Power and Sample Size Calculation for Log-rank Test under a Non-proportional Hazards Model*

Daowen Zhang

Department of Statistics

North Carolina State University

zhang@stat.ncsu.edu

http://www4.stat.ncsu.edu/∼dzhang2/

* Joint work with Hui Quan, Department of Biostatistics & Programming, Sanofi-Aventis

OUTLINE

1. **Motivating example: Rimonabant trial on cardiovascular risk**
2. **Review of the log-rank test statistic**
3. **Distributions of the log-rank test statistic**
4. **Detailed power calculation**
5. **Example and simulation Results**
6. **Summary**

1. Motivating example: Rimonabant trial

- Rimonabant trial: Assess the benefit of Rimonabant on reducing cardiovascular risk.
- Placebo-controlled
- Primary endpoint: time to cardiovascular event; event rates expected to be low in each group
- Log-rank test was proposed to assess the treatment effect.
- Power and sample size consideration should also be based on the log-rank test

- It is straightforward if treatment effect is characterized by

$$\frac{\lambda_1(t)}{\lambda_0(t)} = \mathrm{e}^{\beta},$$

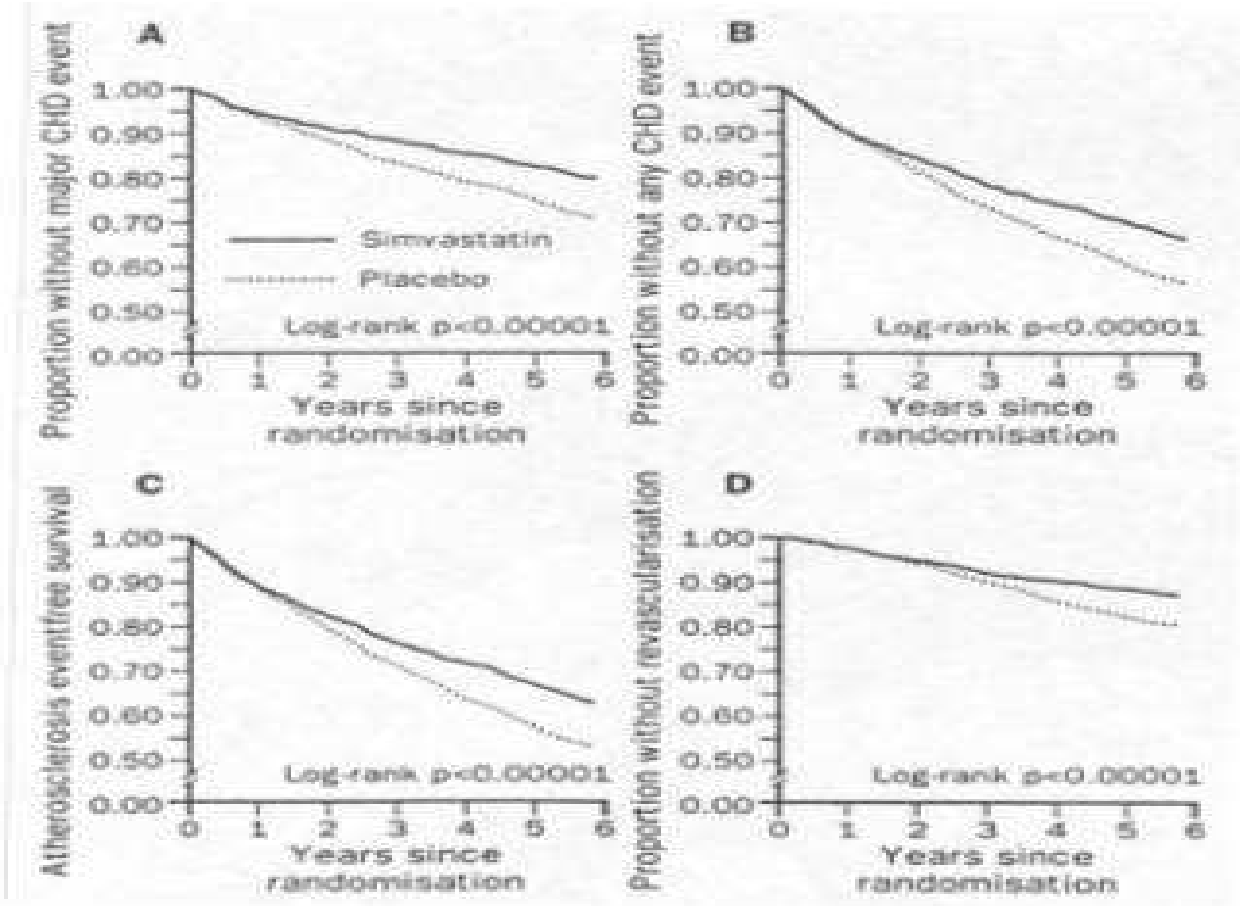
  $\lambda_1(t)$: hazard of cardiovascular event for treatment
  $\lambda_0(t)$: hazard of cardiovascular event for placebo

- If $\beta \approx 0$ and censoring process independent of treatment group, log-rank test statistic T has distribution (Schoenfeld, 1981 *Biometrika*)

$$T \overset{a}{\sim} N(\beta\sqrt{\theta(1-\theta)D}, 1),$$

  θ: Allocation probability to the treatment
  D: Expected total # of deaths (under H_a) from both groups

- Can be used to calculate power and sample size if the treatment effect model (PH model) is reasonable.
- However ...


