AML Survival Analysis

Diarmuid Lenihan

2025-02-15

1 Study Overview

- **Dataset:** AML patient survival data (**n = 200**, **133 events**).
- **Objective:** Assess the impact of **FLT3 mutation** on survival, adjusting for **age, WBC count, BM blast percentage, and cytogenetic risk classification**.
- **Statistical Model:** Cox proportional hazards model with **age stratification** (`strata(age_group)`).
- **Survival Estimation:** Kaplan-Meier curves comparing **FLT3-mutant vs. WT patients** across different **age groups**.

---

2 Load Processed Data

```
# Load the processed dataset
aml_data <- readRDS("processed_aml_data.rds")  # Use read.csv("processed_aml_data.csv") if saved as CSV

# Show structure of the dataset
str(aml_data)
```

```
## 'data.frame':    200 obs. of  19 variables:
##  $ patient_ID                          : Factor w/ 200 levels "TCGA-AB-2802",..: 1 2 3 4 5 6 7 8 9 10
##  $ FLT3_mut_status                     : Factor w/ 2 levels "0","1": 1 1 1 1 1 1 1 1 2 ...
##  $ sex                                 : Factor w/ 2 levels "Female","Male": 2 1 2 2 2 1 2 1 1 2 ...
##  $ race                                : Factor w/ 13 levels "ASIAN","BLACK",..: 13 13 13 13 13 13 13
##  $ FAB                                 : Factor w/ 9 levels "M0","M1","M2",..: 5 4 4 1 2 2 3 3 3 5 .
##  $ age_at_diagnosis                    : int  50 61 30 77 46 68 23 64 76 81 ...
##  $ BM_blast_percentage_at_diagnosis    : int  88 44 82 67 90 91 59 60 48 98 ...
##  $ WBC_count_at_diagnosis              : num  16.9 1 5.7 92 29.4 ...
##  $ risk_cyto_at_diagnosis              : Factor w/ 4 levels "Good","Intermediate",..: 2 1 2 2 1 2 2 1
##  $ risk_molecular_at_diagnosis         : Factor w/ 4 levels "Good","Intermediate",..: 2 1 1 2 1 2 2 1
##  $ date_of_form_completion             : Factor w/ 1 level "14/12/2010": 1 1 1 1 1 1 1 1 1 1 ...
##  $ date_of_initial_pathologic_diagnosis: Factor w/ 10 levels "2001-00-00","2002-00-00",..: 1 1 1 2 2
##  $ days_to_birth_from_diagnosis        : int  -18385 -22584 -11203 -28124 -16892 -25143 -8520 -23679
##  $ days_to_death_from_diagnosis        : Factor w/ 69 levels "[Not Applicable]",..: 37 60 1 48 68 16
##  $ days_to_last_followup_from_diagnosis: Factor w/ 60 levels "[Not Available]",..: 1 1 32 1 1 1 35 1
##  $ vital_status                        : num  1 1 0 1 1 1 0 1 1 1 ...
##  $ survival_time                       : int  365 792 2556 576 944 180 2861 62 31 243 ...
##  $ age_group                           : Factor w/ 3 levels " 50","51-65",..: 1 2 1 3 1 3 1 2 3 3 ...
##  $ survival_time_years                 : num  1 2.17 7 1.58 2.59 ...
```

