Time-To-Event Analysis in Randomized Clinical Trials with Recurrent Events

Lehana Thabane, Professor,
Health Research Methods, Evidence, and Impact,
McMaster University
Vice-President Research, St Joseph's Healthcare Hamilton
(Ontario, Canada)

Society for Clinical Trials

May 21st, 2024

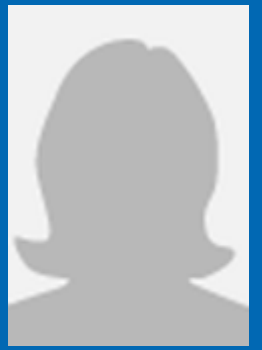
2024 BOSTON

SCT | 45TH ANNUAL MEETING

Speakers

Shun-Fu Lee, Ph.D. Assistant Professor, Health Research Methods, Evidence, and Impact, Faculty of Health Sciences at McMaster, Senior Statistician at Population Health Research Institute, (Ontario, Canada)

Ann-Kathrin Ozga (Ph.D. Senior statistician), University Medical-Center Hamburg Eppendorf, Institute of Medical Biometry and Epidemiology (Hamburg, Germany)



Lu Mao, Ph.D. Associate Professor, Department of Biostatistics and Medical Informatics, School of Medicine and Public Health, University of Wisconsin – Madison **(**Wisconsin, USA)

SCT | 45TH ANNUAL MEETING

Total-event analysis approaches in presence of between-subject heterogeneity in cardiovascular trials with composite outcomes

Shun Fu Lee

Assistant Professor

Health Research Methods, Evidence, and Impact (HEI),

Faculty of Health Sciences at McMaster,
Senior Biostatistician at Population Health Research Institute (PHRI)

Society for Clinical Trials

May 21st, 2024

SCT | 45TH ANNUAL MEETING

Illustration: Time-to-first occurrence of the composite

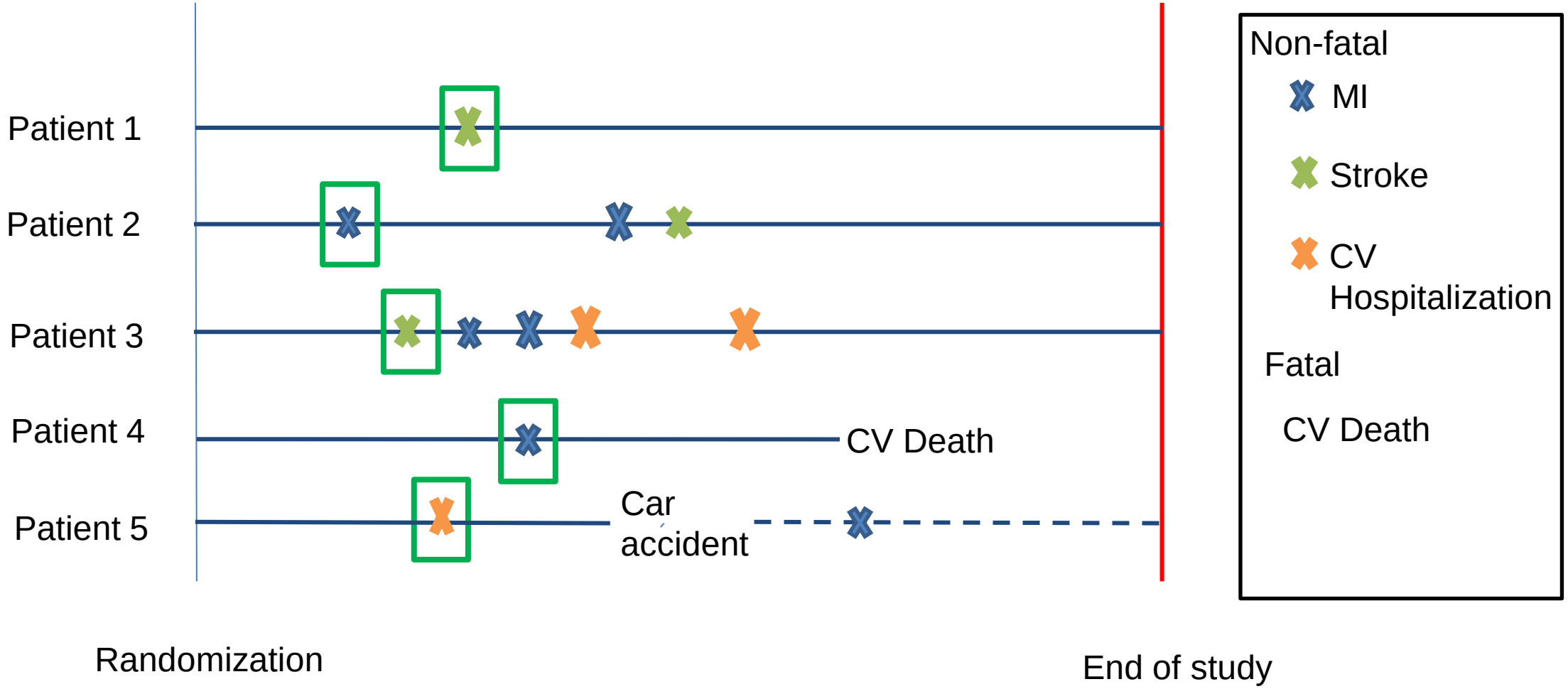
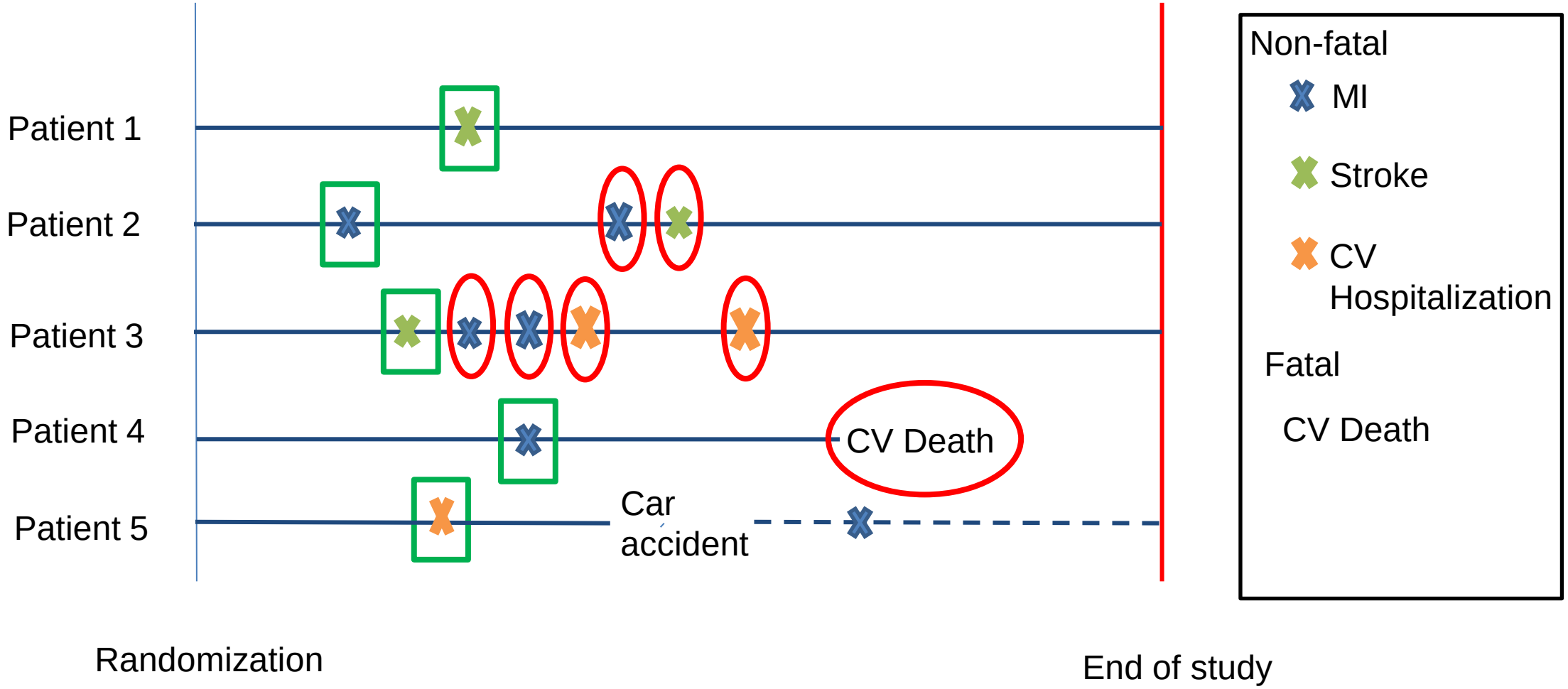


Illustration: Time-to-all occurrence of the composite (Total Event Analysis)



Three PHRI cardiovascular randomized trials:

	ORIGIN (Gerstein, 2012)	COMPASS (Eikelboom, 2017)	TRANSCEND (Yusuf, 2008)
Treatment Groups	Glargine vs Standard	Rivaroxaban+Aspirin vs Aspirin Alone	Telmisartan vs Placebo
N	12,537	18,269	5,926
Mean Follow-up in years (SD)	5.9 (1.4)	1.9 (0.7)	4.6 (1.0)
Composite 1: MI, Stroke, CV death	2054 (16.4%)	875 (4.8%)	824 (13.9%)
Composite 2: MI, Stroke, HF hosp, CV death	2394 (19.1%)	1085 (5.9%)	969 (16.4%)

Objectives

- Describe the increase in incidence after incorporating recurrent events
- What is the between-subject heterogeneity?
- How does the heterogeneity associate with the different analysis approaches?
- How do the treatment effect and confidence interval differ across different analysis approaches after incorporating recurrent events?

Recurrent Event Analysis Approaches

Traditional Cox Proportional Hazard model  Independence Assumption Violation

Extensions of the Cox Models:

- Andersen-Gill (AG) (1982): Assuming recurrent events within subjects are independent.
- Prentice, William and Peterson (PWP) (1981) total time: start with the entry date, conditional risk sets.
- Prentice, William and Peterson (PWP) (1981) gap time: reset the start date to zero after an event occurred, conditional risk sets.
- Wei, Lin Weissfeld (WLW) (1989): cumulative time from randomization to event
- Lin-Wei-Yang-Ying (LWYY) (2000): same estimate as AG but construct CI using a robust sandwich type estimator (SE) to account for the clustering of events in high-risk patients.