A
Proportion without major CHD event
1.00
0.90
0.80
0.70
0.60
0.50
0.00
Simvastatin
Placebo
Log-rank p<0.00001
0 1 2 3 4 5 6
Years since randomisation
B
Proportion without any CHD event
Log-rank p<0.00001
Years since randomisation
C
Atherosclerosis event-free survival
Log-rank p<0.00001
Years since randomisation
D
Proportion without revascularisation
Log-rank p<0.00001
Years since randomisation

- Other issues:
  1. censoring (information cannot be retrieved)
  2. drop-out (information can be retrieved during the study)
- How to handle drop-out?
  1. treat it as censoring: assumption?
  2. conduct ITT analysis: efficiency loss?
- Problem: how to calculate power and sample size for each strategy? which is better?
- Need to investigate the distribution of the log-rank test statistic for our problem

2. Review of the log-rank test statistic

- The (standard) log-rank test statistic

$$T = \frac{U}{\sqrt{\widehat{\mathrm{var}}(U)}},$$

where

$$U = \sum_x \left\{ d_1(x) - n_1(x) \frac{d(x)}{n(x)} \right\}$$

$$\widehat{\mathrm{var}}(U) = \sum_x \frac{n_1(x) n_0(x) d(x) \{n(x) - d(x)\}}{n^2(x) \{n(x) - 1\}}$$

- Under $H_0 : S_1(t) = S_0(t) \Longleftrightarrow H_0 : \lambda_1(t) = \lambda_0(t)$,

$$T \overset{a}{\sim} N(0,1)$$

So reject H_0 if $|T| \geq z_{\alpha/2}$.

- Under $H_a : \lambda_1(t) \neq \lambda_0(t)$ (but $\lambda_1(t) \approx \lambda_0(t)$) (Schoenfeld, 1981 *Biometrika*)

$$T \overset{a}{\sim} N(\phi, 1),$$

where

$$\phi = \frac{\sqrt{n}\int_0^\infty log\{\lambda_1(t)/\lambda_0(t)\}\pi(t)\{1-\pi(t)\}V(t)dt}{[\int_0^\infty \pi(t)\{1-\pi(t)\}V(t)dt]^{1/2}},$$

where $V(t)$ describes process of observing deaths, $\pi(t) \longrightarrow \theta$ if censoring process is the same in both groups.

- Special case: PH alternative

$$H_a : \frac{\lambda_1(t)}{\lambda_0(t)} = \mathrm{e}^{\beta} \quad (\beta \approx 0),$$

  then

$$T \overset{a}{\sim} N(\beta\sqrt{\theta(1-\theta)D}, 1),$$

- Can be used to calculate the power for PH alterative.

3. Distribution of the log-rank test statistic

- It is reasonable to assume the alternative for our problem:

$$H_a : \frac{\lambda_1(t)}{\lambda_0(t)} = \begin{cases} 1 & t \in [0, t_0) \\ \mathrm{e}^{\beta} \quad (\beta \approx 0) & t \in [t_0, \infty) \end{cases}$$

  $\lambda_1(t)$ = hazard of treated group
  $\lambda_0(t)$ = hazard of untreated group

- Distributions of the log-rank test statistic under H_a for two strategies?

  1. Strategy 1: Treat drop-out as censoring
  2. Strategy 2: Conduct ITT analysis

Distribution for Strategy 1

- Direct use of the result of Schoenfeld, 1981 (*Biometrika*) $\Longrightarrow$

$$T \overset{a}{\sim} N(\phi, 1),$$

$$\begin{aligned} \phi &\approx \frac{\sqrt{n}\beta \int_{t_0}^{\infty} \pi(t)\{1-\pi(t)\}V(t)dt}{[\int_0^{\infty} \pi(t)\{1-\pi(t)\}V(t)dt]^{1/2}} \\ &\approx \beta\sqrt{\theta(1-\theta)} \times \frac{\tilde{D}}{\sqrt{D}}, \end{aligned}$$

  D = total expected # of deaths from two groups in the study
  $\tilde{D}$ = total expected # of deaths from two groups after t_0.

- Power $= P[Z > |\phi| - z_{\alpha/2}]$.
- Concern: approximation good enough? better one?

- The use of a series of double expectation theorem leads to

$$\phi \quad \approx \quad \sqrt{\theta(1-\theta)} \times \frac{(1-e^{-\beta})\tilde{D}_1 + (e^{\beta}-1)\tilde{D}_0}{\sqrt{D}}$$

  $\tilde{D}_1 =$ total # of deaths from treated group after t_0
  $\tilde{D}_0 =$ total # of deaths from untreated group after t_0

- Assumption: drop-out independent of the (unerlying) survival time had the patient not dropped out; the same in both groups.

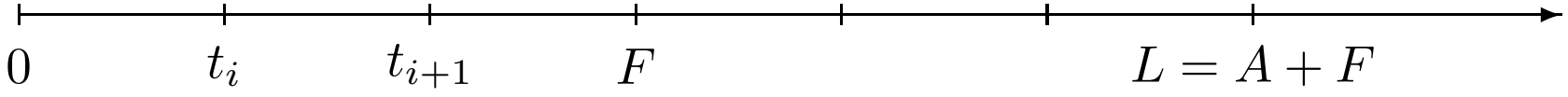
- Let

  $D_1 =$ total expected # of deaths from treated group
  $D_0 =$ total expected # of deaths from untreated group
  $D_1^* =$ total expected # of deaths from treated group before t_0
  $D_0^* =$ total expected # of deaths from placebo group before t_0

$$D = D_0 + D_1, \quad \tilde{D}_1 = D_1 - D_1^*, \quad \tilde{D}_0 = D_0 - D_0^*$$

Distribution for Strategy 2

- Lakatos (1988, *Biometrics*) derived an approx. dist. of the log-rank test under any $H_a : \lambda_1^*(t) \neq \lambda_0^*(t)$ ($\lambda_1^*(t) \approx \lambda_0^*(t)$).
- $\lambda_0^*(t)$ = hazard of the group randomized to placebo
  $\lambda_1^*(t)$ = hazard of the group randomized to treatment
- Partition patient time $[0, L = A + F) = \cup[t_i, t_{i+1})$ with equal width Δ.



A = accrual period, F = follow-up time, L = study length.

