

**USE OF A CONDITIONAL QUANTILES METHOD TO PREDICT FUTURE HEALTH
OUTCOMES BASED ON THE TRAJECTORY OF PEDIATRIC END-STAGE LIVER
DISEASE (PELD) SCORES**

by

YuZhou Liu

B.S in Actuarial Mathematics, University of Pittsburgh, 2011

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

University of Pittsburgh

2013

UNIVERSITY OF PITTSBURGH
GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

YuZhou Liu

It was defended on

April 12, 2013

and approved by

Ruosha Li, Ph.D, Assistant Professor, Department of Biostatistics, Graduate School of Public
Health, University of Pittsburgh

Kehui Chen, Ph.D, Assistant Professor, Department of Statistics and Department of
Psychiatry, University of Pittsburgh

Thesis Advisor: Chung-Chou H. Chang, Ph.D, Associate Professor, Departments of
Medicine and Biostatistics, School of Medicine and Graduate School of Public Health,
University of Pittsburgh

Copyright © by YuZhou Liu

2013

USE OF A CONDITIONAL QUANTILES METHOD TO PREDICT FUTURE HEALTH OUTCOMES BASED ON THE TRAJECTORY OF PEDIATRIC END- STAGE LIVER DISEASE (PELD) SCORES

YuZhou Liu, M.S

University of Pittsburgh, 2013

ABSTRACT

Pediatric patients with advanced liver diseases often require liver transplantation. To maximize the efficiency and effectiveness of this procedure in saving lives, the United Network of Organ Sharing (UNOS) uses the results of the pediatric end-stage liver disease (PELD) scoring system to prioritize the list of pediatric patients who are awaiting a transplant. In the analysis reported here, we used data derived from pediatric patients who had a primary diagnosis of biliary atresia, were awaiting a liver transplant, and had PELD scores reported in the Standard Transplant Analysis and Research (STAR) database of UNOS. We used a conditional distribution quantiles method to predict a patient's future distribution of 90-day PELD scores on the basis of his or her PELD scores in the past 30 days. Because this method takes into account both the scores and the rate of change of scores, it is able to demonstrate how patients with the same current scores may have different distributions of scores in the future. To examine the quality of our predictions, we compared our estimated distribution of 90-day PELD scores with the observed distribution of 90-day PELD scores. We used the diagnostic plot to assess the overall goodness of fit of the model.

Public health significance: Transplantation is an effective treatment for patients with end-stage liver diseases. Use of an efficient and effective method to allocate organs among pediatric candidates remains a major challenge. We proposed a method to more accurately estimate the

future health condition using not only the current information but also the path of reaching the current status. If clinicians and policy makers will adopt the method to improve the current organ allocation policy, organs will be used more efficiently and eventually save more lives.

TABLE OF CONTENTS

PREFACE.....	IX
1.0 INTRODUCTION.....	1
2.0 METHODS	3
2.1 DATA	3
2.2 CONDITIONAL DISTRIBUTION QUANTILES ANALYSIS WHEN COVARIATES ARE FUNCTIONS	5
3.0 RESULTS	8
4.0 DISCUSSION	16
BIBLIOGRAPHY	18

LIST OF TABLES

Table 1. Baseline characteristics for pediatric liver transplant candidates with primary diagnosis of biliary atresia in the analytic and in the stabilized samples.....	9
Table 2. Estimated quantiles $Q(\alpha)$ at different levels of quantile α	14

LIST OF FIGURES

Figure 1. Flowchart of data selection.....	7
Figure 2. The distribution of the PELD scores within the first 3 months after listing.....	9
Figure 3. Observed trajectory of PELD scores between 3 months of listing for nine randomly selected patients. Blue cross indicates the observed PELD scores. Red line indicates a smoothed curve for the observed PELD scores.....	10
Figure 4. Mean PELD score function among all patients at a given time value between 0 and 3 months from listing. The x-axis indicates the days from listing and y-axis indicates the average PELD scores.....	11
Figure 5. Predicted distribution function of month 6 PELD scores. Blue smoothed line indicates the prediction function. Red vertical bar indicates the observed PELD score at month 6	12
Figure 6. Diagnostic plot of the fitted model. Y-axis indicated the predicted quantiles $\bar{I}(\alpha)$ and the x-axis indicated the given quantile level α	13
Figure 7. Two patients with the same score 18 (at month 3) but different scores, 29 and 23, (at month 6).....	14
Figure 8. Two patients with the same score 16 (at month 6) but different scores, 17 and 14, (at month 3).....	15

PREFACE

I would like to thank my advisor Professor Chung-Chou H. Chang for introducing me to this fascinating field. I also wish to thank Professor Chen Kehui and Professor Li Ruosha for being on my thesis committee. In the end, I cannot come this far without my parents' and friends' love and support.

1.0 INTRODUCTION

From 1988 to 2009, the number of liver transplants performed annually in the United States increased from 1,713 to 6,320 [1]. The steady growth in the number of adult and pediatric candidates for liver transplantation was accompanied by changes in the manner of allocating organs to these candidates. Initially, allocation was based on the availability of donated livers in the geographic localities of transplant candidates. Subsequently, it was based on the Child-Turcotte-Pugh scores of candidates [2]. In 2002, however, the United Network for Organ Sharing (UNOS) adopted an allocation system for candidates that was based on the severity of liver disease, as reflected in scores from the model for end-stage liver disease (MELD) and the model for pediatric end-stage liver disease (PELD) [3].

