BIOST 537 Final Project: Group 4

Introduction

Chemotherapy drugs and radiation used in the treatment of acute leukemia damages bone marrow responsible for the creation of new blood cells. Bone marrow transplant is often used to restore stem cell levels in patients after treatment, and successful recovery from bone marrow transplant is dependent on a range of factors present both before the transplant and during the postoperative period. Understanding the impact of these factors on survival and recovery is key to improving patient outcomes. Previous studies have suggested that in allogeneic transplant patient's CMV status may be associated with patient outcomes, while evidence for an association with the donor's CMV status has been inconclusive. To better understand factors associated with patient outcomes, we will be analyzing data from a study of 137 patients enrolled between March 1, 1984 and June 30, 1989 and followed until death or the end of the study period.

The aim of this exploratory analysis is to understand factors related to patient outcomes and differences in baseline factors between different patient groups. We investigate if patients in different disease groups and FAB classifications differ from each other with respect to baseline characteristics and if differences in any of these baseline characteristics are associated with disease-free survival. Post operative factors may also have an impact on patient survival, here we investigate if aGVHD occurrence is associated with disease-free survival and if there are any baseline characteristics associated with survival among those who develop aGVHD. We also investigate if prophylactic use of methotrexate is associated with the development of aGVHD as well as the relationship between the recovery of normal platelet levels with disease-free survival and risk of relapse.

Methods

In our analysis we define disease-free survival as being the length of time an individual is observed to not die or relapse. When examining relapse as the event of interest, we consider a death event to be censoring. We accounted for left truncation when testing associations with baseline variables (delayed entry informed by wait time until transplant) and used an alpha level of .10 to test for significance because our sample size was very small. We did not adjust for left truncation for time-varying covariates as information on their levels were not known until after the operation event. For investigations into associations for baseline and time-varying covariates we utilized Cox Proportional-Hazard Models to best adjust for multiple quantitative covariates. It was observed that methotrexate usage was completely determined by recruitment hospital, with two hospitals giving all patients methotrexate and the two other hospitals giving no patients methotrexate. Therefore, in our models that contain baseline covariates we used hospital site and excluded methotrexate to avoid multicollinearity. Methotrexate usage is however investigated in a subsequent analysis.

To summarize the distribution of disease-free survival time we estimated the survival function with multiple distributions and visually inspected them for the best fit compared with a nonparametric

estimate via a Kaplan-Meier (KM) curve, while adjusting for delayed entry as wait time until transplant. None of the parametric models fit the KM curve well, so we reported a non-parametric estimate of median survival time.

To understand differences in baseline characteristics between those in different disease groups and FAB classification, we generated descriptive statistics for these subgroups to compare values of baseline variables.To investigate the association between baseline variables and disease-free survival we fit a Cox Proportional-Hazards (PH) model while adjusting for left truncation and reported p values and 95% confidence intervals for variables with significant estimates at the alpha=0.10 level.

We fit a Cox PH model using aGVHD as a time-varying covariate to test for associations between developing aGVHD with both disease-free survival and risk of relapse while controlling for disease group, fab classification, patient age and sex, patient and donor CMV status as well as hospital. We then subset our data to include only those who had developed aGVHD and fit a Cox PH model to test for associations between baseline variables and disease-free survival and reported p values and 95% confidence intervals for variables with significant estimates in the same way as was done for the analysis of baseline associations for the whole population.

To investigate if prophylactic use of methotrexate is associated with an increased or decreased risk of developing aGVHD we created an unadjusted nonparametric estimate for aGVHD-free survival probability stratified by methotrexate usage. We fit a Cox PH model to investigate this association, controlling for patient sex, patient CMV status, donor CMV status, disease classification and fab group. Since neither group reached 50% survival and all patients who develop aGVHD do so within the first 100 days after operation we could not report a median survival time.

To investigate the association between recovery of normal platelet levels with improved disease-free survival and decreased risk of relapse, we fit Cox PH models with return to normal platelet levels as a time-varying covariate. We reported estimates of the coefficient for recovery of platelet levels in both models with p values and 95% confidence intervals.

