Nonparametric Tests of Treatment Effect for a Recurrent Event Process that Terminates

Susan Murray, University of Michigan

Joint Work with Nabihah Tayob

July 13 2013

Motivation: Pulmonary Clinical Trial Design

- Common to design pulmonary studies with one year of follow-up for large number of patients
  - Study endpoint is time to some combined endpoint comprised of recurrent events (exacerbations) and terminal events
- Existing designs based on time-to-first-event analysis or a recurrent event analysis
  - Recurrent event analyses often subject to restrictive assumptions on the way recurrent events come in (independence between event gap times)
  - Time-to-first event does not take advantage of events beyond first event
- Statistical goal to develop methodology that will take advantage of additional statistical information without making assumptions about independence between events

Motivation: Pulmonary Clinical Trial Design

- Common to design pulmonary studies with one year of follow-up for large number of patients
  - Study endpoint is time to some combined endpoint comprised of recurrent events (exacerbations) and terminal events
- Existing designs based on time-to-first-event analysis or a recurrent event analysis
  - Recurrent event analyses often subject to restrictive assumptions on the way recurrent events come in (independence between event gap times)
  - Time-to-first event does not take advantage of events beyond first event
- Statistical goal to develop methodology that will take advantage of additional statistical information without making assumptions about independence between events

Motivation: Pulmonary Clinical Trial Design

- Common to design pulmonary studies with one year of follow-up for large number of patients
  - Study endpoint is time to some combined endpoint comprised of recurrent events (exacerbations) and terminal events
- Existing designs based on time-to-first-event analysis or a recurrent event analysis
  - Recurrent event analyses often subject to restrictive assumptions on the way recurrent events come in (independence between event gap times)
  - Time-to-first event does not take advantage of events beyond first event
- Statistical goal to develop methodology that will take advantage of additional statistical information without making assumptions about independence between events

Motivation: Pulmonary Clinical Trial Design

- Common to design pulmonary studies with one year of follow-up for large number of patients
  - Study endpoint is time to some combined endpoint comprised of recurrent events (exacerbations) and terminal events
- Existing designs based on time-to-first-event analysis or a recurrent event analysis
  - Recurrent event analyses often subject to restrictive assumptions on the way recurrent events come in (independence between event gap times)
  - Time-to-first event does not take advantage of events beyond first event
- Statistical goal to develop methodology that will take advantage of additional statistical information without making assumptions about independence between events

Motivation: Pulmonary Clinical Trial Design

- Common to design pulmonary studies with one year of follow-up for large number of patients
  - Study endpoint is time to some combined endpoint comprised of recurrent events (exacerbations) and terminal events
- Existing designs based on time-to-first-event analysis or a recurrent event analysis
  - Recurrent event analyses often subject to restrictive assumptions on the way recurrent events come in (independence between event gap times)
  - Time-to-first event does not take advantage of events beyond first event
- Statistical goal to develop methodology that will take advantage of additional statistical information without making assumptions about independence between events

Motivation: Pulmonary Clinical Trial Design

- Common to design pulmonary studies with one year of follow-up for large number of patients
  - Study endpoint is time to some combined endpoint comprised of recurrent events (exacerbations) and terminal events
- Existing designs based on time-to-first-event analysis or a recurrent event analysis
  - Recurrent event analyses often subject to restrictive assumptions on the way recurrent events come in (independence between event gap times)
  - Time-to-first event does not take advantage of events beyond first event
- Statistical goal to develop methodology that will take advantage of additional statistical information without making assumptions about independence between events

New Way of Looking at This Data Structure

- Short-term windows of treatment effect provide an alternative understanding of treatment effect throughout the trial.
  - Multiple (overlapping) follow-up windows may improve efficiency
  - Time-to-first-event analysis approach in each window uses information beyond initial event while avoiding dependent censoring issues that plague traditional recurrent event methods
  - Terminal events are naturally incorporated into the analysis, in addition to recurrent events

New Way of Looking at This Data Structure

- Short-term windows of treatment effect provide an alternative understanding of treatment effect throughout the trial.
  - Multiple (overlapping) follow-up windows may improve efficiency
  - Time-to-first-event analysis approach in each window uses information beyond initial event while avoiding dependent censoring issues that plague traditional recurrent event methods
  - Terminal events are naturally incorporated into the analysis, in addition to recurrent events

New Way of Looking at This Data Structure

- Short-term windows of treatment effect provide an alternative understanding of treatment effect throughout the trial.
  - Multiple (overlapping) follow-up windows may improve efficiency
  - Time-to-first-event analysis approach in each window uses information beyond initial event while avoiding dependent censoring issues that plague traditional recurrent event methods
  - Terminal events are naturally incorporated into the analysis, in addition to recurrent events

New Way of Looking at This Data Structure

- Short-term windows of treatment effect provide an alternative understanding of treatment effect throughout the trial.
  - Multiple (overlapping) follow-up windows may improve efficiency
  - Time-to-first-event analysis approach in each window uses information beyond initial event while avoiding dependent censoring issues that plague traditional recurrent event methods
  - Terminal events are naturally incorporated into the analysis, in addition to recurrent events

New Way of Looking at This Data Structure

- Short-term windows of treatment effect provide an alternative understanding of treatment effect throughout the trial.
  - Multiple (overlapping) follow-up windows may improve efficiency
  - Time-to-first-event analysis approach in each window uses information beyond initial event while avoiding dependent censoring issues that plague traditional recurrent event methods
  - Terminal events are naturally incorporated into the analysis, in addition to recurrent events

Implication for Study Design

- Ideal for study design where patients are followed longer than one year, with several one year windows available to estimate time to first events
  - Additional efficiency might make studies within institution more feasible, as opposed to a multi-center study
  - Trade-off between watching patients longer and having all research contained within same institution
  - Important to to properly account for potential correlation between events in design phase or study may be underpowered
- New test based on overall estimate of τ-restricted mean survival from combined follow-up window information

