

# Discrete Time-to-event and Score-based Methods with Application to Composite Endpoint for Assessing Evidence of Disease Activity-Free

Macaulay Okwuokenye

Biogen, Cambridge, MA

BASS XXII, Washington DC-Rockville  
November 02, 2015

Study support provided by Biogen. M. Okwukenye is a full-time employee of and holds stock/stock options in Biogen

# Outline

- 1 Background
- 2 Objective
- 3 Method and Notation
  - Estimating Proportion of Patients With(out) the Composite Endpoint
  - Grouped or Life-Table Time-to-Event Method
- 4 Inference on The Composite Endpoint Proportions
- 5 Assessment of Covariate Effect
- 6 Illustration
- 7 Other Discrete time-to-event methods
- 8 Score-Based Method
- 9 Statistical Inference for Score Approach

# Table of Contents

- 1 Background
- 2 Objective
- 3 Method and Notation
  - Estimating Proportion of Patients With(out) the Composite Endpoint
  - Grouped or Life-Table Time-to-Event Method
- 4 Inference on The Composite Endpoint Proportions
- 5 Assessment of Covariate Effect
- 6 Illustration
- 7 Other Discrete time-to-event methods
- 8 Score-Based Method
- 9 Statistical Inference for Score Approach

Composite endpoints are endpoints derived from more than one endpoint. They may be chosen in a trial:

- to capture the disease of interest better than any single endpoint
- to possibly increase power.
- to better characterize a disease that manifests in complex ways
- where no agreement exist among experts on the most relevant efficacy endpoint.

Consider a composite endpoint

- that is a function of outcomes of continuous right censored time-to-event endpoints and another endpoint(s) whose exact time of occurrence is unknown—but only known to have occurred within an interval.

## Background Continued...

E.g., evidence of disease activity (EDA) assessed using a composite endpoint comprising clinical disease activity (relapse and disability worsening) and lesions determined by magnetic resonance imaging (MRI), in relapsing-remitting multiple sclerosis (RRMS):

- The exact time of clinical relapse may be ascertained
- But exact time of lesion occurrence is unknown—MRI lesions are assessed periodically.
- All that may be known is that lesions are present or absent at time of assessment
- MRI lesions could have occurred before, after, or at the same time relative to clinical relapse or disability worsening.

EDA-free  $\implies$  no relapse, no disability worsening, no lesions determined by MRI

**QUESTION:** To what extent does treatment prevent or delay the composite endpoint?

Traditional statistical analyses practice:

- Assume that subjects who discontinued trial without EDA at their last visit are EDA-free [1] [2].
- Create a binary outcome (EDA-free? yes/no)
- Obtain a crude estimate of the proportion of subjects without EDA at the end of a  $t$ -year period.
- Fit a logistic regression to assess treatment effect adjusting for relevant baseline covariates.
- Obtain odds ratio for comparing proportion EDA-free

**Limitations:** Subjects' differential follow-up times are ignored and information on non-completers is not appropriately utilized:

- This can result in selection bias and overestimate of proportion
  - A treatment arm with high early dropout rate may appear artificially beneficial

May compromise randomization:

- Analyses population may not be the intention-to-treat (ITT) population
- Could induce severe selection bias as randomization can no longer be relied upon to guarantee a balance of unmeasured confounders.

Patterns of EDA are not reflected:

- EDA could occur early in a treatment group and latter in the other treatment group.



# Background Continued...

Some ad hoc fixes for analyses issues arising from unknown evidence of EDA status at discontinuation:

- Estimate proportion of subjects without EDA assuming all discontinuation have EDA
- Estimate another proportion, assuming none of the discontinuation have EDA (reference).

## Drawbacks of ad hoc fixes:

- The range of estimates from these two extremes could be quite wide for applicability.
- Classifying patients as having no EDA is impossible at any point without complete knowledge of the status of all the endpoints.

