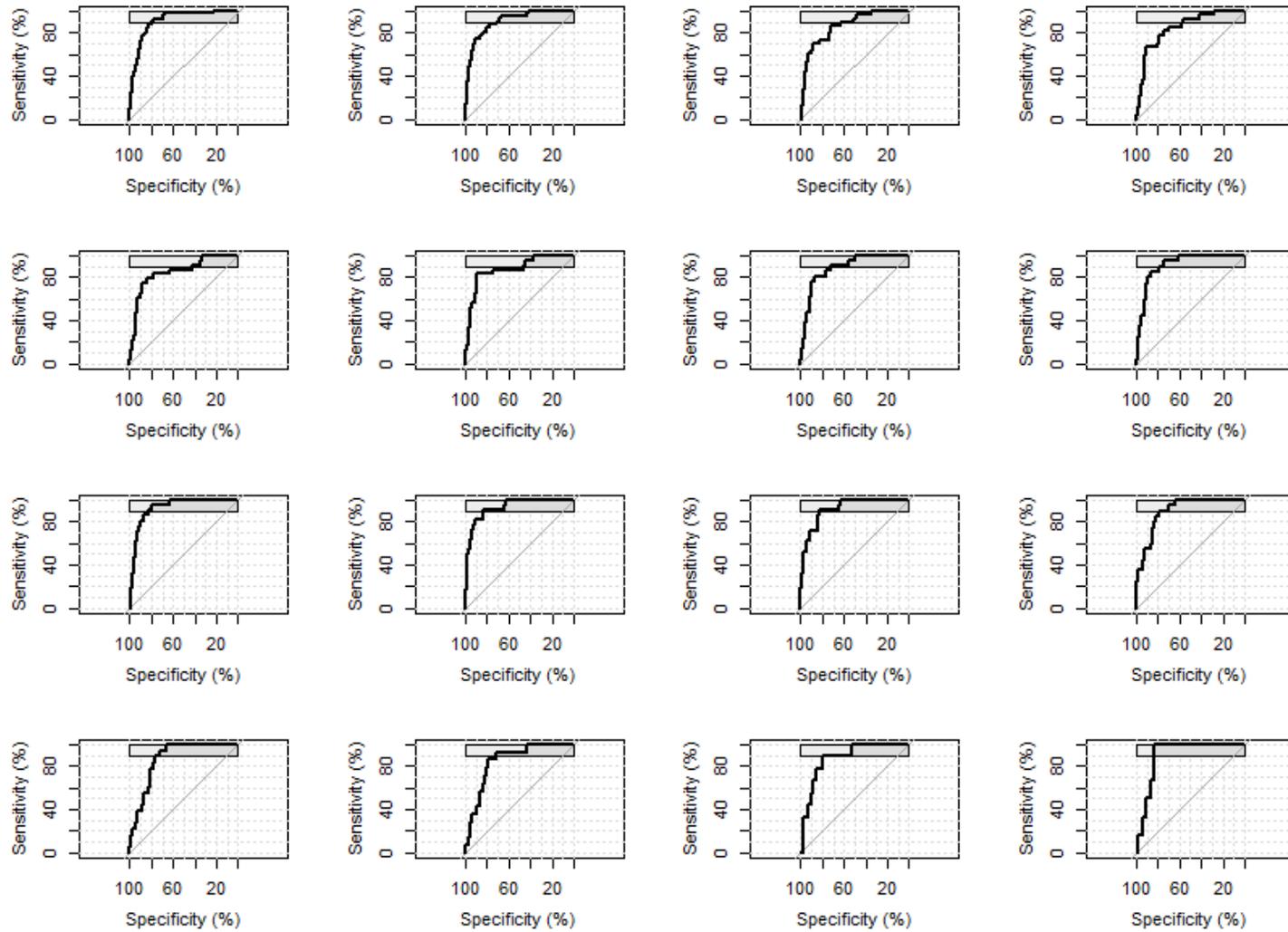


Figure 13: Bootstrapped estimates of dynamic AUC for the landmark supermodel at distinct landmark time points $s \in (s_0, s_{15})$, or simply each of the first sixteen days of the PALF Study.

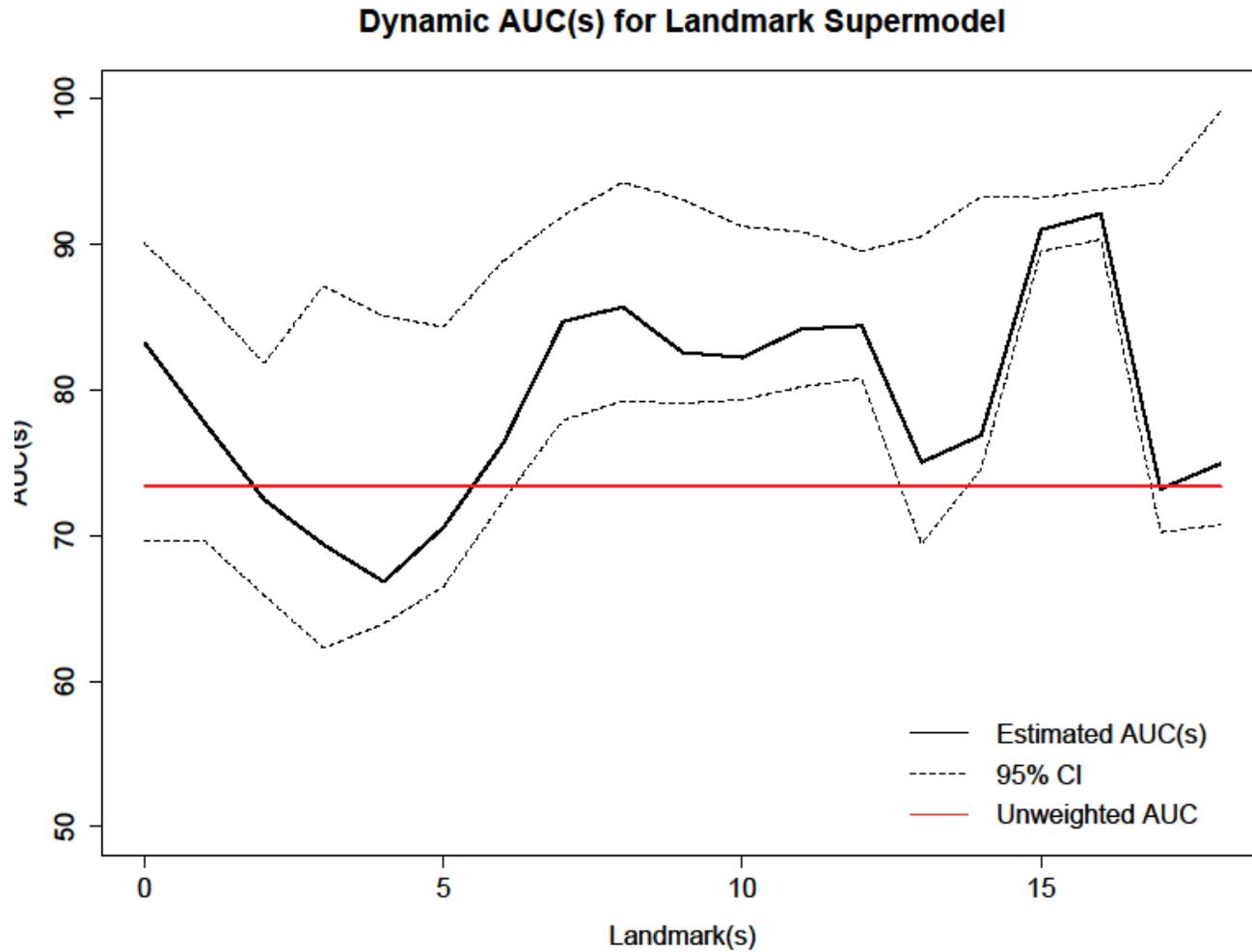


Figure 14: Bootstrapped estimates of dynamic AUC for the landmark supermodel plotted as a function of landmark time point s ; the red line indicates the aggregated (or unweighted) estimate of AUC.