Factors to consider for recurrent events

1. Number of events

2. Starting

3. Spacing/gap

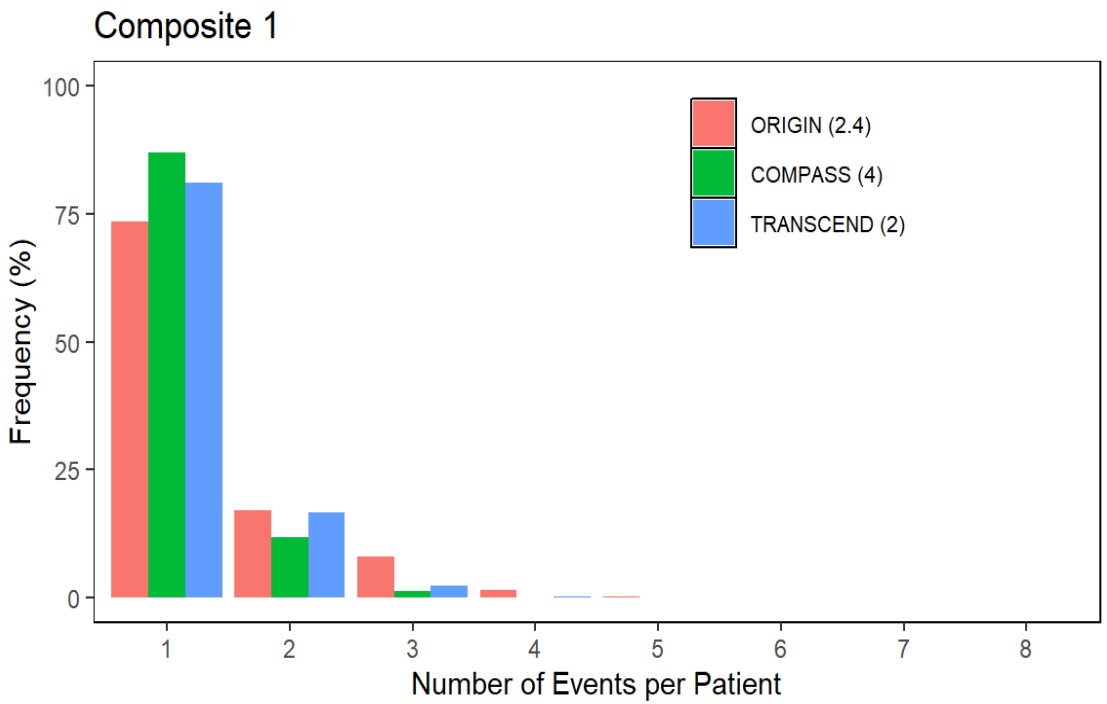
4. Correlation between successive events → Between Subject Heterogeneity

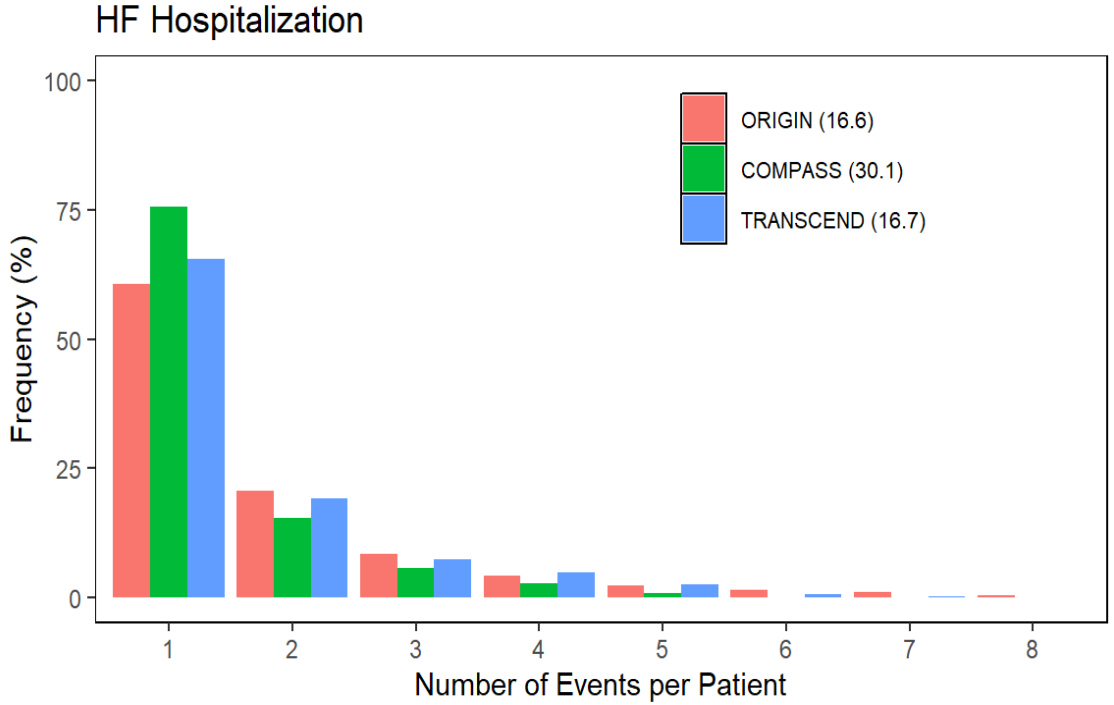
5. Total follow-up time

Cook & Lawless, 2007

Between-Subject Heterogeneity

A measure of patient-level risk of event occurrence relative to the population average.
Large heterogeneity suggests that recurrent events may result from a few patients with many occurrences, rather than from many patients with few occurrences. This measure is associated with the statistical power gained from incorporating recurrent events into a total-event analysis (Claggett et al, 2018)





ORIGIN (N=12,537)

	1st Event			Total Event			Between-subject heterogeneity
	Overall N (/100 person years)	Treatment N (/100 person years)	Control N (/100 person years)	Overall N (/100 person years)	Treatment N (/100 personyears)	Control N(/100 person years)	
Composite 1	2054 (2.90)	1041 (2.94)	1013 (2.85)	2797 (3.78)	1407 (3.81)	1390 (3.75)	2.4
Composite 2	2394 (3.43)	1205 (3.45)	1189 (3.41)	3875 (5.23)	1899 (5.14)	1976 (5.33)	3.4
CV Death	1156 (1.56)	580 (1.57)	576 (1.55)	---	---	---	---
Myocardial Infarction	662 (0.92)	336 (0.93)	326 (0.90)	789 (1.07)	396 (1.07)	393 (1.06)	5.3
Stroke	650 (0.90)	331 (0.91)	319 (0.88)	852 (1.15)	431 (1.17)	421 (1.14)	8.7
HF Hospitalization	653 (0.90)	310 (0.85)	343 (0.95)	1078 (1.46)	492 (1.33)	586 (1.58)	16.6
Non CV Death	760 (1.03)	371 (1.00)	389 (1.05)	---	---	---	---

COMPASS (N=18,269)

	1st Event			Total Event			Between-subject heterogeneity
	Overall N (/100 person years)	Treatment N (/100 person years)	Control N (/100 person years)	Overall N (/100 person years)	Treatment N (/100 personyears)	Control N(/100 person years)	
Composite 1	875 (2.53)	379 (2.18)	496 (2.88)	1006 (2.86)	432 (2.45)	574 (3.28)	4.0
Composite 2	1085 (3.16)	485 (2.81)	600 (3.52)	1403 (3.99)	638 (3.62)	765 (4.37)	5.9
CV Death	363 (1.03)	160 (0.91)	203 (1.16)	---	---	---	NA
Myocardial Infarction	383 (1.10)	178 (1.02)	205 (1.18)	405 (1.15)	188 (1.07)	217 (1.24)	3.7
Stroke	225 (0.64)	83 (0.47)	142 (0.82)	238 (0.68)	84 (0.48)	154 (0.88)	7.8
HF Hospitalization	300 (0.86)	155 (0.89)	145 (0.83)	397 (1.13)	206 (1.17)	191 (1.09)	30.1
Non CV Death	328 (0.93)	153 (0.87)	175 (1.00)	---	---	---	---

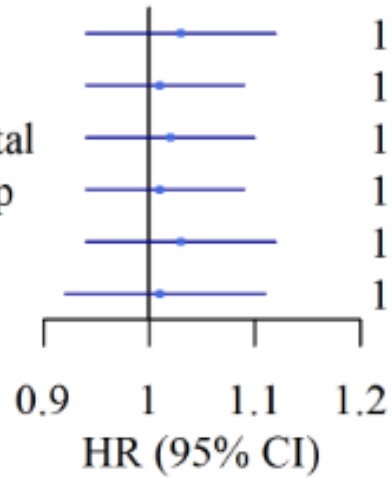
TRANSCEND (N=5,926)

	1st Event			Total Event			Between-subject heterogeneity
	Overall N (/100 person years)	Treatment N (/100 person years)	Control N (/100 person years)	Overall N (/100 person years)	Treatment N (/100 personyears)	Control N(/100 person years)	
Composite 1	824 (3.11)	384 (2.90)	440 (3.33)	1016 (3.69)	474 (3.46)	542 (3.92)	2.0
Composite 2	969 (3.73)	465 (3.58)	504 (3.87)	1417 (5.15)	656 (4.79)	761 (5.51)	3.1
CV Death	450 (1.64)	227 (1.66)	223 (1.61)	---	---	---	---
Myocardial Infarction	263 (0.97)	116 (0.86)	147 (1.09)	285 (1.04)	122 (0.89)	163 (1.18)	4.1
Stroke	248 (0.92)	112 (0.83)	136 (1.01)	281 (1.02)	125 (0.91)	156 (1.13)	2.9
HF Hospitalization	263 (0.97)	134 (1.00)	129 (0.95)	401 (1.46)	182 (1.33)	219 (1.59)	16.7
Non CV Death	263 (0.96)	137 (1.00)	126 (0.91)	---	---	---	---

Treatment Effects for ORIGIN

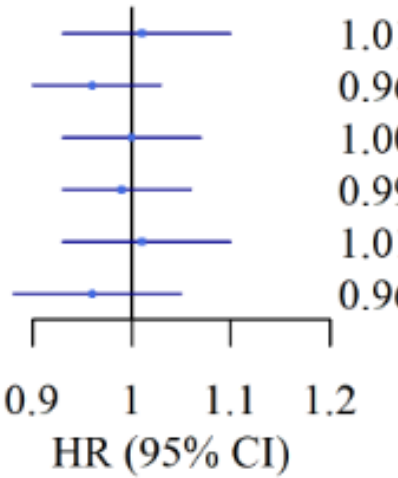
a) ORIGIN Composite 1 (2.4*)

Method	HR (95% CI)	CI Width
Cox	1.03 (0.94 - 1.12)	0.18
AG	1.01 (0.94 - 1.09)	0.15
PWP Total	1.02 (0.94 - 1.10)	0.16
PWP Gap	1.01 (0.94 - 1.09)	0.15
WLW	1.03 (0.94 - 1.12)	0.18
LWYY	1.01 (0.92 - 1.11)	0.19



b) ORIGIN Composite 2 (3.4*)

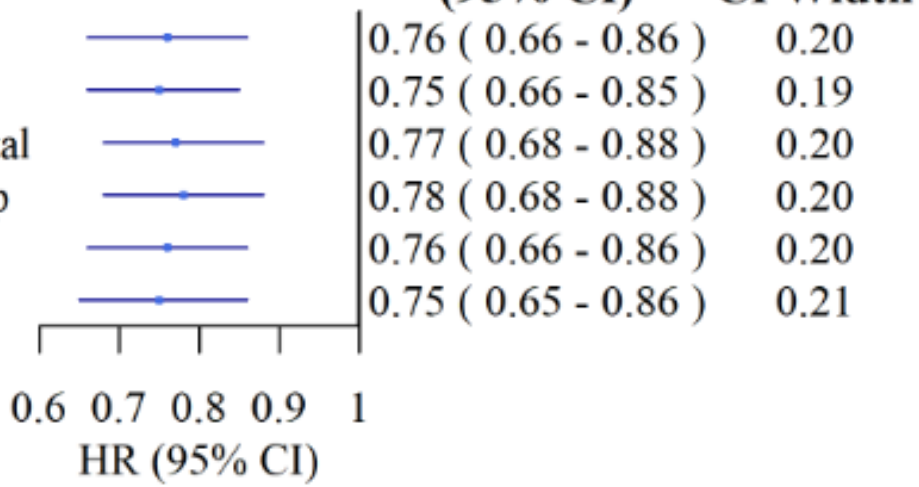
HR (95% CI)	CI Width
1.01 (0.93 - 1.10)	0.17
0.96 (0.90 - 1.03)	0.13
1.00 (0.93 - 1.07)	0.14
0.99 (0.93 - 1.06)	0.13
1.01 (0.93 - 1.10)	0.17
0.96 (0.88 - 1.05)	0.17



Treatment Effects for COMPASS

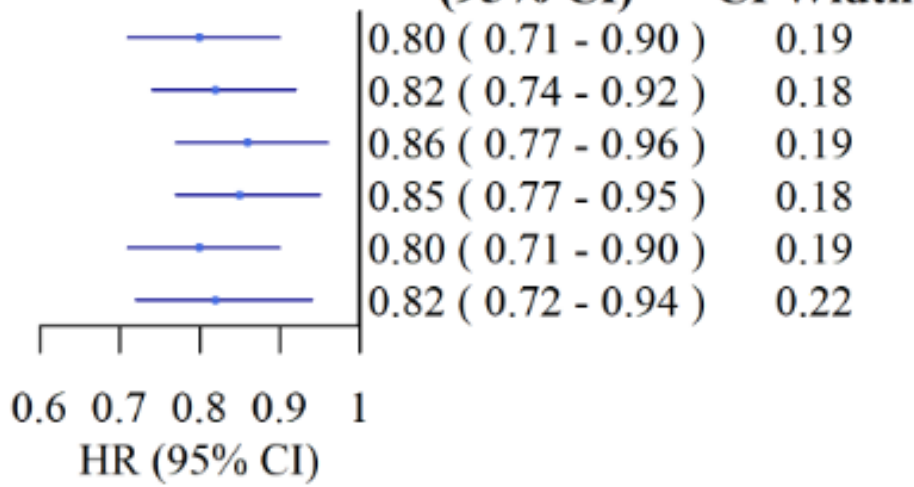
c) COMPASS Composite 1 (4.0*)

