Cox Proportional Hazards Regression

CRP 245 Tutorial

Duke University Clinical Research Training Program

2026-01-22

Table of contents

Introduction

Learning Objectives

After completing this tutorial, you will be able to:

- **Understand** the purpose and assumptions of Cox Proportional Hazards regression.
- **Estimate** Hazard Ratios for treatment effects and patient characteristics.
- **Interpret** hazard ratios for both categorical and continuous predictors.
- **Calculate** hazard ratios for clinically meaningful intervals (e.g., 5-year age change).
- **Visualize** survival curves with Kaplan-Meier plots stratified by predictors.
- **Distinguish** between statistical significance and clinical importance.

0.1 Clinical Context

When treating patients with serious chronic conditions, a critical question is: **"How much does this treatment reduce the risk of a bad outcome?"** Cox Proportional Hazards regression provides a rigorous method to answer this question while accounting for the fact that not all patients experience the outcome during follow-up.

Why Cox Regression?

Unlike simple comparison of event rates, Cox regression:

1. **Accounts for censoring** — patients who don't experience the event by study end still contribute information.
2. **Adjusts for covariates** — we can separate treatment effects from patient characteristics.
3. **Provides Hazard Ratios** — an intuitive measure of relative risk over time.

0.2 Study Overview: The CGD Trial

This tutorial uses data from a landmark randomized, double-blind, placebo-controlled clinical trial investigating recombinant gamma interferon (IFN-) for preventing serious infections in chronic granulomatous disease (CGD).

About CGD:

- CGD is a rare inherited immunodeficiency (1 in 200,000 individuals)
- Characterized by recurrent, life-threatening bacterial and fungal infections
- Primarily affects males (X-linked inheritance pattern)

- Often diagnosed in childhood

Study Design:

- Patients randomized 1:1 to gamma interferon vs. placebo
- Primary endpoint: Time to first serious infection

0.3 Data Dictionary

Table 1: CGD Trial Key Variables

Variable	Description
`id`	Patient identification number
`treat`	Treatment group (0=Placebo, 1=Gamma interferon)
`sex`	Patient sex (0=Male, 1=Female)
`age`	Age at enrollment (years)
`tstop`	Follow-up time: Days to infection or censoring
`status`	Event indicator (1=Infection, 0=Censored/No infection)

---

1 Setup and Data Loading

1.1 Loading Required Packages

We begin by loading the specialized R packages required for survival analysis. Each package serves a specific purpose.

```
# survival: Core package for Cox regression and Kaplan-Meier estimation
# Key functions: Surv(), coxph(), survfit()
if (!requireNamespace("survival", quietly = TRUE)) install.packages("survival")
library(survival)

# survminer: Publication-quality survival plots
# Key function: ggsurvplot()
if (!requireNamespace("survminer", quietly = TRUE)) install.packages("survminer")
library(survminer)
```

1.2 Loading the Data

We load the CGD trial dataset and examine the study population.

```
# Load the trial data
load(url("https://www.duke.edu/~sgrambow/crp241data/cgd.RData"))

# Examine data structure
str(cgd)
```

```
'data.frame':   128 obs. of  10 variables:
 $ id      : int  1 2 3 4 5 6 7 8 9 10 ...
 $ treat   : num  1 0 1 1 0 1 0 1 0 1 ...
 $ sex     : num  0 1 1 1 1 0 1 1 1 1 ...
 $ age     : int  12 15 19 12 17 44 22 7 27 5 ...
 $ height  : num  147 159 171 142 162 ...
 $ weight  : num  62 47.5 72.7 34 52.7 45 59.7 17.4 82.8 19.5 ...
 $ inherit : num  0 0 1 1 1 0 1 1 0 1 ...
 $ propylac: int  0 1 1 1 1 0 1 1 1 1 ...
 $ tstop   : int  219 8 382 388 246 364 292 363 294 371 ...
 $ status  : int  1 1 0 0 1 0 1 0 1 0 ...
```

```
# Summary statistics
summary(cgd)
```

```
       id              treat             sex              age       
 Min.   :  1.00   Min.   :0.0000   Min.   :0.0000   Min.   : 1.00  
 1st Qu.: 32.75   1st Qu.:0.0000   1st Qu.:1.0000   1st Qu.: 7.00  
 Median : 64.50   Median :0.0000   Median :1.0000   Median :12.00  
 Mean   : 64.81   Mean   :0.4922   Mean   :0.8125   Mean   :14.64  
 3rd Qu.: 96.25   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:22.00  
 Max.   :135.00   Max.   :1.0000   Max.   :1.0000   Max.   :44.00  
     height          weight          inherit          propylac      
 Min.   : 76.3   Min.   : 10.40   Min.   :0.0000   Min.   :0.0000  
 1st Qu.:116.5   1st Qu.: 20.68   1st Qu.:0.0000   1st Qu.:1.0000  
 Median :140.8   Median : 34.85   Median :1.0000   Median :1.0000  
 Mean   :140.1   Mean   : 40.56   Mean   :0.6719   Mean   :0.8672  
 3rd Qu.:169.7   3rd Qu.: 59.17   3rd Qu.:1.0000   3rd Qu.:1.0000  
 Max.   :189.0   Max.   :101.50   Max.   :1.0000   Max.   :1.0000  
     tstop           status      
 Min.   :  4.0   Min.   :0.0000  
```

```
 1st Qu.:197.0   1st Qu.:0.0000
 Median :269.0   Median :0.0000
 Mean   :241.1   Mean   :0.3438
 3rd Qu.:304.5   3rd Qu.:1.0000
 Max.   :388.0   Max.   :1.0000
```

