Survival Analysis in Clinical Trials

Michael Martens, PhD

October 27, 2022

Survival Analysis

- Interest lies in the time to an event, T
- Ex: Time to death, disease relapse, discharge from hospital
- Features of survival analysis:
 - Distribution of T is typically skewed (not normal)
 - Presence of censoring: Actual event time may not be observed
 - In clinical trial setting, most censoring is administrative (end of study censoring)
 - Some censoring may result from patient withdrawal or loss to follow-up, but should be infrequent

Basic Quantities of Interest

• **Survival function** S(t): Probability of surviving past time t,

$$S(t) = P(T > t) = 1 - F(t),$$

where F(t) is the cdf of T

• Hazard function $\lambda(t)$: Failure rate at time t among survivors,

$$\lambda(t) = \lim_{s \downarrow 0} \frac{P(t \le T \le t + s | T \ge t)}{s}$$

 Interpretation: Instantaneous failure "probability" at time t among survivors

Basic Quantities of Interest

 One-to-one relationship exists between hazard function and survival distribution:

$$\lambda(t) = \frac{f(t)}{S(t)} = \frac{-S'(t)}{S(t)} = \frac{-d[\log(S(t))]}{dt},$$

where f(t) is the pdf of T

Cumulative hazard function:

$$\Lambda(t) = \int_0^t \lambda(u) du = -\log[S(t)],$$

which relates to the survival function through

$$S(t) = \exp[-\Lambda(t)]$$

- Kaplan-Meier Estimator: Nonparametric estimator $\widehat{S}(t)$ of survival function
- Construction:
 - Divide interval [0, t] into subintervals $[0, t] = \bigcup_j [u_j, u_{j+1}]$
 - Probability of surviving past time t is the product of the chances of surviving in each successive interval
 - Probability of surviving interval $[u_j, u_{j+1}) = 1 Q_j$, where

$$Q_j = rac{\# ext{ who died in } [u_j, u_{j+1})}{\# ext{ at risk in } [u_j, u_{j+1})}$$



- Suppose the events occur at times $t_1 < t_2 < \cdots < t_K$
 - d_j = the number dying at time t_j
 - $n_j =$ the number still at risk (alive and on study) just before time t_j
- Then for t in the interval $[t_k, t_{k+1})$,

$$\widehat{S}(t) = \prod_{j=1}^{k} (1 - Q_j) = \prod_{j=0}^{k} \left(1 - \frac{d_j}{n_j}\right) = \prod_{j=0}^{k} \frac{n_j - d_j}{n_j},$$

where $t_0 = 0$

• $\widehat{S}(t)$ only changes at values of t where a death occurs, so is a step function

 Variance of K-M estimator can be estimated by Greenwood's formula,

$$\widehat{Var}[\widehat{S}(t)] = \widehat{S}(t)^2 \sigma_s^2$$
, where $\sigma_s^2 = \sum_{j=1}^k \frac{d_j}{n_j(n_j - d_j)}$

• For large n, $\widehat{S}(t)$ approximately follows the normal distribution $Normal(S(t),\widehat{Var}[\widehat{S}(t)])$

- Confidence intervals for S(t) can be constructed using Greenwood's formula:
 - Linear Wald CI:

$$\widehat{S}(t) \pm z_{1-lpha/2} \sqrt{\widehat{Var}[\widehat{S}(t)]}$$

 Complementary log-log transformed CI gives better coverage for small n:

$$[\widehat{S}(t)^{1/ heta},\widehat{S}(t)^{ heta}], \quad ext{where } heta = \exp\left[rac{z_{1-lpha/2}\sigma_{ extsf{s}}(t)}{\log \widehat{S}(t)}
ight]$$

Example

- Randomized clinical trial comparing two treatments for psoriasis:
 - PUVA therapy (standard): methoxsalen treatment given, followed by exposure to UVA (longer wavelength UV rays)
 - TL-01 therapy (experimental): lamp for administering shorter wavelength UVB radiation
- Outcome: Number of visits to phototherapy clinic until psoriasis cleared

Example Using SAS

```
data a; set d.puvavstl01trial; run;
ods graphics on;
proc lifetest data=a plots=(s) alpha=0.05;
time visits*clearance(0);
strata group;
survival out=survci conftype=loglog;
run;
proc print data=survci; run;
```

