**Predicting Short-term Outcomes of Liver Transplants with an**

**Emphasis on the role of Body Composition**

By: Sarah Torrence

**Executive Summary:**

Patients with end-stage liver disease require liver transplant surgery from a living or deceased donor. As this is a very serious condition, outcomes can include infections, complications and even death. Most patients spend a few days in the ICU, around a week in the hospital and possibly even time in an inpatient rehabilitation facility post-liver transplant. Being able to predict the outcome status and journey of a liver transplant patient can help with resourcing and reducing costs. Further understanding what factors are associated with liver transplant outcomes can assist in patient recovery and preventative care to decrease the need of liver transplant surgery. The goal of this project is to predict short-term outcomes of liver transplant surgery and understand if body composition is associated with outcomes. For this analysis odds ratios, then converted to probabilities of being in a state, are predicted for each of the 90 days post-liver transplant across four ordinal levels of state: 0 – at home or in an inpatient rehabilitation facility, 1 – in the hospital, 2 – in the ICU or on a ventilator, 3 – dead.

The original data set includes 140 variables for 422 liver transplant patients at Vanderbilt University Medical Center from 2009-2019. This data was cleaned and processed in order to have an observation of each patient on each of the 90 days post liver transplant. Data reduction included removing variables based on domain knowledge, a redundancy analysis using adjusted R2 measures, and a clustering analysis based on Spearman correlation and Hoeffding D. Missing data was imputed using predictive mean matching with 5 iterations of imputation and the final data set used for modeling included 27 variables (1 outcome and 26 predictors) of 37,501 observations.

There were two different models created each with advantages and disadvantages that complement each other. A proportional odds model, with correction for within patient correlated, has an R2 of 0.913 and included splines for all continuous variable and age and sex interaction with skeletal muscle area. The predictors with the largest measures of association (based on χ2 and p-values) with the outcome state are the previous state, MELD score, skeletal muscle area and time. Skeletal muscle area had a statistically significant association with outcome (based on an α = 0.05) and additionally a chunk test of overall body composition showed a significant association with outcome state. Although significant, body composition proved to have a small effect on outcome states. The second model is a shallow neural network with 2 layers (a fully connected layer and an output layer) using the consistent rank logits (CORAL) framework to handle ordinal outcomes. This model has an accuracy of 97.77%, however, does not allow for easy analysis of feature importance. Model output is a 4 x 37,501 matrix with a probability for each of the 4 states for each observation. This can be flexibly used to compare probabilities of different states, averaged over patients, or calculated for specific time ranges allowing the model to be used for many applications. To further understand the relationship between body composition and liver transplant outcomes, a more granular analysis could be performed to look at specific body compositions individually as well as using body composition to predict other outcomes such as infections, complications, or hospital readmission.

**Introduction:**

Having a healthy liver is critical to survival which is why liver transplant surgery is vital for those with end-stage liver disease. There are about 8,000 liver transplant surgeries performed annually in the United States, however, due to a lack of donors there are thousands more who require the surgery that sit on waiting lists. The recovery post liver transplant includes staying in the ICU, once stable spending time in the hospital to recuperate and recover, the possibility of being discharged to an inpatient long term care facility and/or being discharged to a patient’s home. All patients have continued follow up appointments, physical therapy and will take several medications including some they will have to take for the rest of their lives.1 Can outcomes of liver transplants including the recovery process be predicted? If so, what factors contribute to adverse outcomes? In this study, I built two models to predict the status of a patient on a given day for the first 90-day period post liver transplant. Further, I analyzed which factors are the most important in contributing to these predictions, specifically analyzing how different body composition measures can affect short-term liver transplant outcomes. Two different methods were used to perform this analysis: one of a statistical nature (proportional odds regression) and one pertaining to machine learning (neural networks). A secondary interest of this work was to compare and contrast these two methods to understand which might be better at tackling different components of the problem. This research does not include an overall survival analysis of liver transplant patients, but an analysis of predicting short-term outcomes within the first 90 days post-surgery. While extremely important, follow-up post 90 days is out of the scope of this project.

