Correlation of Health-Related Quality of Life After Liver Transplant With the Model for End-Stage Liver Disease Score

Eric T. Castaldo, MD, MPH; Irene D. Feurer, PhD; Robert T. Russell, MD, MPH; C. Wright Pinson, MD, MBA

Objective: To determine whether a correlation exists between the Model for End-Stage Liver Disease (MELD) score and health-related quality of life (HRQOL) after liver transplant (LT).

Design: Prospective cohort.

Setting: University hospital.

Patients: Adult LT recipients (N=209).

Main Outcome Measures: Postoperative HRQOL over a 1-year period after LT as measured via multiple regression–based path analysis testing the effects of the MELD score, preoperative variables, and postoperative variables on scores on the physical component summary and mental component summary scales of the 36-Item Short Form Health Survey and on composite physical and mental HRQOL scores derived from multiple scales.

Results: The MELD score (β=.16), cholestatic cirrhosis (β=.12), autoimmune/metabolic disease (β=.18), neoplasm (β=.23), time after LT (β=.16), and the Karnofsky score (β=.49) had significant effects on the physical component summary scale score. Autoimmune/metabolic disease (β=.16) and the Karnofsky score (β=.25) had significant effects on the mental component summary scale score. The MELD score (β=.15), high school education (β=.15), college education (β=.17), autoimmune/metabolic disease (β=.15), neoplasm (β=.23), time after LT (β=.11), and the Karnofsky score (β=.51) had significant effects on the composite physical HRQOL score. Autoimmune/metabolic disease (β=.23), neoplasm (β=.15), and the Karnofsky score (β=.42) had significant effects on the composite mental HRQOL score.

Conclusions: An increasing MELD score, when computed without any diagnosis-based exception points, was associated with improved physical HRQOL in the first year after LT. The MELD score did not affect mental HRQOL.

Arch Surg. 2009;144(2):167-172

The definitive therapy for end-stage liver disease is hepatic transplant. In 2006, 6650 hepatic transplants were performed in the United States, which is an almost 4-fold increase since 1988. Improvements in surgical technique, preoperative and postoperative care, immunosuppression regimens, antibiotics, organ preservation and allocation, and increasing demand have contributed to its rise. Improved survival has led to increased importance of incorporating health-related quality of life (HRQOL) measures into a comprehensive outcomes assessment strategy as it may influence decisions involving organ transplant.

In 1999, the Institute of Medicine suggested that a continuous disease severity score for patients with chronic liver disease de-emphasizing waiting time could improve the allocation of cadaveric livers. The Model for End-Stage Liver Disease (MELD) was originally designed to predict short-term mortality for patients undergoing a transjugular intrahepatic portosystemic shunt procedure. The model was later validated as a way to predict mortality rates in patients with varying severity of hepatic disease. It comprised 3 factors (serum creatinine level, serum bilirubin level, and the international normalized ratio of the prothrombin time) that were significantly associated with mortality and were modeled as a continuous score reflecting disease severity. The MELD system was implemented in 2002 for purposes of organ allocation, the scores range from 6 (least ill) to 40 (gravely ill).

There are exceptions to the computed MELD score. Patients with hepatocellular carcinoma who meet the Milan crite-
ria are allocated a score of 22 to make them more competitive for a hepatic graft prior to the development of incurable disease. There are also other clinical situations in which exceptions can be made on a case-by-case basis.5

There are limited data describing the relationship between the MELD score and HRQOL. Kanwal et al6 identified a negative correlation between an increasing MELD score and HRQOL in patients with end-stage liver disease. While a stronger association was seen with the Child-Turcotte-Pugh score, the mean MELD score in this population was low at 12. Saab et al7 reported that ascites or encephalopathy significantly influenced HRQOL but found no correlation between the MELD score and HRQOL in patients awaiting liver transplant (LT). However, the mean MELD score was 13.6, which was limited to patients waiting for a transplant. In 2006, Rodriguez et al8 reported that an increasing MELD score was negatively associated with HRQOL after LT, especially with physical functioning.

These data led us to our primary aim, to determine whether the pretransplant MELD score correlates with posttransplant HRQOL. Our hypothesis was that an increasing MELD score would have an inverse relationship with physical and mental HRQOL.