4.0 DISCUSSION

Pediatric acute liver failure is a devastating clinical syndrome for which timely transplantation represents the only suitable treatment option for the majority of patients. Despite an influx of efforts to understand the underlying physiological mechanisms and improve management protocols, the mortality rate remains high. In this study of a multi-center cohort of children with ALF, we developed a dynamic prediction model of pre-transplant survival to mitigate the uncertainty surrounding the complex clinical decision making process. The model performed well, reporting a broad range of AUC estimates. We observed the lowest and highest predictive value of the landmark supermodel occurred on the third and sixteenth day following enrollment, with estimates respective of 66.8 and 92.1. The extent of variability is not surprising given the dynamic nature of the clinical process and constantly evolving risk set.

As time progresses from study enrollment, the dynamic prediction model provides more relevant prognostic information in comparison to traditional survival analyses [and less informative methods of modeling such as logistic regression]. Although we cannot formally test a related hypothesis, the close examination of predictions enables us to support our claim. Figure 11 reveals a rapid surge in discriminative ability beginning in day four and extending to day sixteen, with one noticeable deviation. This late boost in predictive ability illustrates the ability of conditional survival methodology to account for temporal changes in the cohort of interest.

There are several limitations to our approach. The study is retrospective in design, as the analysis is not primary but secondary in form. Specifically, we may not be cognizant of temporal changes in management or measurement practices dating back to 1999. The lack of a validation cohort is another weakness in the analysis. We plan to investigate external generalizability of our model by applying the landmark supermodel to patients of PALF Study from countries other than the United States.

The availability of data, or lack thereof, beyond the first seven days presents a notable limitation of the study as well. For 75% of the patients in our analysis, updated health status data is not available beyond the first week. Resultantly, a majority of the estimates in the latter weeks are subject to a potential attenuation effect due to aging of the time-dependent covariates.

The most significant of limitations pertains to a common difficulty in survival analysis. Our analysis attempts to estimate the likelihood of mortality in the presence of a strong competing risk: emergency transplant. As with any analysis subject to a competing risk, the possibility of biased estimates is non-ignorable. However, the competing risk landmark supermodel methodology is not currently available although its development is long underway. The typical approach in the related literature is to model the outcome as a composite event (group died and received a transplant together). Given the association between pre-transplant death and OLT, the estimation procedure for clustering the events together as a unit is considerably less complicated. Results published in the [Lu et al. \(2013\)](#) secondary evaluation of the Liver Injury Unit measure supports this claim. Notably, the c-index of the Liver Injury Unit for predicting death versus survival (ignoring transplant altogether) is 0.76 based on a convenience sample selected on the basis of valid ammonia data ($n = 276$). We explicitly elected to model deaths alone and not the joint event, in order to mimic the actualization of the medical decision making process in a real

life scenario. When evaluating a PALF patient the physician inherently estimates the probability of death, not the probability of death or receiving a transplant.

On a related note, our study imposed weak requirements for inclusion with respect to missing data. That is, we did not limit the analysis sample size on the basis of the availability of the strongest predictors. The sole inclusion of cases with venous ammonia measurements restricts our sample from 658 patients to roughly half of the original size. This practice inherently introduces bias, typically by reducing heterogeneity in the sample or assuming the availability of powerful yet oft-absent information. Under these conditions we would expect an ill-informed increase in predictive ability.

Regarding novelty, our study may be the first to explicitly assign probabilities of clinical outcomes to liver transplantation. Additionally, we believe our study may be the first apply conditional survival modeling to an acute scenario, regardless of disease. Updated prognosis is particularly relevant to acute conditions, as clinicians are asked to repeatedly assess the appropriateness of high-risk interventions for a given patient. Independent of the short-term success with modeling acute liver failure in children, landmarking modeling appears to be a promising quantitative method for analyzing dynamic changes in risk sets and conditional survival. Further research is warranted.

Public Health Impact:

Dynamic prediction models are of great interest to clinicians and patients alike, enabling well-informed decisions in light of the unpredictable nature of clinical and pathophysiological systems. Extensions of our model may be utilized to facilitate proper allocation of scarce resources, such as donor organs.

APPENDIX A: ALTERNATIVE LANDMARK SUPERMODELS

- i. *Aging Covariate Model* allows for time-varying covariate effects within the prediction window w
- ii. *Derivative-Type Model* incorporates the rate of change in laboratory value vector {ALT, AST, Bilirubin, PT} since the last time of measurement

A.1 AGING COVARIATE EFFECTS (LINEAR)

	coef	exp(coef)	se(coef)	robust se	z	p
ast	0.6583	1.9315	0.1645	0.1446	4.55	5.3e-06
alt	-0.4263	0.6529	0.1560	0.1384	-3.08	2.1e-03
ptp	0.6722	1.9585	0.3356	0.2697	2.49	1.3e-02
tbili	0.6932	2.0001	0.0985	0.1523	4.55	5.3e-06
i.respr	2.1383	8.4852	0.1807	0.2882	7.42	1.2e-13
i.icu	1.0913	2.9781	0.3349	0.5635	1.94	5.3e-02
ptptmins	0.2730	1.3139	0.2024	0.1222	2.23	2.5e-02
asttmins	-0.0836	0.9198	0.0987	0.0461	-1.81	7.0e-02
alttmins	0.0997	1.1048	0.0926	0.0392	2.54	1.1e-02
LM1	2.9070	18.3023	3.2907	1.3163	2.21	2.7e-02
LM2	-3.1275	0.0438	3.5062	1.0512	-2.98	2.9e-03

A.2 AGING COVARIATE EFFECTS (SQUARE)

	coef	exp(coef)	se(coef)	robust se	z	p
ast	0.6105	1.8413	0.1208	0.1319	4.63	3.7e-06
alt	-0.3644	0.6946	0.1155	0.1308	-2.79	5.3e-03
ptp	0.8247	2.2811	0.2475	0.2588	3.19	1.4e-03
tbili	0.6933	2.0003	0.0985	0.1523	4.55	5.3e-06
i.respr	2.1383	8.4852	0.1807	0.2881	7.42	1.2e-13
i.icu	1.0915	2.9788	0.3349	0.5635	1.94	5.3e-02
ptptmins2	0.0881	1.0921	0.0664	0.0395	2.23	2.6e-02
asttmins2	-0.0265	0.9739	0.0324	0.0145	-1.83	6.8e-02
alттmins2	0.0299	1.0303	0.0305	0.0126	2.37	1.8e-02
LM1	2.9683	19.4584	3.2490	1.2783	2.32	2.0e-02
LM2	-2.9925	0.0502	3.4952	1.0079	-2.97	3.0e-03

A.3 AGING COVARIATE EFFECTS (LINEAR) WITH INTERACTIONS

	coef	exp(coef)	se(coef)	robust se	z	p
ast	0.851	2.3419	0.1858	0.1874	4.54	5.6e-06
alt	-0.608	0.5447	0.1843	0.1902	-3.19	1.4e-03
ptp	0.883	2.4183	0.3561	0.2695	3.28	1.0e-03
tbili	0.961	2.6150	0.1357	0.1927	4.99	6.1e-07
i.respr	2.026	7.5862	0.1832	0.2861	7.08	1.4e-12
i.icu	1.173	3.2325	0.3364	0.5693	2.06	3.9e-02
ptptmins	0.248	1.2813	0.2016	0.1167	2.12	3.4e-02
asttmins	-0.101	0.9037	0.0995	0.0487	-2.08	3.7e-02
alттmins	0.118	1.1256	0.0989	0.0466	2.54	1.1e-02
LM1	3.532	34.2009	3.3638	1.5201	2.32	2.0e-02
LM2	-3.985	0.0186	3.5556	1.0987	-3.63	2.9e-04
ptp:tbili	-0.550	0.5767	0.2104	0.2698	-2.04	4.1e-02
ast:tbili	-0.381	0.6831	0.1101	0.1495	-2.55	1.1e-02
alt:tbili	0.308	1.3610	0.1098	0.1546	1.99	4.6e-02

