Assessment of a treatment effect for recurrent event data in the presence of a terminal event

DSBS/FMS, October 23, 2018
Philip Hougaard,
Lundbeck, Denmark

Outline

- What is recurrent events?
- Why consider recurrent events? (Compared to first event)
- Models in absence of mortality
- Models in presence of mortality

Recurrent events

- Occurrence of similar events over time
- Examples: Multiple sclerosis; Heart failure; Hospitalizations

- **Overall frame:**
  Events experienced a few times (excludes epileptic seizures and diabetic hypoglycemia).
  Events experienced and recorded separately (time-stamped).
  Events may have consequences (such as mortality; additional treatment)

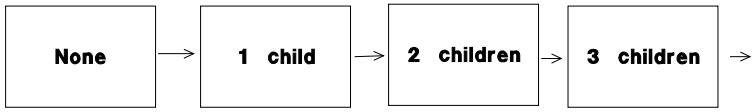
Why use recurrent events?

- **Current analysis method: Time-to-first-event only**
- Methods: Logrank test or Cox model
- Advantage: Statistical independence between events; implying simple analysis. No need to consider consequences such as mortality after an event

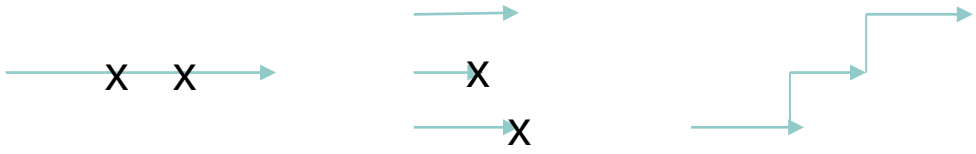
- **Arguments for using all events instead:**
- Better reflection of disease burden (slogan: ”The second event is as bad as the first” both for the individual patient and for the society)
- Time-to-first-event suffers from selection effects
- Potentially better statistical efficiency

Part 1: Mortality not considered

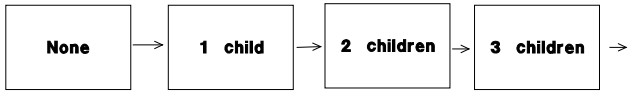
- Multi-state model:



- Three ways to illustrate observed data:
- Running time Gap times Counting process



- Continuous observation (0,C]
- Times of events: $T_1,\ldots T_N$, (N random; N ε {0,1,2,…})



Meaning of T_1?

- T_1: age of woman when having her first child
- For a woman having children, T_1 is well-defined
- For a woman never having children, we formally define $T_1 = \infty$, but can only observe $T_1 > C$, (C is end of observation)
- This solves the issue of old women without children but does not handle mortality
- Is this a problem or just an odd way of saying things?

- A multi-state model does not need to define T_1
- The ”hazard” of giving birth is defined as $\lambda_0(t)=\lim_{\Delta t\to 0}$ Prob{delivering child between age t and t+ Δt $\mid$ Alive at age t and without children}/ Δt
- The event of never having children is reflected by $\lim_{t\to\infty} \int_0^t \lambda_0(u)\, du < \infty$

Choice of time scale

- Possibility 1: **Running time** (time since start – birth or treatment start): $0 < T_1 < T_2 < T_3$ …
- Possibility 2: **Gap times** (time since most recent event)
  Note: First event must be handled differently than later events.
  $\Delta_1 = T_1$, $\Delta_2 = T_2 - T_1$, $\Delta_3 = T_3 - T_2$, …
- **Model overlap:** Constant hazard (exponential gap times with same parameter). Too restrictive (in my mind)
- In a typical drug trial, I recommend the running time approach. One reason is that the accumulated number of events is a meaningful quantity (see later)

Poisson process model

- Hazard of an event conditional on the history until that time is $\lambda(t)$ (independent of history)
- All subjects have the same risk (extendable in regression models)
- Two events in small interval has small probability (no simultaneous events)
- => Independent increments N_t