Implication for Study Design

- Ideal for study design where patients are followed longer than one year, with several one year windows available to estimate time to first events
  - Additional efficiency might make studies within institution more feasible, as opposed to a multi-center study
  - Trade-off between watching patients longer and having all research contained within same institution
  - Important to to properly account for potential correlation between events in design phase or study may be underpowered
- New test based on overall estimate of τ-restricted mean survival from combined follow-up window information

Implication for Study Design

- Ideal for study design where patients are followed longer than one year, with several one year windows available to estimate time to first events
  - Additional efficiency might make studies within institution more feasible, as opposed to a multi-center study
  - Trade-off between watching patients longer and having all research contained within same institution
  - Important to to properly account for potential correlation between events in design phase or study may be underpowered
- New test based on overall estimate of τ-restricted mean survival from combined follow-up window information

Implication for Study Design

- Ideal for study design where patients are followed longer than one year, with several one year windows available to estimate time to first events
  - Additional efficiency might make studies within institution more feasible, as opposed to a multi-center study
  - Trade-off between watching patients longer and having all research contained within same institution
  - Important to to properly account for potential correlation between events in design phase or study may be underpowered
- New test based on overall estimate of τ-restricted mean survival from combined follow-up window information

Implication for Study Design

- Ideal for study design where patients are followed longer than one year, with several one year windows available to estimate time to first events
  - Additional efficiency might make studies within institution more feasible, as opposed to a multi-center study
  - Trade-off between watching patients longer and having all research contained within same institution
  - Important to to properly account for potential correlation between events in design phase or study may be underpowered
- New test based on overall estimate of τ-restricted mean survival from combined follow-up window information

Notation

- Usual set-up for $i = 1, \ldots, n_g$ patients, $g = 1, 2$ groups:
  - D_{gi} is the death time and C_{gi} is the censoring time
  - $T_{gi1} < T_{gi2} < \ldots < T_{giJ_i} = D_{gi}$ are the recurrent event times
  - Observed data: $X_{gij} = \min(T_{gij}, C_{gi})$ and $\delta_{gij} = I(T_{gij} \leq C_{gi})$ for $j = 1, 2, \ldots, \tilde{J}_{gi} \leq J_{gi}$.
- For follow-up windows from $t \in \{t_1, \ldots, t_b\}$, with $t_1 = 0$:
  - $X_{gi}(t)$ is the time to the next event from t and $\delta_{gi}(t)$ is the associated failure indicator variable.
- Counting and at risk processes for window starting at t:

$$N_g(t, u) = \sum_{i=1}^{n_g} I\{X_{gi}(t) \leq u, \delta_{gi}(t) = 1\}$$

$$Y_g(t, u) = \sum_{i=1}^{n_g} I\{X_{gi}(t) \geq u\}$$

Notation

- Usual set-up for $i = 1, \ldots, n_g$ patients, $g = 1, 2$ groups:
  - D_{gi} is the death time and C_{gi} is the censoring time
  - $T_{gi1} < T_{gi2} < \ldots < T_{giJ_i} = D_{gi}$ are the recurrent event times
  - Observed data: $X_{gij} = \min(T_{gij}, C_{gi})$ and $\delta_{gij} = I(T_{gij} \leq C_{gi})$ for $j = 1, 2, \ldots, \tilde{J}_{gi} \leq J_{gi}$.
- For follow-up windows from $t \in \{t_1, \ldots, t_b\}$, with $t_1 = 0$:
  - $X_{gi}(t)$ is the time to the next event from t and $\delta_{gi}(t)$ is the associated failure indicator variable.
- Counting and at risk processes for window starting at t:

$$N_g(t, u) = \sum_{i=1}^{n_g} I\{X_{gi}(t) \leq u, \delta_{gi}(t) = 1\}$$

$$Y_g(t, u) = \sum_{i=1}^{n_g} I\{X_{gi}(t) \geq u\}$$

Notation

- Usual set-up for $i = 1, \ldots, n_g$ patients, $g = 1, 2$ groups:
  - D_{gi} is the death time and C_{gi} is the censoring time
  - $T_{gi1} < T_{gi2} < \ldots < T_{giJ_i} = D_{gi}$ are the recurrent event times
  - Observed data: $X_{gij} = \min(T_{gij}, C_{gi})$ and $\delta_{gij} = I(T_{gij} \leq C_{gi})$ for $j = 1, 2, \ldots, \tilde{J}_{gi} \leq J_{gi}$.
- For follow-up windows from $t \in \{t_1, \ldots, t_b\}$, with $t_1 = 0$:
  - $X_{gi}(t)$ is the time to the next event from t and $\delta_{gi}(t)$ is the associated failure indicator variable.
- Counting and at risk processes for window starting at t:

$$N_g(t, u) = \sum_{i=1}^{n_g} I\{X_{gi}(t) \leq u, \delta_{gi}(t) = 1\}$$

$$Y_g(t, u) = \sum_{i=1}^{n_g} I\{X_{gi}(t) \geq u\}$$

Notation

- Usual set-up for $i = 1, \ldots, n_g$ patients, $g = 1, 2$ groups:
  - D_{gi} is the death time and C_{gi} is the censoring time
  - $T_{gi1} < T_{gi2} < \ldots < T_{giJ_i} = D_{gi}$ are the recurrent event times
  - Observed data: $X_{gij} = \min(T_{gij}, C_{gi})$ and $\delta_{gij} = I(T_{gij} \leq C_{gi})$ for $j = 1, 2, \ldots, \tilde{J}_{gi} \leq J_{gi}$.
- For follow-up windows from $t \in \{t_1, \ldots, t_b\}$, with $t_1 = 0$:
  - $X_{gi}(t)$ is the time to the next event from t and $\delta_{gi}(t)$ is the associated failure indicator variable.
- Counting and at risk processes for window starting at t:

