Survival Analysis of Heart Transplant Patients:

1. Introduction:

This report analyzes the survival outcomes of 103 patients from a heart transplant dataset from MedDataSets, aiming to understand the effect of heart transplantation on patient survival while adjusting for covariates such as age, waiting time, and prior health condition. The analysis includes Kaplan-Meier survival curves, log-rank tests, and Cox proportional hazards models.

Research question: Does receiving a heart transplant significantly improve the survival time of patients compared to not receiving a transplant, after adjusting for other factors like age, wait time, and prior medical conditions?

Sub questions: Does receiving a heart transplant improve survival compared to not receiving one?
How do patient characteristics (e.g., age, prior conditions, wait time) affect survival time?
Is the effect of these covariates on survival consistent over time (i.e., do they satisfy the proportional hazards assumption)?

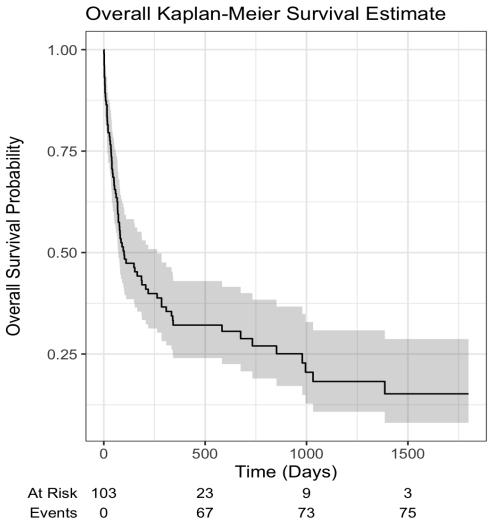
1.1.Dataset Overview:

Variable	Description
id	Unique id to identify patient
acceptyear	The year the patient was accepted for transplantation.
age	The age of the patient at the time of transplantation.
survived	A factor indicating whether the patient survived post-transplant
survtime	The time (in months) the patient survived after the transplant
prior	A factor indicating whether the patient had prior heart conditions
transplant	A factor indicating the type of transplant. Treatment group: control(no transplant) or transplant
wait	Waiting time (days) before treatment or transplant

1.2.Data Preprocessing Steps:

1. **Missing Data Handling:** The variable wait contained 34 missing values, which were imputed using the median wait time computed from available data. No other variables had missing values.
2. **Data Transformation:** Categorical variables (survived and prior) were encoded to numeric to enable use in survival models. Survival time was also converted to numeric.
3. **Summary Statistics:** Of the 103 patients, 75 (approximately 73%) died during the follow-up period. The median survival time for all patients was 90 days, indicating that half of the patients survived less than this duration.
4. **Initial Observation**: The control group (no transplant) has a higher death rate and much shorter median survival compared to the treatment (transplant) group.Treatment patients tend to be older and have longer wait times before transplant.

2.Survival Analysis

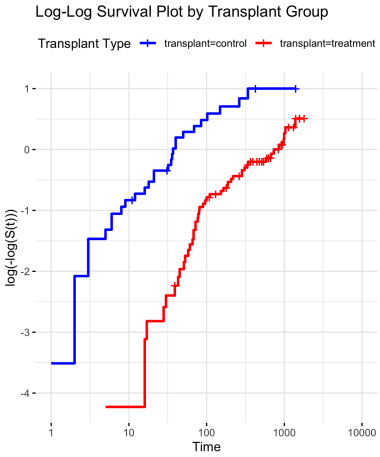


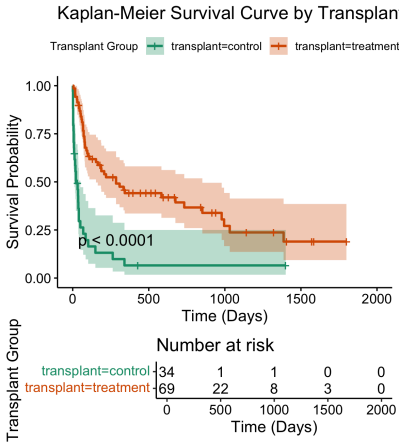
2.1.Kaplan-Meier Survival Analysis:

The overall Kaplan-Meier curve shows steady decline in survival over time. Initially, all 103 patients were at risk, but the survival probability dropped sharply within the first 500 days. By around 1500 days, survival probability falls below 25%.This indicates a high risk of death over time.