The LIFETEST Procedure Stratum 1: GROUP = PUVA

Product-Limit Survival Estimates

Survival Standard Number Number VISITS Survival Failure Failed Left. Error 0.0000 1.0000 49 2,0000 0.9796 0.0204 0.0202 48 2.0000* 47 3.0000* 46 6.0000 0.9583 0.0417 0.0289 45 6.0000* 44 7.0000 3 43 7.0000 42 7.0000 0.8930 0.1070 0.0453 41 8.0000 6 40 8.0000 0.1506 0.0525 0.8494 39

Obs	GROUP	VISITS	_CENSOR_	SURVIVAL	CONFTYPE	SDF_LCL	SDF_UCL	STRATUM
1	PUVA	0		1.00000		1.00000	1.00000	1
2	PUVA	2	0	0.97959	LOGLOG	0.86383	0.99710	1
3	PUVA	2	1	0.97959				1
4	PUVA	3	1	0.97959				1
5	PUVA	6	0	0.95830	LOGLOG	0.84328	0.98941	1
6	PUVA	6	1	0.95830				1
7	PUVA	7	0	0.89296	LOGLOG	0.76158	0.95403	1
8	PUVA	8	0	0.84940	LOGLOG	0.70970	0.92525	1
9	PUVA	9	0	0.76228	LOGLOG	0.61182	0.86074	1
10	PUVA	10	0	0.71872	LOGLOG	0.56523	0.82596	1
11	PUVA	11	0	0.65338	LOGLOG	0.49778	0.77133	1
12	PUVA	12	0	0.63160	LOGLOG	0.47588	0.75252	1
13	PUVA	13	0	0.58805	LOGLOG	0.43293	0.71410	1
14	PUVA	13	1	0.58805				1
15	PUVA	14	1	0.58805		•	•	1
16	PUVA	15	0	0.54100	LOGLOG	0.38663	0.67224	1

- Let $S_i(t)$ be the survival function for treatment group i
- Null hypothesis: $H_0: S_1(t) = S_2(t)$ for all t
- Equivalently, $H_0: \lambda_1(t) = \lambda_2(t)$ for all t

- Log-rank test can be used to compare survival curves:
 - Construct a 2x2 table at each event time t_j

Group	# failing	# surviving	# at risk
1	d_{1j}	$n_{1j}-d_{1j}$	n_{1j}
2	d_{2j}	$n_{2j}-d_{2j}$	n _{2j}
Total	d_j	$n_j - d_j$	nj

• Under H_0 the expected number of failures in group 1 is

$$e_{1j} = n_{1j} \times (d_j/n_j)$$

 Using a Cochran-Mantel-Haenszel test, combine tables by stratifying on t_i's

• Deaths at each time are independent, so d_{1j} has a hypergeometric distribution with mean e_{1j} and variance

$$V_j = \frac{n_{1j}n_{2j}d_j(n_j - d_j)}{n_j^2(n_j - 1)}$$

• Under H_0 , it can be shown that for large samples,

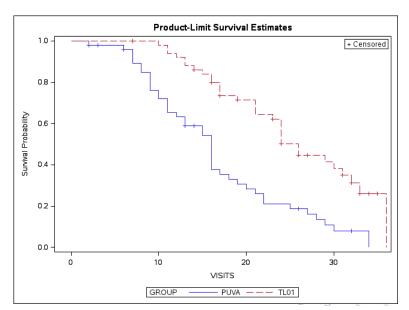
$$U=\sum_j (d_{1j}-e_{1j}) \sim \mathit{N}\Big(0,V=\sum_j V_j\Big)$$
 and so $C=U^2/V\sim \chi_1^2$

 Log-rank test is most powerful when the hazard rates of the two groups are proportional, e.g.

$$\frac{\lambda_1(t)}{\lambda_2(t)} = \eta \text{ for all } t$$

 Loss of power occurs when hazard functions or survival curves cross; better to use alternative test in that case (e.g. weighted log rank test, pointwise comparison of survival curves)

Comparing Two Survival Curves: Example



Test of Equality over Strata

			Pr >
Test	Chi-Square	DF	Chi-Square
Log-Rank	18.9786	1	<.0001
Wilcoxon	20.5522	1	<.0001
-2Log(LR)	8.2238	1	0.0041

