Questions:

1. What is the timeline for our data set? How long after the surgery do we have information tracking patient's body composition? (ex. Paper 1 defines short term as 1 year, intermediate as 3 years and long term as 5 years)

Some patients have info up to 10 years, but not all patients will have this much information. We can decide our own timeframe based on our data and the timespan that still includes all or most of the patients.

1. Many papers discussed the importance of post-transplant intervention, including nutritional therapy, rehabilitation with an evaluation of skeletal muscle mass and muscle functions, and physical activity interventions. Some also discussed inpatient rehabilitation programs: Do we have any information on whether patients did inpatient rehabilitation programs or followed advice from nutritional therapy, rehabilitation, etc.?

We have a yes/no indicator of whether or not a patient did inpatient rehab (check this to make sure), but we don’t have any further information about any kind of rehabilitation.

1. All papers reported many differences in outcomes or body composition between men and women. There also seem to be a larger percentage of men in all of the studies. Do we know any more information as to the cause of these differences and why more men receive liver transplants?

It is not confirmed whether men have liver transplants more often than women. They do have difference in body composition between genders which could cause different outcomes, but we will see what this looks like when we start analysis.

1. Paper 3 uses measures of muscle function (grip strength, Short Physical Performance Battery (SPPB)) - do we have any information on muscle function?

We do not have any information on muscle function

1. How do you study mortality? Is this literally whether someone passed away or not? Wait-list mortality - is this how likely a person on the waitlist will die - like a factor in determining who gets donors?

Mortality is studied commonly using survival analysis and is simply indicated as to whether someone has passed away or not at a given time. This is the same for waitlist mortality but just for patients still on the waitlist.

1. How is a MELD score determined? Is this the only factor in deciding who receives a liver transplant? Do we have this? MELD - model for end-stage liver disease
   * What is the current process for determining who from the waitlist receives a donor?

We currently do not understand the MELD score and this is something we need to look up as it is a variable in our dataset.

1. Are there children in our dataset? Based on my research it seems that there often different causes of LT among children versus adults. Most studies seem to focus on adults.
   * Further do we know the cause of the LT for our data set of patients?

Children typically have the same causes of liver transplant as adults except that they don’t have transplants due to reasons involving alcohol. There are no children in our dataset. We do have information on the cause of the liver transplant for each patient.

1. I can see how things like nutrition and exercise or more lifestyle health markers could be confounding variables. Is there any work done to adjust for these factors? I didn’t see much discussion of this in many of the papers. Some introduced mobility, but that seemed to be about it.

This is a super important factor, but unfortunately we do not have this type of data available.

1. Different studies used different cut off values of skeletal muscle index to determine whether patients had sarcopenia or other disorders. Do we just have body composition variables or do we have diagnosis information on factors like sarcopenia?

We can use the most commonly used cut off values or we can generate our own cut off values from our dataset. This can be used for just a more descriptive statistic that we can report and is helpful to understand our dataset, but for modeling we can ignore cut offs and just use continuous variables.

1. In each study data is collected for different groups of patients or in a different format. Do we have more detailed information on how our data was collected and why?
   * Additionally, it seems that body composition is determined from CT scans in many different ways. How were these variables measured in our data?

Our data is for all liver transplant patients at Vanderbilt during the time period. The body composition variables were measured by standard practice using a software that was built for this purpose.

1. Some of these studies factor costs into their analysis? Is this something that is available to us and if so do we think it could be worth including?

We could possibly collect more data on costs, but this is a entirely different research question and out of scope for this project.

1. In paper 4, they mentioned censoring for death when dealing with hospital stays. What does this mean and how do you censor for death?

Censoring for death in this instance means to remove all patients who passed away before leaving the hospital to not bias the hospital stay length. It can also mean something else to biostatisticians in other contexts.

1. Paper 5 shows a specific process for transplant. The same treatment and process was done for all patients. What can we expect in terms of process/treatment in our data?
   * My understanding is that a liver transplant can be either using a full liver from a deceased donor or part of the liver of a living donor. Do we know the type of liver transplant in our patients? Are there more types than these two that I am missing?

We have an indicator variable as to whether the transplant was from a deceased or living donor.

Notes:

* We have more patients than these studies except #6
* Different studies used different cut off values of skeletal muscle index to determine whether patients had sarcopenia
* In many studies things like age, sex, BMI, etiology, etc. don't seem to have significance in mortality but things like myosteatosis and sarcopenia do - this is on the long term - seems like they aren't associated with short term complications

Data Questions:

* How many patients have data at 1 year, 3 years, 5 years, 10 years??
* Do we have indicator of rehab?
* What are the different causes of liver transplants and what percentage of patients fall under each category?
* How many patients have missing data and for what variables? (percentages)
* Do we have the CT scan images?

Research Questions:

* Do men have liver transplants more often than women? What are differences in genders found generally amongst liver transplant patients?
* What is the MELD score and how is it determined?