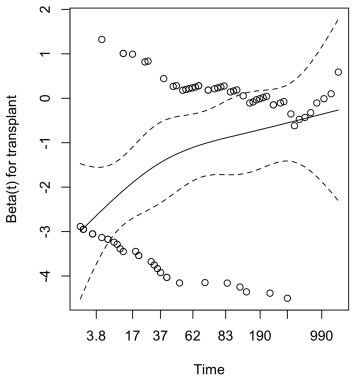
2.2. Survival by Transplant Group:

Kaplan-Meier survival curve stratified by transplant group: compare survival between patients who got a transplant vs those who didn't.The Kaplan-Meier plot shows that patients who received transplants generally survive longer.The log-rank test p-value < 0.001 indicates significant difference in survival and confirms that the survival difference is statistically significant.The log-log survival plots was approximately parallel and suggested the proportional hazards assumption was approximately met for the transplant groups.The death rate was higher in the control group (88%) compared to the treatment group (65%). Median survival was much shorter for control patients (21 days) than those receiving treatment (207 days).Patients receiving the transplant treatment tended to survive longer and had a lower death rate than controls, despite being older and having longer wait times.



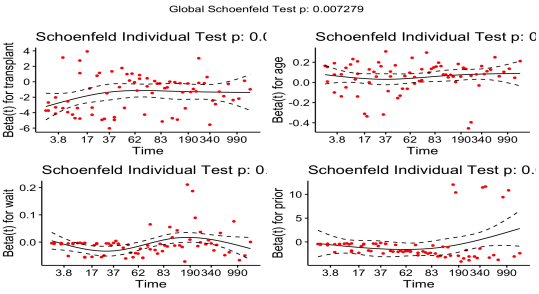


3. Cox Proportional Hazards Models



3.1 Simple Model (transplant only):

Hazard ratio (HR) for transplant vs control: **0.27.**Receiving a transplant reduces hazard of death by about 73%.The low p-value (0.0026) suggests violation of the proportional hazards assumption for the transplant variable.

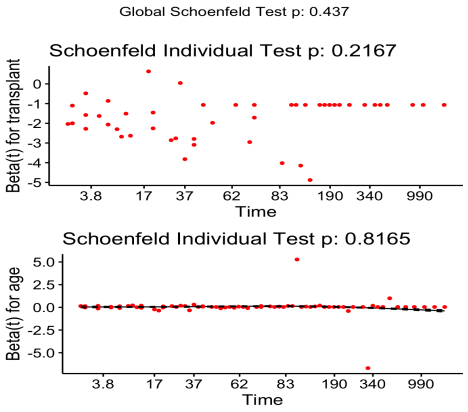


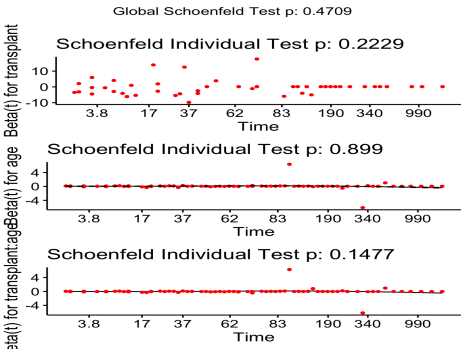
3.2 Multivariable Cox Model:

adjusting for age, wait time, prior condition to study how multiple variables together affect survival. After adjustment, the transplant effect remained protective with similar HR. Age was significantly associated with increased hazard (HR > 1). PH assumption violated ($p < 0.05$) for some variables. Hence, stratify by violating variables. Here, transplant, wait, prior show violation.

3.3 Stratified Cox Model:

Stratified Cox model adjusts for non-proportional hazards in prior and wait.Having a heart transplant lowers the risk of death, but age increases hazard.Age remains significant (HR=1.04 per year, p=0.01).