- Under $H_a : \lambda_1^*(t) \neq \lambda_0^*(t)$ $(\lambda_1^*(t) \approx \lambda_0^*(t))$:

$$T \overset{a}{\sim} N(\phi, 1),$$

$$\phi \approx \frac{\sum D_i \left\{ \frac{\xi_i p_i}{1+\xi_i p_i} - \frac{p_i}{1+p_i} \right\}}{\left\{ \sum D_i \frac{p_i}{(1+p_i)^2} \right\}^{1/2}}$$

1. $D_i = \{n_1(t_i)\lambda_1^*(t_i) + n_0(t_i)\lambda_0^*(t_i)\}\Delta$
   $=$ total expected # of deaths in $[t_i, t_{i+1})$

2. $\xi_i = \lambda_1^*(t_i)/\lambda_0^*(t_i)$

3. $p_i = n_1(t_i)/n_0(t_i)$

4. $n_0(t_i), n_1(t_i)$, number of patients at risk, can be calculated iteratively:

$$n_k(t_{i+1}) = \begin{cases} n_k(t_i)\{1 - \lambda_k^*(t_i)\Delta\} & t_i < F \\ n_k(t_i)\left\{1 - \lambda_k^*(t_i)\Delta - \frac{\Delta}{L - t_i}\right\} & t_i \geq F \end{cases}$$

Assume constant accrual rate in $[0, A]$.

- Need to know the hazard function for each (randomized) group.

- Assume $\lambda_0(t) = \lambda_0 \Longrightarrow \lambda_0^*(t) = \lambda_0$
- Assume drop-out process (has no effect on untreated group)

$$Z \sim exp(\tau)$$

- Then it is reasonable to assume $\lambda_1(t|Z)$ as
  1. Case 1, $Z \leq t_0$: $\lambda_1(t|Z) = \lambda_0$
  2. Case 2, $Z > t_0$:

$$\lambda_1(t|Z) = \begin{cases} \lambda_0 & t \in [0, t_0) \\ \lambda_1 & t \in [t_0, Z) \\ \tilde{\lambda}_1 & t \in [Z, \infty) \end{cases}$$

where $\tilde{\lambda}_1 \in [\lambda_1, \lambda_0]$; e.g.,

$$\tilde{\lambda}_1 = w\lambda_1 + (1 - w)\lambda_0.$$

- The survival function for group randomized to treatment:

$$\begin{aligned} S_1^*(t) &= \mathrm{E}\{I(T \geq t)\} \\ &= \mathrm{E}[\mathrm{E}\{I(T \geq t)|Z\}] \\ &= \mathrm{E}\{S_1(t|Z)\}. \end{aligned}$$

- Case 1, $Z < t_0$:

$$S_1(t|Z) = e^{-\lambda_0 t}$$

- Case 2: $Z \geq t_0$:

$$S_1(t|Z) = e^{-\Lambda_1(t|Z)} = \begin{cases} e^{-\lambda_0 t} & t \in [0, t_0) \\ e^{-\lambda_0 t_0 - \lambda_1(t-t_0)} & t \in [t_0, Z) \\ e^{-\lambda_0 t_0 - \lambda_1(Z-t_0) - \tilde{\lambda}_1(t-Z)} & t \in [Z, \infty) \end{cases}$$

- Can calculate $S_1^*(t)$ and $f_1^*(t)$ and hence

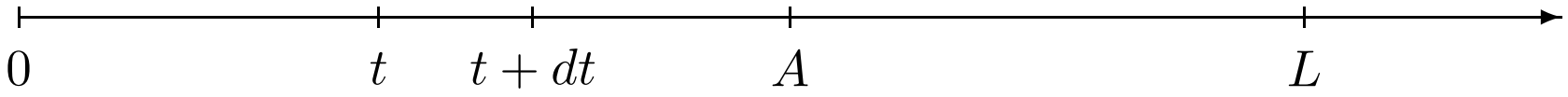
$$\lambda_1^*(t) = \frac{f_1^*(t)}{S_1^*(t)}.$$

- Then can calculate the nc ϕ in $N(\phi, 1)$ for the log-rank test.
- For better numerical accuracy, Δ needs to be small, say, 1/1000, if unit = year.

4. Detailed power calculation for strategy 1

- Some assumptions:
  1. Other than drop-out, end-of-study is the only other censoring (can be relaxed)
  2. $[0, A)$ is the accrual period, a = accrual rate (can be $a(t)$)
  3. F = follow-up period, $L = A + F$ = total study length
  4. $F \geq t_0$.
  5. $\lambda_0(t) = \lambda_0$.

- Consider $[t, t+dt)$ in $[0, A)$:



- Average # of patients entering into study in $[t, t+dt)$:

$$\begin{cases} \theta a dt & \text{treatment group} \\ (1-\theta) a dt & \text{placebo group} \end{cases} \tag{1}$$

- The probability that a patient entering at t is observed to die in the study (i.e., dies before L) is

$$P[T \leq \min(L - t, Z)]$$

- The probability that a patient entering at t is observed to die before t_0 is

$$P[T \leq \min(t_0, Z)]$$

- For placebo group:

$$P[T \leq \min(L-t, Z)] = \mathrm{E}[\mathrm{E}\{I[T \leq \min(L-t, Z)]|Z\}]$$

The inner expectation can be shown to be

$$\mathrm{E}\{I[T \leq \min(L-t, Z)]|Z\} = \begin{cases} 1 - e^{-\lambda_0(L-t)} & Z \geq L-t \\ 1 - e^{-\lambda_0 Z} & Z < L-t \end{cases}$$

$$\Longrightarrow$$

$$P[T \leq \min(L-t, Z)] = \frac{\lambda_0}{\lambda_0+\tau} - \frac{\lambda_0}{\lambda_0+\tau} e^{-(\lambda_0+\tau)(L-t)}$$

- The total expected # of deaths in the study for placebo group:

$$\begin{aligned} D_0 &= \int_0^A a(1-\theta)P[T \le \min(L-t, Z)]dt \\ &= \frac{a(1-\theta)\lambda_0}{\lambda_0+\tau}\left[A - \frac{e^{-(\lambda_0+\tau)L}}{\lambda_0+\tau}\{e^{(\lambda_0+\tau)A}-1\}\right]. \end{aligned}$$

- The total expected # of deaths for placebo group before t_0:

$$\begin{aligned} D_0^* &= \int_0^A a(1-\theta)P[T \le \min(t_0, Z)]dt \\ &= \frac{aA(1-\theta)\lambda_0}{\lambda_0+\tau}\{1 - e^{-(\lambda_0+\tau)t_0}\}. \end{aligned}$$