For pediatric patients (i.e., patients who are 18 years or younger), the PELD score predicts the probability of death during the 90-day period that precedes liver transplantation. The PELD score is calculated on the basis of the patient's age, serum albumin and bilirubin levels, international normalized ratio for prothrombin time (INR), and presence or absence of growth failure. The mathematical formula is as follows [4]:

$$\begin{aligned} \text{PELD score} = & 0.463 * (<1 \text{ year of age}) - 0.687 * \log_e (\text{albumin g/dL}) + 0.480 * \log_e \\ & (\text{total bilirubin mg/dL}) + 1.857 * \log_e (\text{INR}) + 0.667 * (\text{growth failure}). \end{aligned}$$

The main objective of the PELD system is to prioritize patients on the liver transplant waiting list and thereby maximize the efficiency and effectiveness of using transplantation to

save lives. In general, a higher PELD score indicates a more severe risk of mortality for children with chronic liver diseases and therefore a higher priority in receiving a liver transplant.

Section 2 of our article describes how we used the past trajectory of patients' PELD scores to predict the distribution of their PELD scores during the upcoming 90-day period. Specifically, it describes the dataset of transplant candidates that we used and shows how we applied a conditional quantiles method developed by Chen and Müller (2012) to predict the future health status of these candidates. It also describes how we compared our estimated distribution of 90-day PELD scores with the observed distribution of 90-day PELD scores and how we used the diagnostic plot developed by Chen and Müller (2012) to assess the overall goodness of fit of our model. Section 3 of the article presents our results, and Section 4 discusses the implications of the results.

2.0 METHODS

2.1 DATA

At the time of our study, the Standard Transplant Analysis and Research (STAR) database of UNOS included data for 159,207 transplant candidates in the United States.

Our dataset was limited to candidates who met the following criteria: they had a primary diagnosis of biliary atresia and were seeking a liver transplant; they had no previous organ transplant of any type; they were either 18 years or younger between February 2002 (when the MELD and PELD score systems were implemented) and January 10, 2005, or were 12 years or younger between January 11, 2005, and June 2010; and they had never been scored under the MELD system.

Of the 159,207 transplant candidates, 1,812 pediatric met these criteria. Together the 1,812 candidates had 11,635 PELD measurements within the first 3 months of being wait-listed for a liver transplant.

Of the 1,812 candidates, we further excluded 1,245 patients who had been removed from the transplant waiting list within 6 months for any reason, including loss to follow-up, death, or other change in health condition that made them ineligible for a transplant. We also excluded 174 patients who had no outcome data (i.e., no PELD measurement between 3 and 6 months). The final analytic sample included 393 pediatric liver transplant candidates with a primary diagnosis

of biliary atresia and with a total of 1,342 PELD measurements within the first 3 months of listing.

For the stability of the prediction, we created a stabilized sample by excluding patients who had fewer than 4 PELD measurements within the first 3-month interval. Thus, the final sample size included 147 patients with total 852 PELD measurements within the first 3 months of listing (Figure 1).

We used PELD scores as a proxy for the patient's health status. The analysis was designed to predict the future health status of a pediatric liver transplant candidate on the basis of his or her 3-month trajectory of earlier PELD scores. The primary outcome of interest was the PELD score 90 days after the patient's inclusion on the transplant waiting list.

