Bayesian Interim Prediction of Recurrent Events in Clinical Trials

Yangfan Ren (Rice University)
Patrick Schloemer (Bayer AG)
Ming-Dauh Wang (Bayer Healthcare LLC)

Outline

- Background
- Methods
- Simulated case analysis
- Simulations
- Blinded vs. unblinded analysis
- Conclusions

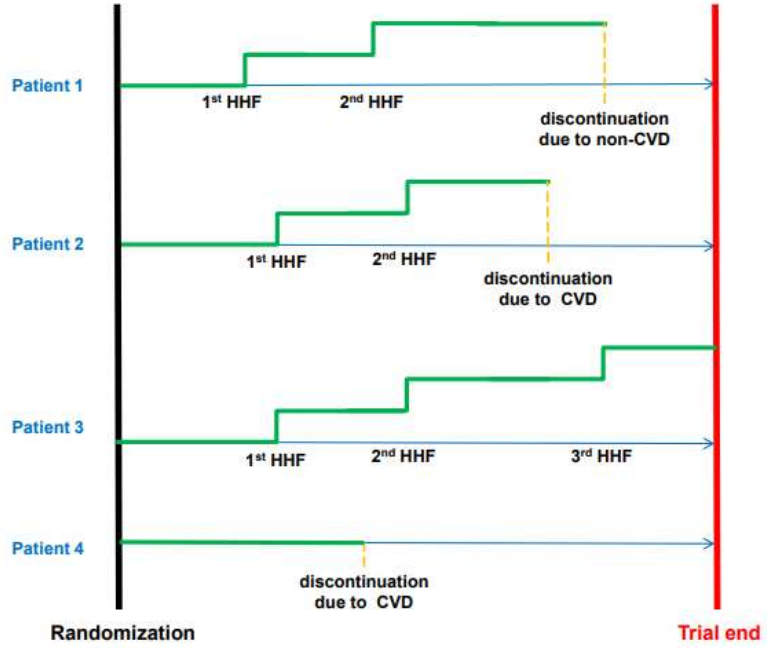
Background

- Recurrent events are repeated occurrences of the same type of events on the same patient over time in clinical trials which can be used as an evaluation of treatment effect.
- Motivated trials:
  - Hospitalization of heart failure (HHF) in heart failure trials – recurrent events.
  - Recurrent hospitalizations are strongly associated with cardiovascular death (CVD) – a terminal event.
- For better trial management, an interest is to predict when a pre-determined number of events can be achieved during an ongoing trial.

RICE

Background

Figure 1: Visualization of four distinct life history processes. CVD: cardiovascular death.



[1] Akacha, M., et al. (2018). Request for CHMP Qualification Opinion: Clinically Interpretable Treatment Effect Measures based on Recurrent Event Endpoints that Allow for Efficient Statistical analyses. Recurrent Event Qualification Opinion Consortium.

Background

- Plenty of research on event projection for time-to-first event available.
- Some of them are done under the Bayesian framework (Donovan et. al, 2006, Aubel et. al, 2020).
- Seminal work on the analysis of recurrent events with a terminal event: Rogers et. al (2016) and Akacha et. al (2018), but not for prediction.
- No known work on prediction of recurrent events in a Bayesian approach so far, especially not in the more complex situation with an associated competing event.

