# BST 210 Project: Survival Analysis

Lauren Mock

11/30/2021

```
library(tidyverse)
library(survival)
library(survminer)
library(pander)
```

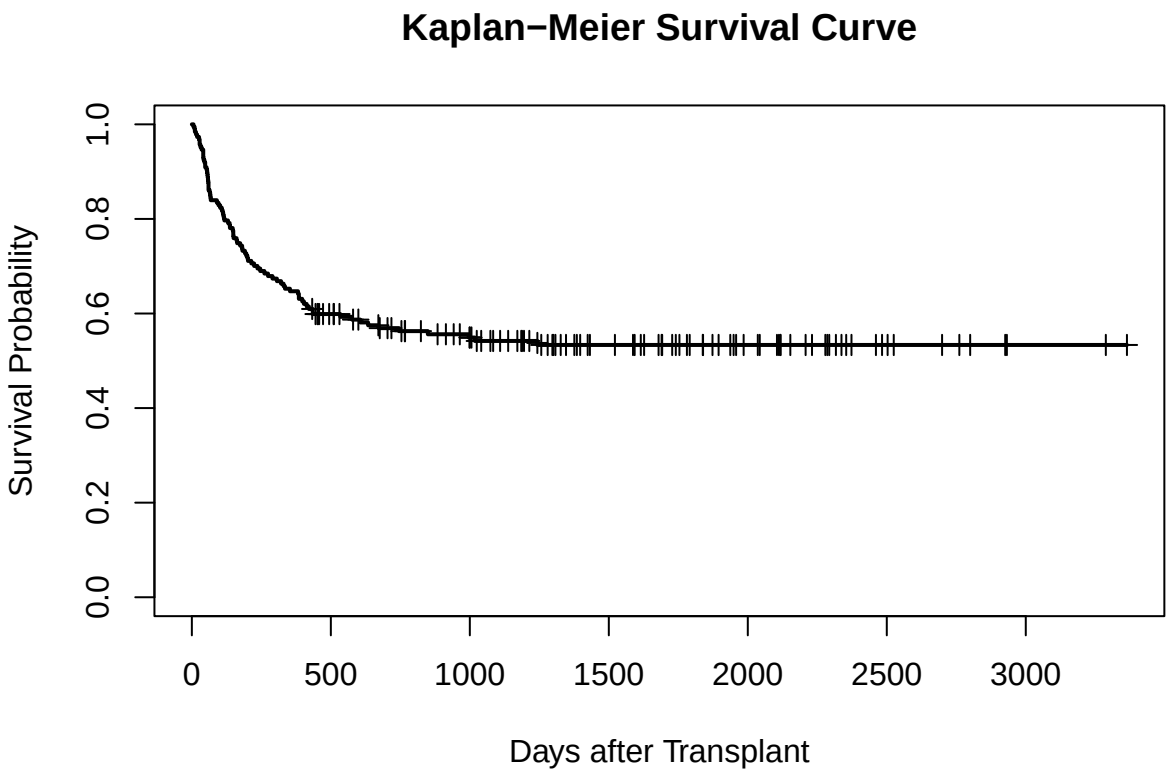
## Stem Cell Source

### Read in Data

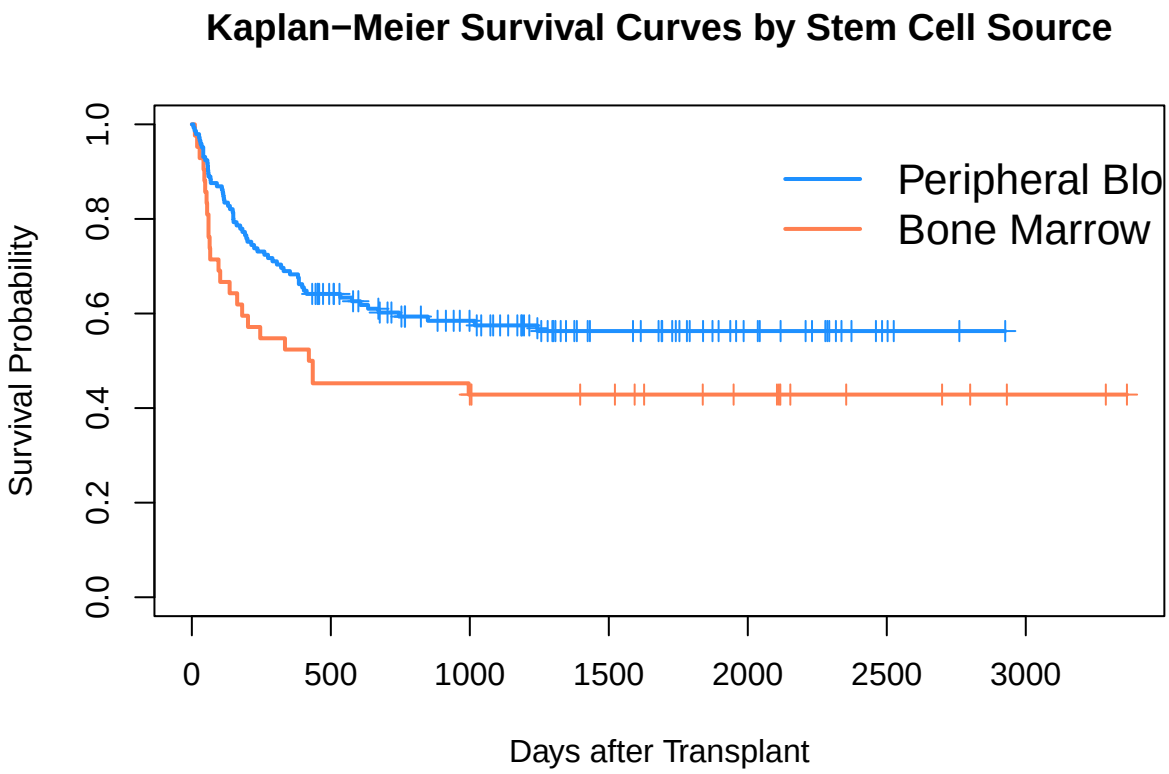
```
bone <- read.csv("data.csv")

# remove HLA_match_raw (was already converted into a different column)
bone <- select(bone, -HLA_match_raw)


# convert numeric columns to numeric
bone_num <- c("donor_age", "recipient_age", "recipient_body_mass", "CD34_x1e6_per_kg...CD34kgx10d6",
              "CD3_x1e8_per_kg", "CD3_to_CD34_ratio", "ANC_recovery", "PLT_recovery",
              "time_to_acute_GvHD_III_IV", "survival_time", "survival_status")

bone[,bone_num] <- lapply(bone_num, function(x) as.numeric(bone[[x]]))
```