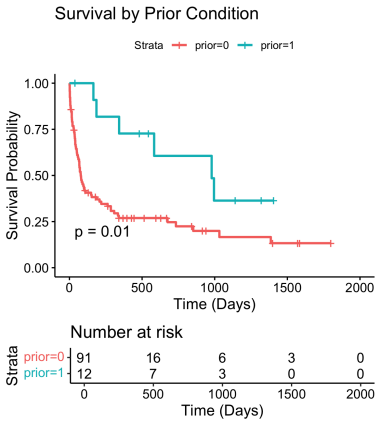


3.4 Interaction Model (Transplant * Age):

The interaction term is not statistically significant (p = 0.49), suggesting no evidence that the transplant effect differs by age. Age was significantly associated with survival (HR = 1.049, p = 0.015), indicating higher risk of death with increasing age.

4. Fit Kaplan-Meier curves by *prior* condition:

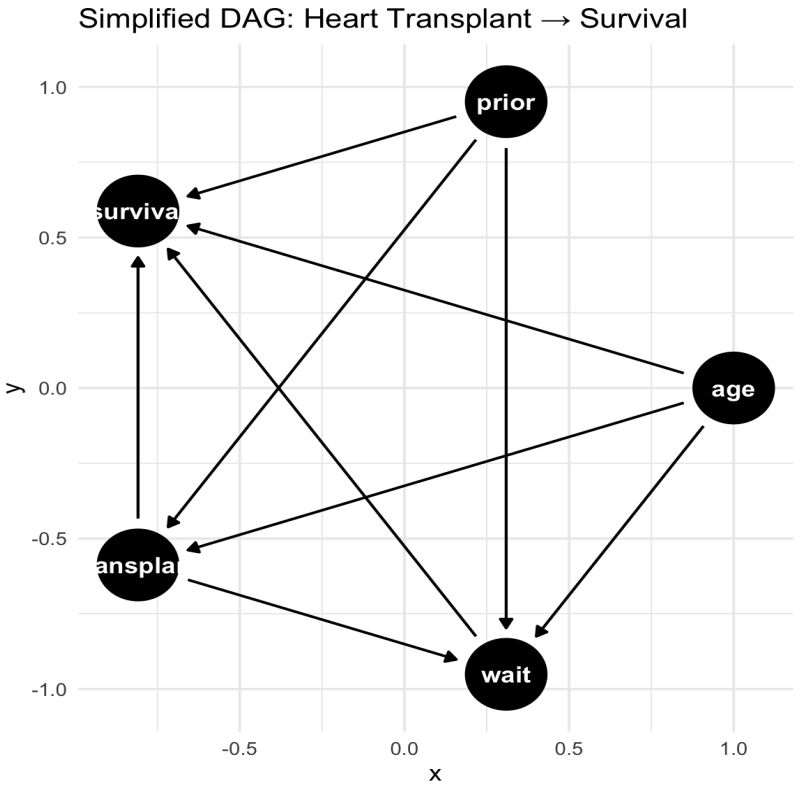
Kaplan-Meier curves by prior health condition showed significantly different survival outcomes (p = 0.01), highlighting prior condition as an important prognostic factor.



5.Compare models by AIC:

Model	df	AIC value
cox_simple	1	572.2967
cox_multi	4	552.8145
cox_strat	2	**194.0033**
cox_interaction	3	195.5448

The **Stratified Cox model** has the lowest AIC and seems like the best model.



6.Interpretation and Conclusions

- Heart transplant significantly improves survival compared to no transplant.
- Older patients have a higher risk of death.
- Violation of PH assumption in transplant variable required stratified modeling.
- Wait time and prior condition do not show significant adjusted effects but impact PH assumption.
- No significant interaction between transplant and age.
- Clinical implication: transplant is strongly beneficial for survival despite older age and longer wait times.

7. Limitations:

- Sample size limited to 103 patients.
- Some covariates show PH assumption violations, complicating model interpretation.