# Table of Contents

- 1 Background
- 2 Objective
- 3 Method and Notation
  - Estimating Proportion of Patients With(out) the Composite Endpoint
  - Grouped or Life-Table Time-to-Event Method
- 4 Inference on The Composite Endpoint Proportions
- 5 Assessment of Covariate Effect
- 6 Illustration
- 7 Other Discrete time-to-event methods
- 8 Score-Based Method
- 9 Statistical Inference for Score Approach

We consider discrete time-to-event statistical analysis methods:

## Advantages of time-to-event methods

Time-to-event method allows

- incorporation of subjects' differential follow-up times and appropriate handling of censoring
- appropriate weighting of loss to follow-up, and incorporation of subjects' information into the analyses for as long as they are known to be in the study.
- the patterns of EDA to be reflected.
- analyses in ITT population

# Table of Contents

- 1 Background
- 2 Objective
- 3 Method and Notation**
  - Estimating Proportion of Patients With(out) the Composite Endpoint
  - Grouped or Life-Table Time-to-Event Method
- 4 Inference on The Composite Endpoint Proportions
- 5 Assessment of Covariate Effect
- 6 Illustration
- 7 Other Discrete time-to-event methods
- 8 Score-Based Method
- 9 Statistical Inference for Score Approach

- Organize event times for the continuous time-to-event endpoints into intervals determined by the scheduled assessment visits of MRI endpoints.
- Create collapsed binary endpoint of any versus none of the events

## Proportion of Patients With the Composite Endpoint

- Apply actuarial (life-table [LT]) or reliability (Kaplan-Meier [KM]) method to estimate and compare the proportion of subjects without the composite endpoint in two treatment regimens.
  - cumulative distribution function cannot be estimated directly in presence of censoring; therefore, estimated as  $1 - \textit{survival}$  function.

## Allure of the proposed approach

- Withdrawals, loss to follow-up, and discontinuations are easily accounted for using any of the time-to-event analyses approaches.
- Results from life-table are easy to understand and interpret

# Method and Notation Continued...

LT and KM method make different assumptions about withdrawals.

## Reliability method

- For tied event and loss to follow-up times, KM method assumes all subjects lost to follow-up were at risk at the time of event.
- Number of withdrawals occurring in an interval are subtracted from the number at risk at the beginning of the interval.

## Actuarial method

- Life-table assumes withdrawals occur uniformly in an interval
- For tied event and lost to follow-up times, LT assumes that  $1/2$  of the subjects that were lost to follow-up were at risk at the time of events.
- $1/2$  number of withdrawals are subtracted from number at risk, as a protection against underestimation and overestimation of the proportion of subjects with the composite endpoint.

### Which is better?

- Driven by the assumptions on dropouts and length of intervals.
- Although KM estimate is the nonparametric MLE of survival function relative to the class of all distributions, it may be more reasonable to consider the number at risk for an interval to be the number at risk at the beginning of the interval minus  $1/2$  the number of withdrawals during the interval.
- Even though KM method is discrete (step function), KM method was developed for survival times on continuous scale with rare ties.
- Life-table is developed for grouped data, where ties are more likely.
- Life-table method recognizes that it is unreasonable to assume that none of the dropouts was at risk in the interval and all were not at risk for the entire interval.

## Method and Notation Continued...

- Let  $N_i$  denote the number at risk at the beginning of each interval;
- $D_i$ , the number known to have the event in the interval;
- $W_i$ , the number who discontinued in the interval not known to have the event at time of discontinuation.
- The conditional probability of having the event in the interval via actuarial estimation is

$$Q_i = \frac{D_i}{N_i - 1/2(W_i)} \quad (1)$$

- and the conditional probability of not having the event in the interval is  $P_i = 1 - Q_i$  given by

$$\frac{N_i - D_i - 1/2(W_i)}{N_i - 1/2(W_i)} \quad (2)$$



- The unconditional probability of not having the event is the product of the  $P'_i$ 's.
- For KM method,  $1/2(W_i)$  is dropped from Equation 2.
- Not considering the number at risk to account for the withdrawals makes the  $Q'_i$ 's smaller, leading to larger  $P'_i$ 's and larger cumulative nonevent rates.
- As the length of subintervals becomes smaller, the actuarial estimate of event probability approaches the KM estimate as a limit, a reason the KM is referred to as the product limit estimate.