- For treatment group:

$$P[T \leq \min(L-t, Z)] = \left(\frac{\tau}{\lambda_0+\tau} - \frac{\tau}{\lambda_1+\tau}\right) e^{-(\lambda_0+\tau)t_0} + \frac{\lambda_0}{\lambda_0+\tau} - \frac{\lambda_1}{\lambda_1+\tau} e^{-(\lambda_0-\lambda_1)t_0-(\lambda_1+\tau)(L-t)}.$$

- The total expected # of deaths in the study for treatment group:

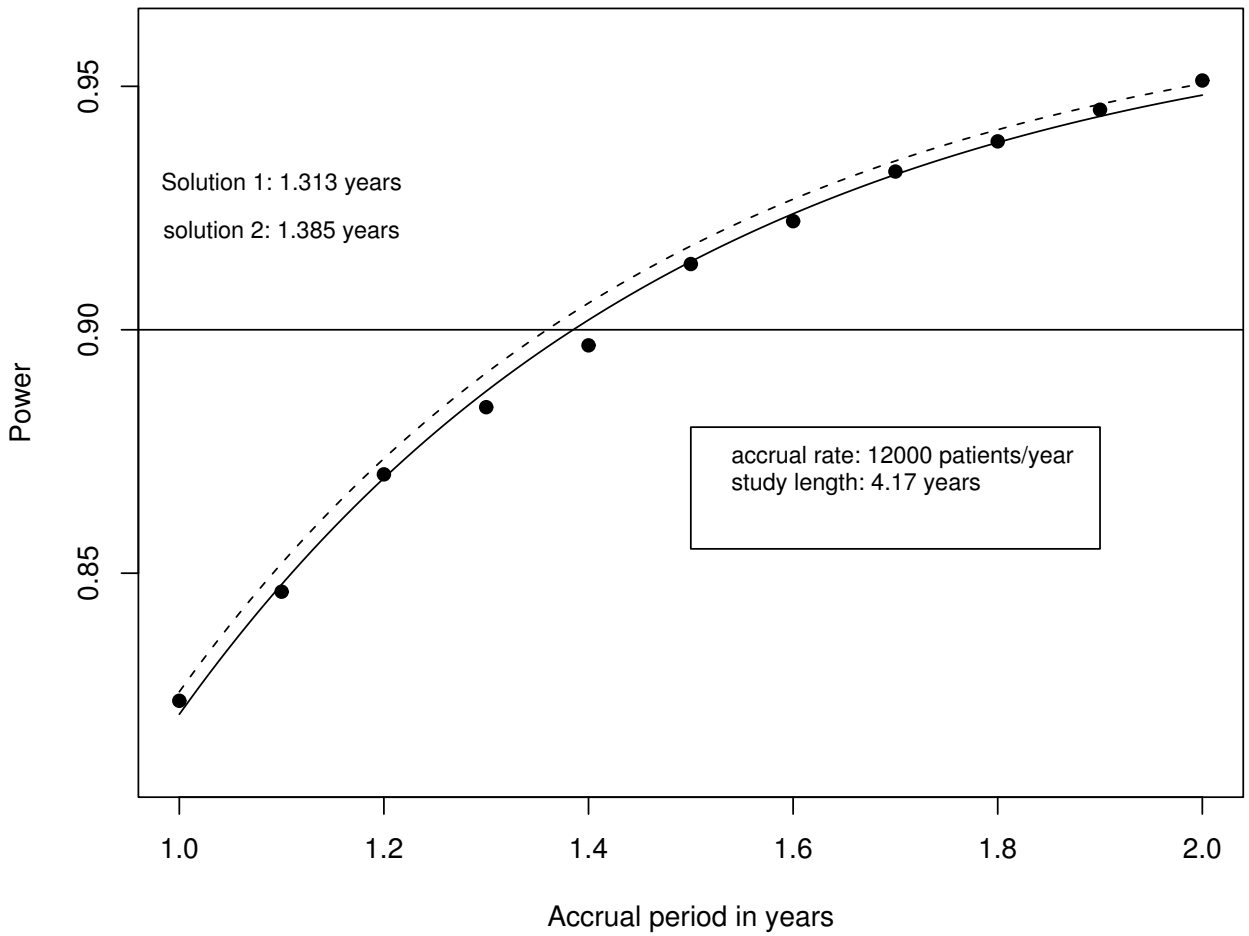
$$D_1 = a\theta \left[KA - \frac{\lambda_1}{(\lambda_1+\tau)^2} e^{-(\lambda_0-\lambda_1)t_0-(\lambda_1+\tau)L} \left\{e^{(\lambda_1+\tau)A} - 1\right\}\right].$$