Method	HR (95% CI)	CI Width
Cox	0.76 (0.66 - 0.86)	0.20
AG	0.75 (0.66 - 0.85)	0.19
PWP Total	0.77 (0.68 - 0.88)	0.20
PWP Gap	0.78 (0.68 - 0.88)	0.20
WLW	0.76 (0.66 - 0.86)	0.20
LWYY	0.75 (0.65 - 0.86)	0.21



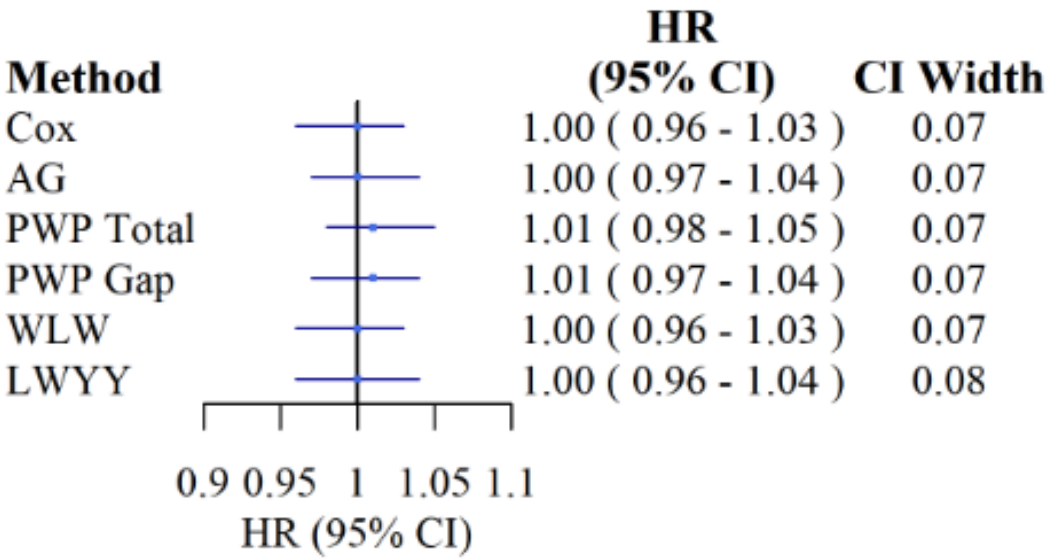
d) COMPASS Composite 2 (5.9*)

	HR (95% CI)	CI Width
	0.80 (0.71 - 0.90)	0.19
	0.82 (0.74 - 0.92)	0.18
	0.86 (0.77 - 0.96)	0.19
	0.85 (0.77 - 0.95)	0.18
	0.80 (0.71 - 0.90)	0.19
	0.82 (0.72 - 0.94)	0.22

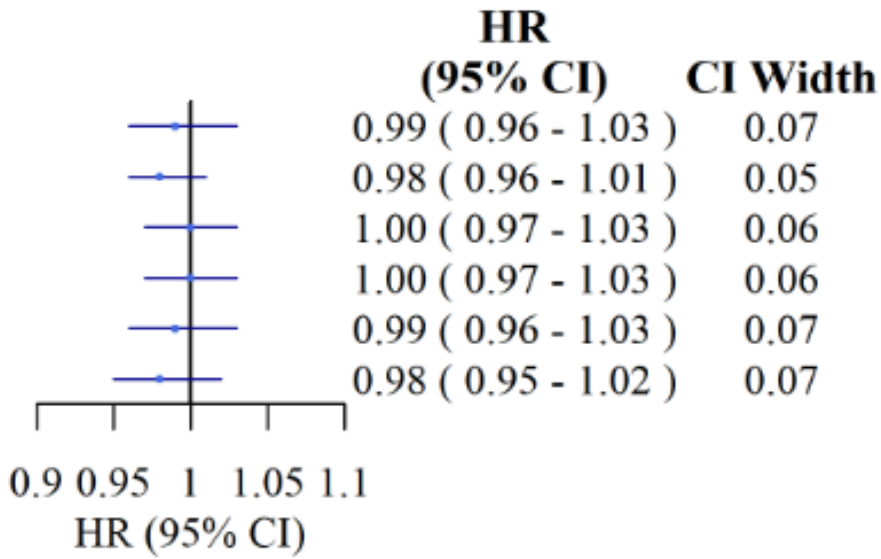


Treatment Effects for TRANSCEND

a) TRANSCEND Composite 1 (2.0*)



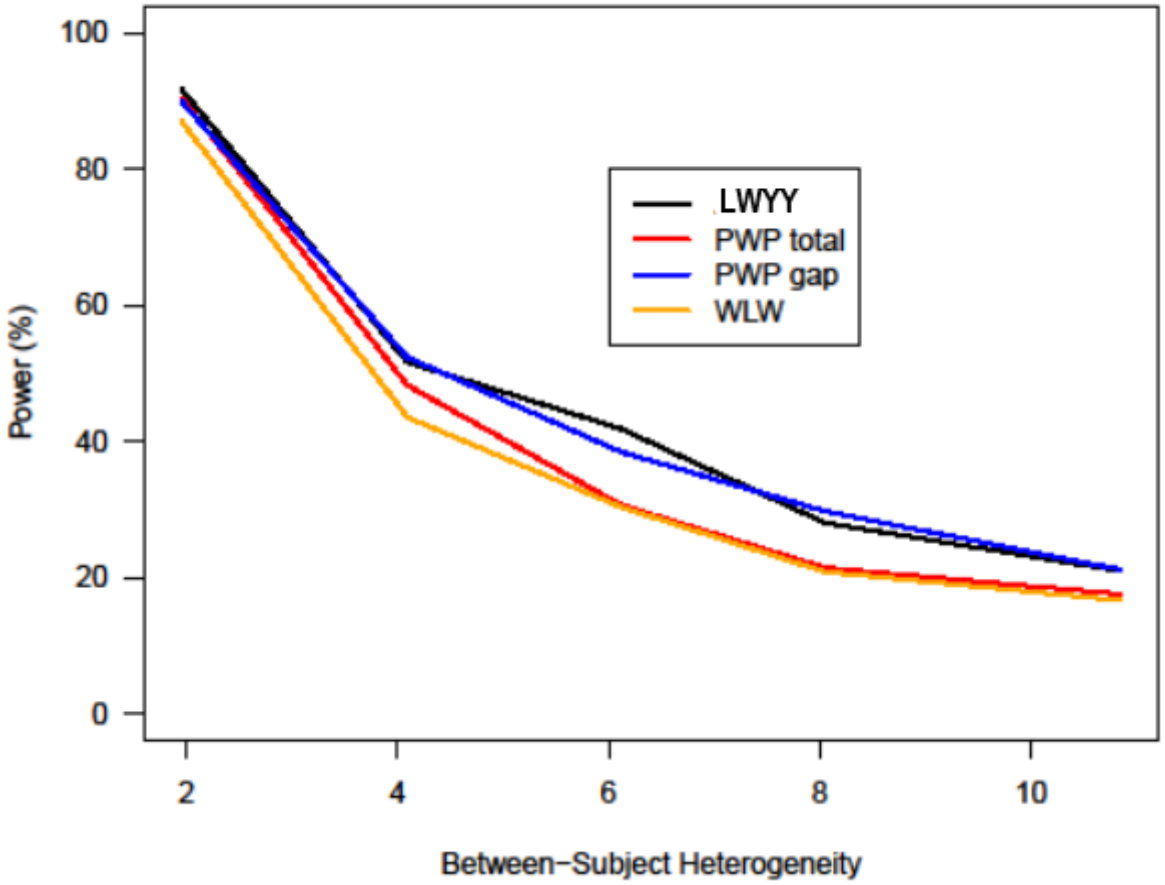
b) TRANSCEND Composite 2 (3.1*)



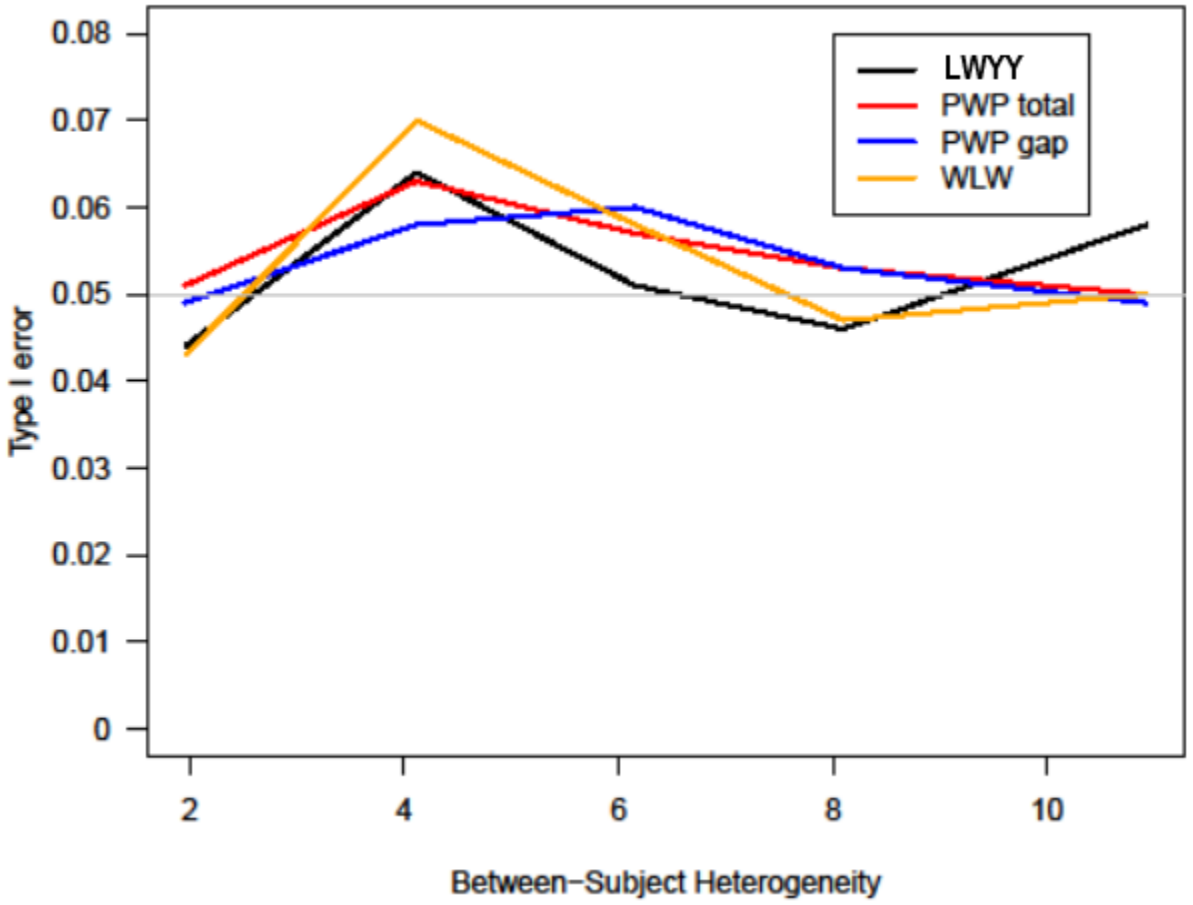
Simulation Parameters

Simulation Parameter	Values
Total N (N)	2000, 5000, 10000
Control Incidence	0.1, 0.15, 0.2
Censoring probability (C)	0.8, 0.9
Heterogeneity (κ)	1, 3, 5, 7, 10, 12
Follow-up years (T)	2, 4
Hazard ratio (HR)	1.11, 1.16, 1.22