***Background:***

There are numerous different causes of chronic liver disease, but the most common is cirrhosis, a disease in which scar tissue replaces normal liver tissue causing malfunction within the liver. This can be the result of many causes including alcoholic steatohepatitis causing damage to the liver due to excessive alcohol consumption, non-alcoholic steatohepatitis causing damage due to fat building up in the liver, autoimmune disorders, and viral diseases such as hepatitis.1 There are numerous studies showing that different body composition or fat measures such as sarcopenic obesity, skeletal muscle index and myosteatosis (skeletal muscle fat infiltration) have an effect on mortality and survival rates at longer term periods such as 1, 3 or 5 years post liver transplant2,3,4. However, there is not as much research on how body composition can affect short-term status outcomes of liver transplants. It is obvious to see how important long-term survival research is to improving liver transplant surgeries and allowing patients to have longer healthier lives post-transplant, but understanding and improving short-term outcomes of liver transplant surgeries can be extremely beneficial as well. Increased ICU length of stay has been associated with increased risk of long-term mortality.5 Additionally, understanding hospital length of stay and inpatient rehabilitation needs can help hospitals with resource allocation and costs as well as give patients recovery and cost expectations. It can also be used in the context of improving interventions pre-transplant to either decrease the need of liver transplant surgery or improve post-surgery outcomes.

**Data:**

The data set used for this research was collected from 422 adult patients who underwent liver transplant surgery between 2009-2019 at Vanderbilt University Medical Center. The original dataset includes 140 variables including demographic information, pre-transplant and post-transplant comorbidities, surgical variables, and outcome information including days in various statuses, infections and mortality. Additionally, pre-transplant and post-transplant follow up CT scans were taken and analyzed for measures of body composition. Some patients have follow-up CT scans for up to 10 years, but many have only a few years of follow-up. It is important to note that the dates of liver transplants span from 01/09/2009 – 12/23/2018 while follow-up data was continued to be collected until summer 2021. As short-term outcomes were the focus of this analysis, many follow up measurements from annual checkups were not included in this work. A full data dictionary can be found at (<https://github.com/sktorre/liver_transplant_research/blob/main/Data%20Dictionary.docx>).

Table 1 shows some descriptive statistics of the 422 liver transplant patients in the data set. Here we can see that 63% are male while 37% are female. The average age is 55.7 and the median hospital length of stay is 8 days. The average MELD score, a score based on results from several lab tests that together, show how well the body is functioning, is 22.1. A higher score means the more likely a patient is to receive a liver from a deceased donor. The three most common etiologies of liver disease are non-alcoholic and alcoholic steatohepatitis (fatty liver disease) as well as viral infections. The average BMI is 29.2 which is in the overweight range.

Table

Description automatically generated

**Table 1:** Descriptive statistics for the 422 liver transplant patients broken down by

gender, age, MELD score, hospital length of stay, etiology of liver disease and BMI

The body composition measurements included in this analysis were all derived from CT scans at the level of third lumbar vertebra and were captured using image analysis software. These measurements include skeletal muscle area (cm2), skeletal muscle mass index (cm2/m2) – skeletal muscle area divided by squared height, visceral adipose tissue (cm2) – fat found predominantly around organs in the abdominal cavity, subcutaneous adipose tissue (cm2) – fat found right under the skin, skeletal muscle density (Hounsfield Units), visceral adipose tissue density (Hounsfield Units) and subcutaneous adipose tissue density (Hounsfield Units). Other important variables in this analysis include categorical variables indicating whether the patient was a smoker, had ascites, had hepatic encephalopathy, as well as surgical variables such as duration of surgery in minutes.

***Cleaning/Processing:***

There were a few categorical variables that were cleaned before processing due to low numbers of individual categories. This included grouping location of discharge into either home or an inpatient long term facility (IPR), making different variations of biliary complications a yes/no indicator of whether a patient had any biliary complications or not and grouping comorbidities pre- and post-transplant into combined measures indicating whether the patient had that comorbidity at any time during the study. These comorbidities include chronic obstructive pulmonary disease, hypertension, diabetes mellitus, coronary artery disease, chronic kidney disease and connective tissue disease.