METHODS

PATIENTS AND DATA ACQUISITION

A single-institution prospective cohort study was performed at Vanderbilt University Medical Center after approval by the institutional review board.

To be eligible, patients underwent an initial hepatic transplant between January 1, 2002, and September 30, 2006. Patients who underwent a second hepatic transplant within 30 days of the initial transplant were included with the date of the first transplant as the date of entry into the study. Any patient requiring a second hepatic transplant more than 30 days after the initial transplant was included if there were posttransplant HRQOL data available prior to their second transplant (the reference transplant date was the date of the first LT). Patients were excluded if they did not complete the Medical Outcomes Study 36-Item Short Form Health Survey (SF-36) at least once during the first posttransplant year. Patients were followed up to 1 year from transplant and/or until the close of the study on October 31, 2006.

Clinical data were collected at the time of transplant and/or during the subsequent hospitalization. The physiological MELD score was calculated just prior to transplant without additional exception points, and this MELD score was used for all of the analyses. Appropriate additional clinical data and the posttransplant course were collected in subsequent clinic visits during follow-up.

HRQOL INSTRUMENTS

AND FUNCTIONAL PERFORMANCE

Posttransplant HRQOL was assessed with the SF-36, the Beck Anxiety Inventory (BAI), the Center for Epidemiologic Studies Depression Scale (CES-D), and the visual analog scale (VAS). These instruments were selected on the basis of their applicability for self-report, published reliability and validity in clinical and nonclinical samples, breadth of application in the literature, and limited response burden to represent a range of HRQOL constructs that are germane to patients undergoing transplantation.

DATA COMPLETION

BY REGRESSION IMPUTATION

Regression imputation, a recognized method for handling missing data,9,10 was performed for cases having complete SF-36 data but incomplete BAI (n=19), CES-D (n=20), and/or VAS (n=27) data. A separate HRQOL data set (n>1500) from our institution comprising solid organ transplant candidates and recipients was used to generate predictive equations for computing BAI, CES-D, and VAS scores on the basis of the 8 SF-36 scales. These equations were used to estimate scores for the BAI, CES-D, and/or VAS for the 52 respondents who failed to complete 1 or more of the surveys in order to permit secondary analyses for models 3 and 4 (described later) having maximum sample sizes.

ENDPOINT VARIABLES FOR HRQOL OUTCOMES

Patient-reported physical and mental HRQOL was treated as 4 separate endpoint measures. The PCS (model 1) and MCS (model 2) scales of the SF-36 were used as the primary HRQOL outcomes, respectively. Then, principal component factor analysis was used to derive a composite physical HRQOL score representing the SF-36 PCS scale and VAS (model 3) and a composite mental HRQOL score representing the SF-36 MCS scale, BAI, CES-D, and VAS (model 4).

STATISTICAL ANALYSIS

Overview of Path Analysis

Separate multivariate models of the 4 HRQOL outcome measures were tested using multiple regression-based, recursive path analysis. Path analysis represents hypothesized systems in which variables relate in causal models. These models reflect combinations of direct and indirect effects, with indirect effects being instances where a variable is hypothesized to influence another through a mediating variable. Path diagrams necessarily represent 2 or more inherent regression equations. Variables on the far left of the diagram are termed exogenous and are recognized in the model to have some degree of intercorrelation. Subsequent variables are termed endogenous, and each serves as a criterion for an inherent regression equation. Saturated models depict all of the direct and indirect effects. Each direct effect is tested for preliminary statistical significance (eg, p < .05) and magnitude (expressed as unstandardized and standardized regression coefficients). Models may be simplified by eliminating nonsignificant effects, and the coefficients are recomputed. The effect of model simplification can be tested via the χ² statistic and goodness-of-fit indices. Then, the total influence of one variable on another, which is the sum of its direct and any indirect effects, may be computed.

Regression imputation, a recognized method for handling missing data,9,10 was performed for cases having complete SF-36 data but incomplete BAI (n=19), CES-D (n=20), and/or VAS (n=27) data. A separate HRQOL data set (n>1500) from our institution comprising solid organ transplant candidates and recipients was used to generate predictive equations for computing BAI, CES-D, and VAS scores on the basis of the 8 SF-36 scales. These equations were used to estimate scores for the BAI, CES-D, and/or VAS for the 52 respondents who failed to complete 1 or more of the surveys in order to permit secondary analyses for models 3 and 4 (described later) having maximum sample sizes.