## Proportion of Subjects Without the Composite Endpoint Across Studies

- Weighting may be applied to obtain pooled estimate of the proportions across studies.
- Denote by  $P_{ji}$  the probability (KM) of the composite endpoint for  $i$ th interval for study  $j$  and  $P_{j'i}$  the probability of the composite endpoint for the  $i$ th interval for study  $j'$ .
- The pooled estimate for interval  $i$ th is

$$\frac{N_{ji} \times P_{ji} + N_{j'i} \times P_{j'i}}{N_{ji} + N_{j'i}}, \quad j \neq j' \quad (3)$$

- Equation (3) above implies weighting the individual study proportion of composite endpoint estimates for interval proportionate to the number at risk for the interval from each study (Other weights, such as inverse variances, may be used).

- The variance of the pooled estimate in Equation (3) is

$$f_{ji}^2 \times \text{Var}(P_{ji}) + f_{jji}^2 \times \text{Var}(P_{jji}), \quad (4)$$

- where

$$f_{ji} = \frac{N_{ji}}{N_{ji} + N_{jji}} \quad \text{for } j = 1, 2, \dots, J \quad (5)$$

- and  $\text{Var}(P_{ji})$  is determined from greenwood [3] formula give by

$$\text{Var}(\hat{P}_{ji}) \approx [\hat{P}_{ji}]^2 \sum_{i=1}^I \frac{D_{ji}}{(N_{ji} - D_{ji}) \times D_{ji}}. \quad (6)$$

# Table of Contents

- 1 Background
- 2 Objective
- 3 Method and Notation
  - Estimating Proportion of Patients With(out) the Composite Endpoint
  - Grouped or Life-Table Time-to-Event Method
- 4 Inference on The Composite Endpoint Proportions**
- 5 Assessment of Covariate Effect
- 6 Illustration
- 7 Other Discrete time-to-event methods
- 8 Score-Based Method
- 9 Statistical Inference for Score Approach

## Inference on The Composite Endpoint Proportions

- Hypothesis of interest: time-to-composite event proportions (or patterns) are equal ( $H_0$ ) versus not equal ( $H_a$ ).
- The null hypothesis is  $H_0 : F_A = F_B$ , versus any appropriate contradiction of the null, where  $F_A$  and  $F_B$  are the cumulative times to composite endpoint distributions for treatment groups A and B, respectively. Alternatively,  $H_0$  may be stated in terms of  $S = 1 - F$ .
- Inference on the difference in cumulative time to composite endpoint patterns between two treatment groups may be obtained using the Cochran-Mantel-Haenzel (CMH) approach[4].
- Inference could be on the difference in cumulative time to non-composite endpoint patterns, if  $H_0$  and  $H_a$  are stated in terms of  $S = 1 - F$ .

## Method and Notation Continued...

Consider a set of multiple 2-by-2 tables whose column is indexed by treatment and its row, by number at risk at the beginning of an interval.

Table 1: Data Set-up in An Interval for Computing Mantel-Haenzel Statistics

Treatment	Event	No Event	At Risk
A	$n_{11}$	$n_{12}$	$n_{1+}$
B	$n_{21}$	$n_{22}$	$n_{2+}$
	$n_{+1}$	$n_{+2}$	$N_T$

For an interval let

- $n_{11}$  denote the number of subjects who have the composite endpoint;
- $n_{12}$  denote those without the composite endpoint among the number of subjects at risk  $n_{1+}$  at the beginning of the interval in treatment group A.
- $n_{21}$  denote the number of subjects with the composite endpoint;