As mentioned previously, the original data set has 422 observations, one for each patient. This means that all longitudinal measurements, dates, and times were in the same row corresponding to that patient. To follow each patient for the first 90 days post-liver transplant, transformation of the data to include 1 row for each day, patient combination up until 90 days with day 0 being the day of the liver transplant surgery was required. The new data set needed to include a variable for *time* (in days) to indicate which day post-surgery each corresponded to and an ordinal variable I called *state* to indicate what status a patient was in on a given day. Based on the available information in the data, 4 different statuses were created: 0 – at home or in an IPR, 1 – in the hospital, 2 – in the ICU or on a ventilator, 3 – dead. The variables used in this process were ICU length of stay (LOS), total hospital LOS (including ICU LOS), days on a ventilator and days from surgery until death. The only information not provided was number of days at home or in an inpatient long term care facility. There is a date of discharge, therefore it was assumed that days at home or in IPR are days post-discharge until death. Once *time* and *state* variables were created, this information was used to create a previous state (*prev\_state*) variable which indicated the previous status of the patient. The data allowed for follow-up far past 90 days for most patients, but as this analysis is focused on the short-term outcomes of liver transplants, it was filtered for only the 90 days post-liver transplant for each patient. This gave 37,980 (422 x 90) observations in total. Finally, as death is an absorbing state (a patient cannot change states once they die), all observations past death were removed for any patients who died within the first 90 days. As previous state is a variable required for modeling, all previous states for days post-death date would be dead giving this variable too much inflated importance. It would affect the analysis of focusing on the transition of states up until death if post-death days were included. This brought the final dataset to 37,501 observations available for modeling. Figure 1 shows the first 60 days after liver transplant for a random sample of 25 patients. As shown with a black square, the last patient in the figure dies on day 9 and follow-up is terminated.

Chart, bar chart

Description automatically generated

**Figure 1:** The post liver transplant journey of a random sample of 25 patients as they transition through the

4 states starting in the ICU/Vent.

**Methodology:**

***Data Reduction/Variable Selection:***

With the original 140 variables, minus a few for processing and adding a few engineered, there was a large selection of features to choose for modeling. This posed as a data reduction challenge as including too many parameters in a model can lead to overfitting. The first form of data reduction was filtering out a majority of variables based on domain knowledge, credibility and the type of analysis being performed. A large portion of variables removed from further analysis were the longitudinal body composition measures taken for up to 10 years post-surgery. As this analysis only includes 90 days post-liver transplant and the first measurements were taken one year after the transplant, these measurements were out of scope. Another variable excluded was the donor risk index. Due to the resident creating this data set establishing his own formula for this measure, it is flawed and not the industry standard donor risk index measure. Additional outcome variables, other than the states derived above, were also excluded such as readmission, infections or other complications.

After manually excluding many variables, there were 29 variables (28 possible predictors and one outcome variable) left. An analysis of missing variables was done to understand if there were any possible predictors missing a large number of values. If so, those variables could be excluded from the analysis. Fortunately, the largest percentage of missing values for any given variable was only 4.5% which is plenty small enough to impute without issues. A redundancy analysis was also performed. If any variables can be easily explained by other variables, it might be redundant to include them in a model. To determine redundancy, the adjusted R2 metric was examined using all other variables to predict each individual variable6. Through this analysis, skeletal muscle mass index was found to be redundant due to it being a derivation of skeletal muscle area and height which are both also included in this set of variables. Many other body composition variables were also explaining much of the variation of other body composition variables. This makes sense as these are all different measures of body weight and fat that are highly correlated, however, due to understanding the impact these measurements have on liver transplant outcomes being a main component of this project, all other body composition measurements were kept as features for modeling.

Diagram, schematic

Description automatically generated

Diagram, schematic

Description automatically generated

**Figure 2:** Hierarchical cluster analysis – the figure on the top shows clusters formed using the Hoeffding

D statistic while the figure below shows clusters formed using Spearman correlations.

The last method of data reduction used was hierarchical cluster analysis to understand how similar variables are to one another6. Using two different similarity measures, the Hoeffding D statistic and Spearman correlations, several clusters were found in the data. As shown in Figure 2, these two methods produce similar results. Based on high Hoeffding D and Spearman correlation measures, weight and subcutaneous adipose tissue form a cluster, height and sex form a cluster, and visceral adipose tissue density and subcutaneous adipose tissue density also form a cluster. Weight is a body composition measure but it not as granular or important to this analysis as the CT scan body measures, therefore weight was removed as a feature for modeling. As mentioned before, although there is similar information captured within some of the body composition measurements, as they are important factors in this analysis, they were all kept as features for modeling. After extensive data reduction analysis due to numerous possible predictors, 26 predictors with of course 1 outcome (state) were kept and used in the modeling process.