A.4 AGING COVARIATE EFFECTS (SQUARE) WITH INTERACTIONS

	coef	exp(coef)	se(coef)	robust se	z	p
ast	0.7914	2.2066	0.1464	0.1750	4.52	6.1e-06
alt	-0.5332	0.5867	0.1438	0.1774	-3.01	2.7e-03
ptp	1.0190	2.7705	0.2752	0.2619	3.89	1.0e-04
tbili	0.9606	2.6132	0.1357	0.1925	4.99	6.1e-07
i.respr	2.0262	7.5850	0.1831	0.2861	7.08	1.4e-12
i.icu	1.1731	3.2320	0.3364	0.5693	2.06	3.9e-02
ptptmins2	0.0812	1.0846	0.0662	0.0376	2.16	3.1e-02
asttmins2	-0.0318	0.9687	0.0326	0.0153	-2.08	3.7e-02
alttmins2	0.0355	1.0361	0.0324	0.0149	2.38	1.7e-02
LM1	3.6495	38.4548	3.3181	1.4716	2.48	1.3e-02
LM2	-3.8593	0.0211	3.5473	1.0513	-3.67	2.4e-04
ptp:tbili	-0.5514	0.5761	0.2102	0.2695	-2.05	4.1e-02
ast:tbili	-0.3804	0.6836	0.1101	0.1496	-2.54	1.1e-02
alt:tbili	0.3071	1.3595	0.1098	0.1548	1.98	4.7e-02

A.5 DERIVATIVE-TYPE MODEL WITH AGING COVARIATES

	coef	exp(coef)	se(coef)	robust se	z	p
ast	0.5512	1.7353	0.0808	0.1243	4.43	9.3e-06
alt	-0.3833	0.6816	0.1152	0.1254	-3.06	2.2e-03
ptp	1.1695	3.2204	0.1784	0.3045	3.84	1.2e-04
tbili	0.7320	2.0792	0.0999	0.1555	4.71	2.5e-06
i.respr	2.0808	8.0108	0.1807	0.2903	7.17	7.6e-13
i.icu	1.1605	3.1915	0.3368	0.5709	2.03	4.2e-02
pc	-0.6532	0.5204	0.3524	0.5028	-1.30	1.9e-01
a2c	-0.6738	0.5098	0.1849	0.2286	-2.95	3.2e-03
alttmins	0.0493	1.0506	0.0600	0.0305	1.62	1.1e-01
LM1	1.6489	5.2014	2.8236	1.1011	1.50	1.3e-01
LM2	-2.3803	0.0925	3.3692	0.9194	-2.59	9.6e-03

A.6 DERIVATIVE-TYPE MODEL WITH AGING COVARIATES AND INTERACTIONS

	coef	exp(coef)	se(coef)	robust se	z	p
ast	0.8281	2.2889	0.1824	0.1808	4.58	4.7e-06
alt	-0.6095	0.5436	0.1789	0.1786	-3.41	6.4e-04
ptp	0.9278	2.5289	0.3527	0.2671	3.47	5.1e-04
tbili	0.9745	2.6498	0.1359	0.1924	5.06	4.1e-07
i.respr	1.9414	6.9684	0.1845	0.2908	6.68	2.5e-11
i.icu	1.2283	3.4153	0.3388	0.5764	2.13	3.3e-02
tc	0.5554	1.7426	0.2843	0.2675	2.08	3.8e-02
a2c	-0.8536	0.4259	0.1767	0.1998	-4.27	1.9e-05
ptptmins	0.2415	1.2731	0.2003	0.1162	2.08	3.8e-02
asttmins	-0.0806	0.9226	0.0983	0.0462	-1.75	8.1e-02
aluttmins	0.1105	1.1168	0.0965	0.0413	2.67	7.5e-03
LM1	2.5321	12.5800	3.3737	1.5240	1.66	9.7e-02
LM2	-3.1070	0.0447	3.5875	1.1542	-2.69	7.1e-03
ptp:tbili	-0.6093	0.5437	0.2102	0.2667	-2.28	2.2e-02
ast:tbili	-0.4000	0.6703	0.1103	0.1527	-2.62	8.8e-03
alt:tbili	0.3100	1.3634	0.1102	0.1538	2.02	4.4e-02

A.7 DERIVATIVE-TYPE MODEL WITH LANDMARK-DEPENDENT EFFECTS

	coef	exp(coef)	se(coef)	robust se	z	p
ptp	1.006	2.7353	0.1820	0.293	3.435	5.9e-04
tbili	0.252	1.2870	0.1457	0.232	1.090	2.8e-01
ast	0.647	1.9090	0.1062	0.143	4.514	6.3e-06
alt	-0.323	0.7236	0.0812	0.123	-2.629	8.6e-03
LM1	2.328	10.2580	2.9118	1.284	1.813	7.0e-02
LM2	-3.551	0.0287	3.3325	1.609	-2.207	2.7e-02
i.respr	2.198	9.0041	0.2315	0.384	5.729	1.0e-08
f2ptp	1.915	6.7858	0.4341	0.658	2.911	3.6e-03
i.icu	1.168	3.2146	0.3415	0.574	2.033	4.2e-02
a2c	-0.720	0.4866	0.1835	0.218	-3.302	9.6e-04
f2ast	-0.451	0.6368	0.2347	0.317	-1.423	1.5e-01
pc	-0.611	0.5426	0.3518	0.499	-1.226	2.2e-01
tc	0.474	1.6071	0.2767	0.271	1.752	8.0e-02
f3i.respr	-0.969	0.3794	0.6642	1.073	-0.903	3.7e-01

**A.8 DERIVATIVE-TYPE MODEL WITH LANDMARK-DEPENDENT EFFECTS
AND INTERACTIONS**

	coef	exp(coef)	se(coef)	robust se	z	p
ptp	1.567	4.7941	0.2288	0.307	5.11	3.2e-07
tbili	0.256	1.2919	0.1573	0.254	1.01	3.1e-01
ast	0.645	1.9063	0.1122	0.153	4.22	2.5e-05
alt	-0.393	0.6752	0.0920	0.140	-2.81	4.9e-03
LM1	2.111	8.2568	2.9097	1.376	1.53	1.3e-01
LM2	-4.363	0.0127	3.3183	1.592	-2.74	6.1e-03
i.respr	2.056	7.8185	0.1877	0.295	6.98	2.9e-12
f2ptp	2.471	11.8334	0.5261	0.980	2.52	1.2e-02
i.icu	1.215	3.3717	0.3419	0.572	2.13	3.4e-02
a2c	-0.701	0.4962	0.1905	0.215	-3.27	1.1e-03
f2ast	-0.499	0.6074	0.2416	0.333	-1.50	1.3e-01
pc	-0.609	0.5441	0.3559	0.494	-1.23	2.2e-01
tc	0.440	1.5534	0.2889	0.281	1.57	1.2e-01
ptp:tbili	-0.802	0.4486	0.2386	0.272	-2.94	3.3e-03
tbili:alt	0.182	1.2001	0.0719	0.122	1.49	1.4e-01
ptp:ast	-0.245	0.7830	0.1104	0.151	-1.62	1.1e-01