### Compare Kaplan-Meier curves for the two stem cell sources

```
# survival object
survival.obj <- Surv(time = bone$survival_time, event = bone$survival_status)

# overall KM
KM.fit1 <- survfit(survival.obj ~ 1, data = bone)

# plot
plot(KM.fit1, xlab = "Days after Transplant", ylab = "Survival Probability", conf.int = FALSE,
mark.time = TRUE, main = "Kaplan-Meier Survival Curve", lwd = 2)
```




```
# KM by treatment
KM.fit2 <- survfit(survival.obj ~ stem_cell_source, data = bone)

# plot
plot(KM.fit2, xlab = "Days after Transplant", ylab = "Survival Probability",
     mark.time = TRUE, conf.int=, col = c("coral", "dodgerblue"),
     main = "Kaplan-Meier Survival Curves by Stem Cell Source", lwd = 2)
legend(x = 2000, y = 0.99,
       legend = c("Peripheral Blood", "Bone Marrow"),
       col = c("dodgerblue", "coral"),
       bty = "n",
       lty = 1:1,
       lwd = 2,
       cex = 1.3)
```




**note: 5 people (out of 187) missing CD3**

**log-rank test to compare these two curves**

```
survdiff(Surv(bone$survival_time, bone$survival_status) ~ stem_cell_source, data=bone)
```

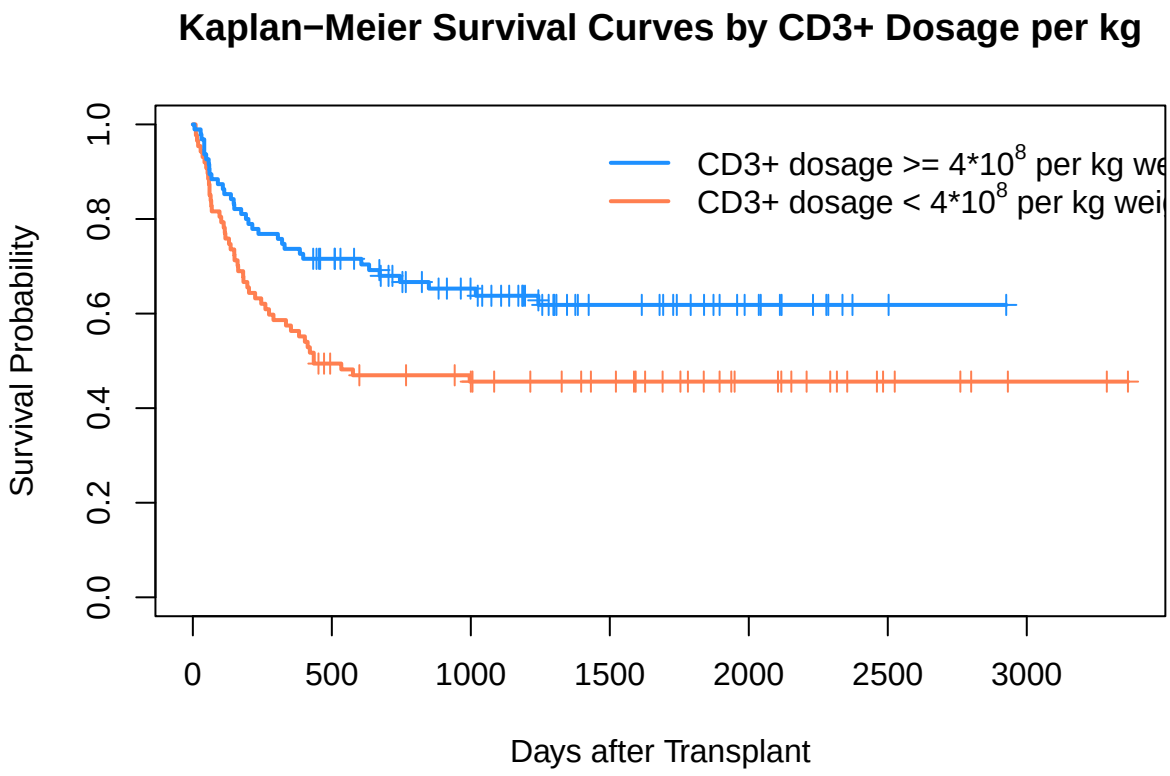
```
## Call:
## survdiff(formula = Surv(bone$survival_time, bone$survival_status) ~
##     stem_cell_source, data = bone)
##
##                                   N Observed Expected (O-E)^2/E (O-E)^2/V
## stem_cell_source=bone_marrow       42       24     16.8     3.065      3.83
## stem_cell_source=peripheral_blood 145       61     68.2     0.756      3.83
##
##  Chisq= 3.8  on 1 degrees of freedom, p= 0.05
```

p-value is 0.05–almost statistically significantly different

## CD3+ dosage

### Compare Kaplan-Meier curves for high vs. low dosage

```
# add new binary columns based on previous research
bone$CD3_over_4 <- ifelse(bone$CD3_x1e8_per_kg >= 4, 1, 0)
bone$CD34_over10 <- ifelse(bone$CD34_x1e6_per_kg...CD34kgx10d6 >= 10, 1, 0)

# model fit
KM.fit4 <- survfit(survival.obj ~ CD3_over_4, data = bone)

# plot
plot(KM.fit4, xlab = "Days after Transplant", ylab = "Survival Probability", mark.time = TRUE, conf.int=
     col = c("coral", "dodgerblue"), lwd = 2,
     main = "Kaplan-Meier Survival Curves by CD3+ Dosage per kg")
legend(x = 1400, y = 1,
       legend = c(expression(paste("CD3+ dosage >= 4*", 10^8, " per kg weight")),
                  expression(paste("CD3+ dosage < 4*", 10^8, " per kg weight"))),
       col = c("dodgerblue", "coral"),
       bty = "n",
       lty = 1:1,
       lwd = 2,
       cex = 1)
```



**log-rank test to compare these two curves**

```
survdiff(Surv(bone$survival_time, bone$survival_status) ~ CD3_over_4, data=bone)
```

```
## Call:
## survdiff(formula = Surv(bone$survival_time, bone$survival_status) ~
##     CD3_over_4, data = bone)
##
## n=182, 5 observations deleted due to missingness.
##
##               N Observed Expected (O-E)^2/E (O-E)^2/V
## CD3_over_4=0 87       47     35.8      3.49      6.28
## CD3_over_4=1 95       34     45.2      2.76      6.28
##
##  Chisq= 6.3  on 1 degrees of freedom, p= 0.01
```

highly significant difference! (p-value 0.01)