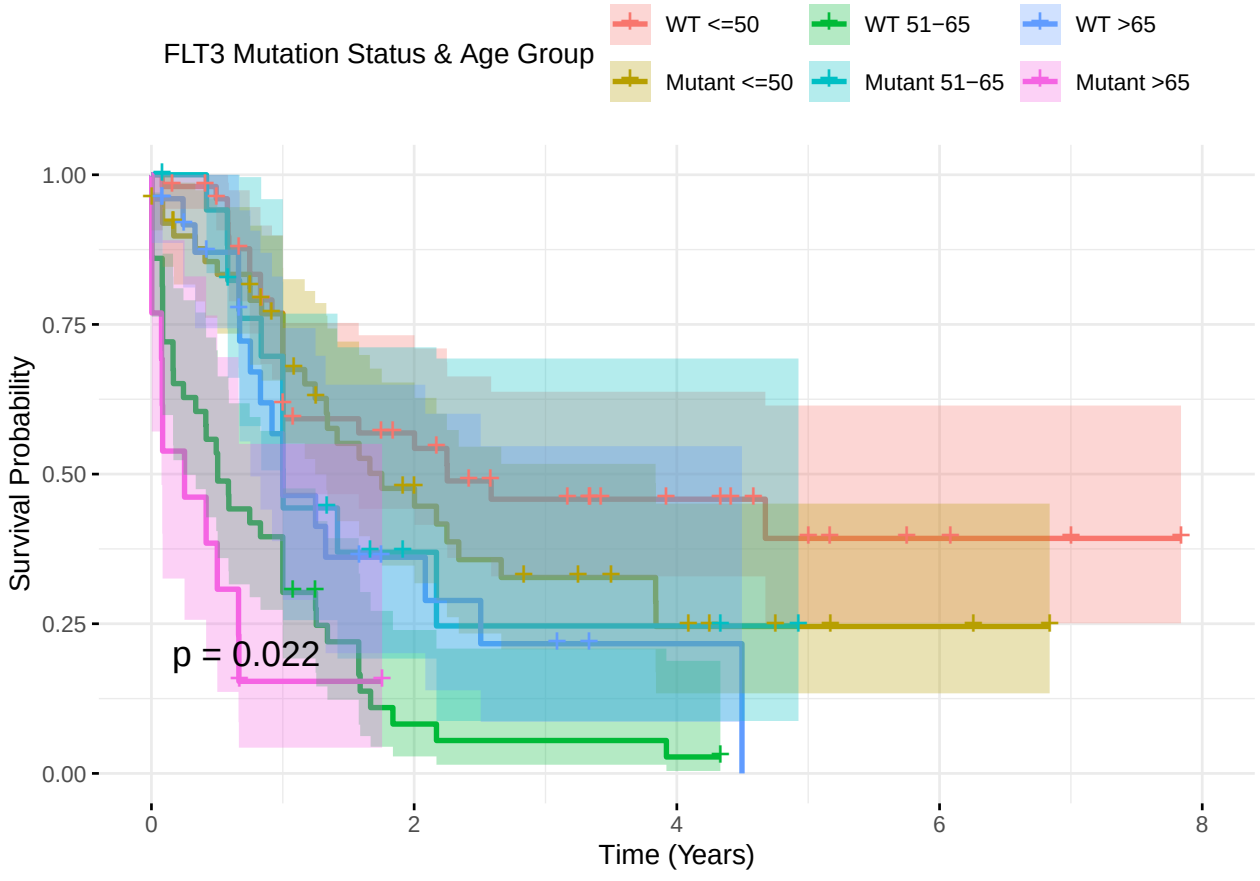
3 Cox Proportional Hazards Model

```r
# Fit the stratified Cox model
cox_model_stratified <- coxph(Surv(survival_time_years, vital_status) ~
                              FLT3_mut_status +
                              BM_blast_percentage_at_diagnosis +
                              WBC_count_at_diagnosis +
                              risk_cyto_at_diagnosis +
                              strata(age_group),
                              data = aml_data)

# Display model summary
summary(cox_model_stratified)
```

```
## Call:
## coxph(formula = Surv(survival_time_years, vital_status) ~ FLT3_mut_status +
##     BM_blast_percentage_at_diagnosis + WBC_count_at_diagnosis +
##     risk_cyto_at_diagnosis + strata(age_group), data = aml_data)
##
##   n= 200, number of events= 133
##
##                                     coef exp(coef) se(coef)     z Pr(>|z|)
## FLT3_mut_status1                   0.446496  1.562827 0.229073 1.949 0.051279
## BM_blast_percentage_at_diagnosis   0.001728  1.001730 0.004774 0.362 0.717352
## WBC_count_at_diagnosis             0.004484  1.004494 0.002050 2.187 0.028720
## risk_cyto_at_diagnosisIntermediate 0.948034  2.580632 0.322146 