ENDPOINT VARIABLES FOR HRQOL OUTCOMES

Patient-reported physical and mental HRQOL was treated as 4 separate endpoint measures. The PCS (model 1) and MCS (model 2) scales of the SF-36 were used as the primary HRQOL outcomes, respectively. Then, principal component factor analysis was used to derive a composite physical HRQOL score representing the SF-36 PCS scale and VAS (model 3) and a composite mental HRQOL score representing the SF-36 MCS scale, BAI, CES-D, and VAS (model 4).

STATISTICAL ANALYSIS

Overview of Path Analysis

Separate multivariate models of the 4 HRQOL outcome measures were tested using multiple regression–based, recursive path analysis. Path analysis represents hypothesized systems in which variables relate in causal models. These models reflect combinations of direct and indirect effects, with indirect effects being instances where a variable is hypothesized to influence another through a mediating variable. Path diagrams necessarily represent 2 or more inherent regression equations. Variables on the far left of the diagram are termed exogenous and are recognized in the model to have some degree of intercorrelation. Subsequent variables are termed endogenous, and each serves as a criterion for an inherent regression equation. Saturated models depict all of the direct and indirect effects. Each direct effect is tested for preliminary statistical significance (eg, p < .05) and magnitude (expressed as unstandardized and standardized regression coefficients). Models may be simplified by eliminating nonsignificant effects, and the coefficients are recomputed. The effect of model simplification can be tested via the χ² statistic and goodness-of-fit indices. Then, the total influence of one variable on another, which is the sum of its direct and any indirect effects, may be computed.

Regression imputation, a recognized method for handling missing data,9,10 was performed for cases having complete SF-36 data but incomplete BAI (n=19), CES-D (n=20), and/or VAS (n=27) data. A separate HRQOL data set (n>1500) from our institution comprising solid organ transplant candidates and recipients was used to generate predictive equations for computing BAI, CES-D, and VAS scores on the basis of the 8 SF-36 scales. These equations were used to estimate scores for the BAI, CES-D, and/or VAS for the 52 respondents who failed to complete 1 or more of the surveys in order to permit secondary analyses for models 3 and 4 (described later) having maximum sample sizes.

ENDPOINT VARIABLES FOR HRQOL OUTCOMES

Patient-reported physical and mental HRQOL was treated as 4 separate endpoint measures. The PCS (model 1) and MCS (model 2) scales of the SF-36 were used as the primary HRQOL outcomes, respectively. Then, principal component factor analysis was used to derive a composite physical HRQOL score representing the SF-36 PCS scale and VAS (model 3) and a composite mental HRQOL score representing the SF-36 MCS scale, BAI, CES-D, and VAS (model 4).

STATISTICAL ANALYSIS

Overview of Path Analysis

Separate multivariate models of the 4 HRQOL outcome measures were tested using multiple regression–based, recursive path analysis. Path analysis represents hypothesized systems in which variables relate in causal models. These models reflect combinations of direct and indirect effects, with indirect effects being instances where a variable is hypothesized to influence another through a mediating variable. Path diagrams necessarily represent 2 or more inherent regression equations. Variables on the far left of the diagram are termed exogenous and are recognized in the model to have some degree of intercorrelation. Subsequent variables are termed endogenous, and each serves as a criterion for an inherent regression equation. Saturated models depict all of the direct and indirect effects. Each direct effect is tested for preliminary statistical significance (eg, p < .05) and magnitude (expressed as unstandardized and standardized regression coefficients). Models may be simplified by eliminating nonsignificant effects, and the coefficients are recomputed. The effect of model simplification can be tested via the χ² statistic and goodness-of-fit indices. Then, the total influence of one variable on another, which is the sum of its direct and any indirect effects, may be computed.
Model Development and Variable Encoding

Our analyses modeled the effects of the MELD score (without diagnosis-related exception points), diabetes (yes or no), education attainment (3 categories), diagnosis (5 categories), length of hospitalization, time after LT, previous graft rejection (yes or no), and the contemporaneous Karnofsky score on the 4 HRQOL outcomes. In those cases having multiple post-transplant data points, the last observation within the first year after transplant was used.