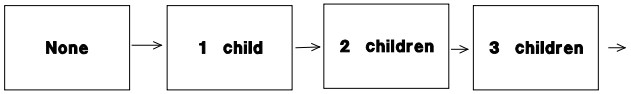
- Derived period count (Poisson distributed) over (0,t], using $\eta=\int_0^t \lambda(u)\, du$,
- $Pr(N_t = n) = \eta^n \exp(-\eta) / n!$
- $E\, N_t = \eta$; $Var(N_t) = \eta$

- Regression model:
- $\lambda(t;z) = \lambda(t) \exp(\beta z)$

Extension to handle dependence within subjects

- **Poisson:**
- $EdN_t = \lambda(t)$ conditional on history at time t (independent increments)

- **Marginal models:**
- $dEN_t = \lambda(t)$
- Use **Poisson** estimate (Nelson-Aalen) but calculate **robust variance estimate** (GEE)
- Details not presented



Multi-state hazard modelling

- The multi-state setup is sufficiently flexible to handle more complex models, with covariates describing the history:

- **First event:**
- $\lambda_0(t;z)= \lambda_0(t)$ (stratified)
- $\lambda_0(t;z)= \lambda_0(t) \exp(\beta z)$

- **After experiencing j events:**
- $\lambda_j(t)$ (stratified) (Markov) (Aalen-Johansen)
- $\lambda_j(t;z)= \lambda_j(t) \exp(\beta z)$ (Markov) (Prentice, Williams and Peterson (1981))
- $\lambda_j(t;z)= \lambda(t) \exp(\beta z + \gamma j)$ (Markov)

- Easy to fit – But can we interpret the results???

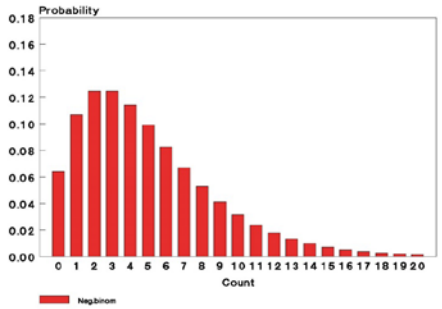
Issues with multi-state models

- **Issue 1: Number of events before time t**
- Example (no covariates): $\lambda_j(t)$ (Markov)
- Prob (no events before t): $\exp(-\int_0^t \lambda_0(u)\,du)$
- Prob ($N_t = 1$): $\int_0^t \lambda_0(u) \exp[-\Lambda_0(u) - \{\Lambda_1(t) - \Lambda_1(u)\}]\,du$)
- Complexity increases with number of events
- No expression for mean or other summaries

- **Issue 2: Which regression model?**
- Estimation easy for many different regression models, but not clear which model gives the best reflection of the treatment effect
- As hazard functions condition on the event history, they suffer from selection effects. Not automatic that comparing, say, $\lambda_j(t)$, between treatments gives a full picture of the usefulness of a treatment

Conditional Poisson (frailty) models

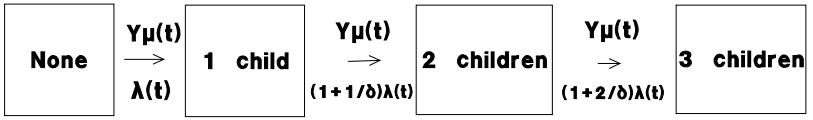
- **Subject differences** can be modelled as frailty model (overdispersion compared to Poisson)
- **Hazard:** Y μ(t), conditional on Y (subject random effect)
  t: time since study start; *not* time since latest event
- Poisson process conditional on Y =>
  N_t is Poisson (Y M(t)| Y)



- Many choices for distribution of Y possible (Hougaard, 2000) but the talk only considers the gamma distribution where the event count follows a negative binomial distribution, Greenwood and Yule (1920)

Relationships (without covariates)

- Poisson model conditional on Y
- Hazard: Y μ(t),
- Mean #events before t: Y M(t)