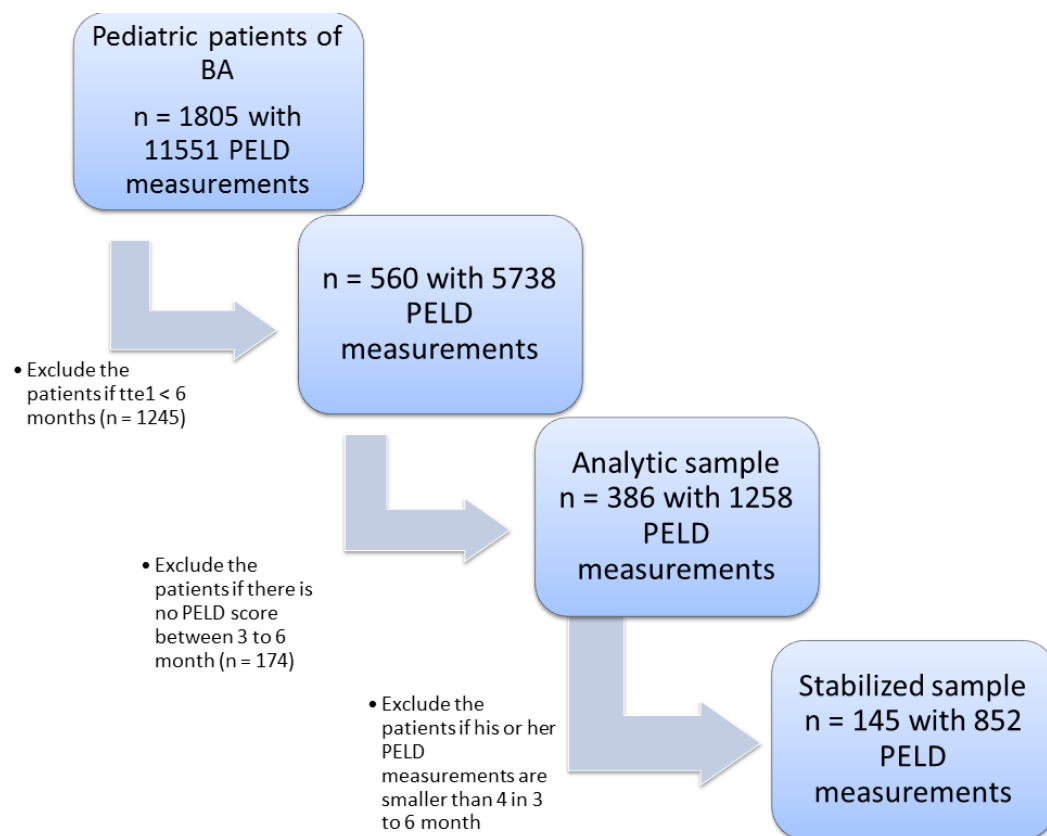


Figure 1. Flowchart of data selection

2.2 CONDITIONAL DISTRIBUTION QUANTILES ANALYSIS WHEN COVARIATES ARE FUNTIONS

The conditional quantiles method has been widely applied to analyze data with functional covariates [7]. Chen and Müller (2012) applied it to predict an individual's adult height based on his or her growth data during childhood or adolescence. For our analysis, we constructed conditional quantiles by taking the entire PELD score history of each patient in the dataset and using the scores as functional covariates (latent covariates). Then we modeled the distribution of the outcome (future health status) based on the functional covariates. The software we used included a package of Chen's code written in MALTAB.

In the model we developed, we define the random process $X(t)$ and the random variable Y as the trajectory of PELD scores in the history and the health outcome at a future time point, respectively. In this case, $t \in [0, T]$ is the time span of the history. The conditional distribution of Y given $\{X(t), t \in [0, T]\}$ is $F(y|X) = P(Y \leq y|X) = E\{I(Y \leq y)|X\}$, where I is an indicator function[5].

In the model developed by Chen and Müller (2012), the relationship between Y and $X(t)$ under the functional generalized linear regression framework takes the following form:

$$F(y|X) = E\{I(Y \leq y)|X\} = g^{-1}\{\beta_0(y) + \int X^c(t)\beta(y, t)dt\}, \quad (1)$$

where g is a known monotonic link function with inverse g^{-1} and where $X^c(t)$ centers the random process $X(t)$ to its mean function $\mu(t)$. Thus, $X^c(t) = X(t) - \mu(t)$; and $\beta(y, t)$ and $\beta_0(y)$ are coefficient functions for a fixed value y . The term $g^{-1}\{\beta_0(y)\}$ can be viewed as a

baseline conditional distribution of Y , which $\beta_0(y)$ is the conditional distribution of Y when the trajectory $X(t)$ equals to its mean trajectory $\mu(t)$ for all $t \in [0, T]$. Coefficient $\beta(y, t)$ is a function of both y and t . When y is fixed, it is a function of t for fixed y . Therefore, $\beta(y, t)$ represents the effect of the random process $X^c(t)$ on the distribution of the outcome Y at time t , and $\int X^c(t)\beta(y, t)dt$ corresponds to the cumulative effect of $X^c(t)$ across all time $t \in [0, T]$.

For a fixed value of y , the indicator $I(Y \leq y)$ has a binomial distribution, with a mean of $F(y|X)$. Therefore, the natural choice of the link function g is the logit function whose inverse, $g^{-1}(z) = \frac{\exp(z)}{1+\exp(z)}$, is the expit function. Under model (1) with the logit link, the conditional quantiles $Q(\alpha|X)$ of Y , can be acquired by inverting the conditional distribution function $F(y|X)$. Note that conditional quantiles is defined as

$$Q(\alpha|X) = \inf\{y: F(y|X) \geq \alpha\}, 0 < \alpha < 1. \quad (2)$$

This means that the conditional quantile $Q(\alpha|X)$ is the minimum value of y in the domain of Y if CDF $F(y|X)$ is larger than or equal to the given quantile value α .

To assess the goodness of fit of our estimated quantiles, we applied the diagnostic method developed by Chen and Müller (2012). For individual i , we let Y_i be the random variable of the month-6 PELD score, we let $\widehat{Q}_i(\alpha)$ be the predicted $100 \times \alpha\%$ quantiles of PELD scores at month 6; we let α be a quantile value, and we define the indicator function $I_i(\alpha)$ as $I_i(\alpha) = I\{Y_i \leq Q_i(\alpha)\}$. A model does not show a lack of fit if the expected value of the indicator function $I_i(\alpha) = I\{Y_i \leq Q_i(\alpha)\}$ is close to α . This means that indicator estimator of probability of Y_i

smaller than its α quantile value is close to α . Note that the mean of indicator function, $E\{I_i(\alpha) | X_i\}$ can be calculated as

$$\bar{I}(\alpha) = \frac{1}{n} \sum_{i=1}^n I\{Y_i \leq Q_i(\alpha)\}.$$

If this value is close to the true quantile value α , the model does not show lack of fit.