## Method and Notation Continued...

- $n_{22}$  denote those without the composite endpoint among the number of subjects at risk  $n_{2+}$  at the beginning of the interval in treatment group B.
- $n_{+1}$  denote number of subjects with the composite endpoint in the interval and  $n_{+2}$  the number without the composite endpoint in the interval.
- $n_{1+} + n_{2+} + n_{+1} + n_{+2} = N_T$
- The Cochran-Mantel-Haenzel statistics (MH) [5] for assessing whether the time-to-composite endpoint pattern differs between the treatment groups A and B is

$$\chi_{MH}^2 = \frac{[|\sum_i n_{11} - (n_{1+} \times n_{+1})/N_T| - 0.5]^2}{V_{MH}} \quad (7)$$

where  $V_{MH} = \sum (n_{1+} \times n_{+1} \times n_{2+} \times n_{+2}) / (N_T^3 - N_T^2)$ .

- Subtraction of 0.5 is to adjust for lack of continuity.

## Assessment of Covariate Effect

- Cox's proportional hazard (Cox's PH) model [6] is commonly used in continuous time-to-event settings.
- A discrete analog of the Cox's PH may be used in the assessment of covariate effect in the case of grouped time-to-event data.
  - This may be easily achieved using a binary regression model with the complementary log-log linearization transformation [7] or using the piecewise Poisson regression model (PPRM) [8].
  - PPRM assumes that the exact time of the events are known; hence, the complementary log-log is favored in the present context since the exact time of event for the MRI outcomes are unknown.



- A statistical model describing association between  $T_i$  the discrete time-to-composite endpoint on an  $i$ th subject with covariate vector  $\mathbf{X}_i$  is

$$\ln[-\ln(1 - P_{ti})] = \beta_0 t + \mathbf{X}'_{ij} \beta \quad (8)$$

- $\mathbf{X}_{ij}$  is a  $P \times 1$  covariate vector that may be constant over time or time-dependent and  $\beta$  is the covariate effect.
- The function  $(1 - P_{ti})$  is the probability of being free of the composite endpoint.
- The grouped-time model in Equation (8) assumes that the events are generated by Cox's proportional hazard model [9, 10].
- Equation 8 yields hazard ratio estimates that are identical to those from the continuous time proportional hazard's model.

## Pooled Estimate from Multiple Studies

- First investigate treatment by study interaction.
- If little or no evidence exists to suggest treatment by study interaction is statistically significant, a fixed effects model that blocks on study may be utilized.
- If evidence exists to suggest between study heterogeneity, then inference may be based on a random effects model.
- Individual patient data (IPD) and Meta-Analysis (MA) are expected to reach similar conclusions [11]

# Table of Contents

- 1 Background
- 2 Objective
- 3 Method and Notation
  - Estimating Proportion of Patients With(out) the Composite Endpoint
  - Grouped or Life-Table Time-to-Event Method
- 4 Inference on The Composite Endpoint Proportions
- 5 Assessment of Covariate Effect
- 6 Illustration**
- 7 Other Discrete time-to-event methods
- 8 Score-Based Method
- 9 Statistical Inference for Score Approach

The nature of the data for present application is confidential. Instead, consider

- a hypothetical RRMS trial conducted for 104 weeks
- 180 patients in drug A; 182 patients in drug B
- exact times of clinical relapse are known
- MRI endpoints (Gd and T2 lesion) are assessed at week 26, week 52, and week 104

If composite endpoint is of interest at the design stage, it may be a good idea to assess MRI more frequently

# Illustration Continued...

**Table 2:** Life-Table Estimates of Evidence of Disease Activity-free Proportion

Interval Weeks	No Event & Discontinued	Events n(%)	Average # at Risk	P(No Event)	SE(P)
Drug A					
(0-26]	10	143	175	0.1829	0.0292
(26-52]	0	5	27	0.1490	0.0275
(52-104)	1	8	21.5	0.0936	0.0232
Drug B					
(0-26]	11	129	176.5	0.2691	0.0334
(26-52]	3	9	40.5	0.2093	0.0314
(52-106)	0	9	30	0.1465	0.0281
Hazard ratio					
Drug B versus A					

Note: Data are contrived for illustrative purpose only and do not represent any data from actual clinical trial. P(no event) and SE(P) are the life-table estimates of proportion of patients without the composite event and the corresponding standard error, respectively.