The primary variable of interest in all of the models was the MELD score as a single composite measure of disease severity. The tested influence of the MELD score on the HRQOL measures in the context of other variables that have been previously shown to influence HRQOL. The exogenous variables were the MELD score, diabetes, cause of liver disease, length of stay, and time after transplant. Endogenous variables were rejection episodes, Karnofsky functional performance, and the 4 post-transplant HRQOL outcome measures.

Each analysis began with a fully saturated model where all of the direct and indirect effects were represented. Simplified models representing those effects having $\alpha < .10$ were tested, and corresponding path coefficients and $P$ values were recomputed. Goodness of fit was tested for each simplified model using the $\chi^2$ test of model fit and the adjusted goodness-of-fit index (AGFI). The $\chi^2$ tests having type I error rates greater than .05 and AGFI values greater than 0.90 were accepted as evidence of a model’s adequate fit to the underlying data.

Statistical analyses were performed using SPSS statistical software version 15.0 and Amos statistical software version 7.0 (SPSS Inc, Chicago, Illinois).

RESULTS

PATIENT POPULATION, DATA COMPLETION, AND STATISTICAL POWER

During the study period, 327 patients underwent LT, 307 of whom survived the perioperative period and returned to the clinic for follow-up. Of these patients, 90 did not complete the surveys within the first post-transplant year and 8 were ineligible owing to incomplete SF-36 data. The final cohort consisted of 209 patients, 157 of whom completed all of the surveys. There were 19 patients with incomplete BAI scores, 20 with incomplete CES-D scores, and 27 with incomplete VAS scores. The full set of 10 covariates, which comprises 3 continuous and 7 indicator variables, resulted in models having adequate sample sizes (>10 cases per covariate) and statistical power (>80%) to detect a moderate (cumulative $R^2 \approx 0.10$) overall effect at $\alpha = .05$.

Demographic data are summarized in Table 1. The mean (SD) MELD score at the time of transplant was 21.4 (7.9), which was lower than the mean (SD) United Network for Organ Sharing listing score of 23.3 (6.4). A MELD exemption was present in 43 patients.

The mean (SD) MELD scores were similar for patients with noncholestatic cirrhosis (22.5 [7.1]), cholestatic cirrhosis (22.5 [7.4]), and autoimmune/metabolic disease (21.3 [7.1]). Patients with acute hepatic failure had a higher mean (SD) MELD score (35.8 [3.3]) and patients with neoplasm had a lower mean (SD) MELD score (13.1 [3.7]).

The HRQOL data were distributed over the first post-transplant year as follows. The 1-month surveys were used in 42 patients, the 3-month surveys were used in 39 patients, the 6-month surveys were used in 40 patients, and the 1-year surveys were used in 88 patients. The mean (SD) time after transplant was 7.7 (5.3) months.

HEALTH-RELATED QUALITY OF LIFE

The first saturated model (Figure, A) tested all of the direct and indirect effects on the SF-36 PCS scale score. Effects with $P > .10$ were subsequently deleted to produce the simplified model (Figure, B). The simplified model demonstrates that cholestatic cirrhosis and acute hepatic failure, when compared with noncholestatic cirrhosis and time after transplant, had a positive effect on the likelihood of subsequent rejection episodes. It also demonstrates that diabetes, length of stay, and graft rejection had negative effects on the Karnofsky score most contemporaneous with HRQOL, whereas time after transplant had a positive effect on the Karnofsky score. Finally, the MELD score, cholestatic cirrhosis, autoimmune/metabolic disease, neoplasm, time after transplant, and the Karnofsky score all had positive effects on the SF-36 PCS scale score. The AGFI was 0.95 and the $\chi^2$ $P$ value of .65 indicated good fit of the model to the underlying data.

All of the statistically significant direct effects composing model 1 (effects on the SF-36 PCS scale score) and the 3 other models (effects on the SF-36 MCS scale score, composite physical HRQOL score, and composite mental HRQOL score) are summarized in Table 2. Effects on mediating variables are identical in all of the models. These are the effects of cholestatic cirrhosis, acute hepatic failure, and time after transplant on rejection and the effects of diabetes, length of stay, time after transplant, and rejection on the Karnofsky score.