Methods

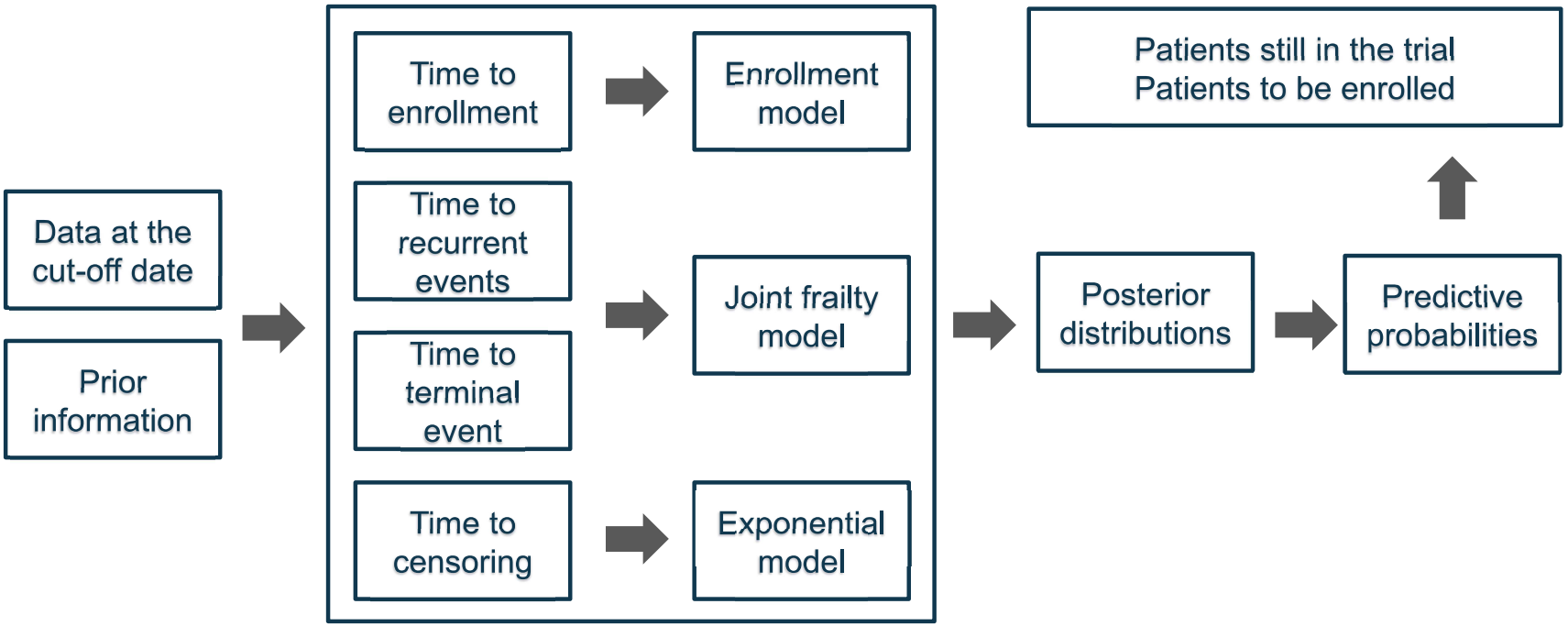


Figure 1: Framework of prediction

Methods

- Uniform enrollment is applied, which can be extended to Poisson-Gamma or site-wise Poisson-Gamma models.
- A joint frailty model [5] for the recurrent and terminal evets is defined through the hazards (unstratified gap-time model) :

$$r_{ij}(t|w_i) = w_i r_0(t),\ \ \lambda_i(t|w_i) = w_i^{\alpha}\lambda_0(t).$$

  - Hazard for recurrent event j of patient i is defined by r_{ij}, hazard for the terminal event of patient i is similarly defined as λ_i.
  - The patient specific frailty w_i follows Gamma($\frac{1}{\theta}, \frac{1}{\theta}$).
  - A parameter α correlating the recurrent and terminal events ($\alpha < 0,\ = 0,\ > 0$ indicate negative, zero, positive correlation).
- Time to censoring is assumed to be independent of recurrent and terminal events and exponentially distributed.

Methods

- Blinded interim prediction
  - Pooled analysis.
  - Bayesian latent class analysis.
- Unblinded interim prediction
  - Issues with unblinding, e.g., type I error control, likely introduction of bias.
  - Would it improve precision?

Simulated case analysis

- Reference: Akacha et al. (2018) - Request for CHMP Qualification Opinion
- Expected enrolled patients: 4350
- Expected recruitment period: 3 years
- Expected follow-up period: 2 years
- Expected events number: 1515
- Cut-off time for interim prediction: 2/3 of the expected events are obtained (1010)
- Uniform enrollment
- Exponential recurrent events – baseline hazard for placebo = 0.16788, HR = 0.7
- Exponential terminal event – baseline hazard for placebo = 0.06036, HR = 0.8
- Exponential censoring – baseline hazard = 0.01716
- Frailty variance = 5.7
- Alpha = 0.75

Simulated case analysis

// At what time, there is a 95% probability of reaching the expected number of recurrent events?