Results and Discussion

Disease-free survival-time distribution

We estimated the median disease-free survival time for patients enrolled in this study to be 510 days (95% CI: 318, 731) for the nonparametric estimator, when accounting for delayed entry. This is fairly early on in the study. None of the parametric estimators fit the data well, so median survival time was only estimated non-parametrically. **Figure 1** shows estimates of the survival function both nonparametrically with a KM curve and with various parametric models.

Differences in baseline variables between disease groups and FAB classifications

We created descriptive statistical tables to summarize the differences in baseline variable values by disease group (**Table 1**) and FAB classification (**Table 2**); where disease group is defined by 1 = “acute lymphoblastic leukemia”, 2 = Low Risk “acute myelocytic leukemia”, 3 = High Risk “acute myelocytic leukemia” and FAB represents the (French-American-British) classification dichotomized to 1 = FAB grade 4, 5 or acute myelocytic leukemia and 0 = otherwise. It can be observed from these tables that those with acute lymphoblastic leukemia were not recruited at the Hahnemann Hospital and had higher proportions at Alfred and St. Vincent, compared to those with acute myelocytic leukemia or AML high risk, who were treated at Hahnemann. Those with FAB classification of 4 or 5 were observed in higher proportions at Alfred and St. Vincent, compared with those classified otherwise. This indicates that both FAB classification and disease group may be associated with recruitment center (and thus methotrexate exposure), and that an analysis of the association of either recruitment center or methotrexate exposure with survival or relapse time may be confounded by disease group or FAB status if they are associated with survival time or relapse.

Baseline Measures Association with Differences in Disease-Free Survival

Based on a Cox PH model, baseline measurements statistically significantly associated with differences in disease-free survival were FAB subtyping of leukemia, the recruitment center, and type of leukemia adjusting for age, sex, cytomegalovirus status, donor age, donor sex, and donor cytomegalovirus status. The hazard of death or relapse is estimated to be approximately 150.2% larger in patients with a FAB grade of 4 or 5 and AML compared to patients without, when controlling for the other baseline factors. The hazard ratio is estimated to be 2.502 (95% CI: 1.438, 4.351), and the associated p-value is 0.001. The hazard of death or relapse is estimated to be approximately 113.0% larger in patients recruited from Alfred compared to patients recruited from Ohio State University, when controlling for the other baseline factors. The hazard ratio is estimated to be 2.130 (95% CI: 1.034, 4.391), and the associated p-value is 0.040. The hazard of death or relapse is estimated to be approximately 66.2% smaller in patients recruited from Hahnemann compared to patients recruited from Ohio State University, when controlling for the other baseline factors. The resulting hazard ratio is estimated to be 0.338 (95% CI: 0.144, 0.793), and the associated p-value is 0.013. The hazard of death or relapse is estimated to be approximately 64.1% smaller in patients with low-risk AML compared to patients with acute lymphoblastic leukemia, when controlling for the other baseline factors. The resulting hazard ratio is estimated to be 0.359 [95% CI (0.177, 0.728)], and the associated p-value is 0.004.

Association of aGVHD with Disease-Free Survival and Risk of Relapse

Based on a Cox PH model using aGVHD as a time-varying covariate, we estimate that individuals who developed aGVHD by a given time have a 64.9% higher hazard of death or relapse and 15.3% smaller hazard of relapse than those who had not experienced acute aGVHD by that time, adjusting for leukemia type, FAB, age, sex, CMV status, donor age, donor sex, donor CMV status, and recruitment center, however these results are not statistically

significant. The estimated hazard of death or relapse and hazard of relapse comparing individuals who experienced acute aGVHD by a given time to those who had not experienced acute aGVHD by that time are 1.649 (95% CI: 0.906, 3.003; p-value= 0.102) and 0.847 (95% CI: 0.318, 2.260; p-value= 0.741), respectively. At the 0.10 significance level, we do not see evidence that aGVHD is associated with either disease-free survival or risk of relapse. Thus, aGVHD may be an important prognostic event, but we are unable to conclude that from these results.