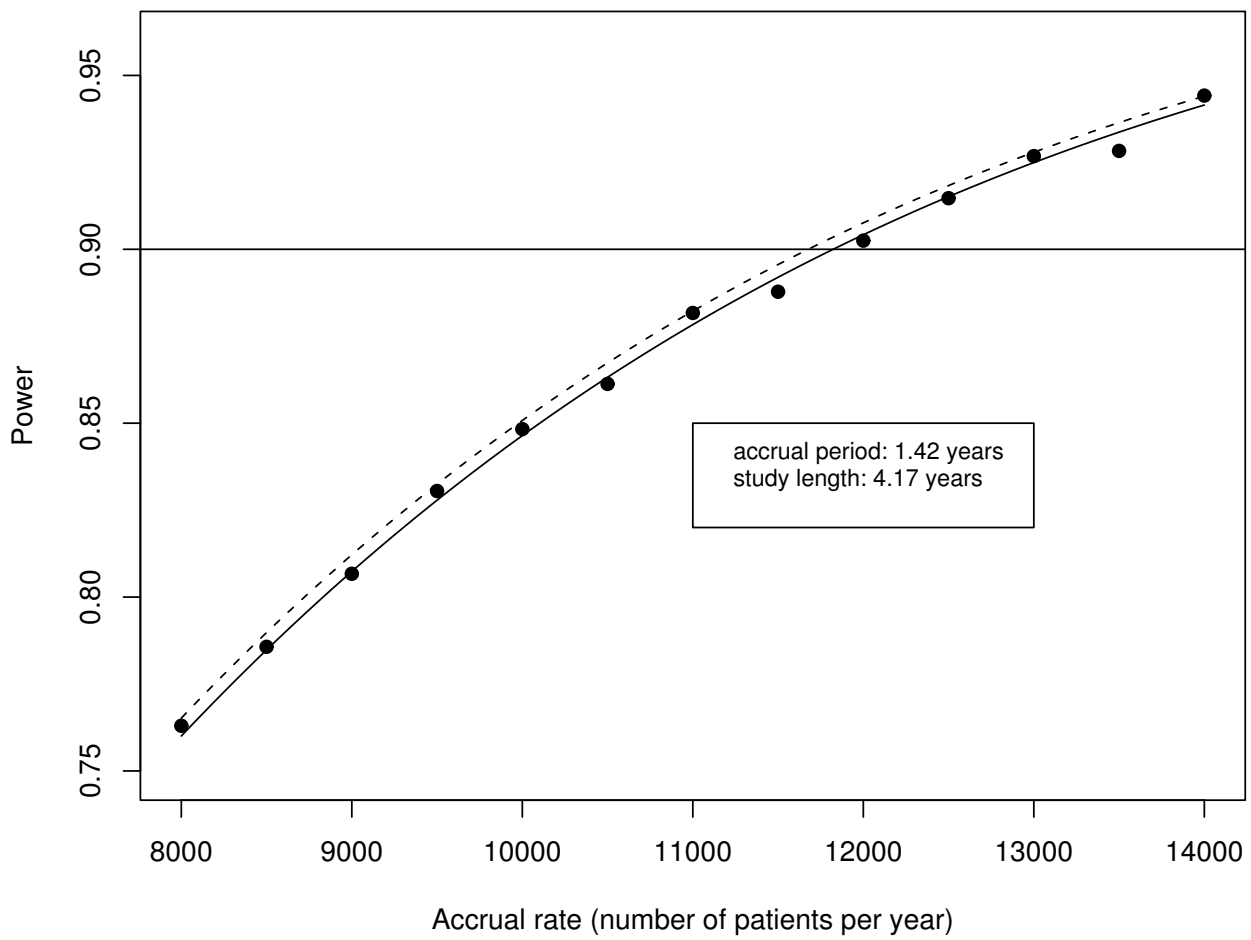
- The total expected # of deaths for treatment group before t_0:

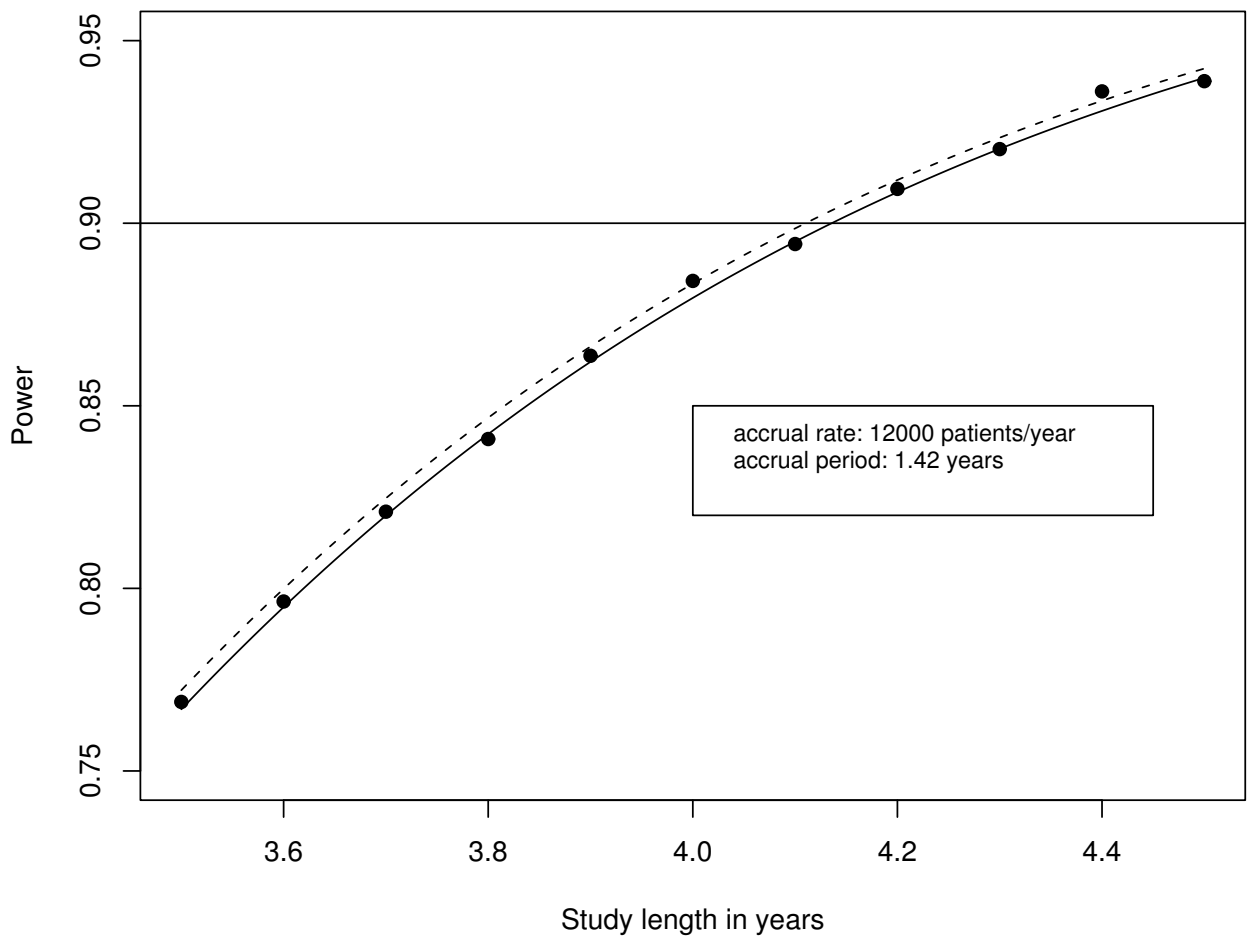
$$D_1^* = \int_0^A a\theta P[T \leq \min(t_0, Z)]dt$$
$$= \frac{aA\theta\lambda_0}{\lambda_0+\tau}\{1 - e^{-(\lambda_0+\tau)t_0}\}.$$