# Cox Proportional Hazards Model

## unadjusted

```
cox_unadj <- coxph(survival.obj ~ as.factor(stem_cell_source) + CD3_x1e8_per_kg,
                       data = bone, ties = "exact")
summary(cox_unadj)
```

```
## Call:
## coxph(formula = survival.obj ~ as.factor(stem_cell_source) +
##     CD3_x1e8_per_kg, data = bone, ties = "exact")
##
##   n= 182, number of events= 81
##    (5 observations deleted due to missingness)
##
##                                                 coef exp(coef) se(coef)      z
## as.factor(stem_cell_source)peripheral_blood -0.04717   0.95393  0.30509 -0.155
## CD3_x1e8_per_kg                             -0.10039   0.90449  0.04215 -2.382
##                                             Pr(>|z|)
## as.factor(stem_cell_source)peripheral_blood   0.8771
## CD3_x1e8_per_kg                               0.0172 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                                             exp(coef) exp(-coef) lower .95
## as.factor(stem_cell_source)peripheral_blood    0.9539      1.048    0.5246
## CD3_x1e8_per_kg                                0.9045      1.106    0.8328
##                                             upper .95
## as.factor(stem_cell_source)peripheral_blood    1.7346
## CD3_x1e8_per_kg                                0.9824
##
```

```
## Concordance= 0.606  (se = 0.033 )
## Likelihood ratio test= 10.56  on 2 df,   p=0.005
## Wald test            = 9.36  on 2 df,   p=0.009
## Score (logrank) test = 9.7  on 2 df,   p=0.008
```

## Confounders?

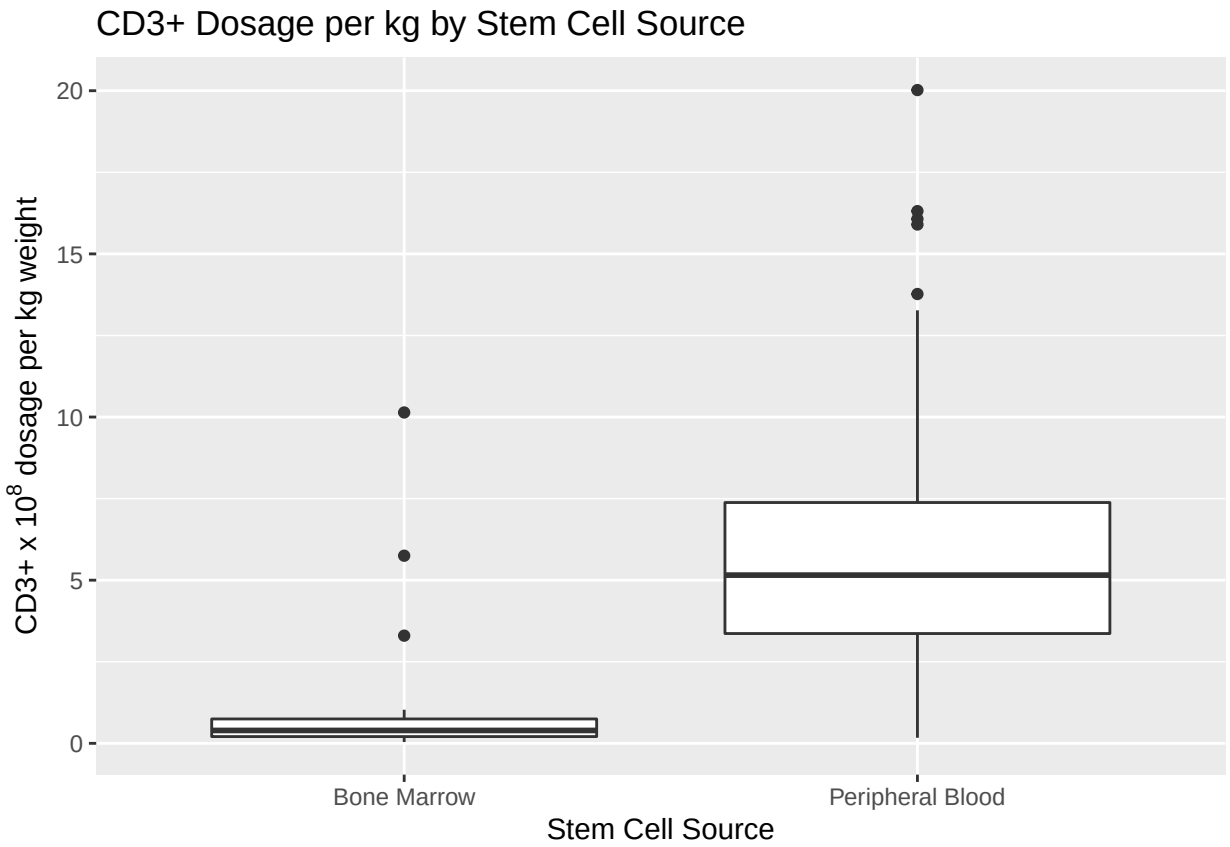
**looking for common causes of stem cell type and survival or common causes of dosage and survival**

did the authors adjust for confounders? no, it was an RCT

possible confounders - always must adjust for age when adjusting for dosage - disease type (Anasetti article says that this could be an effect modifier by interacting with stem cell source)

**look at dosage distribution by source**

```
bone %>% ggplot(aes(x = stem_cell_source, y = CD3_x1e8_per_kg)) +
        geom_boxplot() +
        ggtitle("CD3+ Dosage per kg by Stem Cell Source") +
        xlab("Stem Cell Source") +
        ylab(expression(paste("CD3+ x ", 10^8, " dosage per kg weight"))) +
        scale_x_discrete(labels = c("Bone Marrow", "Peripheral Blood"))
```



- people who get bone marrow always get a much smaller dose
- this could definitely be a confounder, so we'll keep both stem cell source and dose in the model
- interpretation will be "given that someone received this source, what was the effect of the dosage?"

**look at dosage distribution by age**

```
bone %>% ggplot(aes(x = recipient_age_below_10, y = CD3_x1e8_per_kg)) +
        geom_boxplot() +
        ggtitle("CD3+ Dosage per kg by Age") +
        xlab("Stem Cell Source") +
        ylab(expression(paste("CD3+ x ", 10^8, " dosage per kg weight"))) +
        scale_x_discrete(labels = c("Older than 10", "Younger than 10"))
```