# Table of Contents

- 1 Background
- 2 Objective
- 3 Method and Notation
  - Estimating Proportion of Patients With(out) the Composite Endpoint
  - Grouped or Life-Table Time-to-Event Method
- 4 Inference on The Composite Endpoint Proportions
- 5 Assessment of Covariate Effect
- 6 Illustration
- 7 Other Discrete time-to-event methods**
- 8 Score-Based Method
- 9 Statistical Inference for Score Approach

# Interval Censored Method

One could use interval-censored approach to estimate proportion without the composite endpoint.

- Suppose the subjects are assessed periodically for the composite endpoint
- It has been shown that non-parametric cumulative non-composite proportion estimate can only jump in the Turnbull interval  $(j_1; k_1 \dots j_m; p_m)$  that is data driven. See [12] and [13]
- Estimation using this approach requires more complex algorithm because of possible convergence issue.
- Number of unknown parameters might increase as the sample size increases; hence, standard errors of the survival estimates based on the likelihood function are no longer valid
- The ICLIFETEST procedure in SAS computes standard error using multiple imputation and bootstrap methods due to [14].

## Interval Censored Method Continued...

- the hazard ratio comparing risk of composite endpoint between treatment groups can be obtained using Cox's PH model for interval censored data
- See SAS ICLIFETEST and ICPHREG documentation for a software implementation
- A good summary of estimation can also be found in [15]



# Drawback of Collapsed Binary Composite

Some of the drawbacks of collapsed binary endpoint include:

- Overweighting of component(s) of the composites that occur at higher-frequency [16].
- Estimates of treatment effect and results of test are driven by a component with the largest frequencies
  - This could be an issue, particularly, when the overweighted higher-frequency component is of little clinical value.
- It is possible that clinicians are not sure which component of the composite that should be weighted highest.
- Inconsistent treatment effects among the components could pose a challenge because treatment effect may defer for each of the component.

## Drawback of Collapsed Binary Composite Continued...

- It does not discriminate between treatment groups in which subjects experience different number of the individual events.
  - Consider 6 subjects: 3 placebo; 3 active treatment.
    - 2 out of the 3 in placebo experienced event measure by the four endpoints; 2 out of 3 subjects in active treatment experienced only one of the four endpoints.
  - In using collapsed composite endpoint, proportion of subjects with the composite are the same (2/3).
  - However, the level of disease activity (1 vs 4 events) can not be said to same for the two treatment groups.

# Score-Based Method

Denote a possible binary outcome by  $a$  or  $b$  depicting whether or not an event is observed.

- An expression for  $n$  components each with a binary outcome is

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k. \quad (9)$$

- Equation (9) is the Binomial Theorem, where for EDA-free above  $n = 4$ .
- The coefficients from expansion of Equation (9) are related to coefficients obtained from the Pascal Triangle.
- These coefficients enable creation of a metric for severity of disease activity score under EDA-free depending on EDA combinations.
- This metric induces 16 categories of severity of disease activity under EDA-free which could be scored according to perceived severity.

## Score-Based Method Continued...

Under this metric,

- there are four categories for Gd lesions (Gd), T2 lesions (T2), clinical relapse (relapse) and confirmed disability worsening (CDW) only—taken one at a time;
- six categories for Gd, T2, relapse, or CDM taken two at a time, etc.
- EDA-free commands a rank of 1, and the category by taking all four endpoint at a time would command a rank of 16.
- the ranks 2,3,4 5 would apply to the category taking one at a time,
- the ranks 6,7,8,9,10, and 11 would apply to the category taking two at a time,
- the ranks 12, 13, 14, 15 would apply to the category taken three at a time.