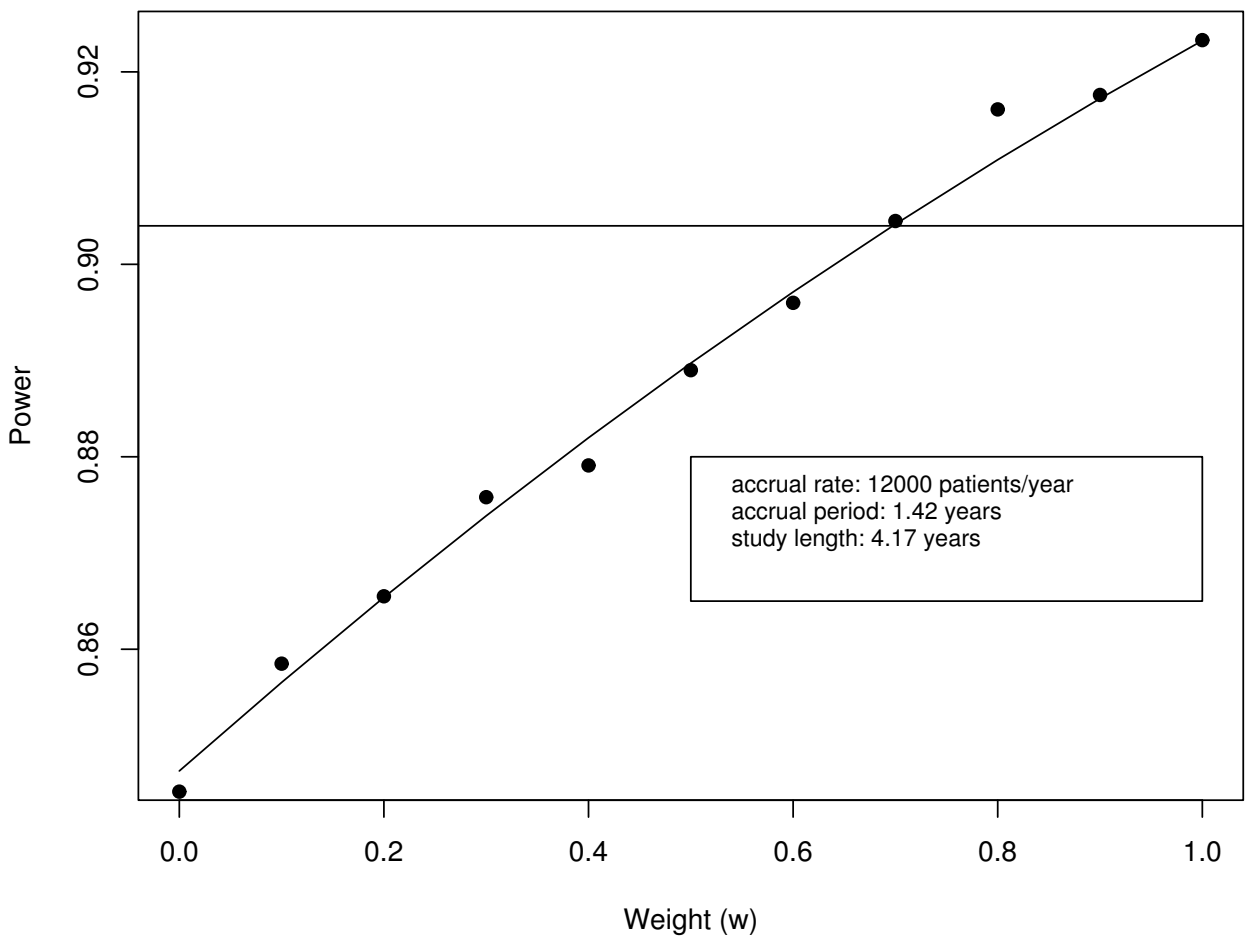
5. Example and simulation results

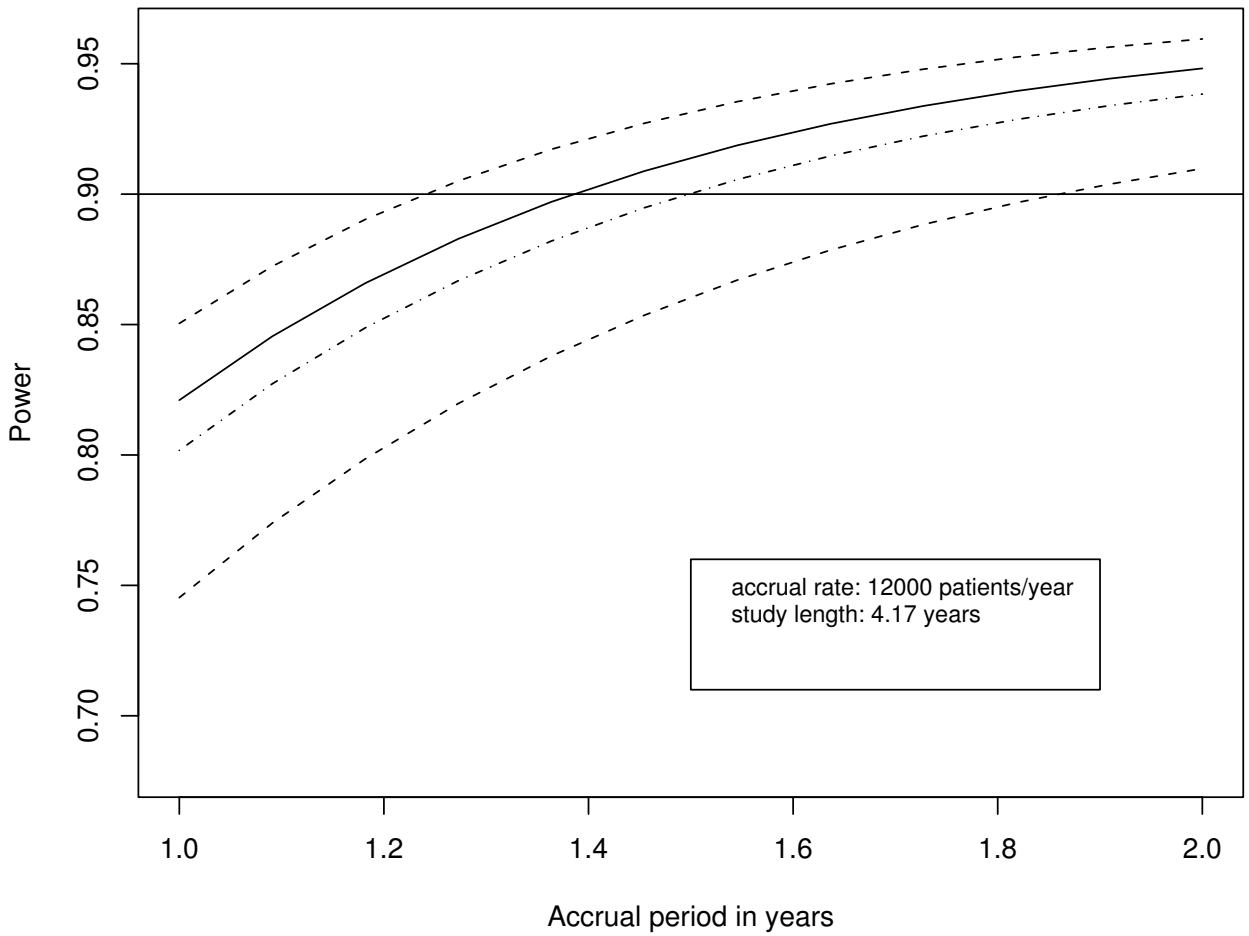
- Expect new treatment takes effect after 1 year $\Longrightarrow t_0 = 1$
- Rate to have cardiovascular risk 0.03 per year ($\lambda_0 = 0.03$)
- Expect 25% reduction when new treatments takes its full effect ($\lambda_1 = 0.0225$).
- Accrual rate $a = 1000$ patients/month