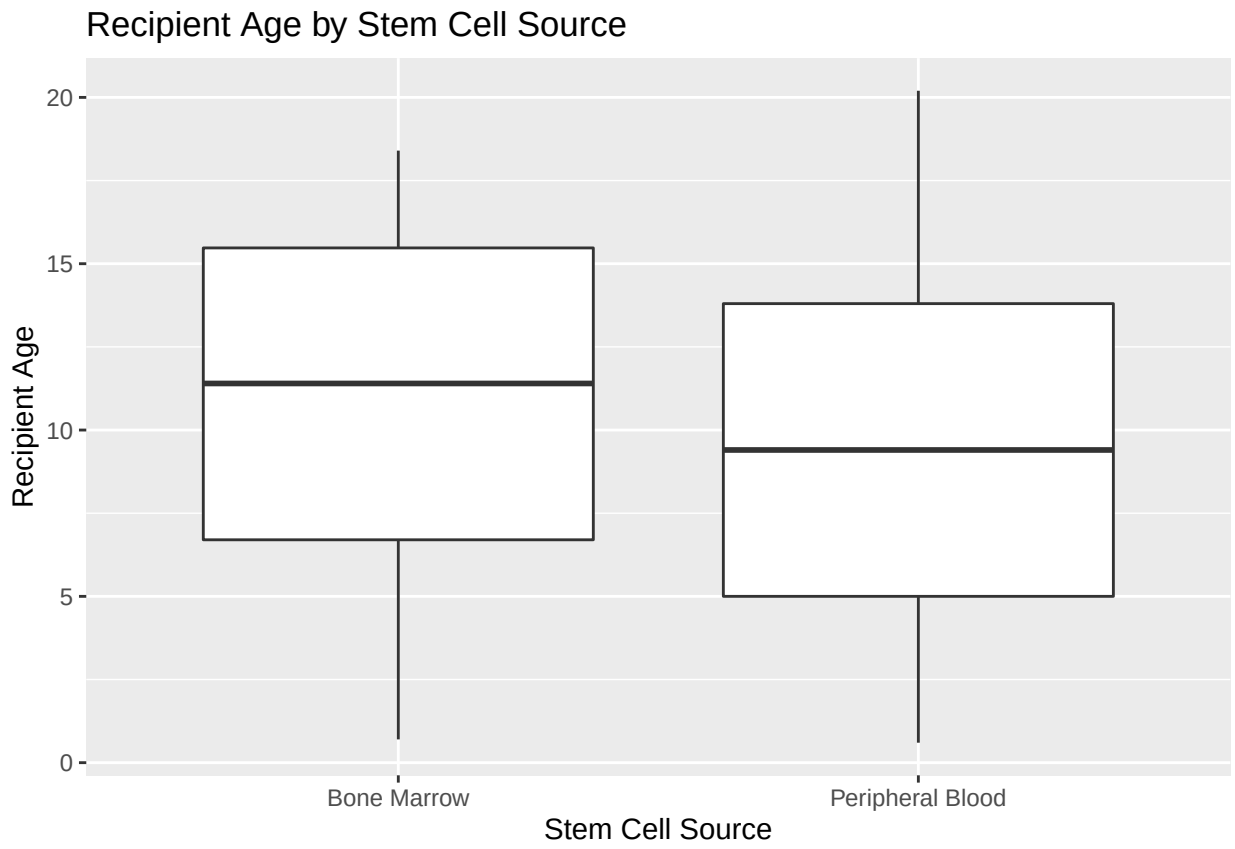
```
## Warning: Removed 5 rows containing non-finite values (stat_boxplot).
```


younger people are more likely to get higher doses (per kg body weight), which makes sense because they are also more likely to get peripheral blood (given in higher doses than bone marrow)

**look at source by recipient age**

```
bone %>% ggplot(aes(x = stem_cell_source, y = recipient_age)) +
        geom_boxplot() +
```

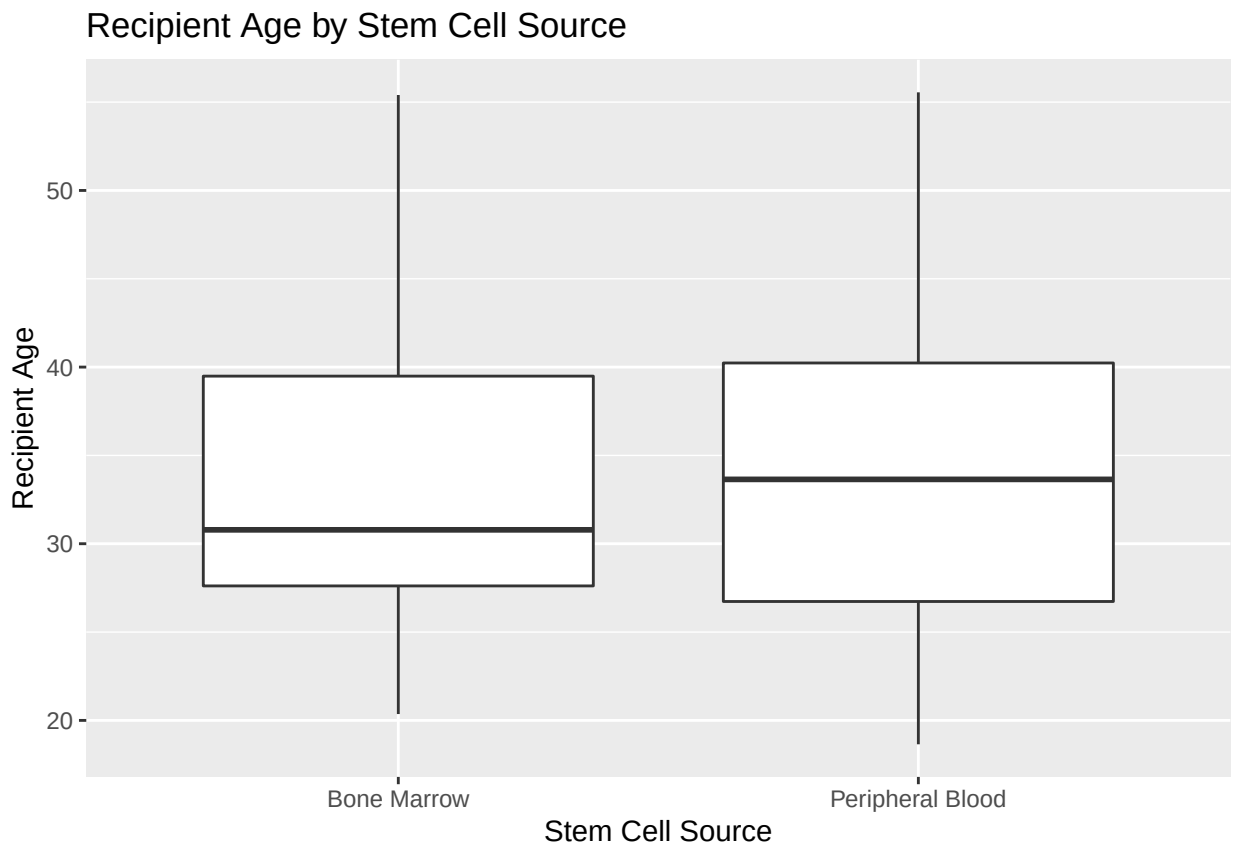
```
        ggtitle("Recipient Age by Stem Cell Source") +
        xlab("Stem Cell Source") +
        ylab("Recipient Age") +
        scale_x_discrete(labels = c("Bone Marrow", "Peripheral Blood"))
```




reasonably similar age distributions

**look at source by donor age**

```
bone %>% ggplot(aes(x = stem_cell_source, y = donor_age)) +
        geom_boxplot() +
        ggtitle("Recipient Age by Stem Cell Source") +
        xlab("Stem Cell Source") +
        ylab("Recipient Age") +
        scale_x_discrete(labels = c("Bone Marrow", "Peripheral Blood"))
```