<table>
<thead>
<tr>
<th>Table 1. Demographic and Clinical Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
</tr>
<tr>
<td>Age, mean (SD), y</td>
</tr>
<tr>
<td>Male, %</td>
</tr>
<tr>
<td>White, %</td>
</tr>
<tr>
<td>MELD score without exception points, mean (SD)</td>
</tr>
<tr>
<td>Listing MELD score with relevant exception points, mean (SD)</td>
</tr>
<tr>
<td>Diabetes, %</td>
</tr>
<tr>
<td>Education, %</td>
</tr>
<tr>
<td>Less than high school diploma</td>
</tr>
<tr>
<td>High school diploma or equivalent</td>
</tr>
<tr>
<td>College</td>
</tr>
<tr>
<td>Diagnosis, %</td>
</tr>
<tr>
<td>Noncholestatic cirrhosis</td>
</tr>
<tr>
<td>Cholestatic cirrhosis</td>
</tr>
<tr>
<td>Autoimmune/metabolic disease</td>
</tr>
<tr>
<td>Acute hepatic failure</td>
</tr>
<tr>
<td>Neoplasm</td>
</tr>
<tr>
<td>Length of stay, mean (SD), d</td>
</tr>
<tr>
<td>Time after LT, mean (SD), mo</td>
</tr>
<tr>
<td>Any rejection episodes, %</td>
</tr>
<tr>
<td>SF-36 PCS scale score, mean (SD)</td>
</tr>
<tr>
<td>SF-36 MCS scale score, mean (SD)</td>
</tr>
</tbody>
</table>

Abbreviations: LT, liver transplant; MCS, mental component summary; MELD, Model for End-Stage Liver Disease; PCS, physical component summary; SF-36, 36-Item Short Form Health Survey.
Autoimmune/metabolic disease, neoplasm, and the Karnofsky score were associated with an improved composite mental HRQOL score. The MELD score had no effect on the composite mental HRQOL score. The AGFI was 0.96 and the chi^2 P value was 0.88.

Summaries of the statistically significant direct and indirect effects on the 4 HRQOL outcome measures are given in Table 3 and Table 4. Values are shown as standardized regression coefficients (β). Total effects denote the additive effect of significant direct and/or indirect paths.

Organ transplantation and its associated HRQOL is a growing field with implications regarding surgical decision making. Our study contributes to this field by examining the relationship among several variables surrounding LT with postoperative, patient-reported HRQOL.

The preoperative MELD score, without exception points, had a direct positive effect on the postoperative physical HRQOL. This means that after controlling for other effects in the model, persons with higher pretransplant MELD scores had better posttransplant physical HRQOL. The effect was small but statistically significant in both models of physical HRQOL and was opposite of what we had expected based on the findings by Rodrigue et al8 in which an increasing MELD score was associated with a worse SF-36 PCS scale score.

An interesting aspect of our model that sheds light on this finding is the inclusion of a diagnosis category. We stratified patients into 5 diagnostic categories, with noncholestatic cirrhosis as our reference category. In this model, patients with neoplasm had the lowest mean MELD score. However, these same patients had better posttransplant physical HRQOL. Additionally, patients with cholestatic cirrhosis and autoimmune/metabolic disease had improved physical HRQOL. This would suggest that the subset of patients who are driving this finding are those with noncholestatic cirrhosis as patients with hepatitis C virus. It is these patients who have worse perceived postoperative physical HRQOL.

The MELD score was not correlated with postoperative mental HRQOL in either of the mental HRQOL models. This finding is slightly different from, but does not refute, the finding by Sainz-Barriga et al16 that patients with more advanced stages of pretransplant liver disease as measured by the Child-Turcotte-Pugh score had a protective effect in terms of the development of psychosocial distress after LT. When comparing these findings with that by Saab et al7 that worse liver disease severity prior to transplant was associated with worse scores on the mental components of the SF-36, it is interesting to note the marked improvement in how patients perceive their mental symptoms after LT. The mean MCS scale score in our population, 48, approximates that of the general population. This is similar to the findings by Moore et al17 and Pinson et al18 that the mean MCS scale score after heart transplant or LT is approximately 48.