**A.9 AGING COVARIATE EFFECT MODEL WITH LANDMARK-DEPENDENT
EFFECTS AND INTERACTIONS**

	coef	exp(coef)	se(coef)	robust se	z	p
ast	0.703	2.02e+00	0.1264	0.158	4.45	8.5e-06
alt	-0.457	6.33e-01	0.1094	0.147	-3.10	1.9e-03
ptp	1.204	3.33e+00	0.2713	0.324	3.72	2.0e-04
tbili	0.237	1.27e+00	0.1545	0.244	0.97	3.3e-01
i.respr	2.151	8.60e+00	0.1880	0.296	7.28	3.3e-13
i.icu	1.184	3.27e+00	0.3438	0.583	2.03	4.2e-02
LM1	3.782	4.39e+01	2.9339	1.365	2.77	5.6e-03
LM2	-7.926	3.61e-04	4.0760	2.608	-3.04	2.4e-03
f2ptp	2.533	1.26e+01	0.5191	0.958	2.65	8.2e-03
ptpbst	9.888	1.97e+04	6.5254	7.600	1.30	1.9e-01
f2ast	-0.741	4.77e-01	0.2909	0.332	-2.23	2.5e-02
altbst	3.414	3.04e+01	2.2920	2.455	1.39	1.6e-01
alt:tbili	0.189	1.21e+00	0.0711	0.120	1.57	1.2e-01
ptp:tbili	-0.840	4.32e-01	0.2433	0.292	-2.88	4.0e-03
ast:ptp	-0.205	8.15e-01	0.1180	0.155	-1.32	1.9e-01

APPENDIX B: R CODE

Brief sample of the R Code utilized to estimate the landmark supermodel and calculate dynamic predictions at the population- and patient-level. Code is an adaptation and extension of the *dynpred* package (Putter, 2011) created and maintained for the purpose of complementing the textbook of van Houwelingen and Putter (2012).

```
rm(list=ls())

#import data from SAS;
#library(Hmisc)
#PALF=sasxport.get('C:/Users/dsd21/desktop/Thesis PC/PALF2_bio_change.xpt')
#save(PALF,file='C:/Users/dsd21/desktop/Thesis PC/PALF2_bio_change.Rda')

load('C:/Users/dsd21/desktop/Thesis PC/PALF2_bio_change.Rda')          #PC
names(PALF)
summary(PALF)
require(dynpred)
require(car)
require(leaps)

#####
### Define parameters
#####

#3140 of ammonia missing among all HF entries
#specify window length
w = 3
lastLM = 21 - w
gap = 1

#unblock these if want to used censored times
#PALF$t.event=PALF$t.eventc
PALF$event1=PALF$event1c
PALF$t.event=PALF$t.eventc
#PALF$event1=PALF$event12

#
```

```

#PALF$t.eventc=.0000000001+PALF$t.eventc
#PALF$t.event=.0000000001+PALF$t.event
PALF$t.eventc=.5+PALF$t.eventc
PALF$t.event=.5+PALF$t.event

#dxi
PALF$dxi=0
PALF$dxi=ifelse(PALF$dx=="Indeterminate", 1, 0)

#modify age
PALF$age = 1
PALF$age[PALF$age.grp=="0-6 months"] <- 0

#transformed in SAS
#log transform
#PALF$ptp=log(PALF$ptp)
#PALF$tbili=log(PALF$tbili)
#PALF$ast=log(PALF$ast)
#PALF$alt=log(PALF$alt)
#PALF$pc=log(PALF$pc)
#PALF$tc=log(PALF$tc)
#PALF$a1c=log(PALF$a1c)
#PALF$a2c=log(PALF$a2c)

par(mfrow=c(2,2))
hist(PALF$ptp)
hist(PALF$tbili)
hist(PALF$ast)
hist(PALF$alt)

#center continuous variables
PALF$ptp = PALF$ptp - mean(PALF$ptp)
PALF$tbili = PALF$tbili - mean(PALF$tbili)
PALF$ast = PALF$ast - mean(PALF$ast)
PALF$alt = PALF$alt - mean(PALF$alt)

#okay to run if not used
PALF$tc=PALF$tc/PALF$tel
PALF$pc=PALF$pc/PALF$tel
PALF$a1c=PALF$a1c/PALF$tel
PALF$a2c=PALF$a2c/PALF$tel

par(mfrow=c(2,2))
hist(PALF$pc)
hist(PALF$tc)
hist(PALF$a1c)
hist(PALF$a2c)

PALF[is.na(PALF)] <- 0

#check vif // no issues

check=lm(t.event ~ ptp + tbili + i.respr + i.icu + ast + alt
          + ptp*tbili +tbili*alt + ptp*ast

```

```

      #+ cluster(ttime)
      , data=PALF)
vif(check)

vcov(check)

#####
#####
### Section 7.2
#####
#####

#prep data with extra stuff for tracking changes in biomarker levels
## Overall survival
palfsurv <- NULL
n <- length(unique(PALF$id))
nid <- length(unique(PALF$id))
n2 <- unique(PALF$id)
for (i in 1:n) {
  tbilii <- PALF[PALF$id==n2[i],]
  n2i <- nrow(tbilii)

  tbilii$Tstart <- tbilii$time
  tbilii$Tstop <- c(tbilii$Tstart[-1],tbilii$event[1])
  tbilii$status <- c(rep(0,nrow(tbilii)-1),tbilii$event1[1])
  #tbilii$TEL <- 0
  #tbilii$TEL <- tbilii$Tstop - tbilii$Tstart
  #tbilii$tc = tbilii$tc/tbilii$TEL
  #tbilii$pc = tbilii$pc/tbilii$TEL
  #tbilii$a1c = tbilii$a1c/tbilii$TEL
  #tbilii$a2c = tbilii$a2c/tbilii$TEL
  palfsurv <- rbind(palfsurv,tbilii)
}

palfsurv[1:20,]

#varified tel is correct
#hist(palfsurv$tel)

#####
Cross-tabulation of first event status and survival status
#####

pslast <- palfsurv[c(which(!duplicated(palfsurv$id))[-1]-1,nrow(palfsurv)),]
table(pslast$status,pslast$event1)

pslast[1:20,]

pred=6

logHR=matrix(0,lastLM+1,pred)

```

```

se=matrix(0,lastLM+1,pred)
lb=matrix(0,lastLM+1,pred)
ub=matrix(0,lastLM+1,pred)

LMdata <- NULL
LMs <- seq(0,lastLM,by=gap)
for (LM in LMs) {
  LMdataLM <- cutLM(data=PALF,outcome=list(time="t.event",status="event1"),
    LM=LM,horizon=LM+w,covs=list(fixed=c("ast","alt","ptp","i.respr","i.icu",
      "ascites","age.grp","dxi","age","dx","tc","pc","a1c","a2c","tel","event2"),
      varying="tbili"),format="long",id="id",rtime=c("ttime"),right=FALSE)

  LMdataLM <- LMdataLM[LM - LMdataLM$time <= 100,] # of wbctime?
  #LMdataLM <- LMdataLM[LM - LMdataLM$time,]
  LMdataLM <- LMdataLM[!is.na(LMdataLM$id),]
  LMdata <- rbind(LMdata,LMdataLM)

  LMcox <- coxph(Surv(LM,t.event,event1) ~ ast + alt + ptp + tbili + i.respr + i.icu
    + cluster(id), data=LMdata, method="breslow",robust=TRUE)
  se[LM+1,] <- (sqrt(diag(LMcox$var)))
  logHR[LM+1,] <- LMcox$coef[1:pred]
}

ub=logHR+1.96*se
lb=logHR-1.96*se

par(mfrow=c(2,3))
for (i in 1:pred) {
  plot(LMs,logHR[i],type="s",lwd=2,
    xlim=c(0,lastLM),ylim=c(-1,4),
    xlab="Time (days)",ylab="Log hazard ratio",lty=1)
  lines(LMs,rep(0,lastLM+1),"s",lty=1,col="red")
  lines(LMs,lb[i],type="s",lty=2)
  lines(LMs,ub[i],type="s",lty=2)
}

#####
Results of landmark analyses for the WBC counts;
### s runs from 0.5 to 3.5 with steps of 0.1; window width w=4
#####

## Simple (ip1)
LMdata$Tstart <- LMdata$LM
LMsupercox0 <- coxph(Surv(Tstart,t.event,event1) ~ ast + alt + ptp + tbili + i.respr + i.icu
  + strata(LM) + cluster(id), data=LMdata, method="breslow",robust=TRUE)
LMsupercox0

## Extended
tt <- sort(unique(LMdata$t.event[LMdata$event1==1]))
dim(LMdata)