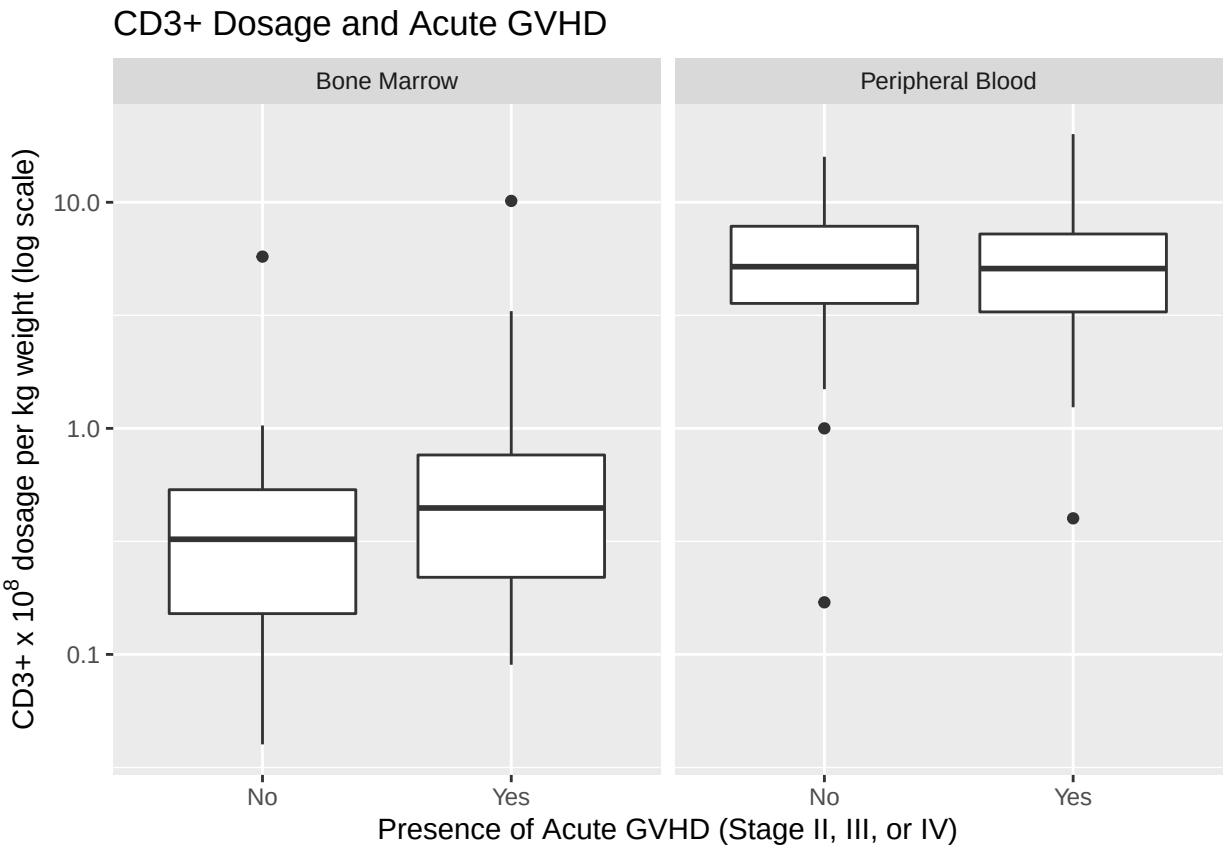



reasonably similar age distributions

**GVHD by dose**

```
bone %>% ggplot(aes(x = acute_GvHD_II_III_IV, y = CD3_x1e8_per_kg)) +
        geom_boxplot() +
        ggtitle("CD3+ Dosage and Acute GVHD") +
        xlab("Presence of Acute GVHD (Stage II, III, or IV)") +
        ylab(expression(paste("CD3+ x ", 10^8, " dosage per kg weight (log scale)"))) +
        scale_x_discrete(labels = c("No", "Yes")) +
        scale_y_log10() +
        facet_grid( ~ stem_cell_source,
                    labeller = as_labeller(c(`bone_marrow` = "Bone Marrow",
                                             `peripheral_blood` = "Peripheral Blood")))
```

```
## Warning: Removed 5 rows containing non-finite values (stat_boxplot).
```




no apparent relationship between high dose of CD3+ and GVHD for either stem cell source

## confounders

### adjusted model 1

```
cox_adj1 <- coxph(survival.obj ~ as.factor(stem_cell_source) + CD3_x1e8_per_kg + recipient_age,
                  data = bone, ties = "exact")
summary(cox_adj1)
```

```
## Call:
## coxph(formula = survival.obj ~ as.factor(stem_cell_source) +
##     CD3_x1e8_per_kg + recipient_age, data = bone, ties = "exact")
##
##   n= 182, number of events= 81
##    (5 observations deleted due to missingness)
##
##                                              coef exp(coef) se(coef)      z
## as.factor(stem_cell_source)peripheral_blood -0.18045   0.83489  0.30931 -0.583
## CD3_x1e8_per_kg                             -0.07077   0.93168  0.04466 -1.584
## recipient_age                                0.04226   1.04317  0.02421  1.745
##                                             Pr(>|z|)
## as.factor(stem_cell_source)peripheral_blood   0.5596
## CD3_x1e8_per_kg                               0.1131
```

```
## recipient_age                                    0.0809 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                                             exp(coef) exp(-coef) lower .95
## as.factor(stem_cell_source)peripheral_blood    0.8349     1.1978    0.4553
## CD3_x1e8_per_kg                                0.9317     1.0733    0.8536
## recipient_age                                  1.0432     0.9586    0.9948
##                                             upper .95
## as.factor(stem_cell_source)peripheral_blood     1.531
## CD3_x1e8_per_kg                                 1.017
## recipient_age                                   1.094
##
## Concordance= 0.615  (se = 0.034 )
## Likelihood ratio test= 13.61  on 3 df,   p=0.003
## Wald test            = 12.44  on 3 df,   p=0.006
## Score (logrank) test = 12.9  on 3 df,   p=0.005
```