- Want to test $H_0: S_1(t) = S_2(t)$ for all t
- Log rank test is most powerful when hazard rates are proportional, i.e. hazard ratio $\eta = \lambda_1(t)/\lambda_2(t)$ is constant
- Under proportionality, can equivalently write null hypothesis as $H_0: \eta = 1$, or $H_0: \theta = 0$ where $\theta = \log(\eta)$

• Under H_0 and for large samples,

$$U = \sum_{j} (d_{1j} - e_{1j}) \sim N(0, V = \sum_{j} V_{j}),$$

where

$$V_j = \frac{n_{1j}n_{2j}d_j(n_j - d_j)}{n_j^2(n_j - 1)}$$

- Under $H_1: \theta = \theta_1, \ U \sim N(\theta_1 V, V)$
- Reject H_0 if $|Z| = |U/\sqrt{V}| > z_{1-\alpha/2}$

• The power to detect $\theta = \theta_1$ is

$$\begin{split} \Pi(\theta_1) &= P\bigg(\frac{U}{\sqrt{V}} > z_{1-\alpha/2} \bigg| \theta = \theta_1 \bigg) + P\bigg(\frac{U}{\sqrt{V}} < -z_{1-\alpha/2} \bigg| \theta = \theta_1 \bigg) \\ &= P\bigg(\frac{U - \theta_1 V}{\sqrt{V}} > z_{1-\alpha/2} - \frac{\theta_1 V}{\sqrt{V}} \bigg| \theta = \theta_1 \bigg) \\ &+ P\bigg(\frac{U - \theta_1 V}{\sqrt{V}} < -z_{1-\alpha/2} - \frac{\theta_1 V}{\sqrt{V}} \bigg| \theta = \theta_1 \bigg) \\ &= P(N(0, 1) > z_{1-\alpha/2} - \theta_1 \sqrt{V}) + P(N(0, 1) < -z_{1-\alpha/2} - \theta_1 \sqrt{V}) \\ &= 1 - \Phi\bigg(z_{1-\alpha/2} - \theta_1 \sqrt{V}\bigg) + \Phi\bigg(-z_{1-\alpha/2} - \theta_1 \sqrt{V}\bigg) \\ &\approx \begin{cases} 1 - \Phi\bigg(z_{1-\alpha/2} - \theta_1 \sqrt{V}\bigg) & \text{if } \theta_1 > 0 \\ \Phi\bigg(-z_{1-\alpha/2} - \theta_1 \sqrt{V}\bigg) & \text{if } \theta_1 < 0 \end{cases} \end{split}$$

Either case implies

$$-z_{1-eta} = z_{1-lpha/2} \pm heta_1 \sqrt{V}, ext{ or } V = rac{(z_{1-lpha/2} + z_{1-eta})^2}{ heta_1^2}$$

- Given a hazard ratio, we can determine the variance V needed for the log-rank test to have (1β) power
- Need to convert that to something more usable (sample size)
- When the number of deaths is small relative to the number at risk $(d_j \text{ vs. } n_j)$, then

$$V = \sum_{j} \frac{n_{1j} n_{2j} d_j (n_j - d_j)}{n_j^2 (n_j - 1)} \approx \sum_{j} \frac{n_{1j} n_{2j} d_j}{n_j^2}$$

• If θ is small and recruitment to each group is balanced then $n_{1j} \approx n_{2j} \approx n_j/2$, and so

$$V \approx \sum_{j} \frac{d_{j}}{4} = d/4,$$

where d is the total number of deaths

• The total number of deaths needed to have $(1 - \beta)$ power to detect a log hazard ratio of θ_1 is

$$d = \frac{4(z_{1-\alpha/2} + z_{1-\beta})^2}{\theta_1^2}$$

 Power for the log-rank test depends on the number of deaths observed



Converting from Survival Differences to Hazard Ratio

- Clinicians often specify treatment effects in terms of improvements in survival probabilities at a fixed time point
- New treatment is expected to increase the survival proportion at time t from $S_2(t)$ under the standard treatment to $S_1(t)$ under the new treatment