***Data Imputation:***

As shown in Figure 3, while about half of the predictors are missing some data, only less than 1% of observations are missing for each of these variables. To utilize these variables as model inputs, there must

Table

Description automatically generated

**Figure 3:** The percentage of missing values for all variables to be used as features

no missing values present. The method of multiple imputation was used to impute missing values. This method takes random draws of a variable conditioning on a set of all other variables which in this case was all the possible predictors and the outcome variable. The method used to calculate imputed values is predictive mean matching which uses a bootstrap sample of donors from the complete observations and predicts the missing value using regression. To account for variation of imputed values, the process ran through 5 iterations creating a new full dataset with imputed values for each iteration with each new iteration using the previous iterations imputed values as a starting point. Technically, the algorithm runs 8 iterations with the first 3 iterations ignored as “burn-in”.6

***Modeling:***

I took two very different approaches to modeling, one statistical and one machine learning, and built two different models to compare which was a better fit for this analysis. There are pros and cons to each of the models and the combination of the two approaches provides for a robust analysis of short-term outcomes of liver transplants based on many factors but more specifically body composition. As mentioned before, the outcome status is ordinal with 4 levels, 0: home/IPR, 1: Hospital, 2: ICU/Vent, 3: Dead. Because the difference between being at home/IPR and being Dead is much different from either being at home/IPR and being in the hospital, the models must account for ordinality and be able to factor in these levels of differences in categories in assessing model performance. Additionally, as the outcome is not a single value for each patient, but a predicted status outcome for each of the 90 days post-liver transplant, the models must be able to handle the dimensions of time, previous day state and intra-patient correlation between observations for the same patient.

***Statistical Approach using Proportional Odds and GEE:***

Methods of ordinal logistic regression have long been used to study ordinal response variables. A common approach to ordinal logistic regression is proportional odds. Proportional odds models are semi-parametric models working as ordinary logistic models for a fixed event (level of ordinal outcome) with the addition of *k* intercepts where *k* is the number of levels in the outcome variable. Model outputs are odds ratios for Y ≥ j for *j* = 1, 2*, … k*. Proportional odds models have the same assumptions as binary logistic models for each *j* cut-off, but also have additional assumptions including the model coefficients being independent of *j* and the proportional odds assumptions that the coefficients that describe the relationship between various levels of *j* are all the same for all levels of *j*.6 This can be verified in several ways, but usually involves plotting the stratum of each level of outcome by each predictor separately in some form. There should be parallelism between each level of outcome as they should have the same relationship across different levels with each predictor. Figure 5 shows an example of what a plot verifying this assumption with respect to time could look like. As shown, the proportional odds assumption does not hold. This is mostly due to all patients starting their journey post-liver transplant in the ICU and death being an absorbing state. Although the proportional odds assumption does not hold, it

is not detrimental to this analysis and a proportional odds model can still be very useful.7

Chart

Description automatically generated with medium confidence

**Figure 5:** Proportion of States across 90 days post-liver transplant showing the proportional odds

assumption does not hold with respect to time

With 26 possible predictors (excluding the patient identifier), 14 categorical and 12 continuous, there are infinite many ways of defining model parameters. Using the following formula, the effective sample size is 39152.34.

**Formula 1**6:

Using a very basic guideline such as the 15:1 rule of thumb of effective *sample size:degrees of freedom* (df), this model can handle a lot of flexibility in terms of parameters without much worry for overfitting. In order to allocate more degrees of freedom to the most important predictors, a comparison of their predictor importance was done by calculating χ2 – df from the variance-covariance matrix after adjusting for within-patient correlation.6 In defining the proportional odds model, continuous covariates with large importance from this analysis have nonlinear effects using restricted cubic splines with 3 knots and categorical variables utilize *k*-1 degrees of freedom for *k* categories. As body composition varies between genders and with age, capturing the interaction between these three factors was important. As there are many body composition variables, skeletal muscle area was chosen to interact with age and sex due to it being the most important body composition measurement from the variable selection analysis.