Above: conditional on Y

Below: unconditional (gamma frailty)

- Model when Y (gamma distributed, δ) is integrated out
- Hazard (first event): $\lambda_0(t) = \mu(t)\, \delta / (\delta + M(t))$
- Hazard (conditional on having j events):
  $\lambda_j(t) = \mu(t)\, (\delta + j) / (\delta + M(t)) = \lambda_0(t)\, (\delta + j) / \delta$
- Mean #events before t: M(t)

Relationships (with covariates)

- Poisson model conditional on Y
- Hazard: Y $\mu(t) \exp(\beta z)$
- Mean #events before t: Y $M(t) \exp(\beta z)$
- Hazard ratio (z vs 0) = Mean ratio = $\exp(\beta z)$

- Model when Y (gamma distributed, δ) is integrated out
- Hazard (first event): $\lambda_0(t;z) = \mu(t) \exp(\beta z) \delta / (\delta + M(t) \exp(\beta z))$
- Hazard (conditional on having j events):
  $\lambda_j(t;z) = \mu(t) \exp(\beta z) (\delta + j) / (\delta + M(t) \exp(\beta z)) = \lambda_0(t;z) (\delta + j) / \delta$
- Mean #events before t: $M(t) \exp(\beta z)$
- Hazard ratio (z vs 0) = $\exp(\beta z) (\delta + M(t)) / (\delta + M(t) \exp(\beta z))$
- Mean ratio = $\exp(\beta z)$

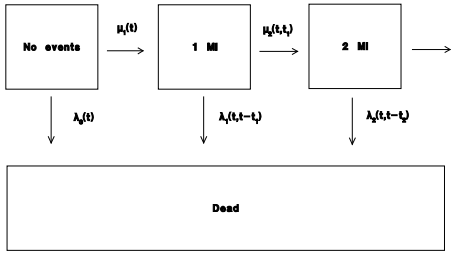
Conclusion from previous slides

- Suppose event occurrence is derived from a patient random effect setup (gamma frailty)

- A hazard ratio treatment effect conditionally => Same effect ratio for mean number of events conditionally => Same effect ratio for mean number of events unconditionally

- Different treatment effect on hazard functions unconditionally on frailty (conditional on history).
  Interpretation: Selection effects make the treatment appear less effective and the treatment effect appears to fade out over time.
  Conclusion: Considering only the time to first event gives an insufficient assessment of the treatment effect.
  Solution: Clinical trial to assess recurrent events rather first event only

Part 2: Recurrent events and mortality



- Mortality may occur in a clinical trial due to age; initial disease; treatment; study duration
- Patients do not get recurrent events after death
- Depending on the statistical approach used, this may make a treatment with high mortality appear as a treatment with few events: Is this preferred or are there approaches, where it does not happen?
- It cannot simply be assumed that recurrent events and death are independent

- The relation ”Mean #events before t = M(t)” is lost as only survivors can get events. Left hand side refers to all patients, but right hand side is conditional on surviving to time u: $(M(t)= \int_0^t \mu(u)du)$

Target for estimation

- Number of events (N_t): as events do not occur after death, the number can be low if mortality is high
- Integrated hazard (M(t)): The hazard refers to the at any time survivor
- Without mortality:
- $EN_t = M(t)$
- With mortality:
- $EN_t < M(t)$
- Q1: Which one do you prefer for assessing a treatment in a clinical trial?
- Q2: Is mortality controlled by the ***same factors*** as the events?: ***Measured*** or ***unmeasured***?

Same factors for mortality and recurrent events (1): Measured

- If ***measured*** (like treatment) and the covariate included in the model, this is not wrong in the technical sense, but looking at recurrent events and deaths separately may not give the full picture

- Treatment effect the same for death as for recurrent events???