$$N_g(t, u) = \sum_{i=1}^{n_g} I\{X_{gi}(t) \leq u, \delta_{gi}(t) = 1\}$$

$$Y_g(t, u) = \sum_{i=1}^{n_g} I\{X_{gi}(t) \geq u\}$$

Notation

- Usual set-up for $i = 1, \ldots, n_g$ patients, $g = 1, 2$ groups:
  - D_{gi} is the death time and C_{gi} is the censoring time
  - $T_{gi1} < T_{gi2} < \ldots < T_{giJ_i} = D_{gi}$ are the recurrent event times
  - Observed data: $X_{gij} = \min(T_{gij}, C_{gi})$ and $\delta_{gij} = I(T_{gij} \leq C_{gi})$ for $j = 1, 2, \ldots, \tilde{J}_{gi} \leq J_{gi}$.
- For follow-up windows from $t \in \{t_1, \ldots, t_b\}$, with $t_1 = 0$:
  - $X_{gi}(t)$ is the time to the next event from t and $\delta_{gi}(t)$ is the associated failure indicator variable.
- Counting and at risk processes for window starting at t:

$$N_g(t, u) = \sum_{i=1}^{n_g} I\{X_{gi}(t) \leq u, \delta_{gi}(t) = 1\}$$

$$Y_g(t, u) = \sum_{i=1}^{n_g} I\{X_{gi}(t) \geq u\}$$

Notation

- Usual set-up for $i = 1, \ldots, n_g$ patients, $g = 1, 2$ groups:
  - D_{gi} is the death time and C_{gi} is the censoring time
  - $T_{gi1} < T_{gi2} < \ldots < T_{giJ_i} = D_{gi}$ are the recurrent event times
  - Observed data: $X_{gij} = \min(T_{gij}, C_{gi})$ and $\delta_{gij} = I(T_{gij} \leq C_{gi})$ for $j = 1, 2, \ldots, \tilde{J}_{gi} \leq J_{gi}$.
- For follow-up windows from $t \in \{t_1, \ldots, t_b\}$, with $t_1 = 0$:
  - $X_{gi}(t)$ is the time to the next event from t and $\delta_{gi}(t)$ is the associated failure indicator variable.
- Counting and at risk processes for window starting at t:

$$N_g(t, u) = \sum_{i=1}^{n_g} I\{X_{gi}(t) \leq u, \delta_{gi}(t) = 1\}$$

$$Y_g(t, u) = \sum_{i=1}^{n_g} I\{X_{gi}(t) \geq u\}$$

Notation

- Usual set-up for $i = 1, \ldots, n_g$ patients, $g = 1, 2$ groups:
  - D_{gi} is the death time and C_{gi} is the censoring time
  - $T_{gi1} < T_{gi2} < \ldots < T_{giJ_i} = D_{gi}$ are the recurrent event times
  - Observed data: $X_{gij} = \min(T_{gij}, C_{gi})$ and $\delta_{gij} = I(T_{gij} \leq C_{gi})$ for $j = 1, 2, \ldots, \tilde{J}_{gi} \leq J_{gi}$.
- For follow-up windows from $t \in \{t_1, \ldots, t_b\}$, with $t_1 = 0$:
  - $X_{gi}(t)$ is the time to the next event from t and $\delta_{gi}(t)$ is the associated failure indicator variable.
- Counting and at risk processes for window starting at t:

$$N_g(t, u) = \sum_{i=1}^{n_g} I\{X_{gi}(t) \leq u, \delta_{gi}(t) = 1\}$$

$$Y_g(t, u) = \sum_{i=1}^{n_g} I\{X_{gi}(t) \geq u\}$$

Avoidance of Dependent Censoring Issue

- For group g, person i, recurrent events $j = 2, 3, \ldots, J_i$
  - Time between recurrent events are $S_{i1} = T_{i1}$ and $S_{ij} = T_{ij} - T_{i(j-1)}$ for $j = 2, 3, \ldots, J_i$.
  - For each gap time S_{ij}, the censoring time $C_i - T_{i(j-1)}$ is a function of the previous event time $\rightarrow$ leads to dependent censoring when analyzing dependent gap times.
  - For our approach of time to first event within each follow-up window starting at t, the censoring time is $C_i - t \rightarrow$ no dependent censoring issue!

Avoidance of Dependent Censoring Issue

- For group g, person i, recurrent events $j = 2, 3, \ldots, J_i$
  - Time between recurrent events are $S_{i1} = T_{i1}$ and $S_{ij} = T_{ij} - T_{i(j-1)}$ for $j = 2, 3, \ldots, J_i$.
  - For each gap time S_{ij}, the censoring time $C_i - T_{i(j-1)}$ is a function of the previous event time $\rightarrow$ leads to dependent censoring when analyzing dependent gap times.
  - For our approach of time to first event within each follow-up window starting at t, the censoring time is $C_i - t \rightarrow$ no dependent censoring issue!

Avoidance of Dependent Censoring Issue

- For group g, person i, recurrent events $j = 2, 3, \ldots, J_i$
  - Time between recurrent events are $S_{i1} = T_{i1}$ and $S_{ij} = T_{ij} - T_{i(j-1)}$ for $j = 2, 3, \ldots, J_i$.
  - For each gap time S_{ij}, the censoring time $C_i - T_{i(j-1)}$ is a function of the previous event time $\rightarrow$ leads to dependent censoring when analyzing dependent gap times.
  - For our approach of time to first event within each follow-up window starting at t, the censoring time is $C_i - t \rightarrow$ no dependent censoring issue!