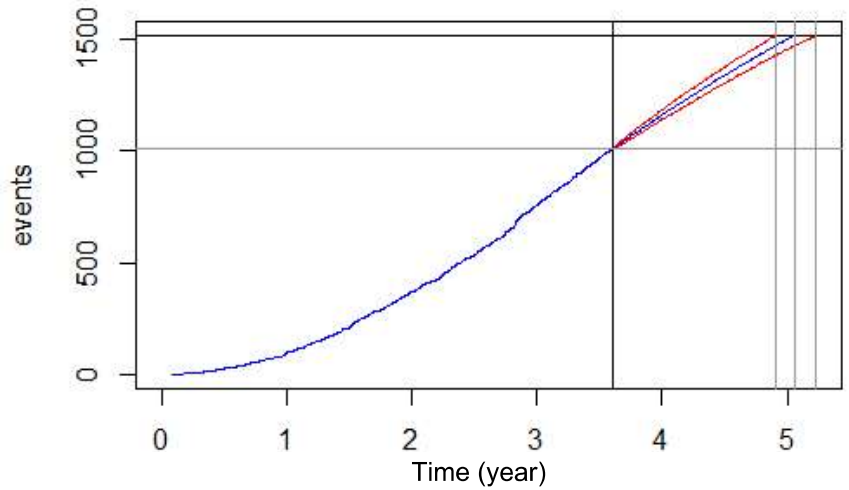


Figure 2: Median observed event number (blue) and 90% predictive interval for median predicted number of events (red)

Simulated case analysis

// How is the overshoot or undershoot in predicted event number?

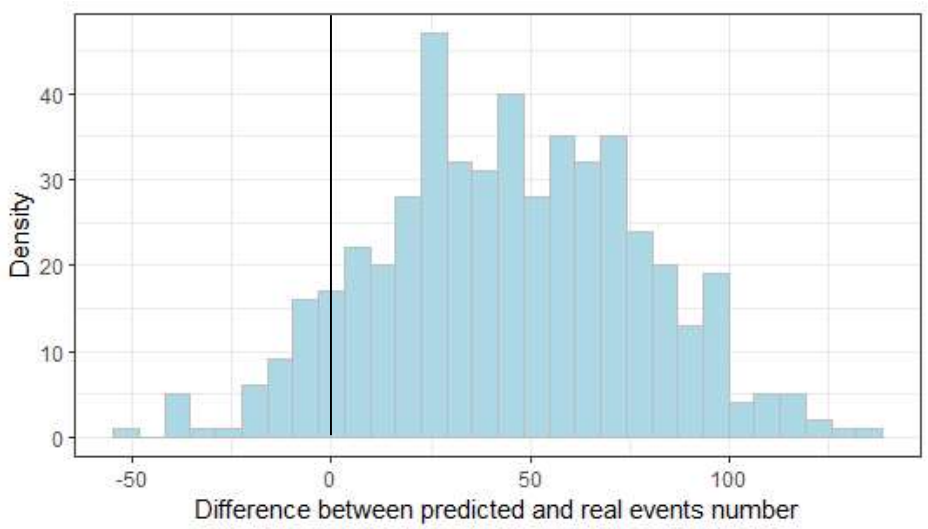


Figure 3: Difference in predicted and observed event numbers at the time when there is 90% predictive probability of achieving expected number or more (targeted number of events: 1515).

Simulated case analysis

// When do we want to do the interim prediction?