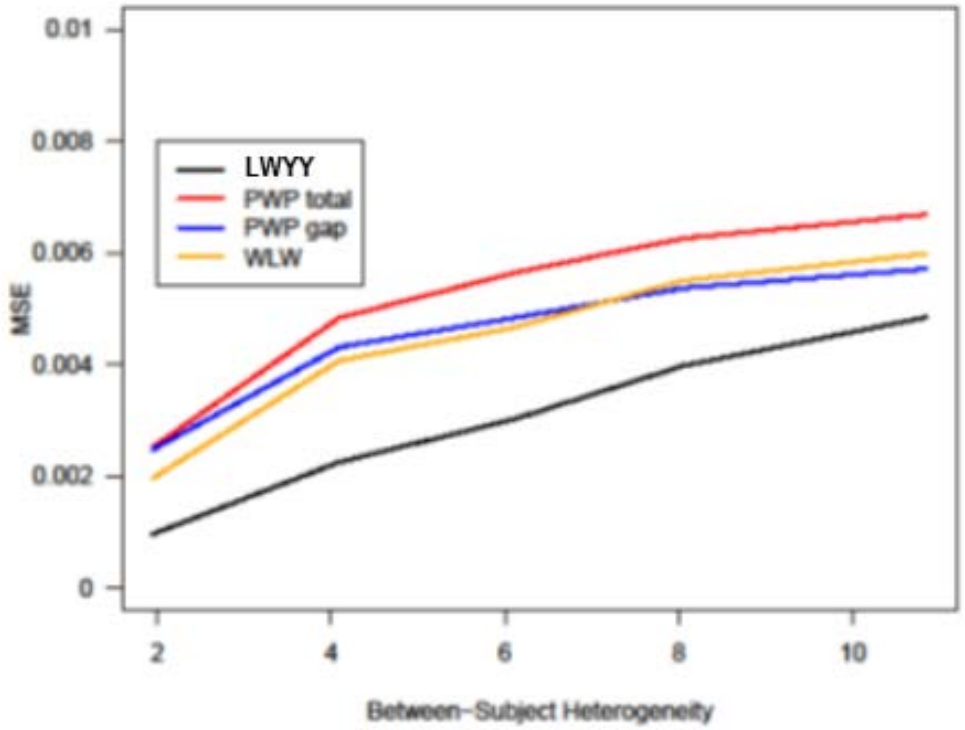
Simulation: Power



Simulation: Type I Error



Simulation: Mean Square Error (MSE)



Conclusion:

- Between-subject heterogeneity, measuring the occurrence of multiple events per participant, aids in estimating efficiency in analyses involving repeated events.
- Heterogeneity tends to be large in cases of a substantial increase in low incidence, with a higher likelihood of having a small number of participants experiencing multiple events.
- Composite outcomes in cardiovascular trials typically exhibit small heterogeneity.
- For small heterogeneity, the treatment effect and confidence interval width remain nearly consistent between time-to-first and total event analysis.
- In cases of substantial heterogeneity, the differences in treatment effects across all methods are minimal and statistically insignificant.
- LWYY has a wider confidence interval with a reasonable performance on simulation.

References:

Cook R, Lawless J. The Statistical Analysis of Recurrent Events. Statistics for Biology and Health. 2007

Claggett B, Pocock S, Wei LJ, Pfeffer MA, McMurray JJV, Solomon SD. Comparison of Time-to-First Event and Recurrent-Event Methods in Randomized Clinical Trials. Circulation. 2018 Aug 7;138(6):570-577.

Gerstein HC, Bosch J, Dagenais GR, Diaz R, Jung H, et al. Basal insulin and cardiovascular and other outcomes in dysglycemia. N Engl J Med. 2012;367:319-28.

Eikelboom JW, Connolly SJ, Bosch J, Dagenais GR, Hart RG, Shestakovska O, et al. Rivaroxaban with or without Aspirin in Stable Cardiovascular Disease. N Engl J Med. 2017;377:1319-30.

Yusuf S, Teo K, Anderson C, Pogue J, Dyal L, Copland I, et al. Effects of the angiotensin-receptor blocker telmisartan on cardiovascular events in high-risk patients intolerant to angiotensin-converting enzyme inhibitors: a randomised controlled trial. Lancet. 2008;372:1174-83.

Andersen PK, Gill RD. Cox's Regression Model for Counting Processes: A Large Sample Study. The Annals of Statistics. 1982;10:1100-20, 21.

Prentice RL, Williams BJ, Peterson AV. On the Regression Analysis of Multivariate Failure Time Data. Biometrika. 1981;68:373-9.

Wei LJ, Lin DY, Weissfeld L. Regression Analysis of Multivariate Incomplete Failure Time Data by Modeling Marginal Distributions. Journal of the American Statistical Association. 1989;84:1065-73.

Lin DY, Wei LJ, Yang I, Ying Z. Semiparametric regression for the mean and rate functions of recurrent events. J R Stat Soc Series B Stat Methodol. 2000;62:711–730.

UKE HAMBURG
Institute of Medical Biometry and Epidemiology

The challenge of time-to-event analysis for multiple events – Comparison of different approaches

Ann-Kathrin Ozga, Ph.D.

iMBE
University Medical Center Hamburg-Eppendorf

Disclosure

No relevant disclosures.

Introduction

Mode of Action and Effects of Standardized Collaborative Disease Management on Mortality and Morbidity in Patients With Systolic Heart Failure

The Interdisciplinary Network for Heart Failure (INH) Study

Christiane E. Angermann, MD*; Stefan Störk, MD, PhD*; Götz Gelbrich, PhD; Hermann Faller, MD; Roland Jahns, MD; Stefan Frantz, MD; Markus Loeffler, MD; Georg Ertl, MD; on Behalf of the Competence Network Heart Failure

Background—Trials investigating efficacy of disease management programs (DMP) in heart failure reported contradictory results. Features rendering specific interventions successful are often ill defined. We evaluated the mode of action and effects of a nurse-coordinated DMP (HeartNetCare-HF, HNC).

Methods and Results—Patients hospitalized for systolic heart failure were randomly assigned to HNC or usual care (UC). Besides telephone-based monitoring and education, HNC addressed individual problems raised by patients, pursued networking of health care providers and provided training for caregivers. End points were time to death or rehospitalization (combined primary), heart failure symptoms, and quality of life (SF-36). Of 1007 consecutive patients, 715 were randomly assigned (HNC: n=352; UC: n=363; age, 69±12 years; 29% female; 40% New York Heart Association class III-IV). Within 180 days, 130 HNC and 137 UC patients reached the primary end point (hazard ratio, 1.02; 95% confidence interval, 0.81–1.30; $P=0.89$), since more HNC patients were readmitted. Overall, 32 HNC and 52 UC patients died (1 UC patient and 4 HNC patients after dropout); thus, uncensored hazard ratio was 0.62 (0.40–0.96; $P=0.03$). HNC patients improved more regarding New York Heart Association class ($P=0.05$), physical functioning ($P=0.03$), and physical health component ($P=0.03$). Except for HNC, health care utilization was comparable between groups. However, HNC patients requested counseling for noncardiac problems even more frequently than for cardiovascular or heart-failure–related issues.

Conclusions—The primary end point of this study was neutral. However, mortality risk and surrogates of well-being improved significantly. Quantitative assessment of patient requirements suggested that besides (tele)monitoring individualized care considering also noncardiac problems should be integrated in efforts to achieve more sustainable improvement in heart failure outcomes.

Clinical Trial Registration—URL: http://www. controlled-trials.com. Unique identifier: ISRCTN23325295.

(***Circ Heart Fail.*** **2012;5:25-35.**)

Introduction

Mode of Action and Effects of Standardized Collaborative Disease Management on Mortality and Morbidity in Patients With Systolic Heart Failure

The Interdisciplinary Network for Heart Failure (INH) Study

Christiane E. Angermann, MD*; Stefan Störk, MD, PhD*; Götz Gelbrich, PhD; Hermann Faller, MD; Roland Jahns, MD; Stefan Frantz, MD; Markus Loeffler, MD; Georg Ertl, MD; on Behalf of the Competence Network Heart Failure

Background—Trials investigating efficacy of disease management programs (DMP) in heart failure reported contradictory results. Features rendering specific interventions successful are often ill defined. We evaluated the mode of action and effects of a nurse-coordinated DMP (HeartNetCare-HF, HNC).

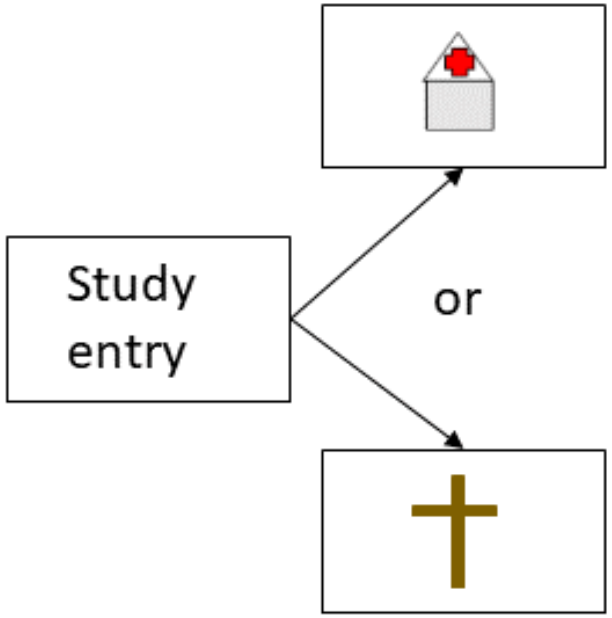
Methods and Results—Patients hospitalized for systolic heart failure were randomly assigned to HNC or usual care (UC). Besides telephone-based monitoring and education, HNC addressed individual problems raised by patients, pursued networking of health care providers and provided training for caregivers. End points were time to death or rehospitalization (combined primary), heart failure symptoms, and quality of life (SF-36). Of 1007 consecutive patients, 715 were randomly assigned (HNC: n=352; UC: n=363; age, 69±12 years; 29% female; 40% New York Heart Association class III-IV). Within 180 days, 130 HNC and 137 UC patients reached the primary end point (hazard ratio, 1.02; 95% confidence interval, 0.81–1.30; P=0.89), since more HNC patients were readmitted. Overall, 32 HNC and 52 UC patients died (1 UC patient and 4 HNC patients after dropout); thus, uncensored hazard ratio was 0.62 (0.40–0.96; P=0.03). HNC patients improved more regarding New York Heart Association class (P=0.05), physical functioning (P=0.03), and physical health component (P=0.03). Except for HNC, health care utilization was comparable between groups. However, HNC patients requested counseling for noncardiac problems even more frequently than for cardiovascular or heart-failure–related issues.

Conclusions—The primary end point of this study was neutral. However, mortality risk and surrogates of well-being improved significantly. Quantitative assessment of patient requirements suggested that besides (tele)monitoring individualized care considering also noncardiac problems should be integrated in efforts to achieve more sustainable improvement in heart failure outcomes.

Clinical Trial Registration—URL: http://www. controlled-trials.com. Unique identifier: ISRCTN23325295.

(***Circ Heart Fail.*** **2012;5:25-35.**)

Introduction



Rectangles: (event) states

: hospitalization

: death

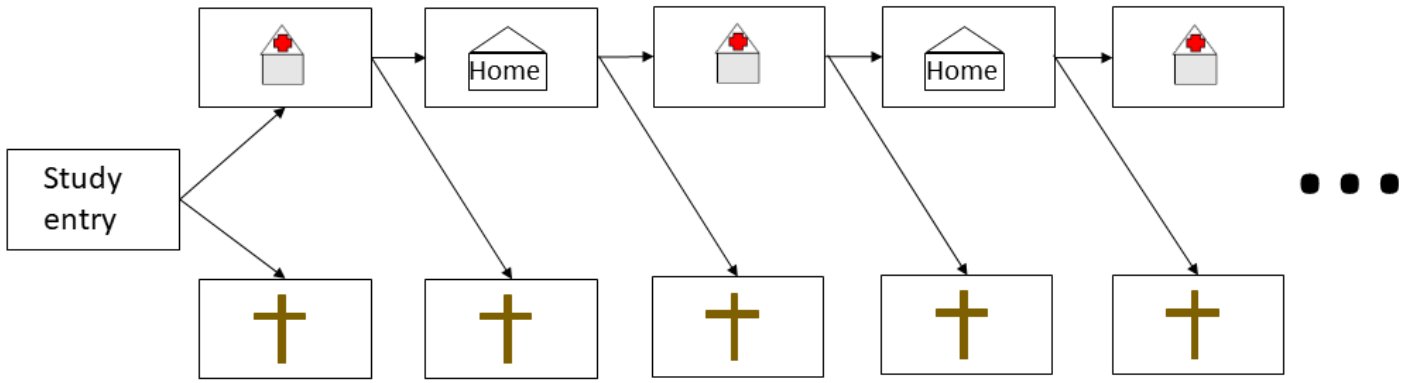
Arrows: transition (hazards)

Often used: Time-to-first event, i.e. Cox proportional hazards (PH) model (assuming no truncation and right-censored data)

Drawbacks:

- Neglecting different clinical relevance of events
- Do not include all events per individual → information loss
- Does not consider competing events?