Avoidance of Dependent Censoring Issue

- For group g, person i, recurrent events $j = 2, 3, \ldots, J_i$
  - Time between recurrent events are $S_{i1} = T_{i1}$ and $S_{ij} = T_{ij} - T_{i(j-1)}$ for $j = 2, 3, \ldots, J_i$.
  - For each gap time S_{ij}, the censoring time $C_i - T_{i(j-1)}$ is a function of the previous event time $\rightarrow$ leads to dependent censoring when analyzing dependent gap times.
  - For our approach of time to first event within each follow-up window starting at t, the censoring time is $C_i - t \rightarrow$ no dependent censoring issue!

Overall estimate of τ-restricted mean survival

- The proposed estimate of the overall τ-restricted mean survival from combined follow-up window information is

$$\hat{\mu}_g^*(\tau) = \int_0^\tau \exp\left\{ - \int_0^s \frac{dN_g(u)}{Y_g(u)} \right\} ds.$$

$$dN_g(t_1, u) + dN_g(t_2, u) + \ldots + dN_g(t_b, u)$$

$$Y_g(t_1, u) + Y_g(t_2, u) + \ldots + Y_g(t_b, u)$$

Overall estimate of τ-restricted mean survival

- The proposed estimate of the overall τ-restricted mean survival from combined follow-up window information is

$$\hat{\mu}_g^*(\tau) = \int_0^\tau \exp\left\{ -\int_0^s \frac{dN_g(u)}{Y_g(u)} \right\} ds.$$

$$dN_g(t_1, u) + dN_g(t_2, u) + \ldots + dN_g(t_b, u)$$

$$Y_g(t_1, u) + Y_g(t_2, u) + \ldots + Y_g(t_b, u)$$

Overall estimate of τ-restricted mean survival

- The proposed estimate of the overall τ-restricted mean survival from combined follow-up window information is

$$\hat{\mu}_g^*(\tau) = \int_0^\tau \exp\left\{ -\int_0^s \frac{dN_g(u)}{Y_g(u)} \right\} ds.$$

$$dN_g(t_1, u) + dN_g(t_2, u) + \ldots + dN_g(t_b, u)$$

$$Y_g(t_1, u) + Y_g(t_2, u) + \ldots + Y_g(t_b, u)$$

Overall estimate of τ-restricted mean survival

- The proposed estimate of the overall τ-restricted mean survival from combined follow-up window information is

$$\hat{\mu}_g^*(\tau) = \int_0^\tau \exp\left\{ -\int_0^s \frac{dN_g(u)}{Y_g(u)} \right\} ds.$$

$$dN_g(t_1, u) + dN_g(t_2, u) + \ldots + dN_g(t_b, u)$$

$$Y_g(t_1, u) + Y_g(t_2, u) + \ldots + Y_g(t_b, u)$$

Overall estimate of τ-restricted mean survival

- The proposed estimate of the overall τ-restricted mean survival from combined follow-up window information is

$$\hat{\mu}_g^*(\tau) = \int_0^\tau \exp\left\{ -\int_0^s \frac{dN_g(u)}{Y_g(u)} \right\} ds.$$

$$dN_g(t_1, u) + dN_g(t_2, u) + \ldots + dN_g(t_b, u)$$

$$Y_g(t_1, u) + Y_g(t_2, u) + \ldots + Y_g(t_b, u)$$

Overall estimate of τ-restricted mean survival

- The proposed estimate of the overall τ-restricted mean survival from combined follow-up window information is

$$\hat{\mu}_g^*(\tau) = \int_0^\tau \exp\left\{ -\int_0^s \frac{dN_g(u)}{Y_g(u)} \right\} ds.$$

$$dN_g(t_1, u) + dN_g(t_2, u) + \ldots + dN_g(t_b, u)$$

$$Y_g(t_1, u) + Y_g(t_2, u) + \ldots + Y_g(t_b, u)$$

Overall estimate of τ-restricted mean survival

- The proposed estimate of the overall τ-restricted mean survival from combined follow-up window information is

$$\hat{\mu}_g^*(\tau) = \int_0^\tau \exp\left\{ -\int_0^s \frac{dN_g(u)}{Y_g(u)} \right\} ds.$$

$$dN_g(t_1, u) + dN_g(t_2, u) + \ldots + dN_g(t_b, u)$$

$$Y_g(t_1, u) + Y_g(t_2, u) + \ldots + Y_g(t_b, u)$$

Consistency of overall estimate of τ-restricted mean survival

- $\hat{\mu}_g^*(\tau) \xrightarrow{p} \mu_g^*(\tau)$
  $= \int_0^\tau \exp\left[- \int_0^s \sum_{k=1}^b \lambda_g(t_k, u) \frac{Pr\{X_{gi}(t_k) \geq u\}}{\sum_{l=1}^b Pr\{X_{gi}(t_l) \geq u\}} du \right] ds,$

  which is the mean of the mixture distribution created from combining times-to-first event across the different follow-up intervals.

Variance of Proposed Estimate

$\sqrt{n}\{\hat{\mu}^*_g(\tau) - \mu^*_g(\tau)\}$ converges in distribution to a normal random variable with finite variance that is estimated by $\hat{\sigma}^2_{*g}$.

Variance calculations involve

- Linearizing components of $\sqrt{n}\hat{\mu}^*(\tau)$ using Taylor series approximations; similar to approach of Woodruff(1971) and more recently Williams (1995)
- Grouping random variables according to correlation structure and identify sums of independent (functions of these) random variables
- Using asymptotic arguments to describe the asymptotic distribution of interest

Variance of Proposed Estimate

$\sqrt{n}\{\hat{\mu}_g^*(\tau) - \mu_g^*(\tau)\}$ converges in distribution to a normal random variable with finite variance that is estimated by $\hat{\sigma}^2_{*g}$.