2.943 0.003252
## risk_cyto_at_diagnosisN.D.         2.170823  8.765498 0.606042 3.582 0.000341
## risk_cyto_at_diagnosisPoor         1.586612  4.887163 0.351287 4.517 6.28e-06
##
## FLT3_mut_status1                   .
## BM_blast_percentage_at_diagnosis
## WBC_count_at_diagnosis             *
## risk_cyto_at_diagnosisIntermediate **
## risk_cyto_at_diagnosisN.D.         ***
## risk_cyto_at_diagnosisPoor         ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                                    exp(coef) exp(-coef) lower .95 upper .95
## FLT3_mut_status1                       1.563     0.6399    0.9975     2.448
## BM_blast_percentage_at_diagnosis       1.002     0.9983    0.9924     1.011
## WBC_count_at_diagnosis                 1.004     0.9955    1.0005     1.009
## risk_cyto_at_diagnosisIntermediate     2.581     0.3875    1.3725     4.852
## risk_cyto_at_diagnosisN.D.             8.765     0.1141    2.6725    28.750
## risk_cyto_at_diagnosisPoor             4.887     0.2046    2.4549     9.729
##
## Concordance= 0.654  (se = 0.028 )
## Likelihood ratio test= 38.11  on 6 df,   p=1e-06
## Wald test            = 35.41  on 6 df,   p=4e-06
## Score (logrank) test = 38.7  on 6 df,   p=8e-07
```