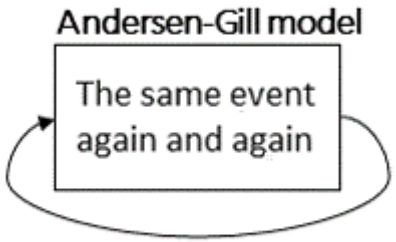
Introduction

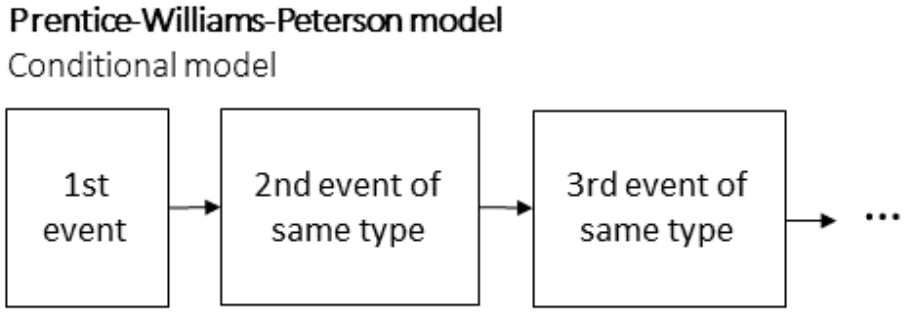


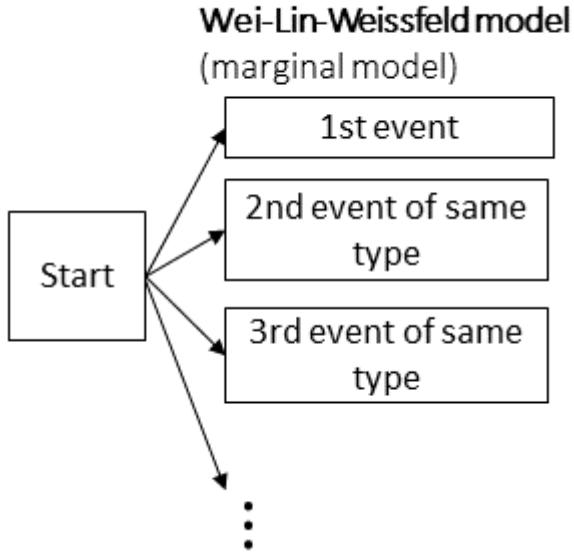
Questions:

- One effect measure for whole disease process?
- One effect measure for each transition?
- One effect measure for each event type?
- Composite endpoint and/or competing event?

Methods
Semiparametric (Cox PH based)







Methods
Semiparametric

Aim: One treatment effect for whole disease process

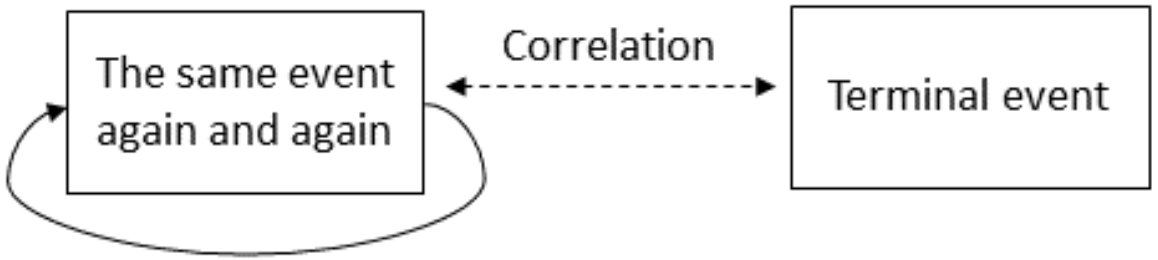
Fatal event death as competing event: Death is censored at that time point (explanation for this for randomized controlled trials see Allison 2018)

Fatal event death within composite endpoint:

- Recommended to used PWP model but AG model also applicable in some scenarios (see Ozga et al. 2018)
- WLW model **not** recommended
- To additionally incorporate clinical relevance of events via weights:
  - Weighted all-cause hazard ratio (Rauch et al. 2018) with extension to composite endpoint with recurrent events (Ozga et al. 2022); also described with parametric estimation
  - Weighted hazard ratio (Lachin et al. 2015, Wei et al. 1984) with extension to composite endpoint with recurrent events (Ozga et al. 2022)

Methods
(Semi)parametric

Joint frailty model



- Cause specific effect estimates, i.e. one for recurrent event and one for terminal/fatal event
- Cox type estimation (or parametric)
- Additional parameter to specify correlation between recurrent and terminal event

Methods
Semiparametric – Further remarks

- Hazard and the hazard ratio from the Cox PH model have a causal interpretation with a view from treatment onset. E.g. the hazard ratio compares the instantaneous risk of a patient getting the treatment with a patient in the control group from the view of time-point 0

- Do need Markov assumption (= means that the transition hazard does not depend on previous states, number of previous states nor on previous event-times)

- Might include number of previous events as covariate (for testing Markov assumption; see Soutinho et al. (2022))

- Violation of the Markov assumption can lead to large bias as shown by a simulation study by Schmeller et al. (submitted 2024) and means loss of causal interpretation for the treatment effect

Methods
Semiparametric

→ To gain effect estimate(s)

→ Other covariates can be included

→ Risk of model miss-specification

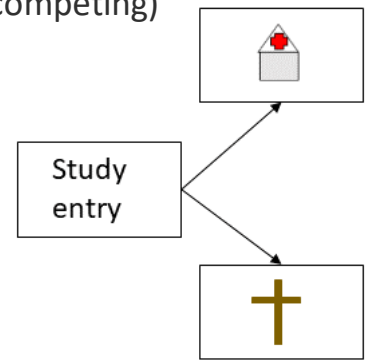
→ Need other methods:

- in case Markov assumption is not fulfilled

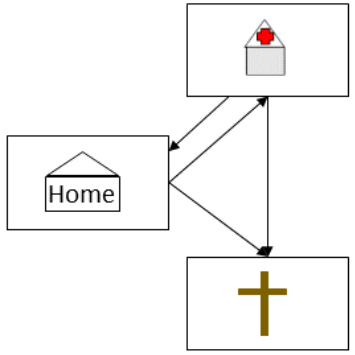
- to described event probabilities/incidences/ more quantities to understand disease process

Methods
Nonparametric – Multistate approach

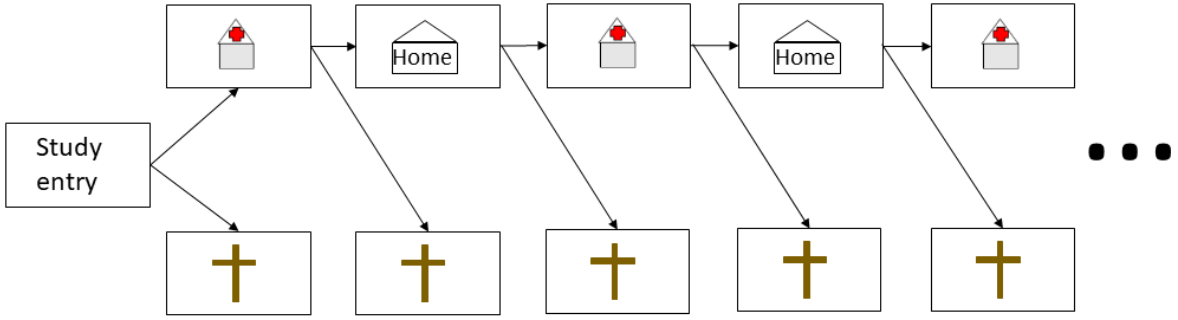
Time-to-first event (composite or competing)



Illness-death model with recovery



Extended multistate model



Aim: Estimation of cumulative hazards, incidences or transition probabilities

Nelson Aalen estimator: Estimation of cumulative (transition) hazards (= total accumulated risk of experiencing the event of interest that has been observed over time)

Aalen-Johansen estimator: Estimation of cumulative incidences or transition probabilities (transition hazards and probabilities are (instantaneous) conditional probabilities and can be interpreted in a causal way from the view of treatment onset)

Methods
Nonparametric – Multistate approach

Time-to-first event:

- Composite endpoint: see example study
- Competing risks**:**
  - e.g. time to first hospitalization of main interest, competing event is death.
  - distinguishes between the type of event
  - gives a deeper understanding about the risk of the first event

Illness death model with recovery:

- Multiple states a patient can be in but the number of states is limited
- Markov assumption (for non-Markov see Niessl et al. 2021)

Extended multistate model:

- One has to specify the number of transitions ahead.
- Markov assumption is less restrictive (different states for the different recurring events, it is only assumed that the transition probability does not depend on how long one has been in the state before)

Methods
Nonparametric

→ To describe cumulative hazards, incidences or transition probabilities

→ To gain a deeper understanding of the disease process

→ Do not gain treatment effect

→ Markov assumption (but partly less restrictive or alternatives available)

Remark
Connection between semiparametric models and multistate approach

AG model: semiparametric treatment estimation within illness death model (with recovery)

PWP model: semiparametric treatment estimation within the extended multistate approach

Further methods
Gosh and Lin (2002)

- Does not require the Markov assumption
- Is a rate model (marginal model)
- Estimates the cause specific rate in a Cox-type form
- Uses inverse probability of censoring or survival weighting to account for a terminal event
- Assumes that there are no gap times between the recurrent events, i.e. after experiencing a recurrent event one is immediately afterwards under risk for a further event

Further methods
Marginal features

- **State occupation probability (via Aalen-Johansen estimator):** estimation of the average length of stay in hospital up to a specific time

- **But**: this estimate and the average length of stay of two groups (e.g. treatment groups) should not be considered alone because a reduction of the mean number can result either because of a reduced event intensity / a longer stay in hospital or an increased mortality.