3.0 RESULTS

Table 1 shows the baseline characteristics of patients of our analytic sample ($n = 386$) and stabilized sample ($n = 145$) of pediatric liver transplant candidates. Although the results for the two samples were similar, here we describe only the results for the stabilized sample.

The stabilized sample of 145 patients had a total of 852 PELD scores. These scores ranged from -10 to 90 , and the mean score was 14 . Figure 2 shows the distribution of scores during the first 3 months after listing. The majority of scores were below 45 .

Figure 3 shows the trajectories of PELD scores from 9 randomly selected patients, along with a smooth curve (using the functional principal component method) that we fitted for each set of scores. For most of these patients, the PELD scores increased linearly during the first 3 months. It was possible to use a quadratic or higher-order smoothing function because of the shorter time intervals or fewer PELD score measurements within the interval. Figure 4 shows the plot of the average observed PELD scores and demonstrates that the PELD scores increased linearly from 14 to 20 during the 3-month interval.

Table 1. Baseline characteristics for pediatric liver transplant candidates with primary diagnosis of biliary atresia in the analytic and in the stabilized samples

	Analytic Sample (n = 386)		Stabilized Sample (n = 145)	
	Mean (SD) or n (%)	Range (min, max)	Mean (SD) or n (%)	Range (min, max)
Age at the time of listing (in year)	0.99 (2.20)	0, 11	0.74 (1.94)	0, 11
Female gender	188 (48.7%)		75 (51.7%)	
White	189 (49.0%)		73 (50.3%)	
Total bilirubin (mg/dL)	9.21 (6.09)	0.2, 32	11.25 (6.43)	0.2, 31.5
Serum albumin (g/dL)	3.22 (0.65)	1.1, 5.4	3.08 (0.71)	1.1, 5.4
INR	1.28 (0.49)	0.70, 6.83	1.36 (0.61)	0.84, 6.83
Presence of growth failure	159 (41.2%)		72 (49.7%)	
Listing PELD score	11.0 (9.5)	-10, 90	14.2 (8.4)	-6, 46
6-month PELD score	15.55(10.17)	-10, 42	18.08 (10.27)	-7, 42
Number of PELD measurements within 3- month of listing	14.11 (9.44)	-10, 90	16.53 (9.02)	-6, 64

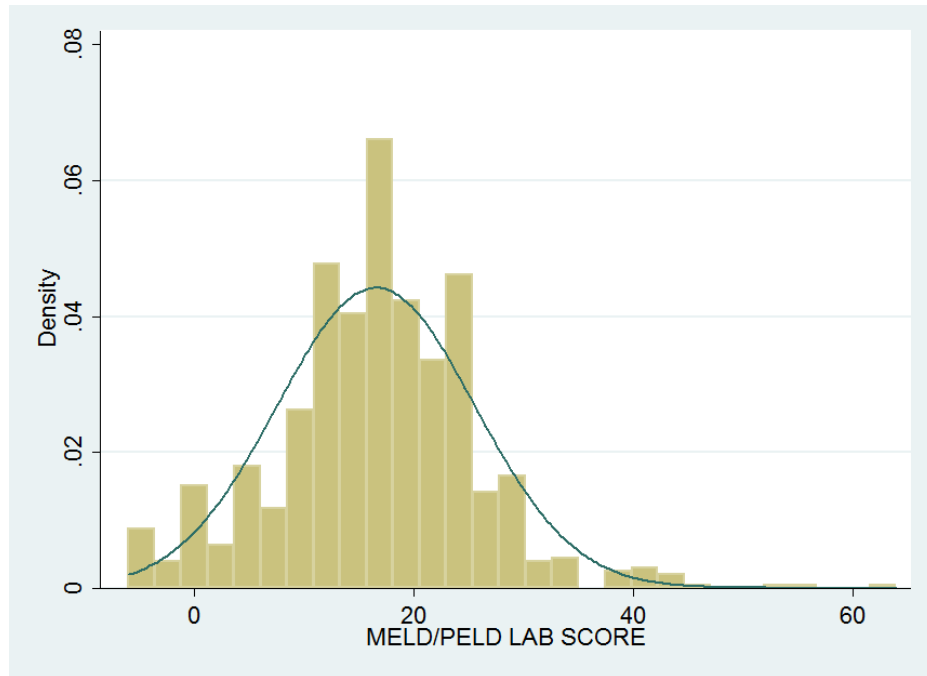


Figure 2. The distribution of the PELD scores within the first 3 months after listing