After fitting a proportional odds model with the chosen parameters described above, a generalized estimating equations method was used to account for within-patient correlation as there are up to 90 observations for each patient. This is done by using the robust sandwich covariance estimator to correct the variance-covariance matrix based on intra-patient correlation after the model is fit.8 Fast backwards variable selection was then used to remove 7 predictors from the model. This method removes entire predictors, not parameters, therefore if a predictor included in an interaction parameter or with multiple parameters to create splines is removed, all parameters including that predictor are removed. The criteria to remove predictors was based on those with the highest p-values. This process selected more variables than I was willing to remove due to domain importance of certain predictors, therefore, some predictors were removed during backward selection but kept in the model. The final model has 16 predictors with 40 degrees of freedom. Time and MELD score are splines with 5 knots, all other continuous predictors are splines with 3 knots and there is interaction between skeletal muscle area and age as well as skeletal muscle area and gender. Bootstrap validation was used to validate the proportional odds model with 100 bootstrap samples. The validation process also used backward step-down variable deletion as it was used in creating the model. Without using backward step-down variable deletion in the validation process, model standard errors and confidence intervals would be too small as there would be no penalization for the variable selection process.6

***Machine Learning Approach using Neural Networks:***

A completely different approach to predicting states for 90 days post-liver transplant is utilizing neural networks. Neural networks are machine learning models that were created to mimic neurons in the human brain to process and understand signals coming from input features. Neural networks are impressively good at classification and more specifically the consistent rank logits (CORAL) framework handles ordinal regression well which is why it made a good choice of models for the analysis. This method can be implemented with any commonly used neural network architecture as only the output layer and cost function need to updated. CORAL turns the *k* states into *k*-1 binary classification problems with monotonic ranking to guarantee these binary tasks output consistently ranked predictions.9 Model outputs are ordinal cumulative logits which can be converted to probabilities using the softmax activation function.

Before fitting a model, the data was split into a training and testing set stratified by patient with 80% of patients in the training set and 20% of patients in the testing set. The split was stratified by patient in order to preserve all observations within the 90-day period for each patient in the same data set. The data was then scaled using a min max scaler to ensure all values were between 0 and 1 for each input feature. The Tensorflow Keras implementation of CORAL10 was used to create a simple neural network with only one fully connected dense layer. The unique patient ID was included in this model to account for within patient correlation. The total number of parameters in the model is 786.

**Results:**

***Statistical Approach:***

The proportional odds model has an adjusted R2 value of 0.913 and a χ2 equal to 31314.66. The rank discrimination index ρ has a value of 0.576. Based on all these measures, this model is robust and explains a high variation in the data and shows strong association between the outcome and at least some of the predictors in the model. Figure 6 shows χ2 and p-values for tests of association between each

Chart, table

Description automatically generated

**Figure 6:** Partial tests of association for each predictor including a total

body composition chunk test

individual predictor and the outcome state. In addition to individual model predictors, it also

includes a chunk test of all the body composition measures. This test has a χ2 value of 36.3 and a p-value of 0.0064 meaning body composition does have a significant association with post-liver transplant outcome states. Based on an α = 0.05, all predictors having a significant association with outcome

Diagram

Description automatically generated

**Figure 7:** Partial effects plots for each predictor in the model.

state are the state on the previous day, MELD scores, skeletal muscle area, time – days since liver transplant, the duration of the surgery, chronic kidney disease diagnosis, diabetes mellitus, and an interaction between skeletal muscle area and age. Although these predictors all show associations with the outcome, Figure 6 shows that the previous state dominates most of the variance explained in the model and has the largest effect on outcomes. This can limit the understanding of other predictors in the model. Figure 7 shows partial effects plots for each of the predictors in the model. Here it is easy to see that even though there are numerous significant predictors, their effect on the outcome state is minimal except for previous state. Bootstrap validation solidifies this as there is a low chance of overfitting. The index corrected measures for adjusted R2, ρ and slope are essentially the same as their original measures with extremely low levels of optimism.

***Machine Learning Approach:***

The neural network has an accuracy of 97.77%, a mean absolute error of 0.023 and a root mean squared error of 0.16. These measures account for the ordinality of the outcome and therefore penalize errors larger for predictions further from the true values. For example, if the model predicts the patient is in the hospital but the patient has died, the error rate will be larger than if the prediction was in the hospital but the patient is actually in the ICU and/or on a ventilator. These measures were calculated using the test set showing this model predicts very accurately on new data. Although the model is accurate, it is very challenging to understand feature importance from a neural network as the parameters are weights sharing information across various predictors that cannot be easily traced and associated with single predictors. This limits our understanding of how and if body composition is important in predicting short-term outcome status of liver transplant patients. Correlation can act as a proxy estimate to give a better understanding of the relationship between body composition and outcome state. Correlation measures of body composition and state are stable across time within this study, however, do not show any strong relationships. For example, at day 10 the strongest relationships between any body composition variables and state are a negative correlation of 0.27 between state and skeletal muscle density and a positive correlation of 0.25 between state and subcutaneous adipose tissue density. Like the results of the proportional odds model this does not show much association between body composition and state.

