Comparison of Bayesian and Frequentist Meta-Analytical Approaches for Analyzing Time to Event Data

By Brenda Crowe
(Joint work with Monica Bennett, Karen Price, James Stamey and John Seaman Jr)

Basel Biometric Section of the Austro-Swiss Region of the International Biometric Society 2014
Comparison of Bayesian and Frequentist Meta-Analytical Approaches for Analyzing Time to Event Data

Monica M. Bennett, Brenda J. Crowe, Karen L. Price, James D. Stamey & John W. Seaman Jr.

\(^a\) Institute for Health Care Research and Improvement, Baylor Health Care System, Dallas, Texas, USA

\(^b\) Eli Lilly and Company, Indianapolis, Indiana, USA

\(^c\) Department of Statistical Sciences, Baylor University, Waco, Texas, USA


To link to this article: http://dx.doi.org/10.1080/10543406.2013.737210
Outline

• Background/motivation
• Simulation study (time to event data)
  • Methods, parameters
  • Software
  • Results
  • Discussion, recommendations
• References
Background/Motivation

- Lots of literature comparing MA methods for binary data
  - E.g., Sweeting et al. (2004, 2006), Bradburn et al. (2007)
- Not much for time-to-event data, though anticipate problems similar to binary data
Background/Motivation

- The guidance requires that the upper limit of the 2-sided 95% confidence interval for the risk ratio be less than 1.8 prior to submission and less than 1.3 after submission.
- Can be shown by performing a meta-analysis of phase 2 and 3 clinical trials and if these are insufficient, a large safety trial must be conducted.
Background/Motivation: Our Research

• Used simulation study to compare the performance of several meta-analytic approaches in the survival analysis context.

• Considered two frequentist approaches and a Bayesian approach with and without informative prior.
Statistical Issues with Meta-analysis of Rare/Sparse Adverse Event Data

• Standard inferences for meta-analysis rely on large sample approximations. They may not be accurate and reliable when number of events is low.
• Zero events observed in one or both treatment arms for some studies
• Low power to detect heterogeneity (especially when the number of studies is modest)
Simulation Study: Meta-analytical approaches for analyzing time to event data
Overview of Methods

1. Standard Cox proportional hazards (CPH)
2. CPH with Firth correction term (penalized likelihood)
3. Bayesian CPH (with and without informative prior)

- All methods model two treatment arms and stratify by study
The proportional hazards survival model for patient $i$ in study $j$ is

$$H_{ij}(t) = \lambda_{0j}(t) \exp(\beta x_{ij})$$

- $i = 1, \ldots, n_s$
- $j = 1, \ldots, s$
- $\lambda_{0j}(t)$ is the baseline hazard for study $j$
- $x_{ij} = 1$ if patient $i$ in study $j$ is on treatment and $x_{ij} = 0$ otherwise
- $\beta$ is the log hazard ratio.
CPH with Firth Correction

- When events are rare the problem of monotone likelihood can be encountered.
  - Estimates may not be available due to lack of convergence.
  - Estimates may be imprecise and have large standard errors.

- Firth (1993) developed a penalization method used to reduce bias in maximum likelihood parameter estimates.

- Heinze and Schemper (2001) adapted the Firth method to be used with the Cox model.
Bayesian CPH

- Basic model assumes constant baseline hazard over time and specifies prior distributions for $\lambda$ and $\beta$.

$$H_{ij}(t) = \lambda_{0j} \exp(\beta x_{ij})$$

$\beta \sim \text{Normal}(\mu, \sigma^2)$

$\lambda_{0j} \sim \text{Gamma}(a, b)$
Study Designs for Simulation

- 3 phase 2 studies:
  - $n_0 = 50$, $n_1 = 150$, duration = 90 days

- 3 phase 3 studies:
  - $n_0 = 250$, $n_1 = 500$, duration = 1 year

- 1 outcome study:
  - $n_0 = 3500$, $n_1 = 3500$, duration = 2 years

- 10% uniform dropout rate for all studies

Included in the 1st meta-analysis study grouping

Included in the 2nd study grouping
Simulation Design/Parameters

• Factorial layout:
  • 2 study groupings per previous slide
  • 3 hazard ratios: 1.0, 1.3, 1.8
  • 3 control event rates: 0.01, 0.02; 0.05 (events/person year)
• 1000 data sets are generated for each of the scenarios.
• Exponential distribution for data generation (constant hazard over time).

• Analysis methods

• Study group 1: 4 analysis methods (CPH, Firth, 2 Bayesian-diffuse and informative prior)

• Study group 2: 3 analysis methods (CPH, Firth, Bayesian-diffuse prior)
Bayesian Parameters

1. Diffuse priors
   - $\Lambda_0 \sim \text{gamma}(0.01, 0.01)$
   - $\beta \sim \text{normal}(0, 1000)$

2. More informative priors
   - Used shape parameter for gamma prior = 0.01, 0.02 and 0.05 for corresponding event rates
   - Rate parameter = 1
   - For log hazard ratio, for $\exp(\beta) = 1.0$, prior mean = 0
   - For $\exp(\beta) = 1.3$, prior mean was 0.25 and for $\exp(\beta) = 1.5$, prior mean was 0.5

Used prior variance of 2 for each. Informative priors were only used for first study grouping.
• PROC PHREG
  proc phreg data=meta.gendata;
    strata=study;
    *use FIRTH option to perform Firth correction;
    model time*event(0) = treatment / firth;
    *use BAYES statement for Bayesian analysis;
    bayes seed=1 initial = NBI= NMC= coeffprior= plots= ;
  run;
R

- `coxph{survival}`
  - `coxph(Surv(time,event)~ treatment + strata(study),
          data=gendata)`

- `coxphf{coxphf}`
  - `coxphf(Surv(time,event)~ treatment + strata(study),
            data=gendata)`

- For the Bayesian methods WinBUGS or OpenBUGS can be used.

- The models for the Bayesian methods are based on the model in the “Leuk: survival analysis using Cox regression” example in WinBUGS.
Simulation Results
Standardized Bias Plots: Meta-analysis of Phase 2 and 3 Trials.

Firth gives best results (closest to zero bias line) in all situations.
Bayes with informative prior has overly high coverage in all scenarios (as do CPH and Firth, but they have less bad). Bayes with diffuse prior has lower coverage than desired, with exception of one scenario (lambda = 0.05), which may be because of the bias seen on previous slide.
Proportion of Upper Bounds Less Than 1.8: Meta-analysis of All Phase 2 and 3 Studies

For true log HR = 0 and 0.262 (HR = 1, 1.3), higher proportions are better. For true HR = 1.8, lower are better. Firth does well/best in all situations.
Standardized Bias Plots: Meta-analysis of all Studies

All methods have std. bias close to zero, with exception of Bayesian method, where drops to -0.1 for HR = 1.8.
95% CI Coverage Plots: Meta-analysis of all Studies

Coverage in most scenarios is between 0.94 and 0.96.

Exceptions are when true log HR = 0. E.g., Bayes and CPH have coverage = 0.935 when baseline event rate is 0.01.
Proportion of Upper Bounds less than 1.8: Meta-analysis of All Studies.

All methods perform well.
References

References

References

Concluding Remarks: Time to Event Data

• Based on the scenarios we studied, the Firth correction to the CPH is a good option for analyzing time-to-event data when the baseline event rate is low.

• For Bayesian method, informative prior reduces the bias of the estimated log HR.
  • However a misspecified prior makes the situation worse (results not shown).

• With larger number of events there is not a big difference between the methods.
The End