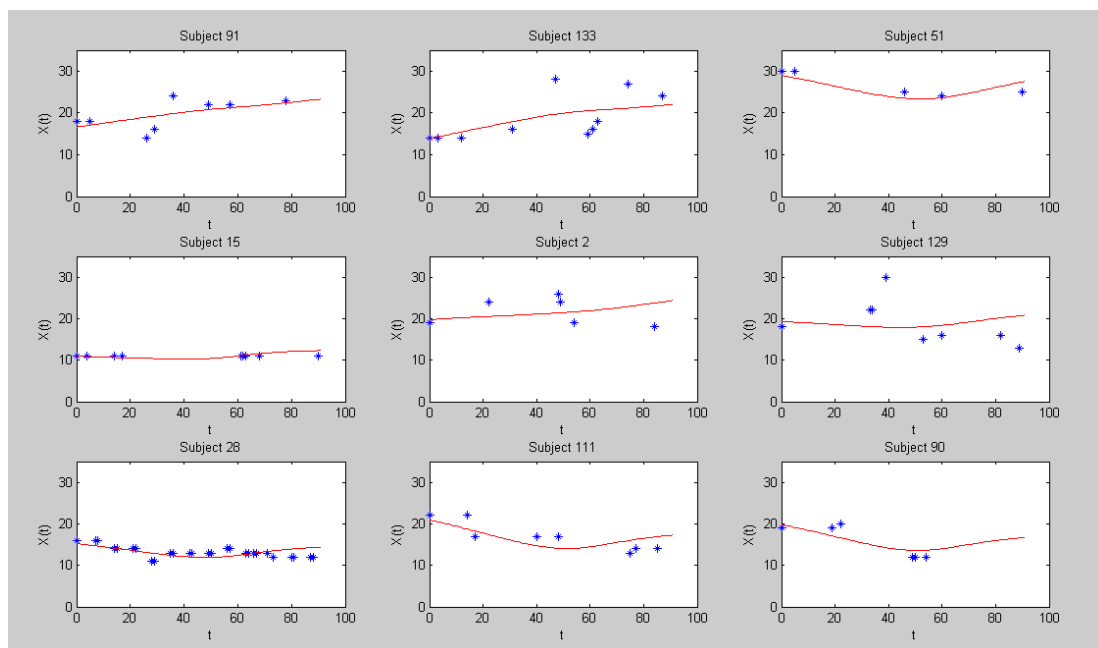


Figure 3. Observed trajectory of PELD scores between 3 months of listing for nine randomly selected patients. Blue cross indicates the observed PELD scores. Red line indicates a smoothed curve for the observed PELD scores

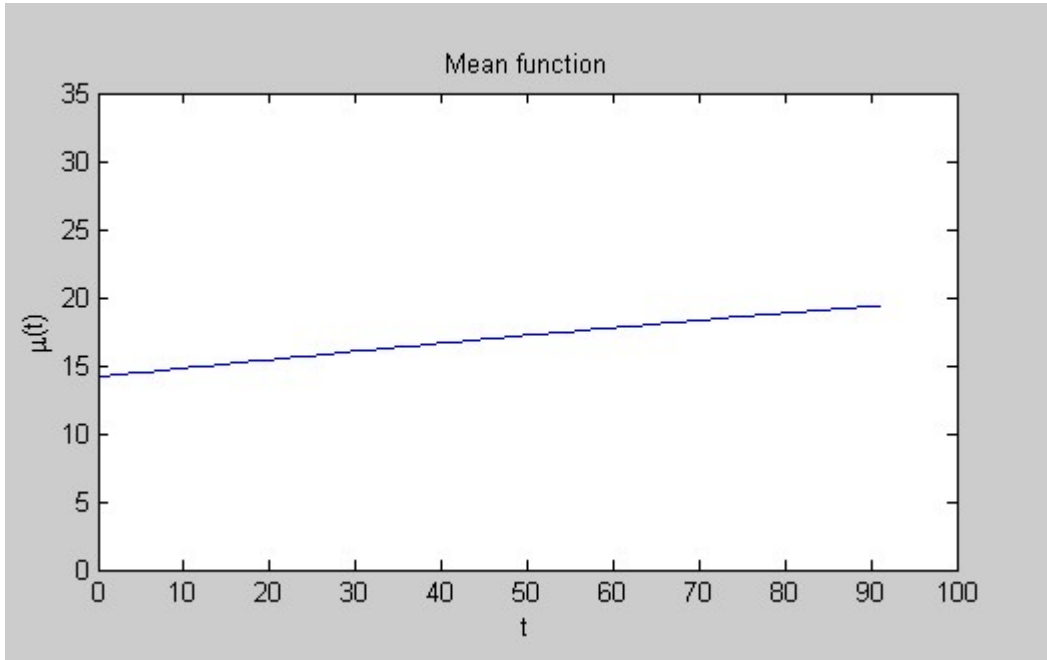


Figure 4. Mean PELD score function among all patients at a given time value between 0 and 3 months from listing.

The x-axis indicates the days from listing and y-axis indicates the average PELD scores

Figure 5 depicts the predicted 6-month quantile PELD scores for 4 patients randomly selected from the group of 9. The prediction for each of the 4 patients is based on his or her trajectory of PELD scores during the first 3 months of listing.

The observed 6-month PELD score for patient #15 and for patient #93 was 11 and 14, respectively. Using the conditional quantile method, patient #15 will have 95% predicted probability of having score 25 or less and patient #93 will have 85% predicted probability of having score 25 or less. This indicates a good prediction for both patients. The observed 6-month PELD scores for patients #2 was 29. His or her predicted probability of having a score no more than 25 is just 65%, which shows a good fit as well. However, for patient #12, the observed 6-

month PELD score was 27 but his or her predicted probability of having a score no more than 25 is about 90%, which does not indicate a good prediction.