The most salient predictor of improved physical and mental HRQOL in all of the models was the Karnofsky score.
most contemporaneous with the HRQOL measure. This has been shown previously in work done with kidney transplant, heart transplant, and LT recipients.\textsuperscript{18-21}

Another predictor of physical and mental HRQOL is time. In both physical HRQOL models, time after transplant had positive direct and indirect effects on physical HRQOL. In both models of mental HRQOL, time after transplant had a positive indirect effect on mental well-being that was mediated by rejection episodes and Karnofsky functional performance. This is consistent with previous reports showing that organ transplant has small, positive effects on mental HRQOL.\textsuperscript{18-21}

In this study, several patients did not complete all of the surveys. To address this, we used regression imputation to estimate respondent answers based on their complete SF-36 data and subsequently computed composite physical and mental HRQOL scores. There was very little difference between these physical or mental composite HRQOL models (models 3 and 4) and their PCS and MCS.
scale counterparts (models 1 and 2). The only additional factor that significantly influenced the composite physical HRQOL score was educational attainment. Diagnosis of neoplasm was the only additional covariate having an effect on the composite mental HRQOL score. When the composite physical and mental HRQOL models were refitted using only those 157 patients with complete survey data, they did not differ from their full-sample (N = 209) counterparts (data not shown).

The implications of this research are 2-fold. First, patients with more advanced disease and higher MELD scores can be expected to have better HRQOL after hepatic transplant than patients with less severe disease as reflected by MELD scores that do not include any diagnosis-related exception points. Even though the finding was relatively small, it was statistically significant with a modest sample and indicates that even the most gravely ill patients, given survival through the perioperative period, can be expected to have no worse physical HRQOL as compared with those who are less ill. Second, having patients with less severe disease undergo transplant early may result in poorer physical HRQOL outcomes in the initial posttransplant year. Our results indicate that the MELD score has not only stratified organ allocation to those most in need, maximizing survival, but has also created a system in which HRQOL is maximized. These findings further support the use of the MELD scoring system for assigning priority in patient selection for hepatic transplant. Although we cannot propose adding HRQOL to the MELD system currently used for organ allocation, it should be considered with additional research.

There are limitations to this study. Perioperative morbidity and mortality preclude some patients from participating, so the potential for a selection bias exists. Some patients choose not to respond and thus could influence the results. In our study, we had a 71% participation rate. Another limitation is the response burden and fatigue inherent with patients completing multiple surveys that may influence some of the HRQOL measures. We attempted to account for this with regression imputation. Finally, one must consider that an acceptable model fit does not imply causality but that the model is one of many possible models that may be fitted to the underlying data.

In conclusion, we determined that an increasing MELD score (without diagnosis-based exception points) prior to LT was associated with improved perceived physical HRQOL in the first year after LT. We also found that the MELD score had no significant effect on mental HRQOL after hepatic transplant.

Accepted for Publication: February 1, 2008.

Correspondence: C. Wright Pinson, MD, MBA, Division of Hepatobiliary Surgery and Liver Transplantation, Vanderbilt University Medical Center, Ste 801 Oxford House, 1313 21st Ave S, Nashville, TN 37232-4733 (wright.pinson@vanderbilt.edu).

Author Contributions: Study concept and design: Castaldo, Feurer, and Pinson. Acquisition of data: Castaldo and Russell. Analysis and interpretation of data: Castaldo and Feurer. Drafting of the manuscript: Castaldo. Critical revision of the manuscript for important intellectual content: Feurer, Russell, and Pinson.

Statistical analysis: Castaldo and Feurer. Administrative, technical, and material support: Castaldo, Feurer, Russell, and Pinson. Study supervision: Pinson.

Financial Disclosure: None reported.

Funding/Support: This work was supported in part by an educational grant from Novartis Pharmaceuticals, Inc.

Previous Presentations: This paper was presented in part at the 115th Scientific Session of the Western Surgical Association; November 6, 2007; Colorado Springs, Colorado.

Additional Contributions: Jerita Payne, MSN, April DeMeers, MSN, Marni Groves, MSN, Julie Dykes, MSN, and Matthew Bumbalough, MSN, were the transplant coordinators in this study, and Hua Ye, BS, and Mindy Stahley, MSN, were the data managers in this study.

REFERENCES