- Quantify the relation between hospitalization and death with a non-parametric summary measure to compare two treatment groups (see Wei et al. 2022):
  - ratio of the expected number of hospitalizations per group divided by the overall survival classifies the number of hospitalizations in relation to the survival
  - interpret the estimates in relation to other estimates changing populations and time periods

Example
Network for Heart Failure (INH) study

- Follow-up: 60 months
- Patients: 1022 (513 in usual care (UC), 509 in treatment (HNC) group)
- Deaths: 663
- Rehospitalizations: 3016 rehospitalizations (with a maximum of 27 events per patient)
- Median follow-up: 2596 days
- Random censoring (=censoring time is stochastically independent of the time to event process)

Example
Network for Heart Failure (INH) study

	AG model - hospitalization alone (HR 95% CI)	PWP model - hospitalization alone (HR 95% CI)	AG model - composite endpoint (HR 95% CI)	PWP model - (weigthed*,^) composite endpoint (HR 95% CI)
Treatment vs. UC	0.86 (0.74, 1.00)	0.88 (0.80, 0.97)	0.86 (0.75, 0.98)	0.89 (0.82, 0.97)
				0.89 (0.82, 0.97)*
				0.89 (0.76, 1.02)^
Adjusted model:				
Treatment vs. UC	0.88 (0.79, 0.98)	0.91 (0.84, 0.99)	0.88 (0.79, 0.97)	0.91 (0.85, 0.98)
Previous events (AG) or entry time (PWP)	1.24 (1.21, 1.27)	1.00 (1.00, 1.00)	1.22 (1.20, 1.25)	1.00 (1.00, 1.00)

→ Seems to be non-Markov in AG model

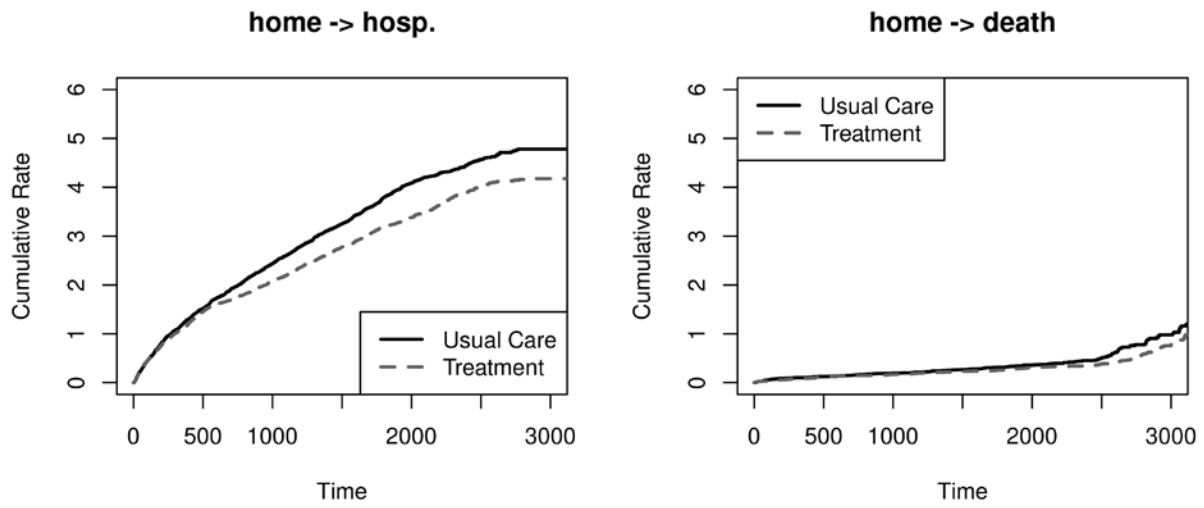
HR= hazard ratio, CI= confidence interval
*weighted Rauch et al.; ^ weighted Wei-Lachin (hospitalization counts as half of death); death: HR (95%CI): 0.90 (0.75, 1.08)

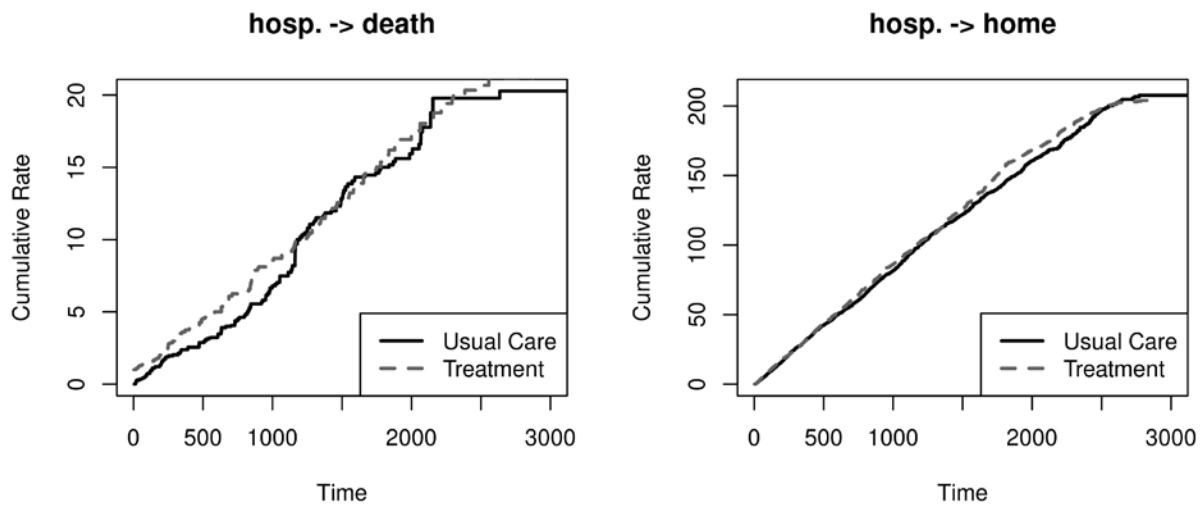
Example
Network for Heart Failure (INH) study

	Gosh and Lin model (RR 95% CI)	**Joint frailty model (HR 95% CI)**
Treatment vs. UC (recurrent event)	0.86 (0.74, 1.01)	0.86 (0.80, 0.92)
Treatment vs. UC (terminal event)	0.82 (0.70, 0.96)	0.86 (0.74, 1.00)

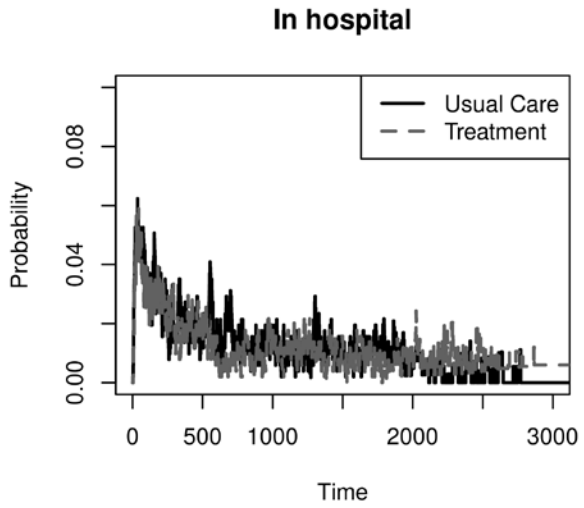
HR= hazard ratio, RR=risk ratio, CI= confidence interval

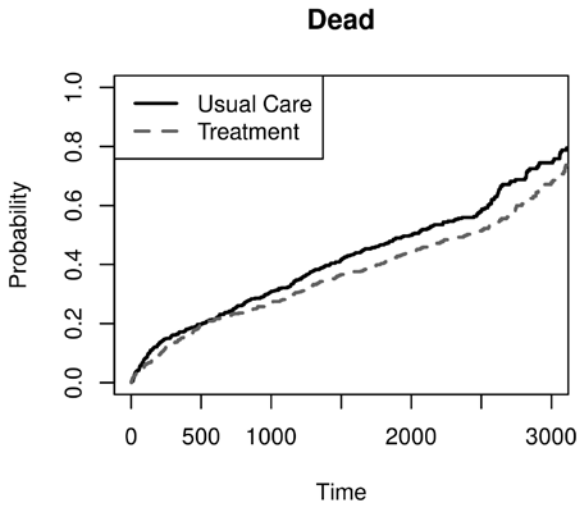
Example – Cumulative hazard (Nelson-Aalen estimate)

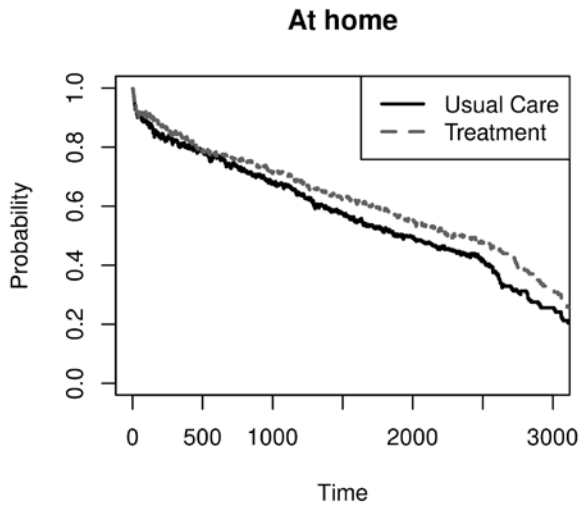




Example - State occupation probabilities in an Illness death model with recovery

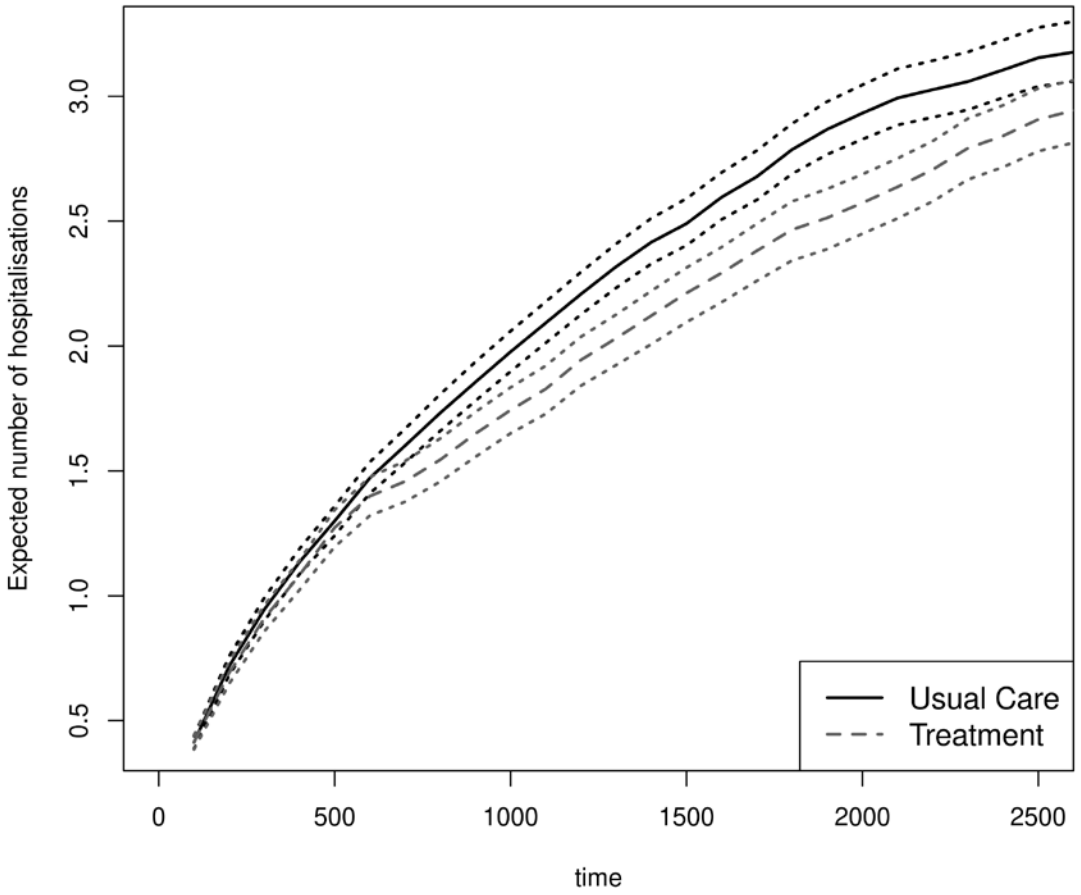






Example

Expected number of hospitalizations with bootstrapped 95% confidence intervals

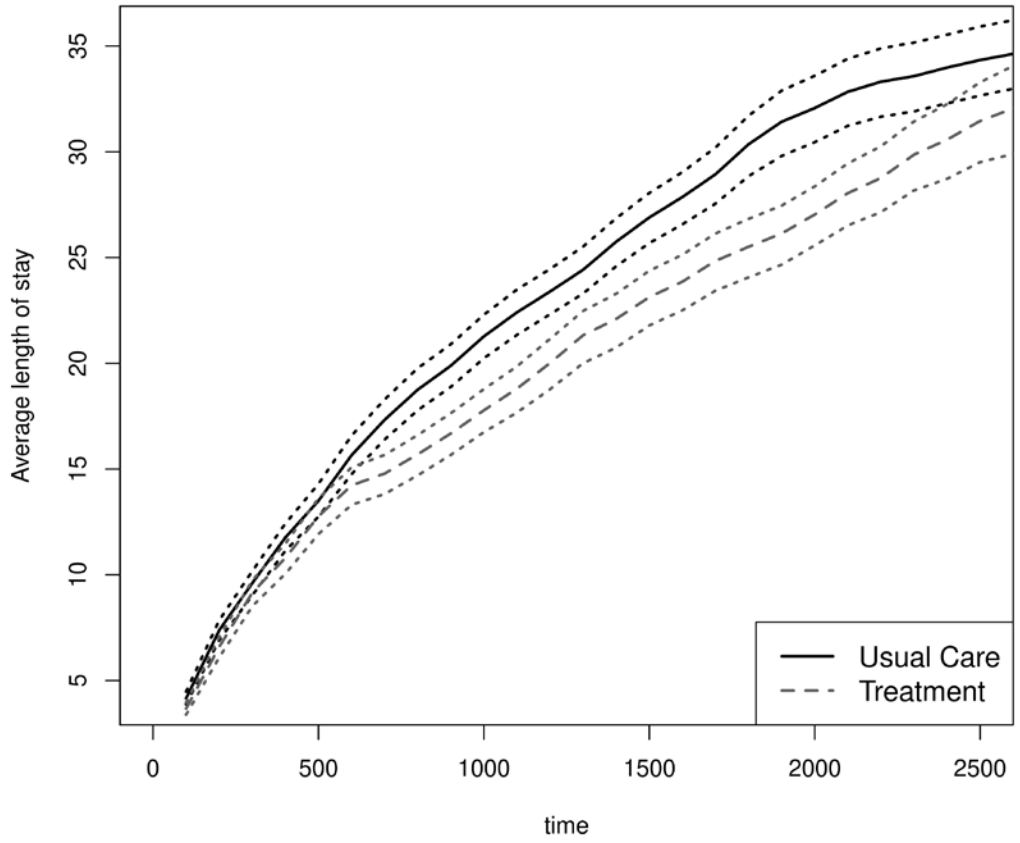


Ratio of the expected number of hospitalizations per group divided by the overall survival (RR):

