Covid-19 Vaccine Analysis

Nathan Dennis

Background

Theoretical $S(t)$:

S(0) = 1

$S(t)$

$S(\infty) = 0$

0   t   ∞

(1)

$$S(t) = P(T > t)$$

- Covid-19 & vaccine process
- Recreated 29148 observations
- Survival time, until "event"
- Days until symptomatic Covid-19
- Placebo vs Experimental
- Censored data (3)
- Survival function, S(t) (1), (2)

(3)

(2)

Weeks →
2   4   6   8   10   12

A $T=5$ — X
B $T=12$ — Study end
C $T=3.5$ Withdrawn
D $T=8$ — Study end
E $T=6$ — Lost
F $T=3.5$ X

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t \mid T \ge t)}{\Delta t}$$
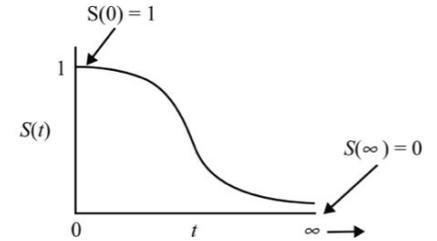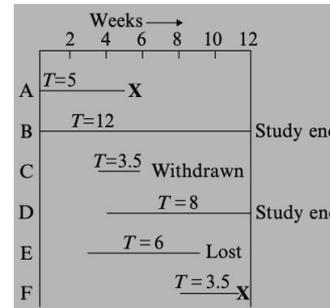
Hazard Function

- Potential per unit time
- Given individual has survived to some time
- Can be greater than 1, no upper bound.
- Greater than or equal to 0
- Experiences event 0.6 times/day or 4.2 times/week
- Synonymous with conditional failure rate

Main Differences with Survival Function

- Not a probability
- Focus on failures
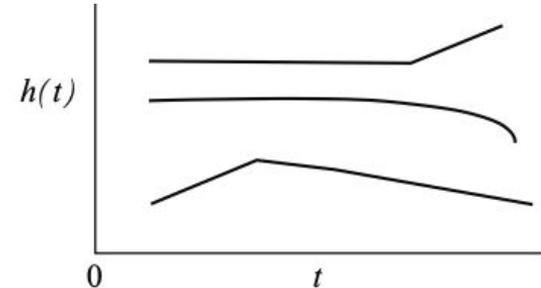
Kaplan-Meier		Cox PH
Non-parametric	Type of model?	Semiparametric
Survival Function	Estimates?	Hazard function
Random censoring	Assumptions?	PH Assumption Random censoring
Survival curve difference	Discovery?	Hazard Ratio, efficacy

$$h(t, \mathbf{X}) = h_0(t)e^{\sum_{i=1}^{p} \beta_i X_i}$$

$$\mathbf{X} = (X_1, X_2, \ldots X_p)$$
explanatory/predictor variables

$$h_0(t) \quad \times \quad e^{\sum_{i=1}^{p} \beta_i X_i}$$

Baseline hazard Involves t but not X's	Exponential Involves X's but not t (X's are time-independent)

Kaplan-Meier

- Significant difference in graphs
- Survival probability nearly 1 for treatment.
- Placebo decreases faster.
- Log rank test: P-Value of <2e-16, reject null hypothesis
- Curves are statistically different.
- Can we learn more?

Probability of COVID-19 Infection
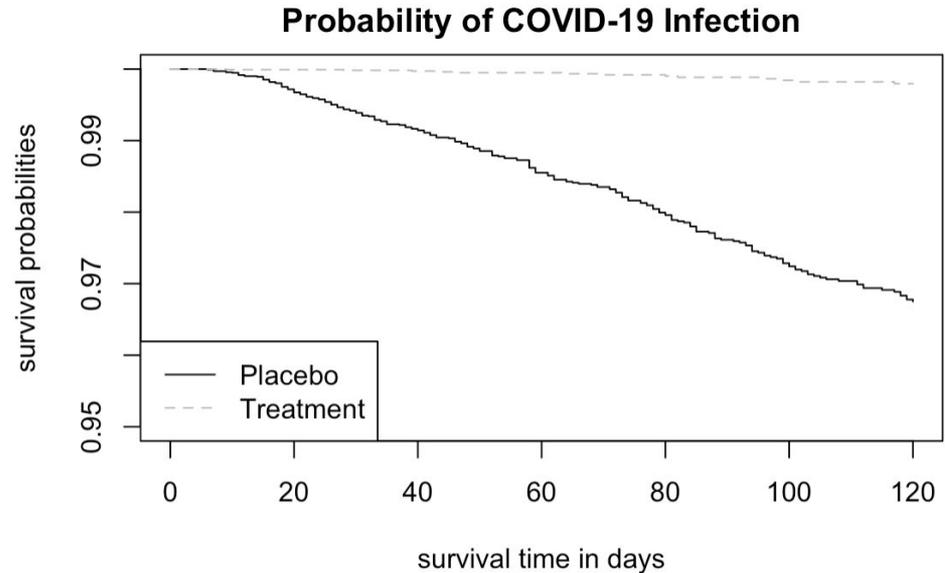
Cox-PH Model

```
             coef exp(coef) se(coef)      z Pr(>|z|)
arm    -2.92154   0.05385  0.28483 -10.26   <2e-16 ***
```

- Hazard Ratio: 0.05385. 95% CI: 0.03081 - 0.09411
- Vaccine efficacy rate ~ 95%, CI: ~ 91% - ~ 97%
- 95% reduction in cases, treatment

- P > 0.05, do not reject the assumption that the hazards are proportional.

```
        chisq df    p
arm      2.06  1 0.15
```

DRP Main Takeaways

- Not all data should be removed.
- Applicable to many fields, focus on medicine.
- Many methods and graphs which may relate to one another.
- Overview on how tests are conducted in the medical field.