Baseline Measures Association with Differences in Disease-Free Survival Among Those Who Developed aGVHD

Baseline measurements statistically significantly associated with differences in disease-free survival among those who developed aGVHD were the recruitment center and type of leukemia adjusting for FAB subtyping, age, sex, donor age, donor sex, and donor CMV status. The hazard of death or relapse is estimated to be approximately 20 times higher in patients with low-risk AML compared to patients with acute lymphoblastic leukemia. The resulting hazard ratio is estimated at 20.086 (95% CI: 1.280, 315.155), and the associated p-value is 0.033. The hazard of death or relapse is estimated to be approximately 31 times higher in patients with high-risk AML compared to patients with acute lymphoblastic leukemia. The resulting hazard ratio is estimated at 31.027 (95% CI: 0.927, 1038.017), and the associated p-value is 0.055.The hazard of death or relapse is estimated to be approximately 98.9% smaller in patients recruited from Hahnemann compared to patients recruited from Ohio State University. The resulting hazard ratio is estimated at 0.011 (95% CI: 0.001, 0.200), and the associated p-value is 0.002. However, these results should be interpreted with caution, as there were only 26 patients who developed aGVHD and thus it is difficult to confirm these associations with such a small sample.

Prophylactic Use of Methotrexate Association with Risk of Developing aGVHD

Table 3 displays descriptive statistics summarizing the differences in baseline variables by prophylactic use of methotrexate. We observe differing distributions of sex, positive CMV status, and donor-positive CMV status across those who did and did not experience prophylactic use of methotrexate. This is an indicator to include these variables in statistical analysis, as they differ between the two groups we are comparing. Additionally, CMV status, donor-CMV status FAB classification, and disease group may impact both risk of aGVHD and whether methotrexate is used as prophylaxis, so they were included as potential confounders. **Figure 2** depicts the nonparametric estimates of aGVHD-free survival probability by methotrexate usage. Here we can see a slight difference in survival time and methotrexate use, but note that the curves do not reach 50% probability. The hazard of developing aGVHD is estimated to be approximately 38.6% smaller in patients with methotrexate usage compared to patients without, adjusting for sex, CMV status, donor CMV status, FAB, and type of leukemia, however this result is not significant. The resulting hazard ratio is estimated at 0.614 (95% CI: 0.230, 1.645). With a p-value of 0.332, we fail to reject the null hypothesis that the association does not differ by methotrexate usage at the 0.10 significance level. Thus, there is no evidence that the hazard of developing aGVHD varies by methotrexate use, adjusting for sex, CMV status, donor CMV

status, FAB, and type of leukemia. There were very few patients (26) that developed aGVHD, and all who developed aGVHD did so very early in the study.

Association of Recovery of Normal Platelet Levels with Disease-Free Survival

Using a Cox PH model with recovery of normal platelets as a time-varying covariate, we estimate that individuals who return to normal platelet counts by a given time have a 62.8% smaller hazard of death or relapse than those who did not return to normal platelet counts by that time, adjusting for FAB subtyping, age, sex, donor age, donor sex, recruitment center, type of leukemia, and donor CMV status. The estimated hazard of death or relapse comparing individuals who return to normal platelet counts by a given time to those who did not return to normal platelet counts by that time is 0.372 (95% CI: 0.186, 0.747; p-value= 0.005). At the 0.10 significance level, we see evidence that recovery of normal platelet levels and disease-free survival are associated, indicating that recovery of normal platelets may be an important prognostic factor. Additionally, we estimate that individuals who return to normal platelet counts by a given time have a 9.6% higher hazard of relapse than those who did not return to normal platelet counts by that time, adjusting for FAB subtyping, age, sex, donor age, donor sex, recruitment center, type of leukemia, and donor CMV status, however this result is not significant. The estimated hazard of relapse comparing individuals who return to normal platelet counts by a given time to those who did not return to normal platelet counts by that time is 1.097 (95% CI: 0.280, 4.294; p-value= 0.895). At the 0.10 significance level, we did not see evidence that recovery of normal platelet levels and relapse risk are associated.