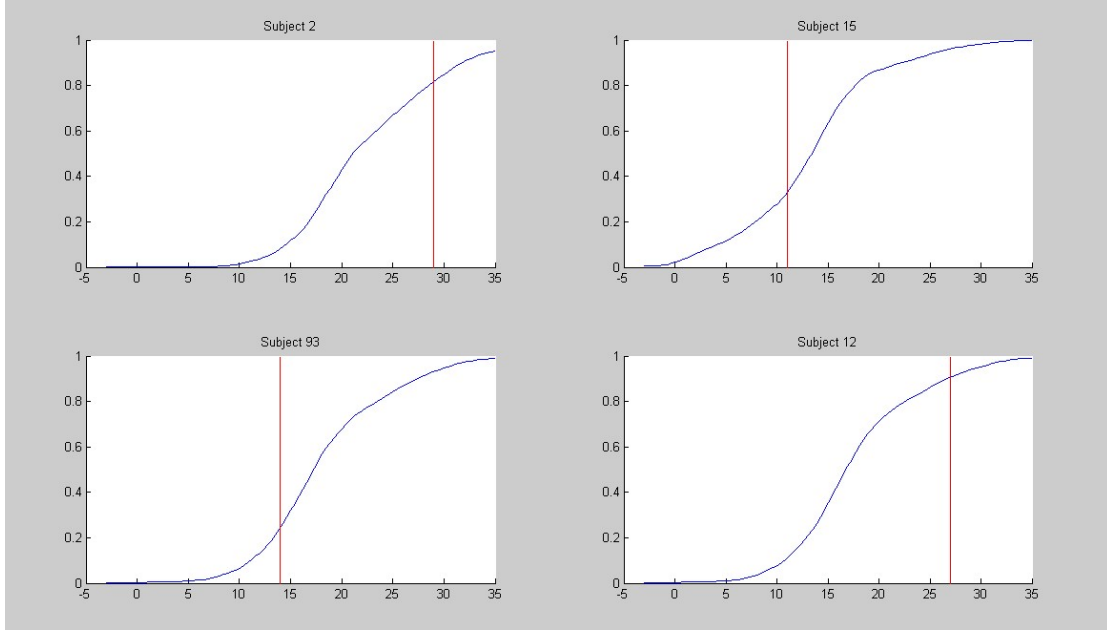


Figure 5. Predicted distribution function of month 6 PELD scores. Blue smoothed line indicates the prediction function. Red vertical bar indicates the observed PELD score at month 6

Figure 6 depicts the diagnostic plot of $\bar{I}(\alpha)$ against α for our fitted model in predicting the distribution quantiles of the 6-month PELD scores $\hat{Q}_t(\alpha)\bar{I}(\alpha)$. The plot is close to the 45-degree identity line, so we have no strong evidence of model lack of fit based on this figure.

We selected patients #2 and #4 to demonstrate that two patients may have similar current PELD score but with different growth paths to this value, their predicted future distributions of PELD scores could be quite different. From Figure 7, both patients #2 and #4 had the same PELD score around month 3 but had different observed PELD scores at month 6: 29 and 18, respectively. Since it is common for clinical exercises to consider a PELD score greater than 25 as a signal of critical condition for a patient, we would like to demonstrate that two patients who have the same PELD scores currently may have different clinical conditions (critical ill vs. not) in month 6.

According to Table 2, the predicted quantiles of the 6-month PELD scores for these two patients score indicate that the probability of having PELD score less than 25 for patient #2 is much smaller than that for patient #4. In other words, we expect that the PELD score, and thus the risk, for patient #2 in month 6 should be higher than the risk of patient #4 based on his or her history of PELD score path even if their current scores are the same. Therefore, we see that the conditional quantile method successfully differentiates the risk levels of the two patients with similar current PELD scores, by taking into account the evolving trends in their historical PELD scores.