```

```
LMdata2 <- survSplit(data=LMdata, cut=tt, end="t.event", start="Tstart", event="event1")
dim(LMdata2)
```

```
LMdata2$tbilitmins <- LMdata2$tbili*(LMdata2$t.event - LMdata2$LM)
LMdata2$tbilitmins2 <- LMdata2$tbili*(LMdata2$t.event - LMdata2$LM)^2
```

```
LMdata2$ptptmins <- LMdata2$ptp*(LMdata2$t.event - LMdata2$LM)
LMdata2$ptptmins2 <- LMdata2$ptp*(LMdata2$t.event - LMdata2$LM)^2
```

```
LMdata2$astmins <- LMdata2$ast*(LMdata2$t.event - LMdata2$LM)
LMdata2$astmins2 <- LMdata2$ast*(LMdata2$t.event - LMdata2$LM)^2
```

```
LMdata2$altmins <- LMdata2$alt*(LMdata2$t.event - LMdata2$LM)
LMdata2$altmins2 <- LMdata2$alt*(LMdata2$t.event - LMdata2$LM)^2
```

```
LMdata2$i.resprmins <- LMdata2$i.respr*(LMdata2$t.event - LMdata2$LM)
LMdata2$i.resprmins2 <- LMdata2$i.respr*(LMdata2$t.event - LMdata2$LM)^2
```

```
LMdata2$i.icutmins <- LMdata2$i.icu*(LMdata2$t.event - LMdata2$LM)
LMdata2$i.icutmins2 <- LMdata2$i.icu*(LMdata2$t.event - LMdata2$LM)^2
```

```
f2 <- function(t) (t/lastLM)
f3 <- function(t) (t/lastLM)^2
```

```
LMdata2$f2ast = LMdata2$ast*f2(LMdata2$LM)
LMdata2$f3ast = LMdata2$ast*f3(LMdata2$LM)
LMdata2$f2alt = LMdata2$alt*f2(LMdata2$LM)
LMdata2$f3alt = LMdata2$alt*f3(LMdata2$LM)
LMdata2$f2ptp = LMdata2$tbili*f2(LMdata2$LM)
LMdata2$f3ptp = LMdata2$tbili*f3(LMdata2$LM)
LMdata2$f2tbili = LMdata2$tbili*f2(LMdata2$LM)
LMdata2$f3tbili = LMdata2$tbili*f3(LMdata2$LM)
LMdata2$f2i.respr = LMdata2$i.respr*f2(LMdata2$LM)
LMdata2$f3i.respr = LMdata2$i.respr*f3(LMdata2$LM)
LMdata2$f2i.icu = LMdata2$i.icu*f2(LMdata2$LM)
LMdata2$f3i.icu = LMdata2$i.icu*f3(LMdata2$LM)
```

```
# ipl
#linear effects
LMsupercox11 <- coxph(Surv(Tstart,t.event,event1) ~ ast + alt + ptp + tbili + i.respr + i.icu
+ strata(LM) + cluster(id), data=LMdata2, method="breslow")
```

```
#polynomial effects
LMsupercox12 <- coxph(Surv(Tstart,t.event,event1) ~ ast + alt + ptp + tbili + i.respr + i.icu
+ strata(LM) + cluster(id), data=LMdata2, method="breslow")
```

```
LMsupercox11
LMsupercox12
```

```
# ipl*
g1 <- function(t) (t/lastLM)
g2 <- function(t) (t/lastLM)^2
```

```

LMdata2$LM1 <- g1(LMdata2$LM)
LMdata2$LM2 <- g2(LMdata2$LM)

LMsupercox21 <- coxph(Surv(Tstart,t.event,event1) ~ ast + alt + ptp + tbili + i.respr + i.icu
  #+ptptmins + 0*ptptmins2 +astmins + 0*astmins2 +altmins + 0*altmins2
  + LM1 + LM2 + cluster(id), data=LMdata2, method="breslow")
LMsupercox22 <- coxph(Surv(Tstart,t.event,event1) ~ ast + alt + ptp + tbili + i.respr + i.icu
  #+0*ptptmins + ptptmins2 +0*astmins + astmins2 +0*altmins + altmins2
  + LM1 + LM2 + cluster(id), data=LMdata2, method="breslow")
LMsupercox21
LMsupercox22

LMsupercox2 <- coxph(Surv(Tstart,t.event,event1) ~ ptp + tbili + i.respr + i.icu + ast + alt
  + f2ptp + f2ast + ptp*tbili + ast*ptp + tbili*alt
  + LM1 + LM2 + cluster(id), data=LMdata2, method="breslow")
LMsupercox2

#####
### Figure 8.5: Time-varying effect of LWBC for the ipl and the ipl*
### models of Table 8.5
#####
par(mfrow=c(1,2))

#plot the linear effect model

tseq <- seq(0,lastLM,by=gap)
plot(tseq,coef(LMsupercox11)[["ast"]]+coef(LMsupercox11)[["astmins"]]*tseq,
  type="l",lwd=2,ylim=c(-2,6),xlab="t-s",ylab="Log hazard ratio",main="Landmark Supermodel Effects (Linear T-
V)",col="green")
lines(tseq,coef(LMsupercox21)[["ast"]]+coef(LMsupercox21)[["astmins"]]*tseq,
  type="l",lwd=2,lty=2,col="green")
lines(c(0,lastLM),rep(coef(LMsupercox0)[["ast"]],2),type="l",lty=3,col="green")

lines(tseq,coef(LMsupercox11)[["ptp"]]+coef(LMsupercox11)[["ptptmins"]]*tseq,
  type="l",lwd=2,lty=1,col="purple")
lines(tseq,coef(LMsupercox21)[["ptp"]]+coef(LMsupercox21)[["ptptmins"]]*tseq,
  type="l",lwd=2,lty=2,col="purple")
lines(c(0,lastLM),rep(coef(LMsupercox0)[["ptp"]],2),type="l",lty=3,col="purple")

lines(tseq,coef(LMsupercox11)[["alt"]]+coef(LMsupercox11)[["altmins"]]*tseq,
  type="l",lwd=2,lty=1,col="blue")
lines(tseq,coef(LMsupercox21)[["alt"]]+coef(LMsupercox21)[["altmins"]]*tseq,
  type="l",lwd=2,lty=2,col="blue")
lines(c(0,lastLM),rep(coef(LMsupercox0)[["alt"]],2),type="l",lty=3,col="blue")

lines(c(0,lastLM),rep(coef(LMsupercox11)[["tbili"]],2),type="l",lty=3,col="red")

#lines(c(0,lastLM),rep(coef(LMsupercox0)[["i.respr"]],2),type="l",lty=3,col="red")
#lines(c(0,lastLM),rep(coef(LMsupercox0)[["i.icu"]],2),type="l",lty=3,col="grey")

legend("topright",lwd=c(2,2,1),lty=1:3,c("Extended ipl","Extended ipl*","Simple ipl"),bty="n")