adding age changed the stem cell source coefficient by $> 10\%$

**adjusted model 2**

check interaction between stem cell source and disease type (Anasetti et al.)

```
cox_adj2 <- coxph(survival.obj ~ as.factor(stem_cell_source):as.factor(disease),
                  data = bone, ties = "exact")
summary(cox_adj2)
```

```
## Call:
## coxph(formula = survival.obj ~ as.factor(stem_cell_source):as.factor(disease),
##     data = bone, ties = "exact")
##
##   n= 187, number of events= 85
##
##                                                                        coef
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)ALL          0.7741
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)ALL     0.4003
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)AML          1.6786
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)AML     0.2383
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)chronic     -0.0497
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)chronic 0.5672
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)lymphoma     1.8165
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)lymphoma 1.8929
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)nonmalignant 1.7978
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)nonmalignant     NA
##                                                                     exp(coef)
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)ALL           2.1686
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)ALL      1.4923
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)AML           5.3583
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)AML      1.2691
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)chronic       0.9515
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)chronic  1.7633
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)lymphoma      6.1500
```

```
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)lymphoma        6.6386
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)nonmalignant         6.0363
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)nonmalignant        NA
##                                                                             se(coef)
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)ALL                  0.4748
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)ALL             0.4188
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)AML                  0.5728
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)AML             0.4747
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)chronic              0.6774
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)chronic         0.4335
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)lymphoma              0.7968
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)lymphoma        0.5213
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)nonmalignant         0.6158
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)nonmalignant    0.0000
##                                                                                    z
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)ALL                   1.630
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)ALL              0.956
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)AML                   2.930
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)AML              0.502
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)chronic              -0.073
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)chronic          1.309
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)lymphoma              2.280
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)lymphoma         3.631
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)nonmalignant          2.919
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)nonmalignant        NA
##                                                                             Pr(>|z|)
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)ALL                0.103045
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)ALL           0.339140
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)AML                0.003386
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)AML           0.615644
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)chronic            0.941514
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)chronic       0.190692
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)lymphoma           0.022633
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)lymphoma      0.000282
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)nonmalignant       0.003506
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)nonmalignant       NA
##
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)ALL
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)ALL
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)AML                 **
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)AML
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)chronic
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)chronic
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)lymphoma            *
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)lymphoma       ***
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)nonmalignant        **
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)nonmalignant
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                                                                             exp(coef)
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)ALL                   2.1686
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)ALL              1.4923
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)AML                   5.3583
```

```
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)AML             1.2691
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)chronic              0.9515
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)chronic         1.7633
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)lymphoma             6.1500
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)lymphoma        6.6386
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)nonmalignant         6.0363
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)nonmalignant       NA
##                                                                            exp(-coef)
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)ALL                   0.4611
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)ALL              0.6701
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)AML                   0.1866
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)AML              0.7880
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)chronic               1.0510
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)chronic          0.5671
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)lymphoma              0.1626
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)lymphoma         0.1506
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)nonmalignant          0.1657
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)nonmalignant        NA
##                                                                            lower .95
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)ALL                  0.8551
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)ALL             0.6567
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)AML                  1.7435
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)AML             0.5005
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)chronic              0.2522
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)chronic         0.7540
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)lymphoma             1.2900
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)lymphoma        2.3896
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)nonmalignant         1.8056
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)nonmalignant       NA
##                                                                            upper .95
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)ALL                   5.500
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)ALL              3.391
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)AML                  16.468
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)AML              3.218
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)chronic               3.590
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)chronic          4.124
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)lymphoma             29.319
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)lymphoma        18.442
## as.factor(stem_cell_source)bone_marrow:as.factor(disease)nonmalignant          20.181
## as.factor(stem_cell_source)peripheral_blood:as.factor(disease)nonmalignant        NA
##
## Concordance= 0.628  (se = 0.032 )
## Likelihood ratio test= 25.32  on 9 df,   p=0.003
## Wald test            = 31.06  on 9 df,   p=3e-04
## Score (logrank) test = 36.76  on 9 df,   p=3e-05
```

definitely some significant results, but need to check that we have enough people in each covariate pattern

```
bone %>%
        group_by(disease, stem_cell_source) %>%
        count() %>%
        pander()
```

**check covariate patterns for disease type and stem cell source interaction**

| disease | stem_cell_source | n |
|---|---|---|
| ALL | bone_marrow | 18 |
| ALL | peripheral_blood | 50 |
| AML | bone_marrow | 6 |
| AML | peripheral_blood | 27 |
| chronic | bone_marrow | 11 |
| chronic | peripheral_blood | 34 |
| lymphoma | bone_marrow | 2 |
| lymphoma | peripheral_blood | 7 |
| nonmalignant | bone_marrow | 5 |
| nonmalignant | peripheral_blood | 27 |

Peripheral blood was more common than bone marrow for all five diseases. Unfortunately, we do not have enough people in each covariate pattern to include this interaction term in our model. For example, all 9 patients who had lymphoma died, so we cannot look at the interaction between lymphoma and stem cell source.

**adjusted model 3**

```
cox_adj3 <- coxph(survival.obj ~ as.factor(stem_cell_source) + CD3_x1e8_per_kg + recipient_age + as.fac
                        data = bone, ties = "exact")
summary(cox_adj3)
```

```
## Call:
## coxph(formula = survival.obj ~ as.factor(stem_cell_source) +
##     CD3_x1e8_per_kg + recipient_age + as.factor(CMV_status),
##     data = bone, ties = "exact")
##
##   n= 182, number of events= 81
##    (5 observations deleted due to missingness)
##
##                                                coef exp(coef) se(coef)      z
## as.factor(stem_cell_source)peripheral_blood -0.15721   0.85453  0.31103 -0.505
## CD3_x1e8_per_kg                             -0.06976   0.93262  0.04392 -1.588
## recipient_age                                0.04044   1.04127  0.02463  1.642
## as.factor(CMV_status)0                      -0.56052   0.57091  0.38849 -1.443
## as.factor(CMV_status)1                      -1.03344   0.35578  0.47783 -2.163
## as.factor(CMV_status)2                      -0.61269   0.54189  0.38210 -1.603
## as.factor(CMV_status)3                      -0.54437   0.58021  0.40414 -1.347
##                                             Pr(>|z|)
## as.factor(stem_cell_source)peripheral_blood   0.6132
## CD3_x1e8_per_kg                               0.1122
## recipient_age                                 0.1006
## as.factor(CMV_status)0                        0.1491
## as.factor(CMV_status)1                        0.0306 *
## as.factor(CMV_status)2                        0.1088
```

```
## as.factor(CMV_status)3                         0.1780
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                                           exp(coef) exp(-coef) lower .95
## as.factor(stem_cell_source)peripheral_blood    0.8545     1.1702    0.4645
## CD3_x1e8_per_kg                               0.9326     1.0722    0.8557
## recipient_age                                 1.0413     0.9604    0.9922
## as.factor(CMV_status)0                        0.5709     1.7516    0.2666
## as.factor(CMV_status)1                        0.3558     2.8107    0.1395
## as.factor(CMV_status)2                        0.5419     1.8454    0.2563
## as.factor(CMV_status)3                        0.5802     1.7235    0.2628
##                                           upper .95
## as.factor(stem_cell_source)peripheral_blood    1.5721
## CD3_x1e8_per_kg                               1.0165
## recipient_age                                 1.0928
## as.factor(CMV_status)0                        1.2225
## as.factor(CMV_status)1                        0.9076
## as.factor(CMV_status)2                        1.1459
## as.factor(CMV_status)3                        1.2811
##
## Concordance= 0.64  (se = 0.033 )
## Likelihood ratio test= 18.34  on 7 df,   p=0.01
## Wald test            = 18.15  on 7 df,   p=0.01
## Score (logrank) test = 18.82  on 7 df,   p=0.009
```