4 Kaplan-Meier Survival Analysis

```
# Fit the survival model
surv_obj <- survfit(Surv(survival_time_years, vital_status) ~ FLT3_mut_status + strata(age_group),
                    data = aml_data)

# Plot survival curves
ggsurvplot(surv_obj,
           data = aml_data,
           conf.int = TRUE,
           pval = TRUE,
           legend.title = "FLT3 Mutation Status & Age Group",
           legend.labs = c("WT  50", "Mutant  50",
                           "WT 51-65", "Mutant 51-65",
                           "WT >65", "Mutant >65"),
           xlab = "Time (Years)",
           ylab = "Survival Probability",
           ggtheme = theme_minimal())
```



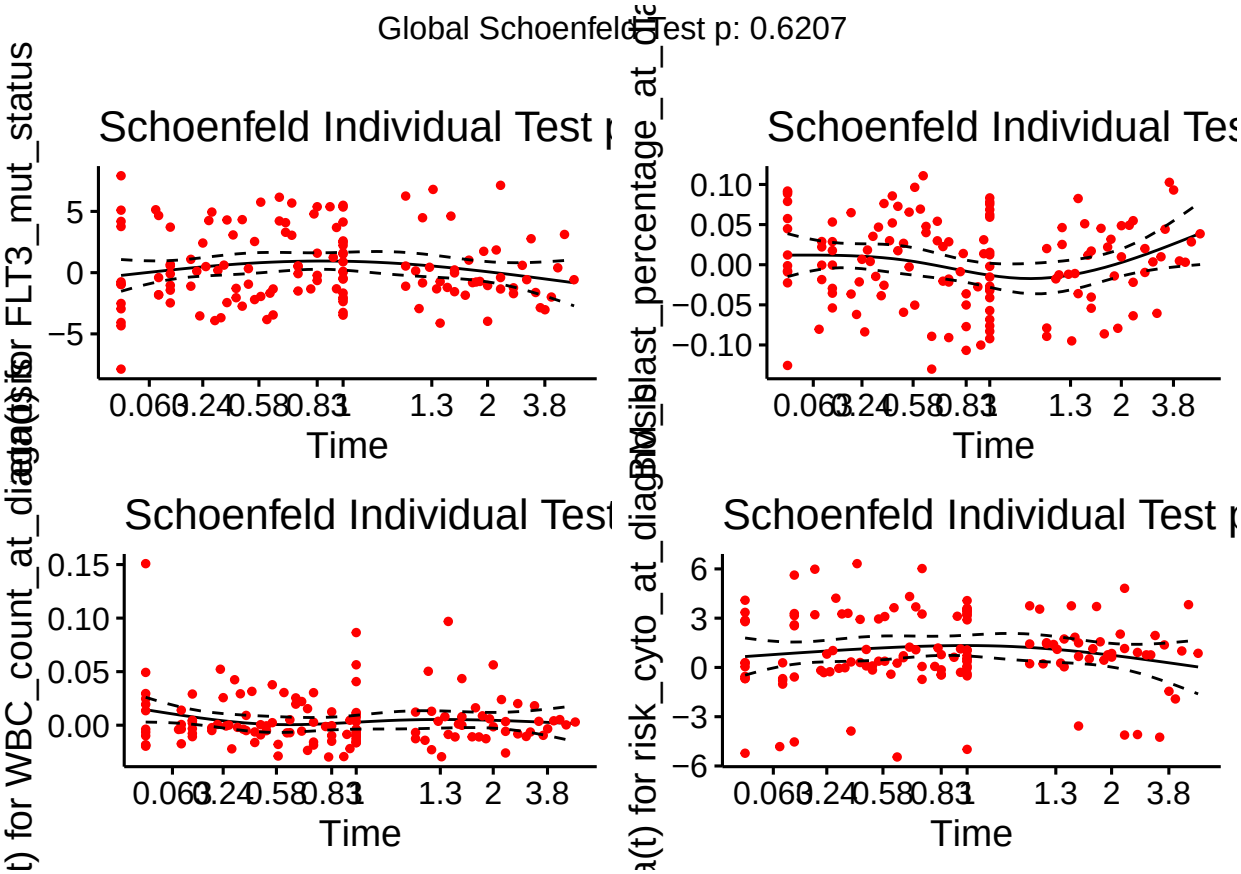
5 Proportional Hazards Assumption Check

```
# Check proportional hazards assumption
ph_test <- cox.zph(cox_model_stratified)
```

```r
print(ph_test)
```

```
##                                   chisq df    p
## FLT3_mut_status                   0.611  1 0.43
## BM_blast_percentage_at_diagnosis  0.444  1 0.51
## WBC_count_at_diagnosis            1.463  1 0.23
## risk_cyto_at_diagnosis            1.355  3 0.72
## GLOBAL                            4.415  6 0.62
```

```r
# Plot Schoenfeld residuals
ggcoxzph(ph_test)
```



---

6 Final Conclusions

- **FLT3 mutation is associated with worse survival** (HR = **1.56**, **p = 0.051**, borderline significant).
- **WBC count at diagnosis is a weak predictor of survival** (**p = 0.0287**).
- **Cytogenetic risk classification is the strongest predictor** (HR = **4.89** for poor-risk, **p < 0.001**).
- **The proportional hazards assumption is satisfied (`p = 0.62`), validating the model.**
- **Survival curves confirm FLT3-mutant patients have significantly lower survival probabilities (`p = 0.022`).**

---

7 Next Steps & Recommendations

- Further explore interactions between FLT3 mutation and chemotherapy response.
- Validate findings with external AML patient datasets.
- Investigate targeted therapies for FLT3-mutant patients.

---

8 References

(Include any relevant citations here.)

Notes

- **Generated using R (`survival`, `ggsurvplot`) and Cox regression modeling.**
- **Kaplan-Meier survival analysis was performed with age stratification (`strata(age_group)`).**
- **Schoenfeld residuals confirmed model validity (`p = 0.62`).**