Variance calculations involve

- Linearizing components of $\sqrt{n}\hat{\mu}^*(\tau)$ using Taylor series approximations; similar to approach of Woodruff(1971) and more recently Williams (1995)
- Grouping random variables according to correlation structure and identify sums of independent (functions of these) random variables
- Using asymptotic arguments to describe the asymptotic distribution of interest

Variance of Proposed Estimate

$\sqrt{n}\{\hat{\mu}_g^*(\tau) - \mu_g^*(\tau)\}$ converges in distribution to a normal random variable with finite variance that is estimated by $\hat{\sigma}^2_{*g}$.

Variance calculations involve

- Linearizing components of $\sqrt{n}\hat{\mu}^*(\tau)$ using Taylor series approximations; similar to approach of Woodruff(1971) and more recently Williams (1995)
- Grouping random variables according to correlation structure and identify sums of independent (functions of these) random variables
- Using asymptotic arguments to describe the asymptotic distribution of interest

Variance of Proposed Estimate

$\sqrt{n}\{\hat{\mu}^*_g(\tau) - \mu^*_g(\tau)\}$ converges in distribution to a normal random variable with finite variance that is estimated by $\hat{\sigma}^2_{*g}$.

Variance calculations involve

- Linearizing components of $\sqrt{n}\hat{\mu}^*(\tau)$ using Taylor series approximations; similar to approach of Woodruff(1971) and more recently Williams (1995)
- Grouping random variables according to correlation structure and identify sums of independent (functions of these) random variables
- Using asymptotic arguments to describe the asymptotic distribution of interest

Variance of Proposed Estimate

$\sqrt{n}\{\hat{\mu}^*_g(\tau) - \mu^*_g(\tau)\}$ converges in distribution to a normal random variable with finite variance that is estimated by $\hat{\sigma}^2_{*g}$.

Variance calculations involve

- Linearizing components of $\sqrt{n}\hat{\mu}^*(\tau)$ using Taylor series approximations; similar to approach of Woodruff(1971) and more recently Williams (1995)
- Grouping random variables according to correlation structure and identify sums of independent (functions of these) random variables
- Using asymptotic arguments to describe the asymptotic distribution of interest

Practical Issues

- Number and spacing of follow-up windows should be chosen to increase the precision of $\hat{\mu}^*(\tau)$.
- We studied the special case of a single terminal event, with closed form asymptotic variance calculations, to aid our understanding of design issues
- For instance, suppose
  - T_i follows an exponential distribution with 1 year restricted mean of 11 months, i.e. $\tau = 1$ year.
  - C_i is independently sampled from a uniform $[2, 3]$.
  - $n = 100$

Practical Issues

- Number and spacing of follow-up windows should be chosen to increase the precision of $\hat{\mu}^*(\tau)$.
- We studied the special case of a single terminal event, with closed form asymptotic variance calculations, to aid our understanding of design issues
- For instance, suppose
  - T_i follows an exponential distribution with 1 year restricted mean of 11 months, i.e. $\tau = 1$ year.
  - C_i is independently sampled from a uniform $[2, 3]$.
  - $n = 100$

Practical Issues

- Number and spacing of follow-up windows should be chosen to increase the precision of $\hat{\mu}^*(\tau)$.
- We studied the special case of a single terminal event, with closed form asymptotic variance calculations, to aid our understanding of design issues
- For instance, suppose
  - T_i follows an exponential distribution with 1 year restricted mean of 11 months, i.e. $\tau = 1$ year.
  - C_i is independently sampled from a uniform $[2, 3]$.
  - $n = 100$

Practical Issues

- Number and spacing of follow-up windows should be chosen to increase the precision of $\hat{\mu}^*(\tau)$.
- We studied the special case of a single terminal event, with closed form asymptotic variance calculations, to aid our understanding of design issues
- For instance, suppose
  - T_i follows an exponential distribution with 1 year restricted mean of 11 months, i.e. $\tau = 1$ year.
  - C_i is independently sampled from a uniform $[2, 3]$.
  - $n = 100$

Practical Issues

- Number and spacing of follow-up windows should be chosen to increase the precision of $\hat{\mu}^*(\tau)$.
- We studied the special case of a single terminal event, with closed form asymptotic variance calculations, to aid our understanding of design issues
- For instance, suppose
  - T_i follows an exponential distribution with 1 year restricted mean of 11 months, i.e. $\tau = 1$ year.
  - C_i is independently sampled from a uniform $[2, 3]$.
  - $n = 100$

Practical Issues

- Number and spacing of follow-up windows should be chosen to increase the precision of $\hat{\mu}^*(\tau)$.
- We studied the special case of a single terminal event, with closed form asymptotic variance calculations, to aid our understanding of design issues
- For instance, suppose
  - T_i follows an exponential distribution with 1 year restricted mean of 11 months, i.e. $\tau = 1$ year.
  - C_i is independently sampled from a uniform $[2, 3]$.
  - $n = 100$