***Comparison of Approaches:***

The two modeling approaches described above are theoretically quite different and while some similar conclusions can be made from both approaches, there are also some major differences. The proportional odds model is a semi-parametric model while neural networks are parametric. A proportional odds model is intended specifically for handling ordinal outcomes where neural networks typically handle classification. However, utilizing CORAL in this instance highlights the true flexibility of applications of neural networks. With a large enough sample size, validating neural networks is most commonly done through a test set as done in this analysis, but with a regression model such as proportional odds model it is more common to use bootstrap validation or cross-validation. A very important difference between these modeling approaches is how predictors are input into the model. A neural network inputs all features the same way in raw form and can handle any nonlinear associations or interactions extremely well. As it is a black box, this is not something that can be understood, but there is no concern for making inaccurate assumptions about predictors and their relationships with the outcome variable. This is a large concern with a proportional odds model as for each predictor it must be manually decided upon how that predictor should added to the model. All nonlinearity, interactions or transformations must be specified even though the true relationship with the outcome is unknown.

The largest limitation to the proportional odds model is its reliance on assumptions, most importantly the proportional odds assumption. While neural networks, do not rely on underlying assumptions, this analysis has shown how difficult it is to understand feature importance from a neural network. This is something that is quite easy to comprehend with a proportional odds model. As far as model predictions, Figure 8 shows the probability of being home/IPR for 90 days post-liver transplant averaging over all patients. The first 10 and the last 30 days are extremely similar for both models, but the proportional odds model has lower probabilities from about day 10 to day 60.

Chart, line chart

Description automatically generated

**Figure 8:** Comparing Model Predictions – Probabilities of being at home/IPR

averaged across all patients

**Conclusion:**

While the focus of this analysis was to create a model to accurately predict short-term outcomes of liver transplants and understand body composition’s role, the models created have a very flexible output that can be used for many purposes. Model outputs are a matrix of odds ratios for each outcome state for each day per patient. These can be converted to probabilities, cumulative probabilities, compared for individual days, across time ranges and could be left as probabilities or convert through classification to predicted states. They can be averaged or looked at more granularly for individual patients as well as understanding quartiles, medians or numerous other statistical values. Using longitudinal and ordinal modeling methods in this analysis allows for a wide range of applications. However, as the main goal of this analysis was to understand the role of body composition in short-term outcome states post-liver transplant, the results are inconclusive. Body composition and more specifically the body composition measure skeletal muscle area is significant in predicting outcome states, however the effect size it has in affecting the outcome states is very minimal.

***Limitations:***

The main limitations to this analysis include the amount of data, data integrity and model limitations. There were only 422 patients in this analysis therefore increasing the sample size would increase the accuracy and legitimacy of results. Additionally, several patient records had incorrect entries or were missing information due to corrupt files. Some observations were cleaned manually to combat this issue, but not all incorrect or missing information could be accounted for. There were also several variables that were thrown out due to legitimacy concerns or miscalculations. This included days a patient was in an inpatient rehabilitation facility. For this analysis days in IPR was combined with days at home, but given the correct information, the analysis could be more granular in separating these categories as there is a clear distinction in outcome state of being at home versus an inpatient rehabilitation facility. Finally, both models have limitations; the proportional odds model relies on the proportional odds assumption which was not met, and the neural network has major limitations in interpreting feature importance.

***Next Steps:***

The most immediate next steps in this analysis include relaxing the proportional odds assumption and creating a deeper neural network. The proportional odds assumption can be relaxed for specific predictors for which the assumption does not hold utilizing a partial proportional odds model.6 This however adds a large amount of complexity to the model as there are additional parameters added for each relaxed predictor based on the number of state outcomes. For example, if the assumption is relaxed for a continuous variable that is added to the model using splines with three knots, there would be an additional 12 parameters added to the model based on just that one predictor. The neural network created for this analysis is extremely shallow with only 2 layers, one fully connected layer and one output layer. The architecture of the network could be deepened and tuned to increase accuracy of results. With the original goal of understanding the role of body composition in liver transplant outcomes, more analysis could be done to understand whether it is associated with short-term outcomes of liver transplants and in what way. The original data included longitudinal measures of body composition for up to 10 years for each patient. These measures could be utilized to understand body composition’s role in long term outcomes of liver transplants. There are also additional liver transplant outcomes not studied in this analysis including rate of infections, readmissions, and other complications that could be associated with body composition measures. Finally, a more granular analysis could be done to understand the effects of specific body composition measures analyzed individually.