Converting from Survival Differences to Hazard Ratio

• Assuming proportional hazards, $\lambda_1(t) = \eta \lambda_2(t)$:

$$S_1(t) = \exp\left[-\int_0^t \lambda_1(u)du\right] = \exp\left[-\int_0^t \eta \lambda_2(u)du\right]$$
$$= \left\{\exp\left[-\int_0^t \lambda_2(u)du\right]\right\}^{\eta}$$
$$= S_2(t)^{\eta}$$

Then

$$\eta = \frac{\log S_2(t)}{\log S_1(t)}$$

• Note that if $\eta > 1$, then $S_1(t) < S_2(t)$ for all t

Example

- Patients suffering from chronic active hepatitis rapidly progress to early death from liver failure
- Interested in designing a trial to compare a new treatment to standard treatment
- New treatment is expected to increase the overall survival probability at 5 years from 35% under the standard treatment to 55%
- Conversion to hazard ratio:

$$\eta = \frac{\log(0.55)}{\log(0.35)} = 0.569$$

Example

 Total number of deaths required to have 90% power to detect this difference in survival curves using a 0.05 level log-rank test:

Using
$$\alpha=0.05$$
, $1-\beta=0.90$, $z_{\alpha/2}=1.96$, $z_{\beta}=1.28$, and $\theta=\log(0.569)=-0.564$,

$$d = \frac{4(1.96 + 1.28)^2}{(-0.564)^2} = 132.005 \approx 133$$

 This is the total number of deaths needed in the entire study, not deaths per treatment arm

Unequal Sample Sizes

• If $n_1=rn_2$, then $n_{1j}\approx rn_{2j}$, and

$$\frac{n_{1j}n_{2j}}{(n_{1j}+n_{2j})^2} \approx \frac{rn_{2j}^2}{(rn_{2j}+n_{2j})^2} = \frac{r}{(r+1)^2}$$

Variance of the log-rank test statistic

$$V \approx \frac{r}{(r+1)^2} \cdot d$$

Required number of deaths is

$$d = \frac{(r+1)^2}{r} \frac{(z_{1-\alpha/2} + z_{1-\beta})^2}{\theta_1^2}$$

- The accrual pattern affects the follow-up time for patients
- Simple accrual pattern: All patients are enrolled instantaneously and followed for fixed time t₀
- If each arm enrolls n patients, expected number of deaths is

$$E(d) = n(1 - S_1(t_0)) + n(1 - S_2(t_0))$$

To obtain a targeted number of deaths, we should enroll

$$n = \frac{d}{2 - S_1(t_0) - S_2(t_0)}$$
 per arm

• Also useful when we don't want to make assumptions about survival curves beyond time t_0 (i.e. ignore follow-up after that which is minimally required)

Example

- We want 90% power to detect an improvement in 5 year survival from 35% to 55%
- Last example determined the target number of deaths is d = 133
- Using a simple accrual pattern where all patients are followed for 5 years regardless of when they enroll,

$$n = \frac{133}{2 - 0.55 - 0.35} = 120.9 \approx 121$$

Need to enroll 121 patients per arm, 242 total

- Staggered entry of patients into study
 - Accrual period, A
 - Minimum follow-up on last patient, F
- Some trials will keep following patients until either they have an event, or trial duration ends
 - Patient's follow-up depends on F and their time of enrollment during the accrual period
 - Enrolled at beginning of trial / accrual period: Follow-up time = F + A
 - Enrolled at end of accrual period: Follow-up time = F

- Required number of patients depends on assumptions about
 - Accrual rate of patients
 - Survival curve for entire study period [0, L = F + A]
- Accrual rate at calendar time u in persons per year

$$a(u) = \lim_{du \to 0} \frac{\text{Expected } \# \text{ patients enrolled in (u,u+du)}}{du}$$

- If $a(u) \equiv a$ we have a constant (uniform) accrual rate
- Total expected number of patients:

$$\int_0^A a(u)du$$

- Survival distribution for treatment group i is $S_i(u)$
- Accrual assumed to be independent of treatment assignment
- Expected number of deaths for treatment group i is

$$E(d_i) = \int_0^A \frac{a(u)}{2} [1 - S_i(L - u)] du$$

To get the desired power, we set

$$E(d_1) + E(d_2) = \frac{4(z_{1-\alpha/2} + z_{1-\beta})^2}{\theta_1^2}$$

- Special case: Exponential survival distribution with parameter λ_i , and uniform accrual
 - Survival distribution for treatment group i

$$S_i(u) = e^{-\lambda_i u}$$

Expected number of deaths for treatment group i

$$E(d_i) = \int_0^A \frac{a}{2} [1 - e^{-\lambda_i (L-u)}] du$$
$$= \frac{a}{2} \Big[A - \frac{e^{-\lambda_i L}}{\lambda_i} (e^{\lambda_i A} - 1) \Big]$$