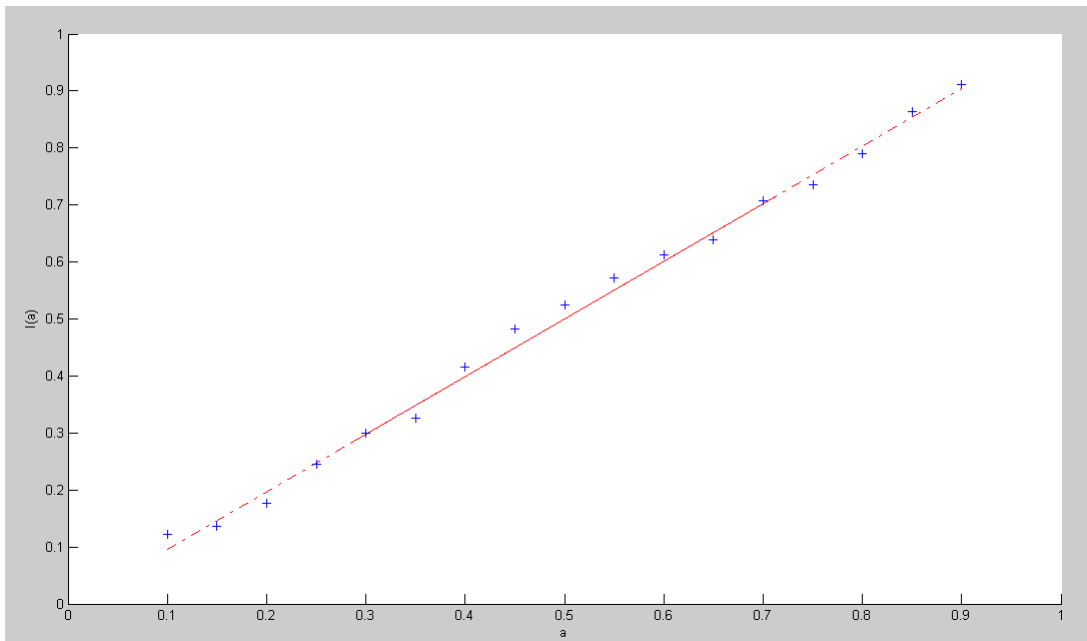


Figure 6. Diagnostic plot of the fitted model. The y-axis indicated the predicted quantiles $\bar{I}(\alpha)$ and x-axis indicated the given quantile level α

Table 2. Estimated quantiles $\hat{Q}(\alpha)$ at different levels of quantile α

	Estimated Quantiles $\hat{Q}(\alpha)$			
Quantile level α	Patient #2	Patient #4	Patient #95	Patient #96
0.1	14.66	12.35	10.82	10.43
0.25	17.73	15.04	13.89	13.51
0.5	21.18	18.50	16.96	16.58
0.75	27.32	23.87	24.64	26.77
0.9	31.93	28.86	26.94	27.71

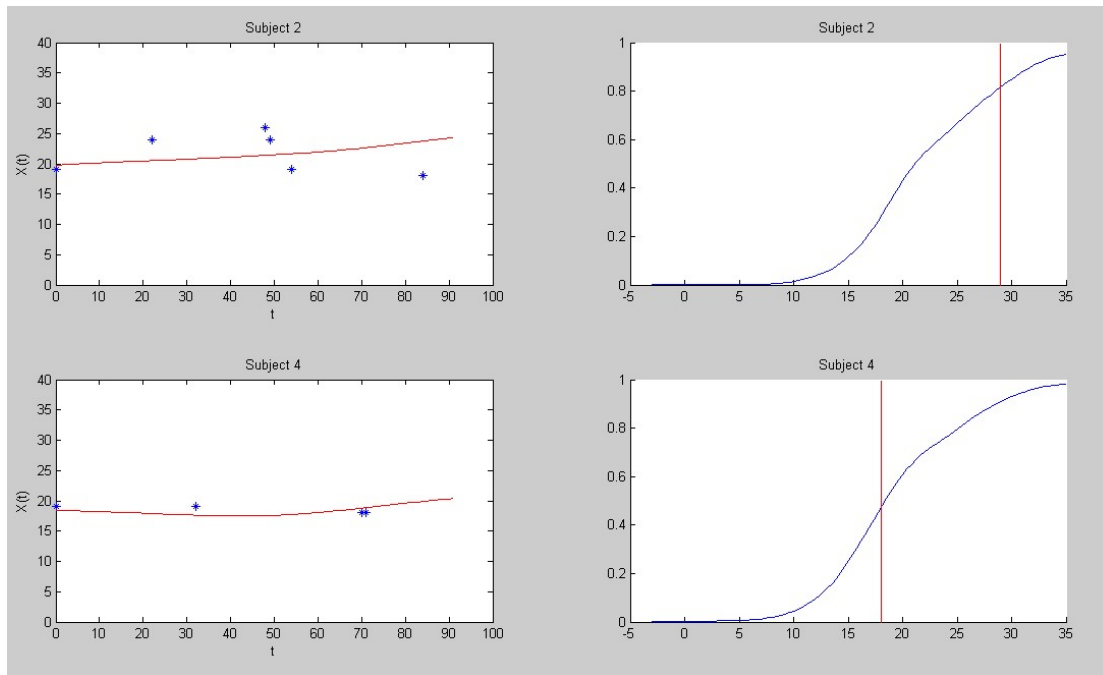


Figure 7. Two patients with the same score 18 (at month 3) but different scores, 29 and 18, (at month 6)

We also give an example to show that two patients with exactly the same expected quantiles of PELD scores in the future may have different scores currently. In Figure 8, both

patients #95 and #96 had the same observed PELD score (=16) at month 6 but had different PELD scores around month 3: 17 and 14, respectively. We can find that even though the current status of these two patients are different, based on their history of PELD scores paths, their predicted distributions of the PELD scores at month 6 are almost identical. From Table 2, we can find that the predicted quantiles are basically the same for these two patients.