***Acknowledgements:***

I could not have completed this project without the help of some amazing individuals. I would first like to thank my academic advisor Dr. Heidi Silver for not only providing me access to this data set and problem, but also assisting me along the way with domain knowledge insights on how the data was collected, what type of similar research has already been done in this field and hypotheses on how variables might interact. I would also like to thank Chiara Di Gravio for her extensive statistical knowledge, consistent support and patience through all my questions. Additionally, I would like to acknowledge Dr. Frank Harrell for pushing me throughout his class and empowering me to explore ordinal longitudinal data outcomes through proportional odds models. Finally, I would like to thank Jesse Blocher and all faculty and peers at the DSI for helping me grow as a data scientist over the last couple of years.

**References:**

1. Mayo Foundation for Medical Education and Research. (2021, June 2). Liver transplant. Mayo Clinic. Retrieved April 15, 2022, from https://www.mayoclinic.org/tests-procedures/liver-transplant/about/pac-20384842.
2. Hegyi PJ, Soós A, Hegyi P, Szakács Z, Hanák L, Váncsa S, Ocskay K, Pétervári E, Balaskó M, Eröss B, Pár G. Pre-transplant Sarcopenic Obesity Worsens the Survival After Liver Transplantation: A Meta-Analysis and a Systematic Review. Front Med (Lausanne). 2020 Dec 16;7:599434. doi: 10.3389/fmed.2020.599434. PMID: 33392221; PMCID: PMC7772841.
3. Kuo SZ, Ahmad M, Dunn MA, Montano-Loza AJ, Carey EJ, Lin S, Moghe A, Chen HW, Ebadi M, Lai JC. Sarcopenia Predicts Post-transplant Mortality in Acutely Ill Men Undergoing Urgent Evaluation and Liver Transplantation. Transplantation. 2019 Nov;103(11):2312-2317. doi: 10.1097/TP.0000000000002741. PMID: 30985575; PMCID: PMC6783339.
4. Irwin NEA, Fabian J, Hari KR, Lorentz L, Mahomed A, Botha JF. Myosteatosis, the More Significant Predictor of Outcome: An Analysis of the Impact of Myosteatosis, Sarcopenia, and Sarcopenic Obesity on Liver Transplant Outcomes in Johannesburg, South Africa. Exp Clin Transplant. 2021 Sep;19(9):948-955. doi: 10.6002/ect.2021.0083. Epub 2021 Aug 9. PMID: https://github.com/ck3734387151.
5. T.A. Williams, K.M. Ho, G.J. Dobb, J.C. Finn, M. Knuiman, S.A.R. Webb, Effect of length of stay in intensive care unit on hospital and long-term mortality of critically ill adult patients, British Journal of Anaesthesia, Volume 104, Issue 4, 2010, Pages 459-464, ISSN 0007-0912, https://doi.org/10.1093/bja/aeq025.
6. Harrell, F. E., Jr. (2016). Regression modeling strategies. Springer International Publishing.
7. Harrell, F. E., Jr. (2022, March 9). Assessing the Proportional Odds Assumption and Its Impact. Assessing the proportional odds assumption and its impact. Retrieved April 15, 2022, from <https://www.fharrell.com/post/impactpo>.
8. Harrell, F. E. (2021, February 27). Longitudinal Ordinal Analysis for Violet2. Longitudinal ordinal analysis for Violet2. Retrieved April 15, 2022, from http://hbiostat.org/proj/covid19/violet2.html#gee-type-proportional-odds-modeling.
9. Cao, W., Mirjalili, V., & Raschka, S. (2019). [Rank-consistent ordinal regression for neural networks](https://arxiv.org/abs/1901.07884). arXiv preprint arXiv:1901.07884, 6.
10. Kennedy, C. (2022) Ordinal regression in Tensorflow Keras, Git-Hub repository, <https://github.com/ck37/coral-ordinal>.