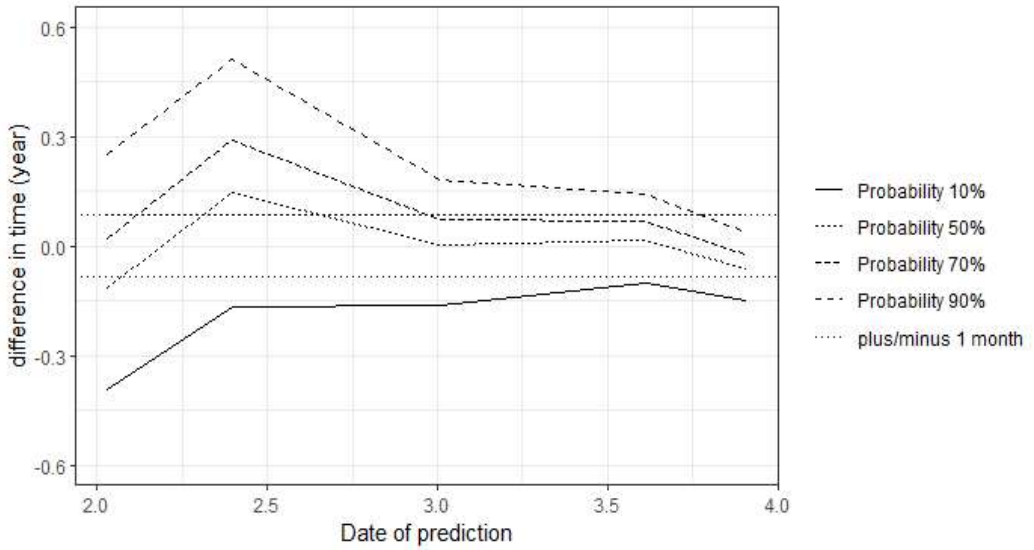


Figure 4: Difference in predicted and observed time with different cut-off date

Simulations

- Reference: Rogers et. al (2016)
- Expected enrolled patients: 300
- Expected recruitment period: 1 year
- Expected follow-up period: 2 years
- Expected events number: 320
- Uniform enrollment
- Exponential recurrent events – baseline hazard for placebo = 0.9, HR = 0.7
- Exponential terminal event – baseline hazard for placebo = 0.37 (annual event rates 31%), HR = 0.8
- Exponential censoring – baseline hazard = 0.05 (annual events rate 5%)
- Frailty variance (θ) = 1
- Alpha = 1

Simulations

// How does the treatment effect for recurrent events affect prediction precision?

Prediction precision

HR for recurrent events	1-mth precision	2-mth precision	3-mth precision	4-mth precision	5-mth precision	6-mth precision
0.8	28%	52%	70%	83%	90%	95%
0.7	21%	38%	58%	73%	83%	91%
0.6	17%	34%	49%	58%	70%	76%

Note:

- Cut-off while observing 67% events.
- The average time period for prediction when HR equals to 0.6 is 1.94, it is 1.57 for HR equals to 0.7 and 1.34 for 0.8.

// The greater the treatment effect for recurrent events, the more uncertain the pooled analysis likely due to higher data disparity in pooled analysis.

Simulations

// Does the treatment effect for terminal event affect prediction precision?

Prediction precision

HR for terminal event	1-mth precision	2-mth precision	3-mth precision	4-mth precision	5-mth precision	6-mth precision
1	17%	30%	47%	59%	68%	75%
0.8	21%	38%	58%	73%	83%	91%
0.6	25%	48%	69%	81%	90%	96%

Note:
- Cut-off while observing 67% events.
- The average time period for prediction when HR equals to 1 is 1.27, it is 1.57 for HR equals to 0.8 and 1.33 for 0.6.

// The greater the treatment effect for terminal event, the more precise the prediction is.

// Interplay between recurrent events and terminal event.

RICE

Simulations

// Does the baseline hazard for recurrent events affect prediction precision?

Prediction precision

Baseline hazard for recurrent events	1-mth precision	2-mth precision	3-mth precision	4-mth precision	5-mth precision	6-mth precision
0.9	21%	38%	58%	73%	83%	91%
1.2	50%	80%	93%	98%	99%	99%

Note:
- Cut-off while observing 67% events.
- The average time period for prediction when baseline hazard equals to 0.9 is 1.57 while it is 1.01 for baseline hazard equals to 1.2.

// Larger baseline hazard for recurrent events result in better prediction precision.

// Recurrent events occur earlier and more patients contribute to the estimates.

Simulations

// Does the baseline hazard for terminal event affect prediction precision?

Prediction precision

Baseline hazard for terminal event	1-mth precision	2-mth precision	3-mth precision	4-mth precision	5-mth precision	6-mth precision
0.29	38%	68%	82%	92%	96%	99%
0.37	21%	38%	58%	73%	83%	91%

Note:

- Cut-off while observing 67% events.
- The average time interval for prediction when baseline hazard equals to 0.29 is 1.31 while it is 1.57 for baseline hazard equals to 0.37.

// Smaller baseline hazard for terminal event results in better prediction precision.

// More terminal events lead to fewer patients contributing to the analysis.

Simulations

// Does the frailty variance affect prediction precision?

Prediction precision

Theta	1-mth precision	2-mth precision	3-mth precision	4-mth precision	5-mth precision	6-mth precision
1	21%	38%	58%	73%	83%	91%
1.5	16%	30%	44%	56%	67%	74%

Note:
- Cut-off while observing 67% events.
- The average time interval for prediction when theta equals to 1 is 1.57 while it is 1.28 for theta equals to 1.5.