Statistical Interpretation

- **Sample Size:** $n = 128$ patients.
- **Treatment Distribution:** 49.2% gamma interferon, 50.8% placebo (balanced).
- **Sex Distribution:** 81.3% male, 18.7% female (reflects X-linked inheritance).
- **Age:** Median 12 years (range: 1-44). This is predominantly a pediatric population.
- **Event Rate:** 44 infections observed (34.4%).
- **Follow-up:** Median 269 days (range: 4-388).

Clinical Context

The young median age (12 years) reflects that CGD typically presents in childhood. The 34% event rate provides adequate statistical power for Cox regression. An 81% male population is expected given the X-linked inheritance pattern of CGD.

---

2 Exploratory Visualization

Before formal regression modeling, we visualize survival differences between treatment arms using Kaplan-Meier curves.

```
# Create Kaplan-Meier survival estimates stratified by treatment
fit.km <- survfit(Surv(tstop, status) ~ treat, data=cgd)

# Generate publication-quality plot with risk table
ggsurvplot(fit.km, data=cgd,
           ggtheme = theme_minimal(),
           legend.labs = c("Placebo","Gamma interferon"),
           risk.table = TRUE,
           xlab = "Days Since Randomization",
           ylab = "Probability of Remaining Infection-Free")
```

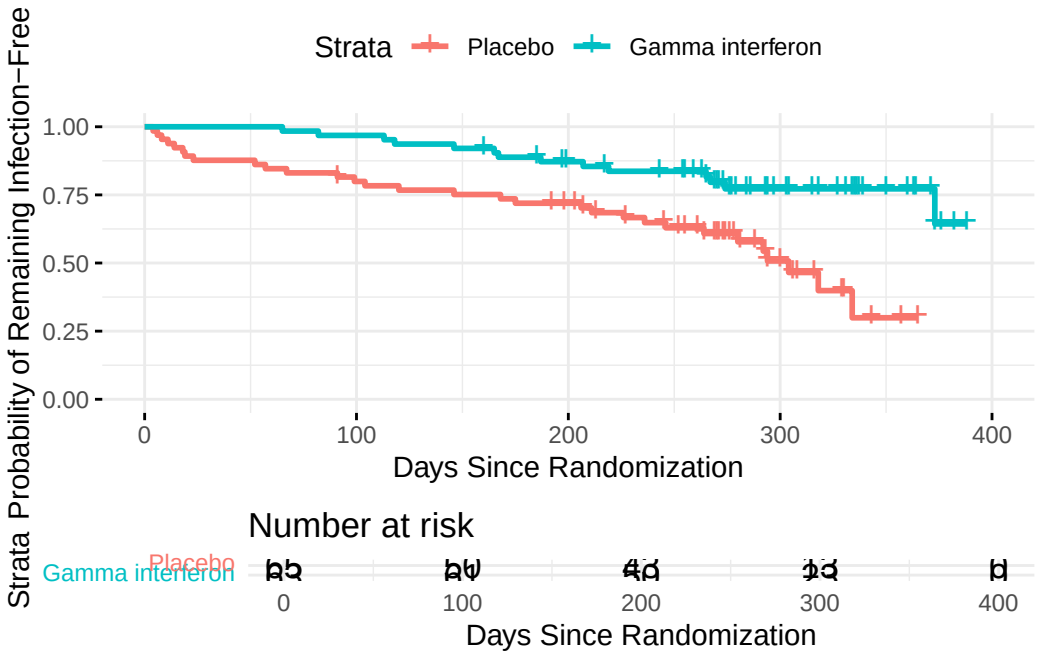

Figure 1: Kaplan-Meier Curves by Treatment Group

Visual Finding

The curves separate early and maintain separation throughout follow-up. Patients receiving gamma interferon have consistently higher probability of remaining infection-free compared to placebo.