Discussion

The results of our Kaplan Meier Curves and Cox Proportional Hazard models indicate that there are factors that could be used for informing patient prognosis. We found that baseline FAB leukemia subtyping, recruitment center, and type of leukemia were associated with disease-free survival. Of these significant findings, FAB leukemia subtyping and type of leukemia should be considered when predicting the likely course of leukemia based on our analysis. We suspect the differences observed at the recruitment sites may be due to the severity of leukemia. Different countries may have differing treatment guidelines, which may explain why the two facilities in Australia (Alfred and St. Vincent) prescribed methotrexate for all their patients in this study, while the two in America (The Ohio State University and Hahnemann) did prescribe methotrexate for any of their transplant patients in this study. This leads to some difficulty in studying the association between methotrexate and disease-free survival, as there may be some other systematic differences (observed or unobserved) in the patients that are treated at each of these hospitals, such as how they are referred to each hospital.

While we found no association between aGVHD and disease-free survival and hazard of relapse, we observed a higher hazard of death or relapse and a lower hazard of relapse for those who developed aGVHD, which seems like a contradiction; however, the likelihood of death when a patient's body is attacked by the donor's stem cells is high, explaining why we would observe a lower hazard of relapse; those who die cannot relapse. We conclude that developing aGVHD and prophylactic use of methotrexate are not associated. We found no

association between recovery of normal platelet levels and disease-free survival. Although there was a higher hazard of relapse among those who recovered their normal platelet counts, we suspect this is because patients in better health live longer and, by default, have a higher chance of relapsing. These associations are also limited by the small sample size, we may have observed a more significant difference if the study had higher power to detect such a difference.

A considerable limitation of these analyses is the small sample size. The small sample size can also explain the discrepancies in some of the computed estimates. For all the factors explored in this analysis, we propose further investigations with a larger sample size to comprehend the factors better. Another limitation to consider is independent left truncation, as leukemia can be deadly for many; left truncation existing within our data and can result in selection bias and an exaggeration of survival time - patients who died from their leukemia before being eligible for a transplant would not have been included. When we could, we adjusted for left truncation using wait time until transplant, however this still would not include patients who died prior to transplant.

Tables and Figures

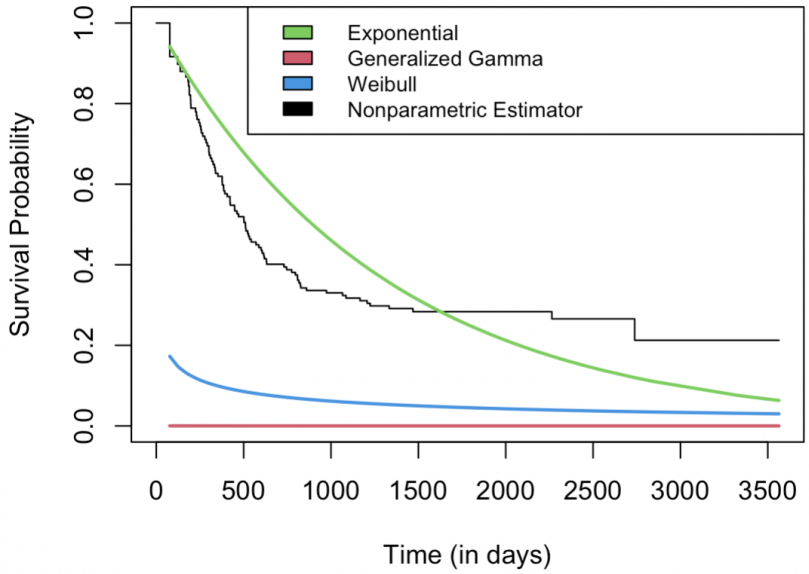

Figure 1. Parametric and nonparametric estimates of the survival function for disease-free survival. It can be observed that none of the parametric estimates have a good fit to the data relative to the nonparametric estimate given by a Km curve.