Example

- We want 90% power to detect an improvement in 5 year survival from 35% to 55%
- F = 5 years of follow-up on last patient
- Accrue uniformly for 3 years; 8 year study duration in total
- Assume exponential survival curves
 - Group 1: $S_1(5) = \exp[-\lambda_1(5)] = 0.35$ so

$$\lambda_1 = \frac{-\log 0.35}{5} = 0.21$$

• Group 2: $S_2(5) = \exp[-\lambda_2(5)] = 0.55$ so

$$\lambda_2 = \frac{-\log 0.55}{5} = 0.12$$

Example

- Targeted number of deaths (from before) is $d_1 + d_2 = 133$
- Expected number of deaths for treatment group 1

$$E(d_1) = \frac{a}{2} \left[3 - \frac{e^{-(0.21)(8)}}{0.21} \left\{ e^{(0.21)(3)} - 1 \right\} \right] = 1.11a$$

Expected number of deaths for treatment group 2

$$E(d_2) = \frac{a}{2} \left[3 - \frac{e^{-(0.12)(8)}}{0.12} \{ e^{(0.12)(3)} - 1 \} \right] = 0.81a$$

Example

Then set

$$E(d_1) + E(d_2) = a(1.11 + 0.81) = 1.92a = 133$$

so a = 69.3 patients per year

- Using 3 years of accrual, we would enroll n = 3a = 208 patients total or 104 patients per arm
- This design requires fewer patients than previous design, where all patients are followed for 5 years maximum
 - Some patients will have more than 5 years of follow-up
 - Greater likelihood of death, so fewer patients needed to get the targeted number of deaths

Example

- Alternative design: A = L = 5 (continue accruing patients until the end of the study)
- Expected number of deaths for treatment group 1

$$E(d_1) = \frac{a}{2} \left[5 - \frac{e^{-(0.21)(5)}}{0.21} \left\{ e^{(0.21)(5)} - 1 \right\} \right]$$
$$= 0.95a$$

Expected number of deaths for treatment group 2

$$E(d_2) = \frac{a}{2} \left[5 - \frac{e^{-(0.12)(5)}}{0.12} \left\{ e^{(0.12)(5)} - 1 \right\} \right]$$
$$= 0.62a$$

Example

Then set

$$E(d_1) + E(d_2) = a(0.952 + 0.62) = 1.57a = 133,$$

so twidehat a = 84.7 patients per year

- Using 5 years of accrual, we would enroll n = 5a = 424 patients total or 212 patients per arm
- Requires much larger sample size (424 vs. 208), but shorter study time (5 years vs. 8 years)

Estimating the Number of Patients

- More generally, numerical solutions may be required when the distribution is not exponential or when accrual is not uniform
- Be careful of exponential assumption:
 - Must be valid for entire study duration L = F + A
 - Unrealistic for many settings (ex. transplant, surgery, cancer therapies)

SAS Examples Using PROC POWER

- Active Hepatitis Example: 90% power to detect an improvement in 5 year survival from 35% on the standard treatment to 55% on the new treatment
 - Follow each patient for 5 years only, ignore staggered accrual: n = 122 per arm
 - Accrue for 3 years, follow-up for 5 years, assume exponential survival distributions: n = 104 per arm
 - Accrue for 2 years, follow-up for 4 years, assume exponential survival distributions: n = 121 per arm
 - Accrue for 2 years, follow-up for 4 years. Assume piecewise linear survival functions:
 - Control: S(2) = 0.7, S(4) = 0.45, S(6) = 0.25
 - Treatment: S(2) = 0.82, S(4) = 0.63, S(6) = 0.45
 - n = 131 per arm