Reading the Risk Table

The "Number at Risk" table shows how many patients remain under observation at each time point. This is critical because:

- Estimates become less reliable as fewer patients remain
- In this trial, adequate patients remain at risk throughout, supporting reliable estimates

3 Question 1: Treatment Effect

Clinical Question: What is the hazard ratio comparing placebo vs. gamma interferon, and is the treatment effect statistically significant?

3.1 Fitting the Cox Model

```r
# Fit Cox Proportional Hazards model for treatment effect
# coxph() estimates the hazard ratio for gamma interferon vs. placebo
mfit <- coxph(Surv(tstop, status) ~ treat, data=cgd)

# Display complete model output
summary(mfit)
```

```
Call:
coxph(formula = Surv(tstop, status) ~ treat, data = cgd)

  n= 128, number of events= 44

         coef exp(coef) se(coef)      z Pr(>|z|)
treat -1.0940    0.3349   0.3348 -3.268  0.00108 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

      exp(coef) exp(-coef) lower .95 upper .95
treat    0.3349      2.986    0.1737    0.6454

Concordance= 0.621  (se = 0.036 )
Likelihood ratio test= 11.8  on 1 df,   p=6e-04
Wald test            = 10.68  on 1 df,   p=0.001
Score (logrank) test = 11.74  on 1 df,   p=6e-04
```

> **!** Understanding the Output Direction
>
> The model compares **gamma interferon** (treat=1) to **placebo** (treat=0, reference). To express the result as "placebo vs. gamma interferon" (the clinical question), we invert the hazard ratio.

3.2 Calculating the Clinical Hazard Ratio

```
# Model gives HR for gamma vs. placebo = 0.3349
# We want HR for placebo vs. gamma = 1/0.3349

# Hazard Ratio: Placebo vs. Gamma Interferon
1/0.3349
```

```
[1] 2.985966
```

```
# 95% Confidence Interval (inverted)
1/0.6454  # Lower bound
```

```
[1] 1.549427
```

```
1/0.1737  # Upper bound
```

```
[1] 5.757052
```

Statistical Interpretation

- **Hazard Ratio (Placebo vs. Gamma):** 2.99 [95% CI: 1.55, 5.76]
- **P-value:** 0.001 (highly statistically significant)
- **Concordance:** 0.621 (moderate model discrimination)

Clinical Interpretation

In plain language: Patients receiving placebo have approximately **3 times the risk** of developing a serious infection compared to those receiving gamma interferon.
What this means for practice:

- The 95% CI [1.55, 5.76] is entirely above 1.0, confirming a true protective effect
- The p-value of 0.001 provides strong statistical evidence
- This is both statistically significant AND clinically meaningful
- For every patient who develops an infection on gamma interferon, about 3 develop infections on placebo

---

4 Question 2: Sex Effect

Clinical Question: Is there an association between patient sex and infection risk?

```
# Fit Cox model examining sex effect
# sex = 0 (male) is the reference; sex = 1 (female) is compared to it
mfit2.cgd <- coxph(Surv(tstop, status) ~ sex, data=cgd)
summary(mfit2.cgd)
```

```
Call:
coxph(formula = Surv(tstop, status) ~ sex, data = cgd)

  n= 128, number of events= 44

      coef exp(coef) se(coef)     z Pr(>|z|)
sex 0.2127    1.2370   0.4123 0.516    0.606

    exp(coef) exp(-coef) lower .95 upper .95
sex     1.237     0.8084    0.5514     2.775

Concordance= 0.507  (se = 0.031 )
Likelihood ratio test= 0.28  on 1 df,   p=0.6
Wald test            = 0.27  on 1 df,   p=0.6
Score (logrank) test = 0.27  on 1 df,   p=0.6
```

Statistical Interpretation

- **Hazard Ratio (Male vs. Female):** 1.24 [95% CI: 0.55, 2.78]
- **P-value:** 0.606 (NOT statistically significant)
- **Concordance:** 0.507 (essentially no discriminative ability)

Clinical Interpretation

Although males show a 24% higher point estimate of risk, this difference is **not statistically significant** ($p = 0.606$). The wide confidence interval [0.55, 2.78] includes 1.0, indicating substantial uncertainty.
Clinical implication: There is no evidence to suggest that treatment decisions should differ based on patient sex in this CGD population.