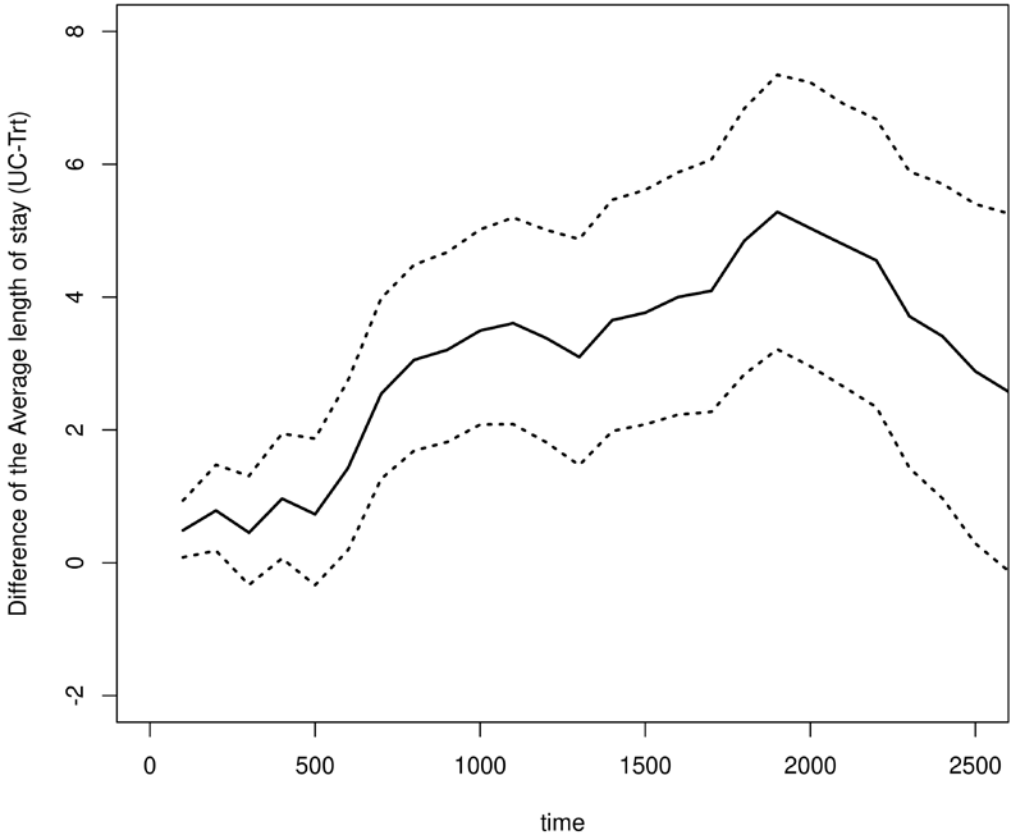
Time [days]	RR
500	1.166
1000	0.992
1500	1.006
2000	0.999

Example

Average length of stay in the hospital with bootstrapped 95% confidence intervals.



Difference of the average length of stay in the hospital from the UC group minus the treatment group (Trt) with bootstrapped 95% confidence intervals.



Summary

Methods for multiple time-to-event analysis are available, like:

- Semiparametric (weighted) AG or (weighted) PWP model
- Nonparametric multistate approaches

and **should best be reported both to gain a complete picture of the disease process.**

Alternative approaches:

- Rate/marginal models (if non Markov); e.g. Gosh & Lin, negative binomial model, Lin-Wei-Yang-Ying model
- Joint frailty models (estimating different treatment effects for recurrent and fatal event)
- Parametric approaches (if no PH)

Further research

The German Research Foundation supports the project entitled:

Methods for planning and analyzing studies with multiple time to event endpoints, taking into account recurring and competing events

Topics:

- Evaluation and development of analysis methods for multiple time-to-event endpoints. Focus on more complex data structures like clustered data, more than two event types, ….
- Planning a study with multiple time-to-event endpoints (e.g. methods for sample size calculation)

References

Allison P. For Causal Analysis of Competing Risks, Don't Use Fine & Gray's Subdistribution Method. 2018. https://statisticalhorizons.com/for-causal-analysis-of-competing-risks/ (accessed April 2024)

Andersen PK, Gill RD. Cox's Regression Model for Counting Processes: A Large Sample Study. The Annals of Statistics. 1982; 10(4):1100- 1120.

Andersen PK, Angst J, Ravn H. Modeling marginal features in studies of recurrent events in the presence of a terminal event. Lifetime Data Anal. 2019 Oct;25(4):681-695. doi: 10.1007/s10985-019-09462-4.

Angermann CE, Stoerk S, Gelbrich G, Faller H, Jahns R, Frantz S, Loeffler M, Ertl G; Competence Network Heart Failure. Mode of action and effects of standardized collaborative disease management on mortality and morbidity in patients with systolic heart failure: the Interdisciplinary Network for Heart Failure (INH) study. Circ Heart Fail. 2012;5(1):25-35. doi: 10.1161/CIRCHEARTFAILURE.111.962969.

Beyersmann J, Schumacher M, Allignol A. Competing Risks and Multistate Models with R. Heidelberg: Springer. 2012.

Erdmann A, Beyersmann J, Bluhmki E. Comparison of nonparametric estimators of the expected number of recurrent events. Pharm Stat. 2023; doi: 10.1002/pst.2356.

Ghosh D, Lin DY. Marginal Regression Models For Recurrent and Terminal Events. Statistica Sinica. 2002; 12(3): 663–688.

Lachin JM, Bebu I. Application of the wei lachin multivariate one-directional test to multiple event-time outcomes. Clin Trials. 2015; 12(6):627–33.

Lin D Y, Wei LJ, Yang I, Ying, Z. Semiparametric regression for the mean and rate functions of recurrent events. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2000; 62(4), 711–730.

Niessl A, Allignol A, Beyersmann J, Mueller C. Statistical inference for state occupation and transition probabilities in non-markov multi-state models subject to both randomleft-truncation and right-censoring. Econometrics and Statistics. 2023; 25: 110-124. Doi:10.1016/j.ecosta.2021.09.008.

Ozga AK, Kieser M, Rauch G. A systematic comparison of recurrent event models for application to composite endpoints,. BMC Med Res Methodol. 2018; 18(2):1–12.

Ozga AK, Rauch G. Weighted composite time to event endpoints with recurrent events: comparison of three analytical approaches. BMC Med Res Methodol. 2022;22(1):38; 1-14. doi: 10.1186/s12874-022-01511-1.

Prentice R L, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. Biometrika. 1981;68(2): 373-379

Rauch G, Kunzmann K, Kieser M, Wegscheider K, Koenig J, Eulenburg C. A weighted combined effect measure for the analysis of a composite time-to-first-event endpoint with components of different clinical relevance. Stat Med. 2018; 37(5):749–67.

Rogers J K, Yaroshinsky A, Pocock SJ, Stokar D, Pogoda J. Analysis of recurrent events with an associated informative dropout time: Application of the joint frailty model. Statistics in Medicine. 2016; 35(13): 2195- 2205.

Schmeller S, Erdmann A, Beyersmann J. Angermann CE, Ozga AK. The challenge of time-to-event analysis for multiple events: a guidance. (Submitted)

Soutinho, G., Meira-Machado, L. Methods for checking the Markov condition in multi-state survival data. Comput Stat. 2022: 37: 751–780. doi:10.1007/s00180-021-01139-7

Wei J, Mütze T, Jahn-Eimermacher A, Roger J. Properties of Two While-Alive Estimands for Recurrent Events and Their Potential Estimators. Statistics in Biopharmaceutical Research.2023; 15(2): 257–267. doi:10.1080/19466315.2021.1994457

Wei L, Lachin J. Two-sample asymptotically distribution-free tests for incomplete multivariate observations. J Am Stat Assoc. 1984; 79(387):653–61.

Wei LJ, Lin DY, Weissfeld L. Regression Analysis of Multivariate Incomplete Failure Time Data by Modeling Marginal Distributions. Journal of the American Statistical Association. 1989; 84(804): 1065-1073.

Acknowledgment:

German research foundation (Deutsche Forschungsgemeinschaft (DFG))
Sandra Schmeller (Ph.D. student; figures & work in progress on nonparametric methods)
Sinan Necdet Cevirme (Master student; data simulation of complex time-to-event data)
Alidan Duoerkongjiang (Ph.D. student; working on DFG project)
Prof. Christiane Angermann & Susanne Sehner (example data)

Contact:

