Survival Analysis

Mingwei Fei

December 11, 2022

1 Sample Size

The $ln(HR)$ follows a normal distribution, we use this to calculate the sample size.

$$ln(\hat{\Delta}) \sim N\left(ln(\Delta), \frac{1}{d_1} + \frac{1}{d_2}\right)$$
$$\left(\frac{1}{d_1} + \frac{1}{d_2}\right)^{-1} = \left[\frac{(z_{\alpha/2} + z_\beta)^2}{(ln\Delta_0)^2}\right]$$

where d_i is the number of observed events.

If hazard ratio set at 2.1, then

$$\left(\frac{1}{d_1} + \frac{1}{d_2})\right)^{-1} = \left[\frac{(1.96 + 0.842)^2}{(ln2.1)^2}\right] = 14.26$$
$$\frac{1}{d_1} + \frac{1}{d_2} = \frac{1}{11.7} = 0.07, \qquad d_1 = d_2 = 28.5$$

The one-sided significance level 0.25, power is 0.8. Note that $Z_{\alpha/2}$ is the z score for the probability $1 - \alpha/2$, and z_β is the z score for the probability $1 - \beta$. Assume the overall event and censored rate is 20%, then the sample size is $57/0.2 = 285$. The total number in the paper is 276.

1.1 Non-inferiority margin Hazard ratio $\Delta_0 = 2.1$

The assumption is that control group (C) event rate 10% and treatment group (T) event rate 20% at 6 months. Assume survival function is an exponential distribution:

$$S_t(t) = exp(-\lambda_1 t), \qquad t = 0.5, S_t = 0.8, -\lambda_1 = ln(0.8)/0.5$$
$$S_c(t) = exp(-\lambda_2 t), \qquad t = 0.5, S_c = 0.9, -\lambda_2 = ln(0.9)/0.5$$
$$\Delta_0 = \frac{\lambda_1}{\lambda_2} = \frac{ln(0.8)}{ln(0.9)} = 2.117$$

1.2 Hazard ratio actual = 0.55

The control group survival 76.8% and treatment group survival 86.2% at 6 months. Assume survival function is an exponential distribution:

$$
\begin{aligned}
S_t(t) &= exp(-\lambda_1 t), \qquad t = 0.5, S_t = 0.862, -\lambda_1 = ln(0.862)/0.5 \\
S_c(t) &= exp(-\lambda_2 t), \qquad t = 0.5, S_c = 0.768, -\lambda_2 = ln(0.768)/0.5 \\
HR &= \frac{\lambda_1}{\lambda_2} \\
&= \frac{ln(0.862)}{ln(0.768)} = 0.56
\end{aligned}
$$

2 Sample Size Formula

The test hypothesis is

$$
\begin{aligned}
H_0 :& \lambda_1 = \lambda_2 \\
H_1 :& \lambda_1 \neq \lambda_2
\end{aligned}
$$

Or equivalently, in terms of hazard ratio, $\Delta = \lambda_1/\lambda_2$

$$
\begin{aligned}
H_0 :& \Delta = 1 \\
H_1 :& \Delta \neq 1
\end{aligned}
$$

A much simpler and quite accurate approximation for a reasonably large number of events is based on the approximate normality of th natural logarithm of the estimated hazard ratio in each treatment group:

$$ln(\hat{\lambda}_i) \sim N(ln\lambda_i, \frac{1}{d_i})$$

where d_i is the number of observed events. Thus, the $ln\Delta = ln\lambda_1 - ln\lambda_2$ also follows a normal distribution with variance $\frac{1}{d_1} + \frac{1}{d_2}$.

$$
\begin{aligned}
ln(\hat{\Delta}) &\sim N\left(ln(\Delta), \frac{1}{d_1} + \frac{1}{d_2}\right) \\
\left(\frac{1}{d_1} + \frac{1}{d_2}\right)^{-1} &= \left[\frac{(z_{\alpha/2} + z_\beta)^2}{(ln\Delta_0)^2}\right]
\end{aligned}
$$

The calculation of sample size follows

$$
\begin{aligned}
Z &= \frac{ln(\hat{\Delta})}{\sigma}, \qquad \sigma = \sqrt{\frac{1}{d_1} + \frac{1}{d_2}}, \qquad \delta = ln(\Delta_0) \\
&P(Z \geq Z_{1-\alpha/2} | H_0) \leq \alpha/2 \\
&P(Z \leq Z_\beta | H_1 = \delta) \geq \beta
\end{aligned}
$$

So we set Z satisfy the below equation

$$\frac{ln(\hat{\Delta})}{\sigma} = Z_{1-\alpha/2}, \qquad \text{H }_0$$

$$\frac{ln(\hat{\Delta}) - \delta}{\sigma} = Z_{\beta}, \qquad \text{H}_1$$

So we have

$$ln(\hat{\Delta}) = Z_{1-\alpha/2}\sigma, \qquad ln(\hat{\Delta}) = Z_{\beta}\sigma + \delta, \qquad Z_{1-\alpha/2}\sigma = Z_{\beta}\sigma + \delta$$

$$\sigma = \frac{\delta}{Z_{1-\alpha/2} - Z_{\beta}}, \qquad \frac{1}{d_1} + \frac{1}{d_2} = \frac{\delta^2}{(Z_{1-\alpha/2} + Z_{1-\beta})^2}$$