- Study length (L = 50) months
- Expect 10% (per year) drop-out rate
- Significance level $\alpha = 0.05$; targeted power = 0.9
- How long should the accrual period (A) be? And sample size?

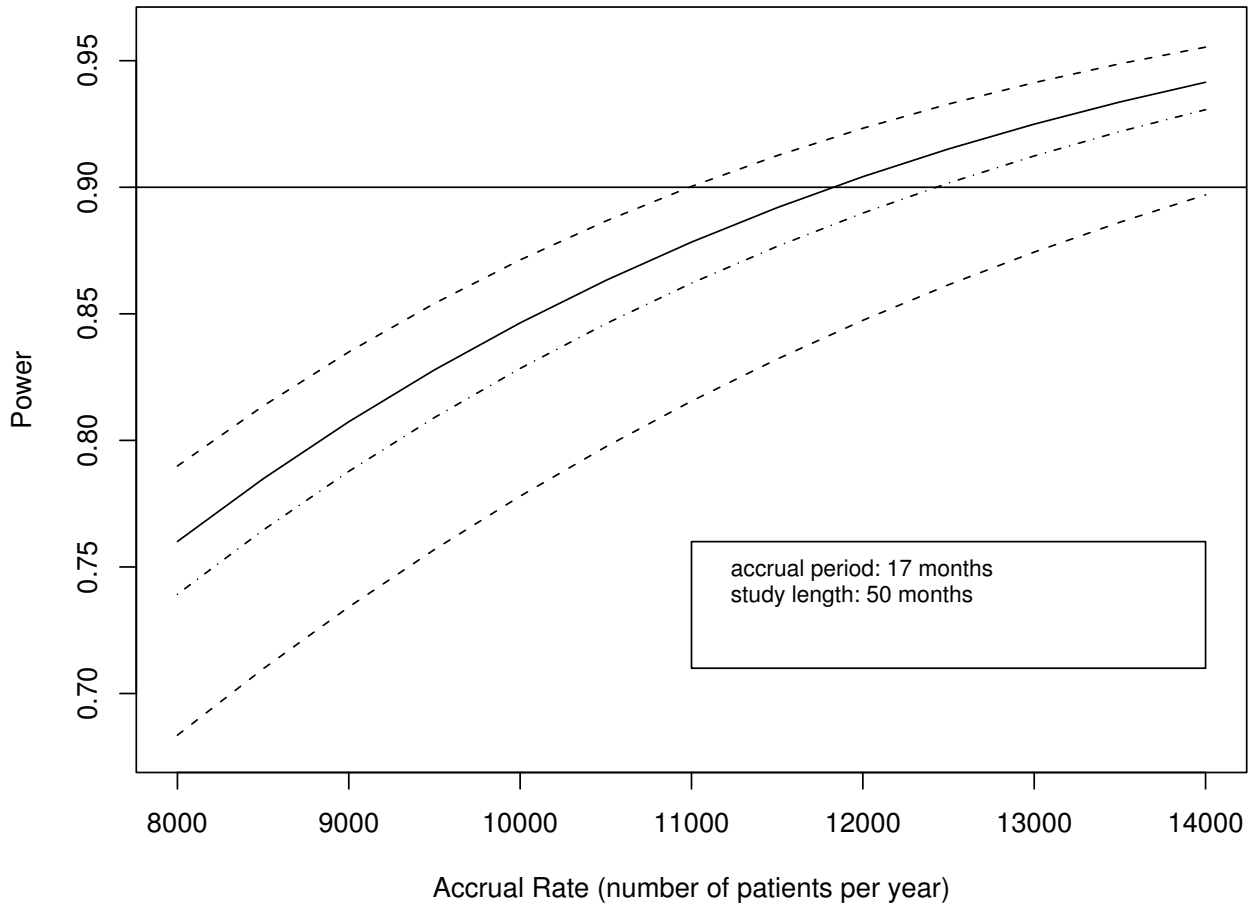

0.95
0.90
0.85
Power
Solution 1: 1.313 years
solution 2: 1.385 years
accrual rate: 12000 patients/year
study length: 4.17 years
1.0
1.2
1.4
1.6
1.8
2.0
Accrual period in years