Same factors for mortality and recurrent events (2): Unmeasured

- If ***unmeasured***, a joint model including frailties may be considered, like Y_1 for recurrent events and Y_2 for death
- Frailty Y_1 is well-defined but not necessarily Y_2 (because death is univariate) so the model has to be chosen with care:
- Model 1 (Rogers et al, 2014): Shared $Y_1 = Y_2$
- Model 2 (Rogers et al, 2016): $Y_1^{\alpha} = Y_2$

Summing up: 3 problems

- 1. Events do not occur after death, which could make a high-mortality treatment appear as giving few events
- 2. Events and deaths cannot be assumed independent
- 3. Treatment effect might differ between events and deaths

- Potential solutions:
- 1. Assess integrated hazards instead of realized events
- 2. Make a joint frailty model
- 3. Evaluate treatment effects separately

- Neither of these solutions fit into the estimand concept

Application: Rogers et al (2016) CHARM-added

	Candesartan	Placebo
Randomized	1276	1272
Follow-up time	3848 y	3713 y
CV deaths	302	347
Other deaths	75	65
First HF hospitalization	323	380
Later HF hospitalizations	266	433
Primary composite	497	561

Analysis results

- Protocol (composite): HR = 0.833 (CI: 0.74-0.94), p=0.003
- Cox CV-death: HR = 0.842 (0.72-0.98), p=0.029
- JFM CV-death: HR = 0.839 (0.70-1.01), p=0.065
- JFM HF hospitalizations: HR: 0.650 (0.53-0.80), p<0.0001
- θ= 3.751; α=0.688
- Results compatible with the theory:
- Different treatment effects (hospitalization vs. CV-death)
- Different treatment effects (first vs. later hospitalization due to selection)
- Dependence between hospitalization and CV-death (α)

Estimands

- **Without mortality:**
- Mean number of events before time t0
- Treatment effect:
  Ratio of means (In conditional Poisson model same as ratio of integrated hazards)

- **With mortality:**
- Mean number of events before time t0 (not preferred as events do not occur after death)
- Treatment effect:
  In conditional Poisson model: ratio of integrated hazards
  In marginal model: ratio of integrated hazards

Conclusion

- Relevant to consider recurrent events in clinical trials: Reflection of disease burden; no selection effects; more information
- There are statistical approaches to handle recurrent events: both model-based and non-model-based
- The choice is really between:
- *First event only:* Throw data away and get a simple conclusion
- *Recurrent events:* Use all data and get a more informative conclusion (which, of course, is more difficult to express)
- Mortality makes statistical calculations more difficult but we still have tools

References

- **Hougaard** (2000) Analysis of multivariate survival data, Springer
- **Prentice, Williams and Peterson** (1981) On the regression analysis of multivariate failure time data. Biometrika 68, 373-379
- **Rogers, Pocock et al** (2014) Analyzing recurrent hospitalizations in heart failure: a review of statistical methodology, with application to CHARM-Preserved. Eur. J. Heart Failure 16, 33-40
- **Rogers, Yaroshinsky et al** (2016) Analysis of recurrent events with an associated informative dropout time: Application of the joint frailty model. Stat. Med. 35, 2195-2205

Contact information

- Philip Hougaard
- Biometrics, Lundbeck, Denmark
- phou@lundbeck.com

Backup slides

Censoring patterns

- The Poisson, multi-state and frailty models allow censoring depending on the realized number of events
- Censoring not allowed to depend on the unobserved frailty
- Marginal models require censoring completely independent of recurrent event process

Application: Mammary tumors (rats)

- Female rats injected with carcinogen day 0
- Survivors (day 60) treated with retinoid (n=23) or control (n=25)
- Time in days to appearance of tumor(s) until day 182 (same period!)

- Poisson (day 60-182): R: 61/23=2.65; C: 149/25=5.96. Ratio of means: 0.445 (0.068)

- Negative binomial (gamma frailty): Ratio of means 0.445 (0.094). Significantly better fit

- Data: Gail, Santner and Brown (1980)
- Analysis: Hougaard (2000, p. 336)