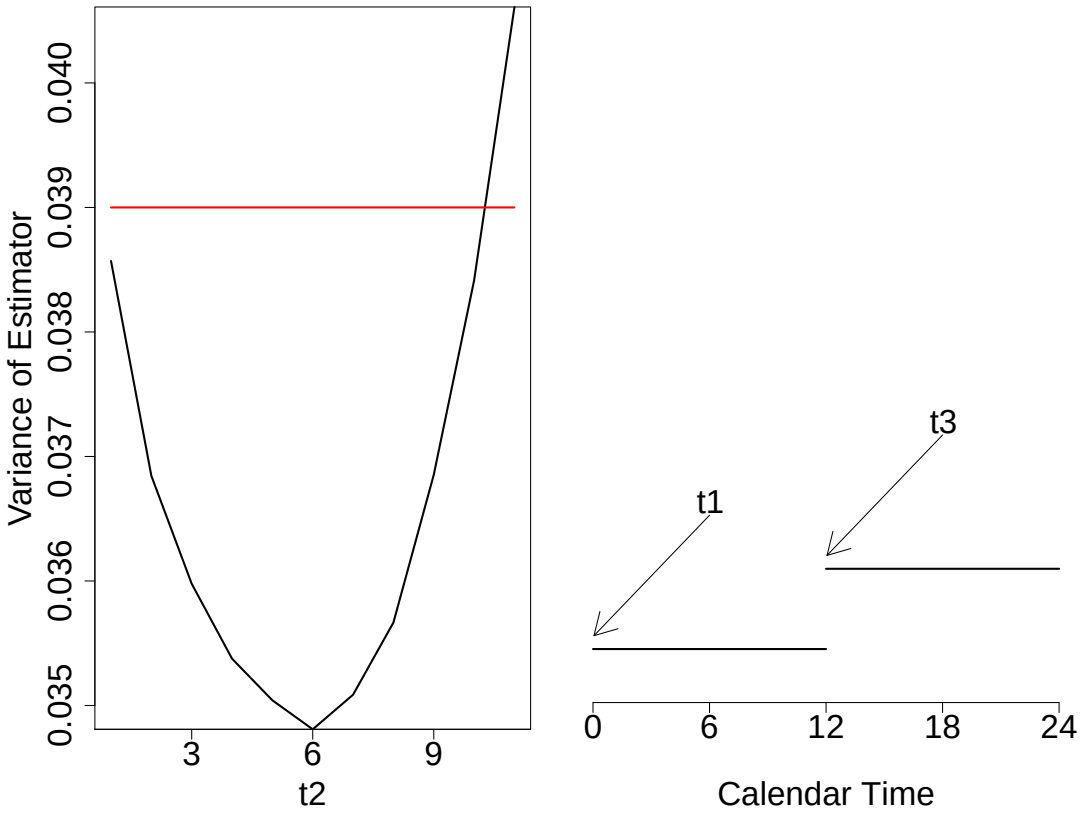

Figure: Variance of $\hat{\mu}^*(1)$ for 3 year study, $t_1 = 0$, $t_3 = 12$ months and varying t_2.

Table: Variance by number (and overlap) of follow-up windows

	Number of Windows	$\{t_1, \ldots, t_b\}$	Variance	ARE
$t_b = 0$	1	0	0.071	1.00
$t_b = 12$	2	0, 12	0.039	1.82
	3	0, 6, 12	0.035	2.03
	5	0, 3, 6, 9,12	0.035	2.03
$t_b = 24$	3	0, 12, 24	0.030	2.37
	5	0, 6, 12, 18,24	0.026	2.73
	9	0, 3, 6, 9, 12, 15, 18, 21, 24	0.025	2.84

Recommendation

- Spacing of $t_1, \ldots, t_b$ at 6-month intervals when estimating an overall 1 year restricted mean.
- In general, we propose using intervals starting from $t_k = (k-1)\frac{\tau}{2}$ for $k = 1, \ldots, b$, with t_b less than $A - \tau$, where A is the length of the study.

Recommendation

- Spacing of $t_1, \ldots, t_b$ at 6-month intervals when estimating an overall 1 year restricted mean.
- In general, we propose using intervals starting from $t_k = (k-1)\frac{\tau}{2}$ for $k = 1, \ldots, b$, with t_b less than $A - \tau$, where A is the length of the study.

Two-Sample Test

- A nonparametric test statistic comparing the overall τ-restricted mean survival in two independent groups of size n_1 and n_2 respectively is

$$\mathscr{T}_* = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \{\hat{\mu}_1^*(\tau) - \hat{\mu}_2^*(\tau)\}$$

  which has a mean zero Normal limiting distribution with variance $\pi_2 \sigma_{*1}^2 + \pi_1 \sigma_{*2}^2$ under the null hypothesis $H_0 : \mu_1^*(\tau) = \mu_2^*(\tau)$.

Simulation Study Design

- Compare performance of $\mathscr{T}_*$ to
  - $\mathscr{T}_{LR}$= Logrank test (time to first event analysis)
  - $\mathscr{T}_{PM}$= Proportional rates model of Lin et al. (2000)
  - $\mathscr{T}_{GL}$= Ghosh and Lin (2000) statistic with equal weights for recurrent and terminal endpoints.

- 36-month study: Recruit 30% of patients at $t = 0$, 70% uniformly through 12 months; $n_1 = n_2 = 100$
- Group 1 gap time incidence rate: $1/12$
- Group 2 gap time incidence rate: $1/12 * \alpha$, $\alpha = \{1, 0.8, 0.6\}$
- Group 1 and 2 hazards for death: $1/36$ and $1/36 * \alpha$
- Events either independent or given correlation structure with correlation between gap times, $\rho_1 \approx 0.5$, correlation between gap times and death time, $\rho_2 \approx 0.3$.
- $\tau = 12$ months, $\{t_1 = 0, t_2 = 6, t_3 = 12, t_4 = 18, t_5 = 24\}$

Simulation Study Design

- Compare performance of $\mathscr{T}_*$ to
  - $\mathscr{T}_{LR}$= Logrank test (time to first event analysis)
  - $\mathscr{T}_{PM}$= Proportional rates model of Lin et al. (2000)
  - $\mathscr{T}_{GL}$= Ghosh and Lin (2000) statistic with equal weights for recurrent and terminal endpoints.

- 36-month study: Recruit 30% of patients at $t = 0$, 70% uniformly through 12 months; $n_1 = n_2 = 100$
- Group 1 gap time incidence rate: $1/12$
- Group 2 gap time incidence rate: $1/12 * \alpha$, $\alpha = \{1, 0.8, 0.6\}$
- Group 1 and 2 hazards for death: $1/36$ and $1/36 * \alpha$
- Events either independent or given correlation structure with correlation between gap times, $\rho_1 \approx 0.5$, correlation between gap times and death time, $\rho_2 \approx 0.3$.
- $\tau = 12$ months, $\{t_1 = 0, t_2 = 6, t_3 = 12, t_4 = 18, t_5 = 24\}$

Simulation Study Design

- Compare performance of $\mathscr{T}_*$ to
  - $\mathscr{T}_{LR}$= Logrank test (time to first event analysis)
  - $\mathscr{T}_{PM}$= Proportional rates model of Lin et al. (2000)
  - $\mathscr{T}_{GL}$= Ghosh and Lin (2000) statistic with equal weights for recurrent and terminal endpoints.