- Score should reflect perceived order of importance of the individual components
- The assignment of the scores may be accomplished with clinical input from project clinician or other clinical experts.
- If the project clinician can not decide, the average rank of each category may be assigned to each member of the group.
- One would still be able to estimate crude proportion of subjects with without evidence of disease activity
  - This is the proportion of subjects with score of 1.

# Table of Contents

- 1 Background
- 2 Objective
- 3 Method and Notation
  - Estimating Proportion of Patients With(out) the Composite Endpoint
  - Grouped or Life-Table Time-to-Event Method
- 4 Inference on The Composite Endpoint Proportions
- 5 Assessment of Covariate Effect
- 6 Illustration
- 7 Other Discrete time-to-event methods
- 8 Score-Based Method
- 9 Statistical Inference for Score Approach

# Statistical Inference for Score Approach

- If the data were cross-tabulated as a 2 (treatment)-by-16 scores, then one could compute the mean score (and variance) for each row and statistically compare the rows.
  - This can be done with the row mean scores resulting from the PROC FREQ in SAS software with CHM option.
  - The resulting test statistics would be a Chi-square with one degree of freedom.
- For large samples, the same inference can be achieved with analysis of variance [17] using PROC GLM with the model statement `score=treatment+error`.
  - The inference is based on the  $F$ -distribution with  $d$  degree of freedom ( $df$ ), where  $d$  is the  $df$  associated with the mean square error (MSE).
  - The  $F$ -distribution with one  $df$  for the numerator approaches the Chi-square with one  $df$  when the denominator  $df$  approaches infinity
  - The inference would be toward one treatment group having greater (or less) average score than the other.

# Statistical Inference for Score Approach Continued...

- If there is concern that distribution of the scores is not normal to rely on inference using GLM, inference based on a permutation/randomization test or other appropriate non-parametric method may be pursued

Score method for inference is based on the implied assumption that

- there is no ordering in the occurrence of events—that is, the individual event can happen in any order
  - Modification is warranted if the events can happen in a particular order
- occurrence of all the four events is more severe than the occurrence of three, which is more severe than occurrence of two, etc.



# Performing Permutation Test

What if residuals,  $e_i = y_i - x_i\beta_i$ , are not normal?

- Outline of steps on how one may proceed with permutation test when residuals are not normal or variance are not homogeneous
  - Permutation test stems from the idea that if the null hypothesis is true, changing the exposure would have no effect on the outcome.
  - By randomly rearranging the exposures we create up as many data sets as we like.
  - If the null hypothesis is true, the rearranged data sets should look like the real data, otherwise they should look different from the real data.
  - The ranking of the original test statistics among the permuted test statistics gives a p-value.
- The steps are:

- 1 Combine data from treatment group A and B (call this sample S).
- 2 From S, randomly sample NA subjects (NA=180 from previous example) and call this new group A. The rest is new group B.
  - All possible assignment of NA patients in A and NB patients in B are produced and represents the sample space for the permutation distribution.
  - The p-value is the probability of observing the original data or data more extreme. One locates the mean difference corresponding to the original data ( $\bar{X}_A - \bar{X}_B$ ) and the tail area to its right is the p-value.
- 3 Estimate treatment effect (mean difference) using new group A and new group B. No matching here.
- 4 Using the same sample as in step 3 above, genetically match patients on covariates in new group A and new group B. Then estimate treatment effect.
- 5 Form difference in estimated treatment effects from step 3 (no matching) and step 4 (genetic matching). See [18] .

- Repeat steps 2 to 5 for say 1,000 times, producing 1,000 differences between estimates of treatment effect with matching and no matching. Note: There are  $\binom{NA+NB}{NA} = (NA+NB)! / NA! NB!$  possible samples.
- Sort the 1,000 differences from step above and plot frequency distribution and generate QQ plot.
- The above will produce the permutation distribution of  $\mu_A - \mu_B$ , from which you can obtain the point estimate and p-value.
- Matching is only necessary if the groups are imbalance.