> **Power Consideration**
>
> With only 18.7% female patients, statistical power to detect a sex difference is limited. Larger studies would be needed to definitively evaluate sex as a predictor.

---

5 Question 3: Visualizing Sex Effect

Clinical Question: Do the Kaplan-Meier curves confirm the non-significant Cox regression result?

```
# Create KM estimates stratified by sex
mfit.sex <- survfit(Surv(tstop, status) ~ sex, data=cgd)

# Generate plot
ggsurvplot(mfit.sex, data=cgd,
           ggtheme = theme_minimal(),
           legend.labs = c("Male","Female"),
           risk.table = TRUE,
           xlab = "Days Since Randomization",
           ylab = "Probability of Remaining Infection-Free")
```

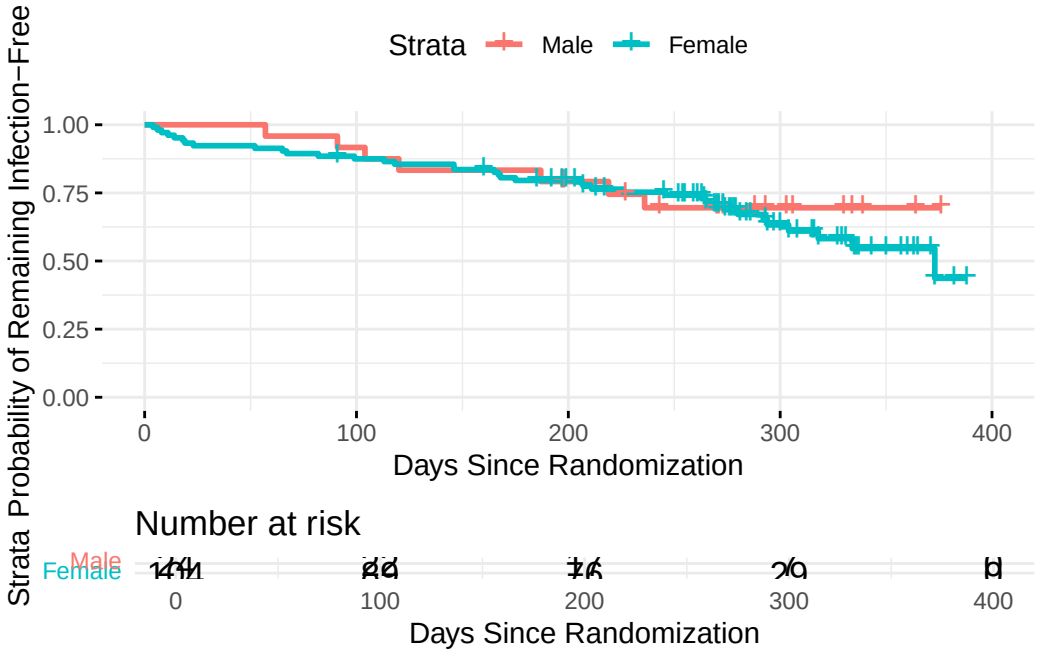

Figure 2: Kaplan-Meier Curves by Patient Sex

Statistical Interpretation

The curves overlap substantially throughout follow-up. There is no consistent visual separation between males and females, consistent with the non-significant p-value of 0.606.

Clinical Context

The overlapping curves visually confirm what the Cox model told us numerically: there is no strong evidence that sex predicts infection risk in CGD patients. The small female sample (n=24 vs. n=104 males) limits our ability to draw definitive conclusions about sex differences.

---

6 Question 4: Age Effect

Clinical Question: Is there an association between patient age and infection risk?