```

```

#plot the non-linear model

tseq <- seq(0,lastLM,by=gap)
plot(tseq,coef(LMsupercox12)[["ast"]]+coef(LMsupercox12)[["asttmins2"]]*tseq^2,
     type="l",lwd=2,ylim=c(-2,6),xlab="t-s",ylab="Log hazard ratio",main="Landmark Supermodel Effects (Non-
linear T-V)",col="green")
lines(tseq,coef(LMsupercox22)[["ast"]]+coef(LMsupercox22)[["asttmins2"]]*tseq^2,
     type="l",lwd=2,lty=2,col="green")
lines(c(0,lastLM),rep(coef(LMsupercox0)[["ast"]],2),type="l",lty=3,col="green")

lines(tseq,coef(LMsupercox12)[["ptp"]]+coef(LMsupercox12)[["ptptmins2"]]*tseq^2,
     type="l",lwd=2,lty=1,col="purple")
lines(tseq,coef(LMsupercox22)[["ptp"]]+coef(LMsupercox22)[["ptptmins2"]]*tseq^2,
     type="l",lwd=2,lty=2,col="purple")
lines(c(0,lastLM),rep(coef(LMsupercox0)[["ptp"]],2),type="l",lty=3,col="purple")

lines(tseq,coef(LMsupercox12)[["alt"]]+coef(LMsupercox12)[["alttmins2"]]*tseq^2,
     type="l",lwd=2,lty=1,col="blue")
lines(tseq,coef(LMsupercox22)[["alt"]]+coef(LMsupercox22)[["alttmins2"]]*tseq^2,
     type="l",lwd=2,lty=2,col="blue")
lines(c(0,lastLM),rep(coef(LMsupercox0)[["alt"]],2),type="l",lty=3,col="blue")

lines(c(0,lastLM),rep(coef(LMsupercox12)[["tbili"]],2),type="l",lty=3,col="red")

#lines(c(0,lastLM),rep(coef(LMsupercox0)[["i.respr"]],2),type="l",lty=3,col="red")
#lines(c(0,lastLM),rep(coef(LMsupercox0)[["i.icu"]],2),type="l",lty=3,col="grey")

legend("topright",lwd=c(2,2,1),lty=1:3,c("Extended ipl","Extended ipl*","Simple ipl"),bty="n")

#####
Baseline hazard and landmark effects in proportional
### baselines landmark supermodel
#####

means <- LMsupercox2$means
means

ndata <- data.frame(ast=0,alt=0,ptp=0,i.respr=0,i.icu=0,tbili=0,tbilitmins=0, ptptmins=0,
                    asttmins=0,alttmins=0,LM1=0,LM2=0,f2ptp=0,f3ptp=0,f2ast=0,f3ast=0)
sf2 <- survfit(LMsupercox2, newdata=ndata)
Haz0 <- data.frame(time=sf2$time,surv=sf2$surv);
Haz0$Haz <- -log(Haz0$surv)

par(mfrow=c(1,2))
par(mar=c(5,4,4,1.6)+0.1)
plot(Haz0$time, Haz0$Haz, type="s", lwd=2, xlab="Time (days)", ylab="Cumulative hazard")

par(mar=c(5,3.6,4,2)+0.1)

```

```

plot(LMs, exp(LMsupercox2$coef["LM1"]*g1(LMs)+ LMsupercox2$coef["LM2"]*g2(LMs)) , type="l", lwd=2,
xlab="Landmark (s)", ylab="exp(theta(s))")
par(mfrow=c(1,1))

```

```

#####
### Dynamic predictions of death within w = 3 years in the
### proportional baselines landmark supermodel for different trajectories
### of laboratory values
#####

tt <- PALF$t.event[PALF$event1==1]
tt <- sort(unique(c(0,tt,tt-w)))
tt <- tt[tt>=0]
tt=seq(0,lastLM,by=1)
nt <- length(tt)

# Custom-made fuction
Fwpredict <- function(bet, Haz0, xdata1, xdata2, xdata3, xdata4,xdata5, xdata6, tt)
{
  nt <- length(tt)
  Haz0$haz <- diff(c(0,Haz0$Haz))
  Fw <- data.frame(time=tt,Fw=NA)
  for (i in 1:nt) {
    sfi <- Haz0 # local copy
    tti <- tt[i]
    sfi$haz <- sfi$haz *
      exp(bet[1]*xdata1[i] + bet[7]*xdata1[i]*g1(tti) #ptp
          + bet[2]*xdata2[i] #tbili
          + bet[3]*xdata3 + bet[4]*xdata4
          + bet[5]*xdata5[i] + bet[8]*xdata5[i]*g1(tti) #ast
          + bet[6]*xdata6[i] #alt
          + bet[11]*xdata1[i]*xdata2[i] + bet[12]*xdata1[i]*xdata5[i] +bet[13]*xdata2[i]*xdata6[i]
          #interaction
          + bet[9]*g1(tti) + bet[10]*g2(tti))
    sfi$Haz <- cumsum(sfi$haz)
    tmp <- evalstep(sfi$time,sfi$Haz,c(tti,tti+w),subst=0)
    Fw$Fw[i] <- 1-exp(-(tmp[2]-tmp[1]))
  }
}

```

```

}
return(Fw)
}

```

```

base=subset(PALF,tttime==0)
means_b = base$means
ri00=subset(PALF,i.respr==0 & i.icu==0)           #no resp, no icu
means_00 = ri00$means
ri01=subset(PALF,i.respr==0 & i.icu==1)           #no resp, yes icu
means_01 = ri00$means
ri10=subset(PALF,i.respr==1 & i.icu==0)           #yes resp, no icu
means_10 = ri00$means
ri11=subset(PALF,i.respr==1 & i.icu==1)           #yes resp, yes icu
means_11 = ri00$means

```

```

# First basic, not taking into account biomarker measurements
cbas <- coxph(Surv(t.event,event1) ~ i.respr+i.icu, data=pslast, method="breslow")
sf <- survfit(cbas)
Fwbas <- Fwindow(sf,width=w)
Fwbas <- subset(Fwbas,time>=0.0)
Fwbas <- subset(Fwbas,time<=(lastLM+.01))

```

```

tt <- seq(0,lastLM,by=gap)
nt <- length(tt)
xdata1 <- rep(0,nt)           #pt
xdata2 <- rep(0,nt)           #bili
xdata3 = means[3]             #respr
xdata4 = means[4]             #icu
xdata5 <- rep(0,nt)           #ast
xdata6 <- rep(0,nt)           #alt
Fw0 <- Fwpredict(LMsupercox2$coef, Haz0, xdata1, xdata2, xdata3, xdata4, xdata5, xdata6, tt)

```

```

xdata1 <- rep(max(PALF$ptp)/3,nt)
xdata2 <- rep(max(PALF$tbili)/3,nt)
xdata3 = means[3]             #resp
xdata4 = means[4]             #icu
xdata5 <- rep(max(PALF$ast)/3,nt)   #ast
xdata6 <- rep(max(PALF$alt)/3,nt)   #alt
Fw1 <- Fwpredict(LMsupercox2$coef, Haz0, xdata1, xdata2, xdata3, xdata4, xdata5, xdata6, tt)

```

```

xdata1 <- (1-(tt/lastLM))*max(PALF$ptp)*2/3
xdata2 <- (1-(tt/lastLM))*max(PALF$tbili)*2/3
xdata3 = means[3]