Follow each patient for 5 years only, ignore staggered accrual: n = 122 per arm

```
proc power;
twosamplesurvival test=logrank
sides=2
alpha=0.05
curve("Standard")=5:0.35
curve("Proposed")=5:0.55
groupsurvival="Standard"|"Proposed"
accrualtime=0
followuptime=5
npergroup=.
power=0.9;
run:
```

Log-Rank Test for Two Survival Curves

Fixed Scenario Elements

Method	Lakatos normal approximation
Number of Sides	2
Accrual Time	0
Follow-up Time	5
Alpha	0.05
Group 1 Survival Curve	Standard
Form of Survival Curve 1	Exponential
Group 2 Survival Curve	Proposed
Form of Survival Curve 2	Exponential
Nominal Power	0.9
Number of Time Sub-Intervals	12
Group 1 Loss Exponential Hazard	0
Group 2 Loss Exponential Hazard	0

${\tt Computed}\ {\tt N}\ {\tt Per}\ {\tt Group}$

Actual N Per Power Group 0.901 122

Accrue for 3 years, follow-up for 5 years, assume exponential survival distributions: n = 104 per arm

```
proc power;
twosamplesurvival test=logrank
sides=2
alpha=0.05
curve("Standard")=5:0.35
curve("Proposed")=5:0.55
groupsurvival="Standard"|"Proposed"
accrualtime=3
followuptime=5
npergroup=.
power=0.9;
run:
```

Log-Rank Test for Two Survival Curves

Fixed Scenario Elements

Method	Lakatos normal approximation
Number of Sides	2
Accrual Time	3
Follow-up Time	5
Alpha	0.05
Group 1 Survival Curve	Standard
Form of Survival Curve 1	Exponential
Group 2 Survival Curve	Proposed
Form of Survival Curve 2	Exponential
Nominal Power	0.9
Number of Time Sub-Intervals	12
Group 1 Loss Exponential Hazard	0
Group 2 Loss Exponential Hazard	0

${\tt Computed}\ {\tt N}\ {\tt Per}\ {\tt Group}$

Actual N Per Power Group 0.902 104

Accrue for 2 years, follow-up for 4 years, assume exponential survival distributions: n = 121 per arm

```
proc power;
twosamplesurvival test=logrank
sides=2
alpha=0.05
curve("Standard")=5:0.35
curve("Proposed")=5:0.55
groupsurvival="Standard"|"Proposed"
accrualtime=2
followuptime=4
npergroup=.
power=0.9;
run:
```

Log-Rank Test for Two Survival Curves

Fixed Scenario Elements

Method	Lakatos normal approximation
Number of Sides	2
Accrual Time	2
Follow-up Time	4
Alpha	0.05
Group 1 Survival Curve	Standard
Form of Survival Curve 1	Exponential
Group 2 Survival Curve	Proposed
Form of Survival Curve 2	Exponential
Nominal Power	0.9
Number of Time Sub-Intervals	12
Group 1 Loss Exponential Hazard	0
Group 2 Loss Exponential Hazard	0

${\tt Computed}\ {\tt N}\ {\tt Per}\ {\tt Group}$

Actual N Per Power Group 0.901 121

Accrue for 2 years, follow-up for 4 years. Assume piecewise linear survival functions

```
proc power;
twosamplesurvival test=logrank
sides=2
alpha=0.05
curve("Standard")=(2 to 6 by 2):(0.7 0.45 0.25)
curve("Proposed")=(2 to 6 by 2):(0.82 0.63 0.45)
groupsurvival="Standard"|"Proposed"
accrualtime=2
followuptime=4
npergroup=.
power=0.9;
run:
```

Log-Rank Test for Two Survival Curves

Fixed Scenario Elements

Method	Lakatos normal approximation
Number of Sides	2
Accrual Time	2
Follow-up Time	4
Alpha	0.05
Group 1 Survival Curve	Standard
Form of Survival Curve 1	Piecewise Linear
Group 2 Survival Curve	Proposed
Form of Survival Curve 2	Piecewise Linear
Nominal Power	0.9
Number of Time Sub-Intervals	12
Group 1 Loss Exponential Hazard	0
Group 2 Loss Exponential Hazard	0

${\tt Computed}\ {\tt N}\ {\tt Per}\ {\tt Group}$

Actual N Per Power Group 0.900 131