# Acknowledgment

We thank Annie Zhang, Amy Pace, Fabio Pellegrini, Carmen Castrillo-Viguera, Shibeshih Belachew, Christophe Hotermans, and Leslie Meltzer for their input in assigning scores, and WM Biostatistics team.

# References I

- [1] L. Kappos, E. Radue, P. O'Connor, and et al., "Fingolimod treatment increases the proportion of patients who are free from disease activity in multiple sclerosis: Results from a phase 3, placebo-control study (freedom)," *Neurology*, vol. 76 (Suppl 4), no. A563, 2011.
- [2] R. Nixon, N. B. D. Tomic, N. Sfikas, G. Cutter, and G. Giovannoni, "No evidence of disease activity: Indirect comparisons of oral therapies for the treatment of relapsing-remitting multiple sclerosis," *Advance Therapy*, 2014.
- [3] M. Greenwood, "The natural duration of cancer," *Reports on Public Health and Medical Subjects (London: Her Majesty's Stationary Office)*, vol. 33, pp. 1–26, 1926.
- [4] N. Mantel and W. Haenszel, "Statistical aspects of the analysis of data from retrospective studies of disease," *Journal of the National Cancer Institute*, vol. 22, no. 4, pp. 719–748, 1959.
- [5] N. Mantel, "Chisquare test with one degree of freedom: Extension of the mantel-haenszel procedure," *Journal of the American Statistical Association*, vol. 58, pp. 690–700, 1963.
- [6] D. R. Cox, "Regression models and life-tables (with discusion)," *Journal of the Royal Statistical Society, Serries B*, vol. 34, pp. 187–220, 1972.
- [7] D. W. Hosmer and S. Lemeshow, *Applied Survival Analysis: Regression Modeling of Time to Event Data*. John Wiley and Sons, Inc, 1999.

# References II

- [8] M. E. Stokes, C. S. Davis, and G. G. Koch, *Categorical Data Analysis Using the SAS System*, second edition ed. Cary, NC: SAS Institute Inc, 2000.
- [9] D. Hedeker, O. Siddiqui, and F. B. Hu, "Random-effects regression analysis of correlated grouped-time survival data," *Statistical Methods in Medical Research*, vol. 9, pp. 161–179, 2000.
- [10] R. L. Prentice and L. A. Gloeckler, "Regression analysis of grouped survival data with application to breast cancer data," *Biometrics*, 34, vol. 34, pp. 57–67, 1978.
- [11] D.-G. D. Chen and K. E. Peace, *"Applied Meta-Analysis with R"*, 1st ed. Chapman and Hall/CRC, 2013.
- [12] R. Peto, "Experimental survival curves for interval-censored data," *Applied Statistics*, vol. 22, pp. 86–91, 1973.
- [13] B. W. Turnbull, "The empirical distribution function with arbitrarily grouped, censored, and truncated data," *Journal of the Royal Statistical Society*, vol. Series B, 38, pp. 290–295, 1976.
- [14] J. Sun, "Variance estimation of a survival function for interval-censored survival data," *Statistics in Medicine*, vol. 20, pp. 1249–1257, 2011.
- [15] C. Guo, Y. So, and G. Johnston, "Analyzing interval-censored data with the iclifetest procedure," 2014.

# References III

- [16] E. J. Mascha and D. I. Sessler, “Design and analysis of studies with binary- event composite endpoints: Guidelines for anesthesia research,” *Anesthesia and Analgesia*, vol. 112, no. 6, pp. 1461–1470, June 2011.
- [17] G. A. Mack and J. H. Skillings, “A friedman-type rank test for main effects in a two-factor anova,,” *Journal of the American Statistical Association*, vol. 75, pp. 947–951, 1980.
- [18] K.-T. Tsai and K. Peace, “Analysis of subgroup data of clinical trials,,” *Journal of Causal Inference*, vol. 1, no. 2, 2013.