```

```

xdata4 = means[4]
xdata5 <- (1-(tt/lastLM))*max(PALF$ast)*2/3
xdata6 <- (1-(tt/lastLM))*max(PALF$alt)*2/3
Fw2 <- Fwpredict(LMsupercox2$coef, Haz0, xdata1, xdata2, xdata3, xdata4, xdata5, xdata6, tt)

plot(tt,Fw0$Fw,type="l",lwd=2,ylim=c(0,.10),xlab="Landmark(s)",ylab="Probability of Death within 3
days",col="blue")
# main="Landmark Supermodel Prediction")
lines(tt,Fw1$Fw,lwd=2,lty=3)
lines(tt,Fw2$Fw,lwd=2,lty=2)
lines(Fwbas$time,Fwbas$Fw,lwd=2,col=8)

legend("topright",c("x(s)=Mean Lab Values","x(s)=Elevated but Stable Biomarker Values",
"x(s)=Highly Elevated but Decreasasing Biomarker Values","Model without Biomarker
Data"),lwd=2,lty=c(1,3,2,1),col=c(1,1,1,8),bty="n")

#hist(pslast$Tstart)

summary(pslast$Tstart)

#####
### Automated procedure to generate trajectories for all patients and AUC estimates
#####

par(mfrow=c(1,1))

#number of randomly drawn patients
#z=round(runif(1,min=1,max=16))
z=8
p <- matrix(NA,length(unique(LMdata$Sid)),lastLM+1)
death <- matrix(NA,length(unique(LMdata$Sid)),lastLM+1)

maxp <- NULL
deathf <- NULL
event <- NULL
tx <- NULL
p1 <- NULL
p0 <- NULL
p3 <- NULL
p6 <- NULL
p9 <- NULL
p12 <- NULL
p15 <- NULL
p18 <- NULL

```

```

graph.row=round(sqrt(z))
#par(mfrow=c(graph.row,ceiling(z/graph.row)))
nid <- length(unique(LMdata$Sid))
n2 <- unique(LMdata$Sid)

pdf() #ouputs as Rplot in Documents/R
par(mfrow=c(3,2))

#for (ii in 1:z) {
for (ii in 1:nid) {

u=round(runif(1,min=1,max=nid))
#temp=PALF[PALF$Sid==n2[u],]      #z random sample
temp=LMdata[LMdata$Sid==n2[ii],]  #entire sample

xdata1=temp$ptp
xdata2=temp$tbili
xdata3=temp$i.respr[1]
xdata4=temp$i.icu[1]
xdata5=temp$ast
xdata6=temp$alt
tt=temp$LM

Fw <- Fwpredict(LMsupercox2$coef, Haz0, xdata1, xdata2, xdata3, xdata4, xdata5, xdata6, tt)
maxp[ii] <- max(Fw$Fw)
last=nrow(temp)
p[ii,1:length(Fw$Fw)]=Fw$Fw
p1[ii] = Fw$Fw[last]
death[ii,1:length(Fw$Fw)]=temp$event1
deathf[ii] <- temp$event1[last]
tx[ii] <- temp$event2[last]
#event[ii] <- temp$event[1]

Outcome=factor(deathf)
levels(Outcome) = c('Censored','Deceased')
#Competing.Risk=factor(event)
#levels(Competing.Risk) = c('Censored','Deceased','Transplanted')

color="darkseagreen"
if (deathf[ii]==1) {color="darkred"}
if (tx[ii]==1) {color="blue"}      #disable this to get rid of transplant indication

style="1"
if (length(Fw$Fw)==1) {style=NULL}

```

```

#plotname=paste("predictions_ext/", Outcome[ii],"/id_", n2[ii], "_",round(1000*maxp[ii])/1000, ".pdf", sep="")

plot(tt,Fw$Fw,type=style,lwd=2,xlim=c(0,max(tt)),ylim=c(0,.85),xlab="Landmark (s)",ylab="Probability of Death
within 3 days",
  main=paste0("Patient ID = ",n2[ii]," (",Outcome[ii],")"),col=color)
}

dev.off()

table(event)

#create new outcome vector to match length of probabilities!!
m=max(maxp)
PRI=100*(p/m) #PALF risk index
#PRI=exp(p/max(p)) #PALF risk index          #modified PRI

par(mfrow=c(1,1))
hist(p)

require(ggplot2)
require(AUC)
require(pROC)

pv=as.vector(p)
#pv=na.omit(pv)

pdf('outcome_p.pdf')
Proximal.Outcome=factor(death)
levels(Proximal.Outcome) = c('Censored','Deceased')
qplot(pv, geom="density", fill=Proximal.Outcome, alpha=I(.4),
  main="Distribution of 3-day Probabilities at Enrollment", xlab="P[Death within 3 days]",
  ylab="Density")
#fill=death[c((s-1)+seq(1,length(n2)*(lastLM),by=lastLM))]
dev.off()

pdf('outcome_logp.pdf')
Proximal.Outcome=factor(death)
levels(Proximal.Outcome) = c('Censored','Deceased')
qplot(log(pv), geom="density", fill=Proximal.Outcome, alpha=I(.4),
  main="Distribution of 3-day Probabilities at Enrollment", xlab="P[Death within 3 days]",
  ylab="Density")
#fill=death[c((s-1)+seq(1,length(n2)*(lastLM),by=lastLM))]
dev.off()

#pROC commands
roc(death,pv)

```

```

roc1 = roc(death,pv, percent=TRUE,
# arguments for auc
partial.auc=c(100, 90), partial.auc.correct=TRUE,
partial.auc.focus="sens",
# arguments for ci
ci=TRUE, boot.n=100, ci.alpha=0.9, stratified=FALSE,
# arguments for plot
plot=TRUE, auc.polygon=TRUE, max.auc.polygon=TRUE, grid=TRUE,
print.auc=TRUE, show.thres=TRUE)
CI_all=c(roc1$ci[1],roc1$ci[2],roc1$ci[3])

t1=0; t2=lastLM #all probs, or up to given time t2
t1=1; t2=4 #specificity range
roc2 = roc(tdeath,tpv, percent=TRUE,
# arguments for auc
partial.auc=c(100, 90), partial.auc.correct=TRUE,
partial.auc.focus="sens",
# arguments for ci
ci=TRUE, boot.n=100, ci.alpha=0.9, stratified=FALSE,
# arguments for plot
plot=TRUE, auc.polygon=TRUE, max.auc.polygon=TRUE, grid=TRUE,
print.auc=TRUE, show.thres=TRUE)

par(mfrow=c(4,4))
days=lastLM+1
CI <- matrix(NA,days,3)

for (ti in 1:days) {
d=ti-1; t1=d; t2=d+1 #pick single day
tdeath=death[(1+(658*t1)):(658*(t2+1))]
tpv=pv[(1+(658*t1)):(658*(t2+1))]
roc2 = roc(tdeath,tpv, percent=TRUE,
# arguments for auc
partial.auc=c(100, 90), partial.auc.correct=TRUE,
partial.auc.focus="sens",
# arguments for ci
ci=TRUE, boot.n=100, ci.alpha=0.9, stratified=FALSE,
# arguments for plot
plot=TRUE, auc.polygon=TRUE, max.auc.polygon=TRUE, grid=TRUE,
print.auc=TRUE, show.thres=TRUE)
CI[ti,]=roc2$ci
}

plot(LMs,CI[,2],type="l",xlab="Landmark(s)",ylab="AUC(s)",lwd=2,
main="Dynamic AUC(s) for Landmark Supermodel",ylim=c(50,100))
lines(LMs,CI[,1],type="l",lwd=1,lty=2)
lines(LMs,CI[,3],type="l",lwd=1,lty=2)
lines(LMs,rep(roc1$ci[2],length(LMs)),col="red",lwd=2)
legend("bottomright",c("Estimated AUC(s)","95% CI","Unweighted AUC"),
lwd=1,lty=c(1,2,1),col=c(1,1,"red"),bty="n")