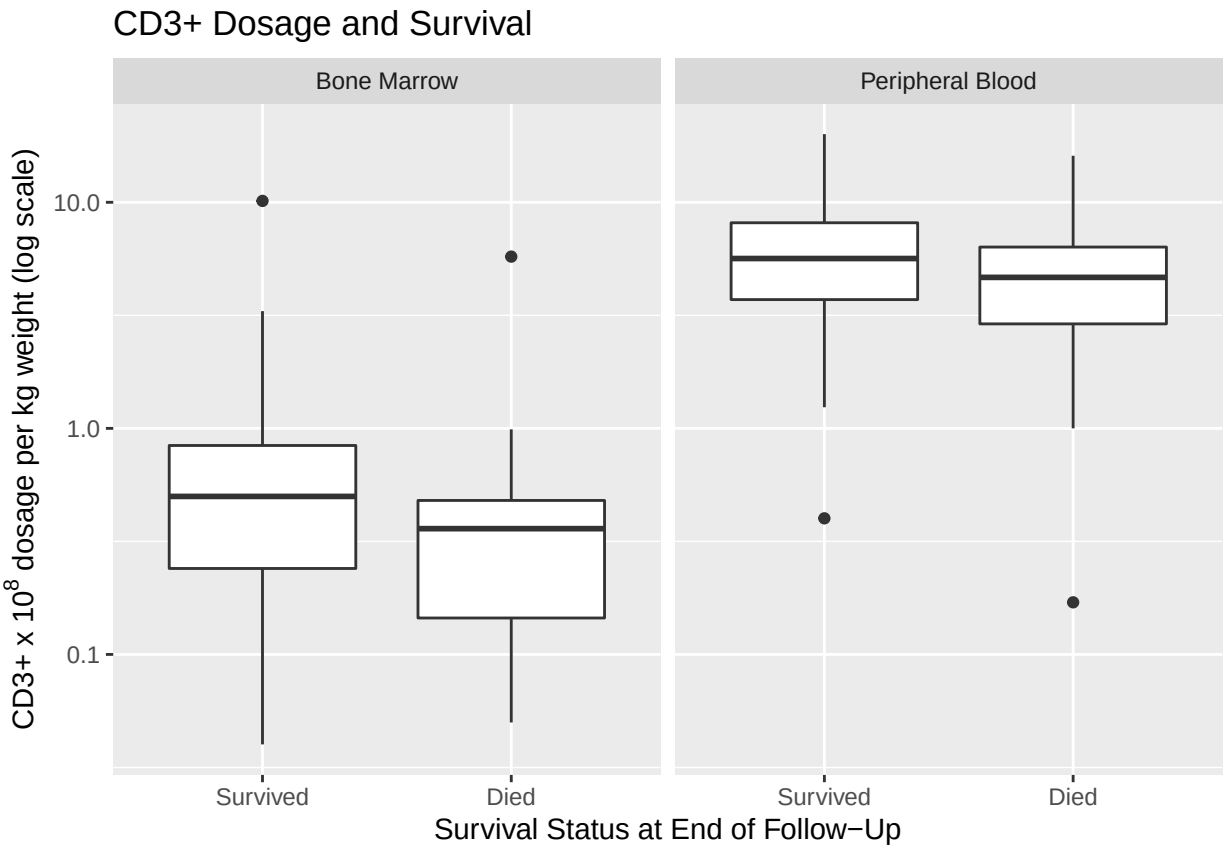
adding CMV status changes stem cell source coefficient less than 10%

**final model = adjusted model 1**

### does the benefit of a higher dose of CD3+ just keep increasing?

```
bone %>% ggplot(aes(x = as.factor(survival_status), y = CD3_x1e8_per_kg)) +
        geom_boxplot() +
        ggtitle("CD3+ Dosage and Survival") +
        xlab("Survival Status at End of Follow-Up") +
        ylab(expression(paste("CD3+ x ", 10^8, " dosage per kg weight (log scale)"))) +
        scale_x_discrete(labels = c("Survived", "Died")) +
        scale_y_log10() +
        facet_grid( ~ stem_cell_source,
                    labeller = as_labeller(c(`bone_marrow` = "Bone Marrow",
                                             `peripheral_blood` = "Peripheral Blood")))
```

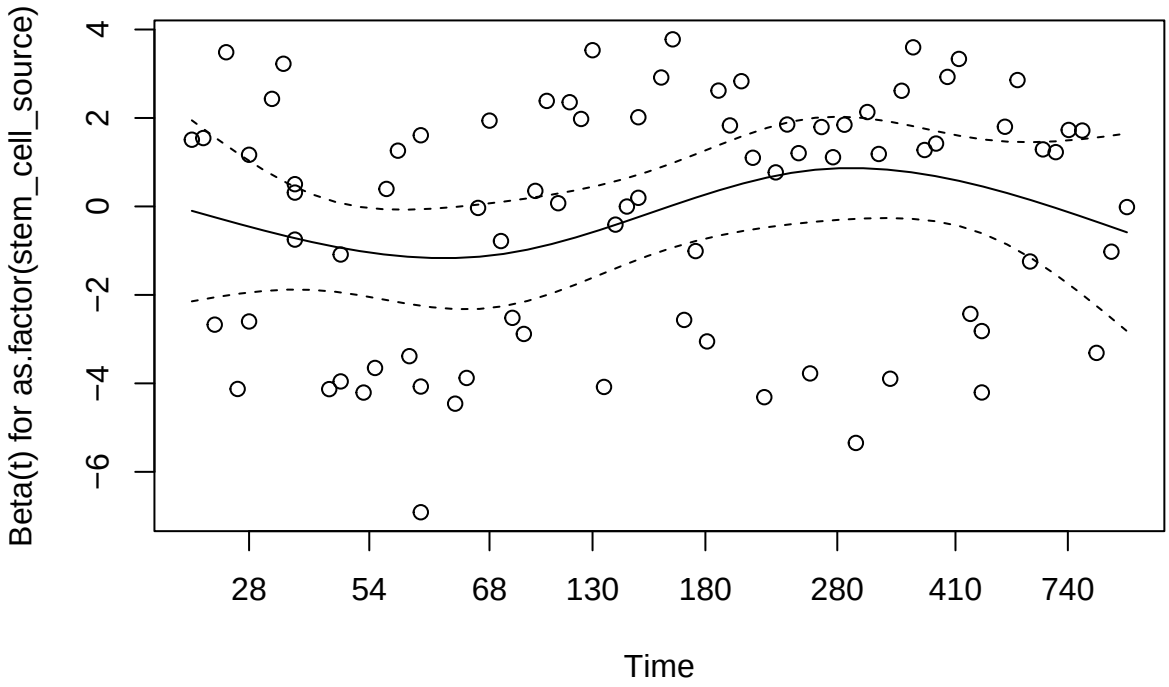
```
## Warning: Removed 5 rows containing non-finite values (stat_boxplot).
```