When age is a **continuous** predictor, the hazard ratio represents the change in risk per 1-unit (1-year) increase.

```r
# Fit Cox model with age as continuous predictor
mfit.age <- coxph(Surv(tstop, status) ~ age, data=cgd)
summary(mfit.age)
```

```
Call:
coxph(formula = Surv(tstop, status) ~ age, data = cgd)

  n= 128, number of events= 44

        coef exp(coef) se(coef)      z Pr(>|z|)
age -0.02121   0.97901  0.01682 -1.261    0.207

    exp(coef) exp(-coef) lower .95 upper .95
age     0.979      1.021    0.9473     1.012

Concordance= 0.57  (se = 0.048 )
Likelihood ratio test= 1.69  on 1 df,   p=0.2
Wald test            = 1.59  on 1 df,   p=0.2
Score (logrank) test = 1.6  on 1 df,   p=0.2
```

Statistical Interpretation

- **Coefficient:** -0.02121 (negative = protective direction)
- **Hazard Ratio (per 1-year):** 0.979 [95% CI: 0.947, 1.012]
- **P-value:** 0.207 (NOT statistically significant)

Clinical Interpretation

For each **1-year increase** in age, the hazard of infection decreases by approximately **2.1%**. This suggests a trend toward lower risk in older patients.
However: This effect is not statistically significant (p = 0.207). The confidence interval [0.947, 1.012] crosses 1.0, so we cannot rule out no effect.
Possible biological rationale: Immune systems may mature with age, providing better defense even in CGD patients. Alternatively, this may reflect survivor bias (more severe cases may not survive to older ages).

---

7 Question 5: Clinically Meaningful Age Effect

Clinical Question: What is the hazard ratio for a **5-year** increase in age?

A 1-year change is often too small to be clinically meaningful. We can calculate the hazard ratio for larger, more relevant intervals.

Mathematical Principle

For continuous predictors:

$$HR_{5-year} = e^{5 \times \beta}$$

If the HR for a 1-year change is e^{β}, then for a 5-year change it becomes $(e^{\beta})^5 = e^{5\beta}$.

```
# Calculate 5-year hazard ratio
# Coefficient for age = -0.02121
exp(5 * -0.02121)
```

```
[1] 0.8993797
```

```
# Calculate 95% Confidence Interval for 5-year change
exp(5 * confint(mfit.age))
```

```
        2.5 % 97.5 %
age 0.7627288 1.0605
```

Statistical Interpretation

- **5-Year Hazard Ratio:** 0.90 [95% CI: 0.76, 1.06]
- **Interpretation:** 10% reduction in risk per 5-year increase in age
- **Statistical significance:** Still non-significant (CI crosses 1.0)

Clinical Interpretation

A 5-year increase in age is associated with a **10% reduction** in infection risk (HR = 0.90). For example, a 15-year-old would have approximately 10% lower risk than a 10-year-old.
Caveats:

- The 95% CI [0.76, 1.06] still crosses 1.0
- We cannot conclude that age is a statistically significant predictor

- The direction is clinically intuitive (older children may have more mature immune function)
- Age should NOT be used as the sole factor for treatment decisions

---

8 Summary and Key Concepts

8.1 Hazard Ratio Reference Guide

Table 2: Interpreting Hazard Ratios

HR Value	Interpretation
HR = 1.0	No effect (reference)
HR > 1.0	Increased risk (harmful exposure)
HR < 1.0	Decreased risk (protective exposure)
HR = 2.0	2x the risk (100% increased risk)
HR = 0.5	0.5x the risk (50% decreased risk)

Key Findings from This Analysis

1. **Treatment Effect (Q1):** Placebo patients had **3x the infection risk** vs. gamma interferon (HR = 2.99, p = 0.001). This is both statistically significant AND clinically meaningful.

2. **Sex Effect (Q2-Q3):** No significant association between sex and infection risk (HR = 1.24, p = 0.606).

3. **Age Effect (Q4-Q5):** A trend toward lower risk in older patients (10% reduction per 5 years), but not statistically significant (p = 0.207).

Key Teaching Points

☐ **Cox regression** provides hazard ratios adjusted for time-to-event data and censoring.
☐ **Hazard Ratio > 1** means increased risk; **HR < 1** means decreased risk.
☐ **Continuous predictors:** The HR represents the change per 1-unit increase; multiply the coefficient for larger intervals.
☐ **Statistical significance** ($p < 0.05$) is not the same as **clinical importance**—

always consider effect size.

☐ **Kaplan-Meier plots** visualize what Cox regression quantifies.

---

9 Practice Exercises

1. **Interpretation Practice:** If a Cox model for a new drug gives HR = 0.65 [95% CI: 0.45, 0.95] compared to standard treatment, what would you tell a patient about the drug's benefit?
2. **Continuous Predictor:** If a model estimates the HR for BMI as 1.03 per unit, what is the HR for a 10-unit increase in BMI? Is this clinically concerning?
3. **Critical Appraisal:** In the sex analysis, the p-value was 0.606. Does this mean there is definitively no sex difference? What would you need to increase confidence in this conclusion?

This tutorial was developed for CRP 245 at Duke University.