```

```
#select landmark time point to plot
s=8
#pdf('outcome_p.pdf')
Proximal.Outcome=factor(death[,s])
levels(Proximal.Outcome) = c('Censored','Deceased')
qplot(log(p[,s]), geom="density", fill=Proximal.Outcome, alpha=I(.4),
      main="Distribution of 3-day Probabilities at Enrollment", xlab="P[Death within 3 days]",
      ylab="Density")
#fill=death[c((s-1)+seq(1,length(n2)*(lastLM),by=lastLM))]
#dev.off()
```

BIBLIOGRAPHY

- C.G. Antoniadis, et al. Actin-free Gc globulin: a rapidly assessed biomarker of organ dysfunction in acute liver failure and cirrhosis. *Liver Transpl*, 13:1254-1261, 2007.
- ANZLTR 2010. *Australia and New Zealand Liver Transplant Registry Annual Report: 2010*.
- N. Azhar, et al. Analysis of serum inflammatory mediators identifies unique dynamic networks associated with death and spontaneous survival in pediatric acute liver failure. *PLoS One*, 8(11):e78202, 2013.
- W. Bernal, et al. Blood lactate as an early predictor of outcome in paracetamol-induced acute liver failure: a cohort study. *Lancet*, 359:558-563, 2002.
- P.Y. Chung, et al. Serum phosphorus levels predict clinical outcome in fulminant hepatic failure. *Liver Transpl*, 9:248-253, 2003.
- D.R. Cox. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B*, 34(2):187-220, 1972.
- R. Dent, et al. Triple-negative breast cancer: clinical features and patterns of recurrence. *Clinical Cancer Research*, 13(15):4429, 2007.
- D. Devictor, et al. Acute liver failure in neonates, infants and children. *Expert Rev Gastroenterol Hepatol*, 5:717-729, 2011.
- K.A. Eagle, et al. A validated prediction model for all forms of acute coronary syndrome: estimating the risk of 6-month postdischarge death in an international registry. *JAMA*, 291(22):2727-2733, 2004.
- L.C. Ee, et al. Acute liver failure in children: a regional experience. *J Paediatr Child Health*, 39:107-110, 2005.
- S. Feng. Long-term management of immunosuppression after pediatric liver transplantation. *Curr Opin Organ Transplant*, 13(5):506-512, 2008.
- S.M. Friedman, et al. Falls and fear of falling: which comes first? A longitudinal prediction model suggests strategies for primary and secondary prevention. *J Am Geriatr*, 50:1329-1335, 2002.
- F.H. Garrison. *History of Medicine*. Saunders Company. Philadelphia, PA, 1966.

- A. Goldhirsch, et al. Meeting highlights: international consensus panel on the treatment of primary breast cancer. *Journal of the National Cancer Institute*, 87(19):1441, 1995.
- R.J. Gray. Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis. *Journal of the American Statistical Association*, 87(420):942-951, 1992.
- D.M. Green and J.A. Swets. *Signal Detection Theory and Psychophysics*. Wiley and Sons, New York, NY, 1966.
- R. Hashemi, et al. A latent process model for joint modeling of events and marker. *Lifetime Data Anal*, 9(4):331-343, 2003.
- P.J. Heagerty and Y. Zheng. Survival model predictive accuracy and ROC curves. *Biometrics*, 61:92-105, 2005.
- R. Henderson, et al. Joint modelling of longitudinal measurements and event time data. *Biostatistics*, 1(4):465-480, 2000.
- E.L. Kaplan and P. Meier. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*, 53:457-481, 1958.
- D.Y. Lin and L.J. Wei. The robust inference for the Cox proportional hazards model. *J Am Stat Assoc*, 84:1074-1078, 1989.
- E. Liu, et al. Characterization of acute liver failure and development of a continuous risk of death staging system in children. *Journal of Hepatology*, 44:134-141, 2006.
- B.R. Lu, et al. Evaluation of the liver injury unit scoring system to predict survival in a multinational study of pediatric acute liver failure. *J Pediatr*, 162:1010-1016, 2013.
- S.V. McDiarmid et al. Liver transplantation for status 1: the consequences of good intentions. *Liver Transpl*, 13:699-707, 2007.
- I.M. Murray-Lyon, et al. Prognostic value of serum alpha-fetoprotein in fulminant hepatic failure including patients treated by charcoal haemoperfusion. *Gut*, 17:576-580, 1976.
- M.A. Nicolaie, et al. Dynamic prediction by landmarking in competing risks. *Statistics in Medicine*, 32(12):2013-2047, 2013.
- J.G. O'Grady, et al. Early indicators of prognosis in fulminant hepatic failure. *Gastroenterology*, 97:439-445, 1989.
- L. Parast, et al. Landmark risk prediction of residual life for breast cancer survival. *Statistics in Medicine*, 32(20):3459-3471, 2013.

- H. Putter. *dynpred: Functions for dynamic prediction in survival analysis*, 2011. URL <http://cran.r-project.org/web/packages/dynpred/dynpred.pdf>. R package version 3.0.3.
- J. Rajanayagam, et al. Artificial neural network is highly predictive of outcome in paediatric acute liver failure. *Pediatr Transplantation*, 17:535-542, 2013.
- J. Rajanayagam, et al. Pediatric acute liver failure: etiology, outcomes, and the role of serial pediatric end-stage liver disease scores. *Pediatr Transplantation*, 17:362-368, 2013.
- M. Sanchez and D.E. D'Agostino. Pediatric end-stage liver disease in acute liver failure to assess poor prognosis. *J Pediatr Gastroenterol Nutr*, 54:193-196, 2012.
- F.V. Schiodt, et al. Predictive value of actin-free Gc-globulin in acute liver failure. *Liver Transpl*, 13:1324-1329, 2007.
- Split Research Group. Studies of pediatric liver transplantation (SPLIT): Year 2000 outcomes. *Transplantation*, 72:463-476, 2001.
- R.H. Squires, et al. Acute liver failure in children: the first 348 patients in the pediatric acute liver failure study group. *J Pediatric*, 148:652-658, 2006.
- R.H. Squires. Acute liver failure in children. *Semin Liver Dis*, 28:153-166, 2008.
- E.W. Steyerberg. *Clinical Prediction Models*. Springer, New York, NY, 2009.
- A.A. Tsiatis and M. Davidian. Joint modeling of longitudinal and time-to-event data: an overview. *Statistica Sinica*, 14:809-834, 2004.
- H.C. van Houwelingen. Dynamic prediction by landmarking in event history analysis. *Scand J Statist*, 34:70-85, 2007.
- H.C. van Houwelingen and H. Putter. Dynamic predicting by landmarking as an alternative for multi-state modeling: an application to acute lymphoid leukemia data. *Lifetime Data Anal*, 14:447-463, 2008.
- H.C. van Houwelingen and H. Putter. *Dynamic Prediction in Clinical Survival Analysis*. CRC Press, Boca Raton, FL, 2012.
- H. Visser, et al. How diagnose rheumatoid arthritis early: a prediction model of persistent (erosive) arthritis. *Arthritis & Rheumatism*, 46(2):357-365, 2002.
- M.S. Wulfsohn and A.A. Tsiatis. A joint model for survival and longitudinal data measured with error. *Biometrics*, 53(1):330-339, 1997.
- Y. Yamagishi, et al. A new prognostic formula for adult acute liver failure using computer tomography-derived hepatic volumetric analysis. *J Gastroenterol*, 44:615-623, 2009.

S.E. Yantorno, et al. MELD is superior to King's College and Clichy's criteria to assess prognosis in fulminant hepatic failure. *Liver Transpl*, 13:822-828, 2007.

M.B. Zaman, et al. MELD score as a prognostic model for listing acute liver failure patients for liver transplantation. *Transplant Proc*, 38:2097-2098, 2006.