// Larger patient heterogeneity makes estimation more difficult and decreases prediction precision.

Simulations

// Does alpha (correlation between recurrent events and terminal event) affect prediction precision?

Prediction precision

Alpha	1-mth precision	2-mth precision	3-mth precision	4-mth precision	5-mth precision	6-mth precision
1/1.5	28%	56%	75%	86%	94%	97%
1	21%	38%	58%	73%	83%	91%
1.5	21%	35%	49%	59%	70%	79%

Note:
- Cut-off while observing 67% events.
- The average time interval for prediction when alpha equals to 1/1.5 is 1.28, it is 1.57 for alpha equals to 1 and 1.49 for alpha equals to 1.5.

// Smaller α (>0) leads to smaller frailty variance for terminal event, which results in better prediction precision.

Bayesian latent class model

- Makes by-treatment prediction by first estimating patient assignments without unblinding.
- Model:
  - Modifications on joint frailty model.
  - Subjects are randomly assigned to treatment group with probability π and control group with probability $1-\pi$.
  - $g_i \sim Bern(\pi)$, $\pi \sim Beta(1,1)$.
  - Recurrent events model: $r_{ij}(t|w_i) = w_i r_{g_i}(t)$.
  - Terminal event model: $\lambda_i(t|w_i) = w_i^{\alpha}\lambda_{g_i}(t)$.
  - If $g_i = 0$, baseline hazard $r_{g_i} = r_0$, $\lambda_{g_i} = \lambda_0$. Otherwise, $r_{g_i} = r_1$, $\lambda_{g_i} = \lambda_1$.
  - Parameters to estimate: $r_0, r_1, \lambda_0, \lambda_1, \psi, \alpha, \pi$.

Blinded vs. unblinded analysis

Prediction precision

Model	1-mth precision	2-mth precision	3-mth precision	4-mth precision	5-mth precision	6-mth precision
Pooled	21%	38%	58%	73%	83%	91%
Latent	20%	40%	56%	69%	81%	90%
Unblinded	24%	45%	60%	73%	83%	89%

// Pooled analysis performs at least as well or even better in general compared to Bayesian latent class and unblinded models.

RICE

Conclusions

- Interim prediction of recurrent events is useful for trial management with selection of a reasonable timing.
- Pooled analysis is generally good compared to Bayesian latent class model and unblinded prediction.
- Bayesian analysis provides straightforward interpretations to help decision making.
- Potential extensions include more general models for enrollment and time to events, stratified recurrent event models, incorporation of covariates and event reporting lag, and use of informative priors.

RICE

Reference

[1] Akacha, M., Binkowitz, B., Bretz, F., Fritsch, A., Hougaard, P., Jahn-Eimermacher, A., Mendolia, F., Ravn, H., Roger, J., Schloemer, P., Schmidli, H. & Wei, J. (2018). Request for CHMP Qualification Opinion: Clinically Interpretable Treatment Effect Measures based on Recurrent Event Endpoints that Allow for Efficient Statistical analyses. Recurrent Event Qualification Opinion Consortium.

[2] Aubel, P., Antigny, M., Fougeray, R., Dubois, F., & Saint-Hilary, G. (2021). A Bayesian approach for event predictions in clinical trials with time-to-event outcomes. Statistics in Medicine, 40(28), 6344-6359.

[3] Bagiella, E., & Heitjan, D. F. (2001). Predicting analysis times in randomized clinical trials. Statistics in medicine, 20(14), 2055-2063.

[4] Mark Donovan, J., Elliott, M. R., & Heitjan, D. F. (2006). Predicting event times in clinical trials when treatment arm is masked. Journal of biopharmaceutical statistics, 16(3), 343-356.

[5] Rogers, J. K., Yaroshinsky, A., Pocock, S. J., Stokar, D., & Pogoda, J. (2016). Analysis of recurrent events with an associated informative dropout time: application of the joint frailty model. Statistics in medicine, 35(13), 2195-2205.

[6] Ying, G. S., & Heitjan, D. F. (2008). Weibull prediction of event times in clinical trials. Pharmaceutical Statistics: The Journal of Applied Statistics in the Pharmaceutical Industry, 7(2), 107-120.

Thank you!

Questions?