- 36-month study: Recruit 30% of patients at $t = 0$, 70% uniformly through 12 months; $n_1 = n_2 = 100$
- Group 1 gap time incidence rate: $1/12$
- Group 2 gap time incidence rate: $1/12 * \alpha$, $\alpha = \{1, 0.8, 0.6\}$
- Group 1 and 2 hazards for death: $1/36$ and $1/36 * \alpha$
- Events either independent or given correlation structure with correlation between gap times, $\rho_1 \approx 0.5$, correlation between gap times and death time, $\rho_2 \approx 0.3$.
- $\tau = 12$ months, $\{t_1 = 0, t_2 = 6, t_3 = 12, t_4 = 18, t_5 = 24\}$

Simulation Study Design

- Compare performance of $\mathscr{T}_*$ to
  - $\mathscr{T}_{LR}$= Logrank test (time to first event analysis)
  - $\mathscr{T}_{PM}$= Proportional rates model of Lin et al. (2000)
  - $\mathscr{T}_{GL}$= Ghosh and Lin (2000) statistic with equal weights for recurrent and terminal endpoints.

- 36-month study: Recruit 30% of patients at $t = 0$, 70% uniformly through 12 months; $n_1 = n_2 = 100$
- Group 1 gap time incidence rate: $1/12$
- Group 2 gap time incidence rate: $1/12 * \alpha$, $\alpha = \{1, 0.8, 0.6\}$
- Group 1 and 2 hazards for death: $1/36$ and $1/36 * \alpha$
- Events either independent or given correlation structure with correlation between gap times, $\rho_1 \approx 0.5$, correlation between gap times and death time, $\rho_2 \approx 0.3$.
- $\tau = 12$ months, $\{t_1 = 0, t_2 = 6, t_3 = 12, t_4 = 18, t_5 = 24\}$

Simulation Study Design

- Compare performance of $\mathscr{T}_*$ to
  - $\mathscr{T}_{LR}$= Logrank test (time to first event analysis)
  - $\mathscr{T}_{PM}$= Proportional rates model of Lin et al. (2000)
  - $\mathscr{T}_{GL}$= Ghosh and Lin (2000) statistic with equal weights for recurrent and terminal endpoints.

- 36-month study: Recruit 30% of patients at $t = 0$, 70% uniformly through 12 months; $n_1 = n_2 = 100$
- Group 1 gap time incidence rate: $1/12$
- Group 2 gap time incidence rate: $1/12 * \alpha$, $\alpha = \{1, 0.8, 0.6\}$
- Group 1 and 2 hazards for death: $1/36$ and $1/36 * \alpha$
- Events either independent or given correlation structure with correlation between gap times, $\rho_1 \approx 0.5$, correlation between gap times and death time, $\rho_2 \approx 0.3$.
- $\tau = 12$ months, $\{t_1 = 0, t_2 = 6, t_3 = 12, t_4 = 18, t_5 = 24\}$

Simulation Study Design

- Compare performance of $\mathscr{T}_*$ to
  - $\mathscr{T}_{LR}$= Logrank test (time to first event analysis)
  - $\mathscr{T}_{PM}$= Proportional rates model of Lin et al. (2000)
  - $\mathscr{T}_{GL}$= Ghosh and Lin (2000) statistic with equal weights for recurrent and terminal endpoints.

- 36-month study: Recruit 30% of patients at $t = 0$, 70% uniformly through 12 months; $n_1 = n_2 = 100$
- Group 1 gap time incidence rate: $1/12$
- Group 2 gap time incidence rate: $1/12 * \alpha$, $\alpha = \{1, 0.8, 0.6\}$
- Group 1 and 2 hazards for death: $1/36$ and $1/36 * \alpha$
- Events either independent or given correlation structure with correlation between gap times, $\rho_1 \approx 0.5$, correlation between gap times and death time, $\rho_2 \approx 0.3$.
- $\tau = 12$ months, $\{t_1 = 0, t_2 = 6, t_3 = 12, t_4 = 18, t_5 = 24\}$

Simulation Study Design

- Compare performance of $\mathscr{T}_*$ to
  - $\mathscr{T}_{LR}$= Logrank test (time to first event analysis)
  - $\mathscr{T}_{PM}$= Proportional rates model of Lin et al. (2000)
  - $\mathscr{T}_{GL}$= Ghosh and Lin (2000) statistic with equal weights for recurrent and terminal endpoints.

- 36-month study: Recruit 30% of patients at $t = 0$, 70% uniformly through 12 months; $n_1 = n_2 = 100$
- Group 1 gap time incidence rate: $1/12$
- Group 2 gap time incidence rate: $1/12 * \alpha$, $\alpha = \{1, 0.8, 0.6\}$
- Group 1 and 2 hazards for death: $1/36$ and $1/36 * \alpha$
- Events either independent or given correlation structure with correlation between gap times, $\rho_1 \approx 0.5$, correlation between gap times and death time, $\rho_2 \approx 0.3$.
- $\tau = 12$ months, $\{t_1 = 0, t_2 = 6, t_3 = 12, t_4 = 18, t_5 = 24\}$