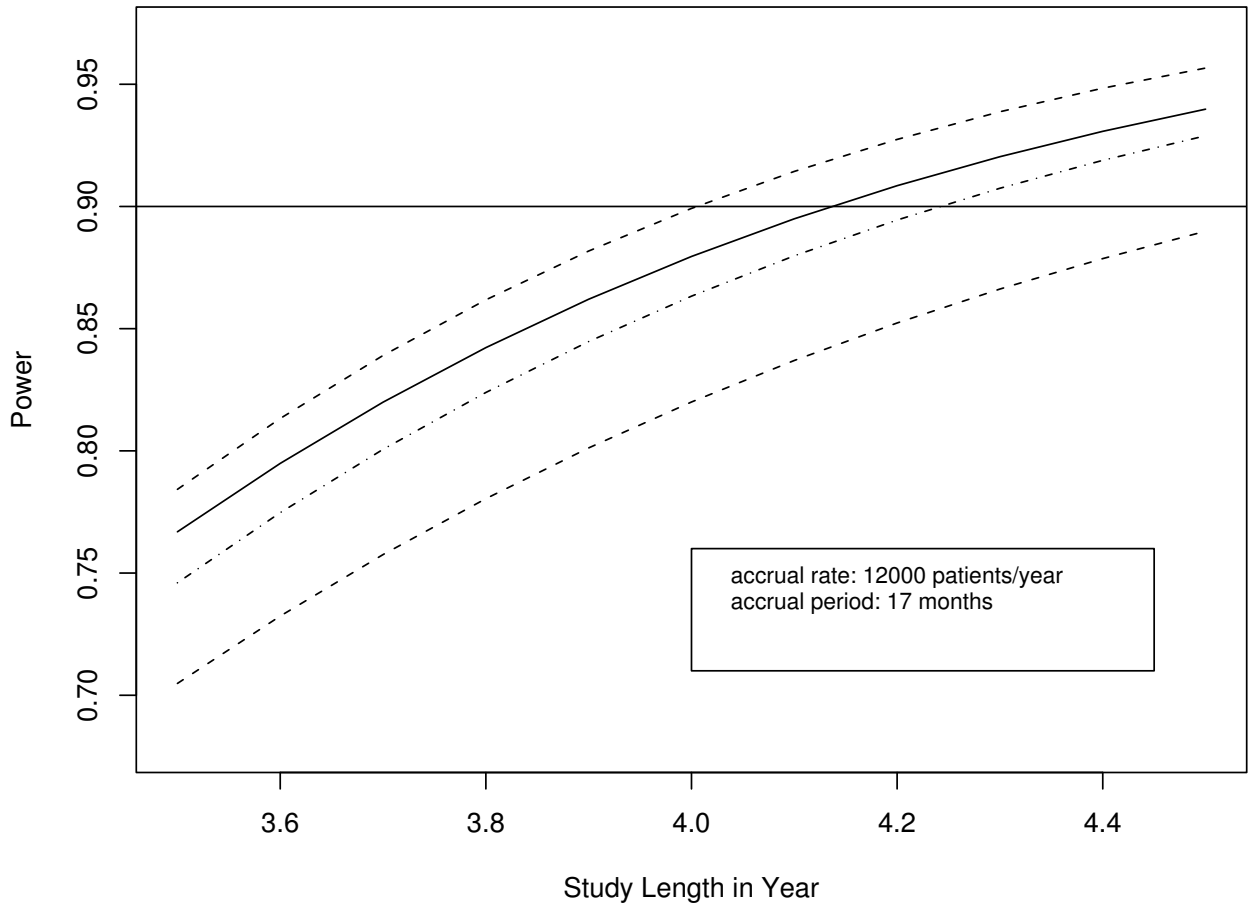

Power
0.75
0.80
0.85
0.90
0.95
8000
9000
10000
11000
12000
13000
14000
Accrual rate (number of patients per year)
accrual period: 1.42 years
study length: 4.17 years


0.95
0.90
0.85
0.80
0.75
Power
accrual rate: 12000 patients/year
accrual period: 1.42 years
3.6
3.8
4.0
4.2
4.4
Study length in years


Power
0.86
0.88
0.90
0.92
0.0
0.2
0.4
0.6
0.8
1.0
Weight (w)
accrual rate: 12000 patients/year
accrual period: 1.42 years
study length: 4.17 years


Power
0.70
0.75
0.80
0.85
0.90
0.95
1.0
1.2
1.4
1.6
1.8
2.0
Accrual period in years
accrual rate: 12000 patients/year
study length: 4.17 years


Power
0.70
0.75
0.80
0.85
0.90
0.95
8000
9000
10000
11000
12000
13000
14000
Accrual Rate (number of patients per year)
accrual period: 17 months
study length: 50 months


Power
0.70
0.75
0.80
0.85
0.90
0.95
3.6
3.8
4.0
4.2
4.4
Study Length in Year
accrual rate: 12000 patients/year
accrual period: 17 months

software: S-plus function *logrankpower(*

alpha=0.05,	signifance level of the log-rank test
lambda0=,	hazard for placebo
lambda1=,	hazard for treatment
t0=0,	t_0 used in the formula
wt=0.5,	weight for residual treatment effect
tau=0,	drop-out rate
acrate=,	accrual rate
acperiod=,	accrual period
slength=,	study lenght (slength-acperiod>t0)
theta=0.5,	allocation prob
nsub=1000,	number of sub-intervals for ITT analysis
itt=F	flag for ITT analysis)

6. Discussion

- Delayed treatment effect + drop-outs present challenge to statisticians
- Proposed two strategies:
  1. Treat drop-outs as censored observations
     (a) Assumption: drop-out process independent of (underlying true) time to event
     (b) Drop-out processes almost the same in both groups.
     (c) Calculation straightforward
     (d) Don't need to specify the hazard for untreated group
  2. Conduct ITT analysis:
     (a) May be what regulatory agencies want
     (b) May have enough power only if residual treatment effect is

relatively large (70% in our example)

(c) Can be computationally intensive (small Δ)

(d) Have to specify the hazard for untreated group

- Derived formula easy to use; confirmed by simulation to have good statistical properties
- Can include other censoring