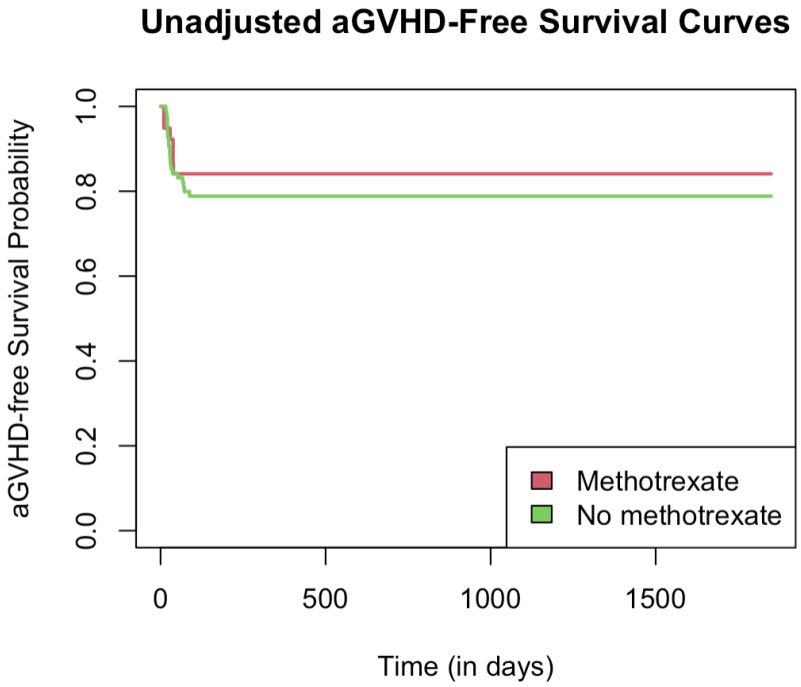

Figure 2. Nonparametric estimates of aGVHD-free survival probability stratified by Methotrexate usage. It can be observed that when no other variables are adjusted for methotrexate exposure significantly impacts survival time.

Table 1: Baseline characteristics by disease group; 1 = acute lymphoblastic leukemia, 2 = acute myelocytic leukemia (AML), 3 = AML high risk

Patient Male (prop)	0.68	0.56	0.53
CMV Positive (prop)	0.39	0.48	0.60
Patient Age (mean)	24.42	29.41	30.44
Donor Age (mean)	26.79	28.07	29.93
Donor Male (prop)	0.68	0.63	0.62
Donor CMV Positive (prop)	0.45	0.41	0.42
Hospital 1 (prop)	0.55	0.50	0.62
Hospital 2 (prop)	0.21	0.09	0.09
Hospital 3 (prop)	0.24	0.13	0.16
Hospital 4 (prop)	0.00	0.28	0.13

Table 2: Baseline characteristics by FAB group; 1 = FAB grade 4 or 5, 0 = Otherwise

FAB Group	0.00	1.00
Count	92.00	45.00
Patient Male (prop)	0.61	0.53
CMV Positive (prop)	0.48	0.53
Patient Age (mean)	28.60	27.89
Donor Age (mean)	29.00	26.96
Donor Male (prop)	0.63	0.67
Donor CMV Positive (prop)	0.48	0.31
Hospital 1 (prop)	0.52	0.62
Hospital 2 (prop)	0.15	0.07
Hospital 3 (prop)	0.20	0.11
Hospital 4 (prop)	0.13	0.20

Table 3: Baseline characteristics by prophylactic use of methotrexate; 1 = Yes, 0 = No

Methotrexate Use	0.00	1.00
Count	97.00	40.00
Proportion Male	0.62	0.50
Proportion CMV Positive	0.42	0.68
Mean Age	27.10	31.42
Mean Donor Age	27.10	31.30
Proportion Male Donor	0.63	0.68
Proportion Donor CMV Positive	0.38	0.52