Simulation Results: Scenario 1

Table: Scenario 1 Power Results; 500 Simulation Iterations

	$\rho_1 = 0, \rho_2 = 0$				$\rho_1 = 0.5, \rho_2 = 0.3$			
α	$\mathscr{T}_*$	$\mathscr{T}_{LR}$	$\mathscr{T}_{PM}$	$\mathscr{T}_{GL}$	$\mathscr{T}_*$	$\mathscr{T}_{LR}$	$\mathscr{T}_{PM}$	$\mathscr{T}_{GL}$
1.0	0.042	0.048	0.044	0.046	0.036	0.044	0.050	0.044
0.8	0.520	0.306	0.648	0.560	0.356	0.302	0.316	0.224
0.6	1.000	0.948	1.000	0.996	0.960	0.922	0.860	0.784

- For proportional intensity setting with independent events, $\mathscr{T}_{PM}$ outperforms its competitors.
- Correlation between an individual's events reduces that statistical information in the data considerably
  - $\mathscr{T}_*$ performs best in this case; no clear winner between $\mathscr{T}_{PM}$ and $\mathscr{T}_{LR}$

Scenario 2: Treatment effect delayed by 6 months; otherwise same as scenario 1

Table: Scenario 2 Power Results; 500 Simulation Iterations

	$\rho_1 = 0, \rho_2 = 0$				$\rho_1 = 0.5, \rho_2 = 0.3$			
α	$\mathscr{T}_*$	$\mathscr{T}_{LR}$	$\mathscr{T}_{PM}$	$\mathscr{T}_{GL}$	$\mathscr{T}_*$	$\mathscr{T}_{LR}$	$\mathscr{T}_{PM}$	$\mathscr{T}_{GL}$
1.0	0.042	0.072	0.058	0.042	0.052	0.056	0.062	0.046
0.8	0.344	0.120	0.396	0.326	0.212	0.132	0.168	0.114
0.6	0.934	0.342	0.956	0.898	0.736	0.404	0.544	0.396

- For this non-proportional intensity setting, $\mathscr{T}_{PM}$ has more comparable power to its competitors when events are independent.
- Correlation between an individual's events again reduces power considerably with highest power for $\mathscr{T}_*$, followed by $\mathscr{T}_{PM}$

Scenario 3: Treatment effect vanishes after 12 months; otherwise same as scenario 1

Table: Scenario 3 Power Results; 500 Simulation Iterations

	$\rho_1 = 0, \rho_2 = 0$				$\rho_1 = 0.5, \rho_2 = 0.3$			
α	$\mathscr{T}_*$	$\mathscr{T}_{LR}$	$\mathscr{T}_{PM}$	$\mathscr{T}_{GL}$	$\mathscr{T}_*$	$\mathscr{T}_{LR}$	$\mathscr{T}_{PM}$	$\mathscr{T}_{GL}$
1.0	0.068	0.064	0.070	0.074	0.048	0.050	0.058	0.054
0.8	0.152	0.230	0.220	0.218	0.092	0.186	0.130	0.096
0.6	0.484	0.662	0.608	0.592	0.334	0.524	0.300	0.230

- Time-to-first-event analysis outperforms all other competitors
- With independent events, $\mathscr{T}_{PM}$ still outperforms $\mathscr{T}_*$, $\mathscr{T}_{GL}$.
- Correlation adversely affects all methods, no clear winner.
  - $\mathscr{T}_*$ doesn't benefit from additional follow-up information (statistical noise added)

Summary

- Hard to beat proportional mean model when events are all independent...Hard to beat our method when correlated.
- Choice of method depends on willingness to accept independence assumptions
  - If design a clinical trial assuming independence, and really have correlation, then power is sorely underestimated. (Less so if use our method)
  - If design a clinical trial based on our method allowing for correlation, power will only go up if less correlated than planned (conservative choice)
- Also considering two-sample tests based on τ-year hazard differences during different follow-up intervals, mean number of events in τ-year intervals and area under restricted mean residual lifetime curves (work in progress)

Summary

- Hard to beat proportional mean model when events are all independent...Hard to beat our method when correlated.
- Choice of method depends on willingness to accept independence assumptions
  - If design a clinical trial assuming independence, and really have correlation, then power is sorely underestimated. (Less so if use our method)
  - If design a clinical trial based on our method allowing for correlation, power will only go up if less correlated than planned (conservative choice)
- Also considering two-sample tests based on τ-year hazard differences during different follow-up intervals, mean number of events in τ-year intervals and area under restricted mean residual lifetime curves (work in progress)

Summary

- Hard to beat proportional mean model when events are all independent...Hard to beat our method when correlated.
- Choice of method depends on willingness to accept independence assumptions
  - If design a clinical trial assuming independence, and really have correlation, then power is sorely underestimated. (Less so if use our method)
  - If design a clinical trial based on our method allowing for correlation, power will only go up if less correlated than planned (conservative choice)
- Also considering two-sample tests based on τ-year hazard differences during different follow-up intervals, mean number of events in τ-year intervals and area under restricted mean residual lifetime curves (work in progress)

Summary

- Hard to beat proportional mean model when events are all independent...Hard to beat our method when correlated.
- Choice of method depends on willingness to accept independence assumptions
  - If design a clinical trial assuming independence, and really have correlation, then power is sorely underestimated. (Less so if use our method)
  - If design a clinical trial based on our method allowing for correlation, power will only go up if less correlated than planned (conservative choice)
- Also considering two-sample tests based on τ-year hazard differences during different follow-up intervals, mean number of events in τ-year intervals and area under restricted mean residual lifetime curves (work in progress)

Summary

- Hard to beat proportional mean model when events are all independent...Hard to beat our method when correlated.
- Choice of method depends on willingness to accept independence assumptions
  - If design a clinical trial assuming independence, and really have correlation, then power is sorely underestimated. (Less so if use our method)
  - If design a clinical trial based on our method allowing for correlation, power will only go up if less correlated than planned (conservative choice)
- Also considering two-sample tests based on τ-year hazard differences during different follow-up intervals, mean number of events in τ-year intervals and area under restricted mean residual lifetime curves (work in progress)

Thank you!