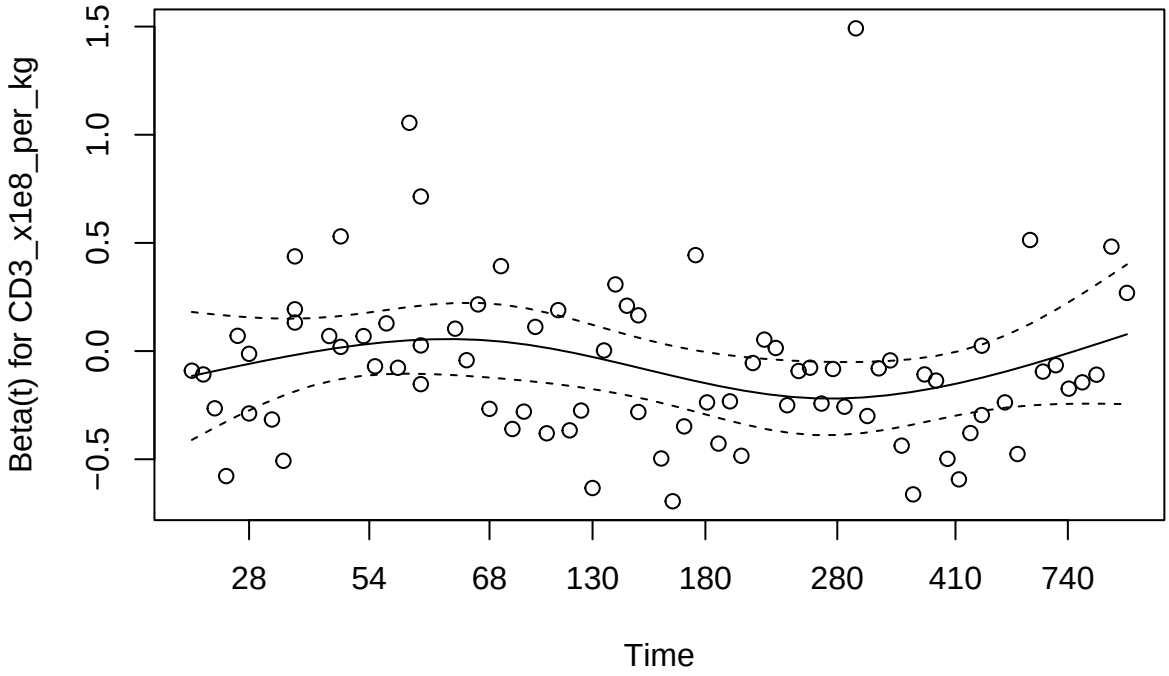


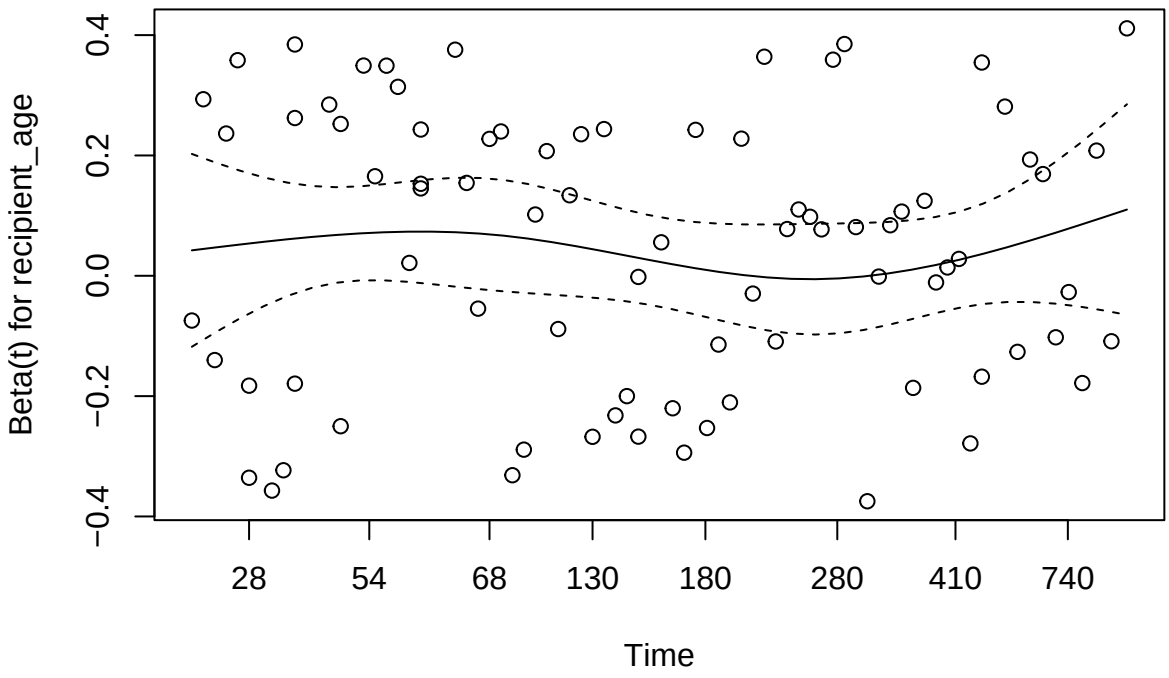

People who survived seem to have gotten higher CD3+ doses among both stem cell source types (but this may not be statistically significant).

### check Schoenfeld residuals for proportional hazards assumption

```
# create plot of schoenfeld resids
wt_sch_dose <- cox.zph(cox_adj1)
plot(wt_sch_dose) # slight decrease then increase over time
```


Beta(t) for as.factor(stem_cell_source)
4
2
0
−2
−4
−6
28
54
68
130
180
280
410
740
Time


Beta(t) for CD3_x1e8_per_kg
−0.5
0.0
0.5
1.0
1.5
28
54
68
130
180
280
410
740
Time




```
# check summary to see if problematic
wt_sch_dose
```

```
##                                chisq df    p
## as.factor(stem_cell_source) 1.357532  1 0.24
## CD3_x1e8_per_kg             0.000467  1 0.98
## recipient_age               0.000396  1 0.98
## GLOBAL                      2.116315  3 0.55
```