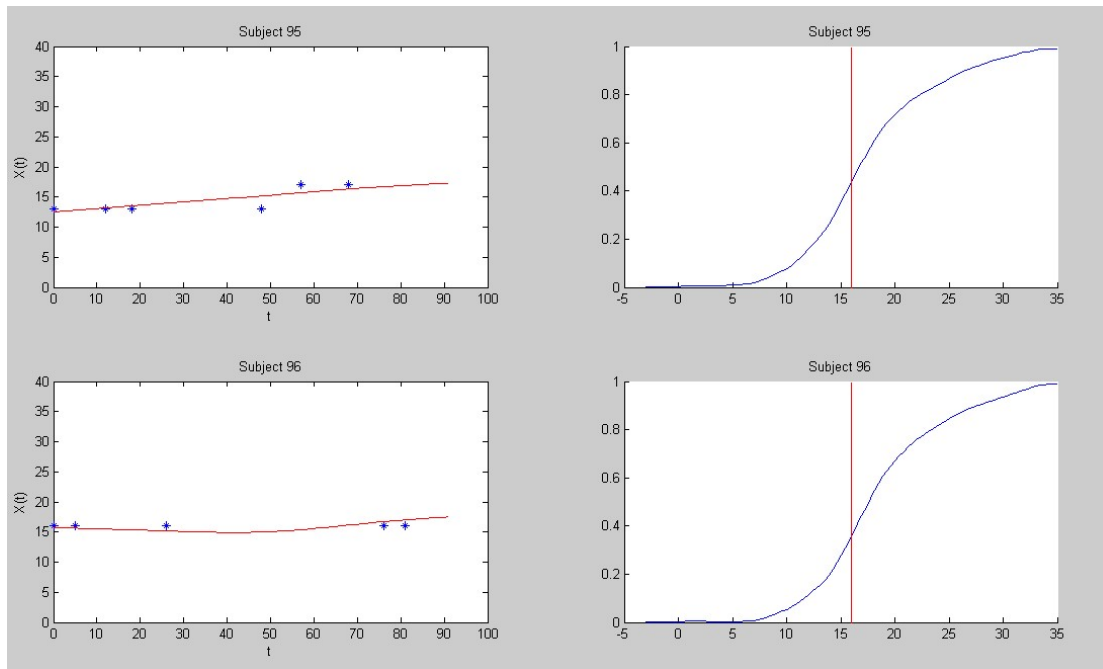


Figure 8. Two patients with the same score 16 (at month 6) but different scores, 17 and 14, (at month 3)

4.0 DISCUSSION

The PELD score is used to predict a patient's mortality risk in 3 months for children are waiting liver transplantation. In the current PELD system, we used cross-sectional health condition to predict the future health status. In order to make more accurate prediction, we propose to use the history of PELD score path instead of the PELD score at one-time point in predicting future health status. To apply the conditional quantile method developed by Chen and Müller (2012), we are able to estimate the distribution of future PELD scores given the history of PELD values.

As what we have seen in the results section, two patients may have similar health status currently, but they may still reach different degrees of disease progression in the future. This is because their trajectories of the PELD history measurements could be distinct. On the other hand, two patients may have similar predicted distribution of the future PELD scores; their trajectories of the PELD history measurements could be very different. Our analysis results and the diagnostic plots show that prediction of future health status based on the change history of the PELD scores is necessary and reasonable.

There are several limitations of this study. First, we excluded patients who had history of PELD scores less than 4 measurements so the estimated results had less bias. In the original dataset, 8% (145 out of 1805) of the patients had less than 4 PELD measurements. Second, the range of PELD scores was wide. The estimation could be unstable if a patient's PELD scores constantly stay too high or too low. To overcome this issue, we may stratify the sample into

different ranges of PELD scores. Third, the method is sensitive to outliers. One abnormal PELD score may influence the estimated distribution significantly and decrease the accuracy of prediction. Therefore, model development by removing outliers is necessary. Finally, our study only explored the condition of one covariate. In the future, more covariates could be included in predicting future health status.

Despite the limitations mentioned above, estimation of future health status using all information in the history provides more accurate prediction, which is important for liver transplant clinicians and policy makers to improve allocation rules. In the future, we will apply this method to study all repeated measured lab data (e.g., total bilirubin level and serum albumin level) and check the impact of the patterns of these data on the changes in health conditions. We will also extend our method to predict the 90-day waitlist mortality, which was the PELD score originally designed for. In summary, PELD scores serve an important role in pediatric liver transplant allocation system, which they can be improved by using a method that can more accurately predict the future health condition.

BIBLIOGRAPHY

1. Wertheim JA, Petrowsky H, Saab S, Kupiec-Weglinski JW, Busuttil RW. Major challenges limiting liver transplantation in the United States. *American Journal of Transplantation*, 2011. **11**(9): 1773-1784.
2. Freeman RB, Jr., Wiesner RH, Harper A, McDiarmid SV, Lake J, Edwards E, Merion R, Wolfe R, Turcotte J, Teperman L. UNOS/OPTN Liver Disease Severity Score, UNOS/OPTN Liver and Intestine, and UNOS/OPTN Pediatric Transplantation Committees. The new liver allocation system: moving toward evidence-based transplantation policy. *Liver transplantation*, 2002. **8**(9): 851-858.
3. McDiarmid SV, Merion RM, Dykstra DM, Harper AM. Selection of pediatric candidates under the PELD system. *Liver transplantation*, 2004. **10**(10 Suppl 2): S23-S30.
4. Kato K. Estimation in functional linear quantile regression. *The Annals of Statistics*, 2012. **40**(6): 3108-3136.
5. Chen K, Müller H-G. Conditional quantile analysis when covariates are functions, with application to growth data. *Journal of the Royal Statistical Society: Series B*, 2012. **74**(1): 67-89.