a.ozga@uke.de

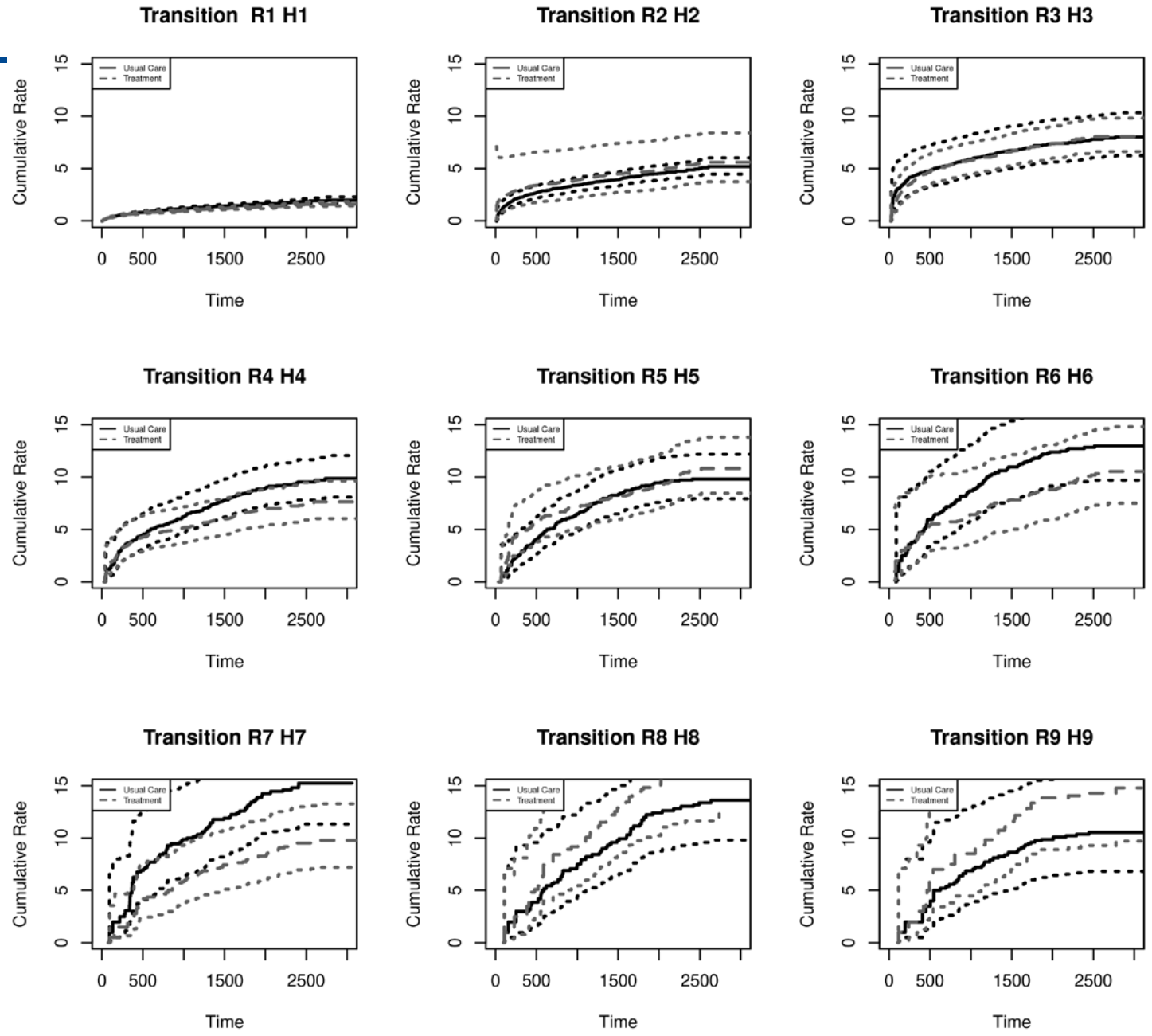
Appendix: Methods
Semiparametric

	Andersen-Gill (AG) model (1982)	**Prentice-Williams-Peterson (PWP) model (1981)**	**Wei-Lin-Weissfeld (WLW) model (1989)**
Model	Cox PH model	stratified Cox PH model	stratified Cox PH model
Time model	time since study start	time since study start or time since last event	time since study start
Event (in)dependence	all events treated as independent	*k+1st* event can only occur if *kth* has occurred; ordered events	no assumptions
Risk set	individual remains in risk set after event occurrence	only individuals with *k* -1 events considered at risk for the *k*th event	all individuals at all event times at risk
Effect measure	mixed effect*	mixed effect	mixed effect

*i.e. a combination of direct and indirect effects (effect by the treatment and the effect due to previous events)

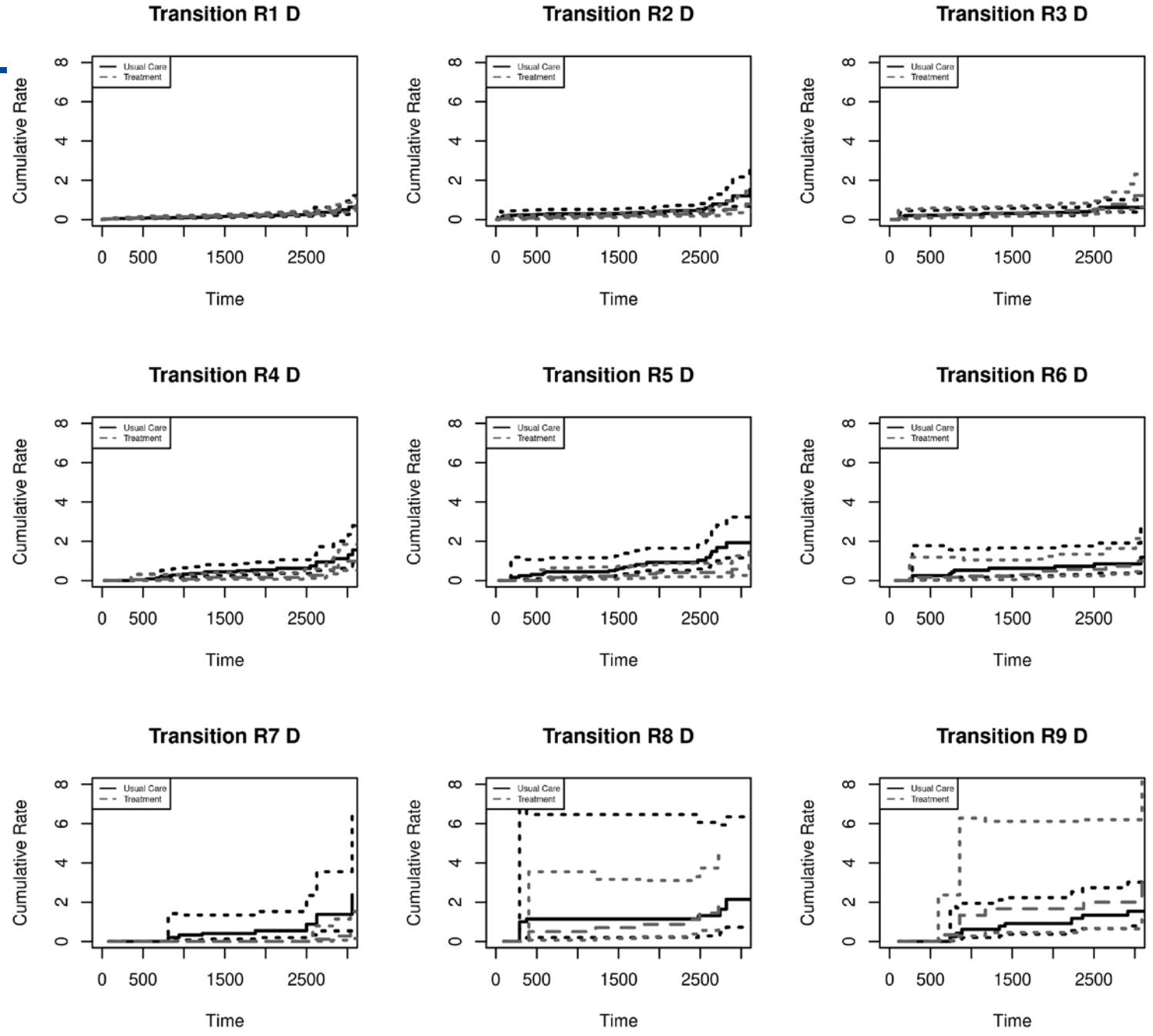
Appendix: Example

Nelson-Aalen estimates in the progressive multistate model. Transitions: At home to Hospitalization



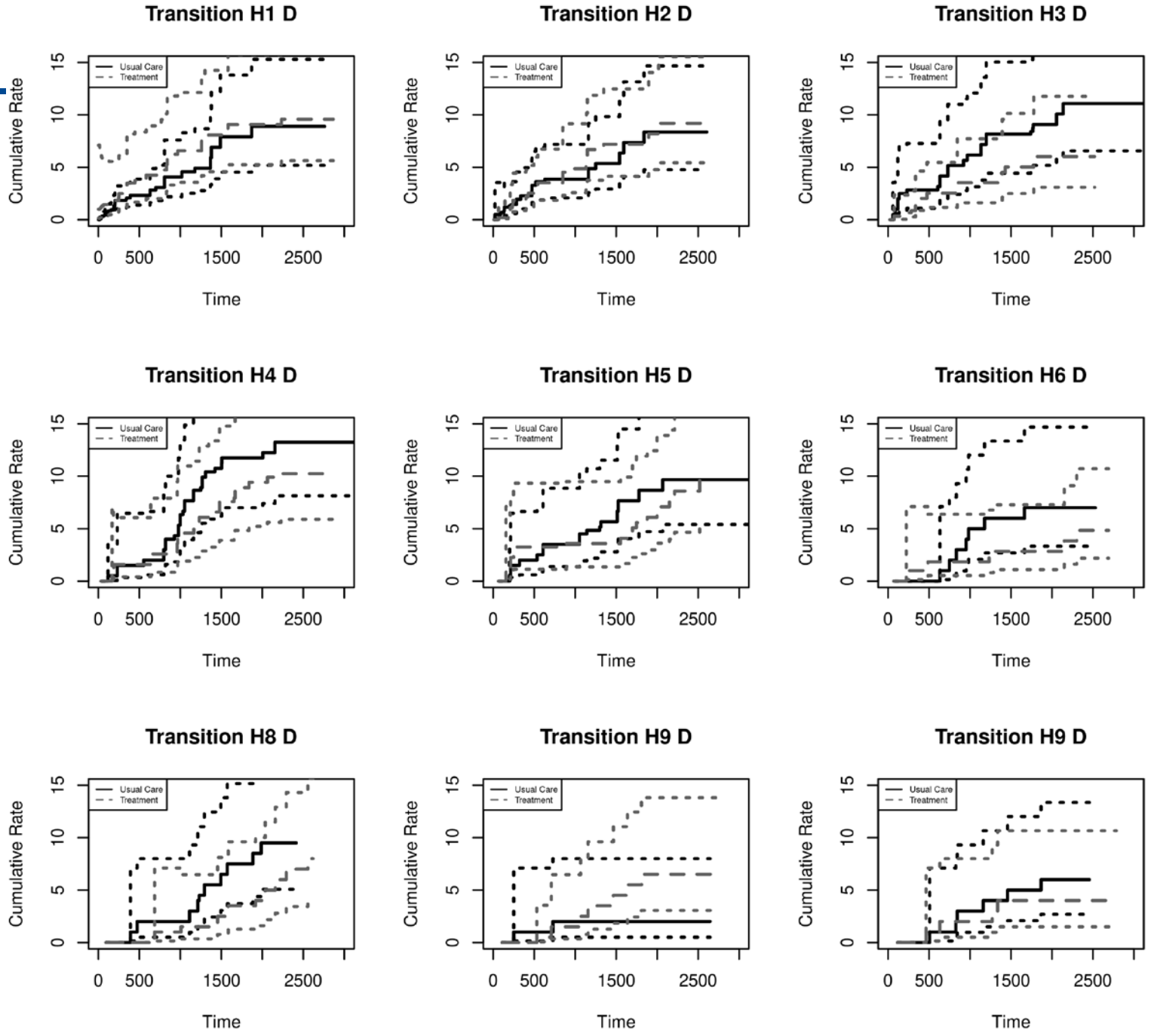
Appendix: Example

Nelson-Aalen estimates in the progressive multistate model. Transitions: At home to Death



Appendix: Example

Nelson-Aalen estimates in the progressive multistate model. Transitions: Hospital to Death



Appendix: Example

Cox Model									
Endpoint:	**overall survival**								
	HR	p	95% CI						
trt. vs. usual care	0.820	0.011	[0.70, 0.96]						
Endpoint:	**hospital free survival**			**Hospitalization**			**Death before hosp.**		
	HR	p	95% CI	HR	p	95% CI	HR	p	95% CI
trt. vs. usual care	0.904	0.133	[0.79, 1.03]	0.872	0.060	[0.76, 1.01]	1.125	0.516	[0.79, 1.60]

While-alive estimands

Summarizing patient experience under differential length of exposure

Lu Mao

lmao@biostat.wisc.edu

Department of Biostatistics & Medical Informatics

University of Wisconsin-Madison

Motivating example

A cardiovascular trial (HF-ACTION) (O'Connor et al. 2009)

- **Subpopulation:** 741 heart failure patients
  - Median and max follow-up 2.5 and 3.9 years, respectively
- **Treatment arms**
  - Exercise training + usual care $(n_1 = 364)$
  - Usual care alone $(n_0 = 377)$
- **Endpoints:** Death and repeated hospitalizations

	Exercise training	**Usual care**
Death rate	13.5%	19.9%
Avg # hospitalization (SD)	1.8 (2.1)	2.0 (2.1)

A fundamental question

How do we measure treatment effect on recurrent hospitalizations when patients survive different lengths in different arms?

Longer survivors tend to experience more events…

Mathematical notation ($a = 1$: Treatment; $a = 0$: Control)

- $D^{(a)}$: Survival time
- $N_D^{(a)}(t) = I\left(D^{(a)} \leq t\right)$: Counting process for death
- $N^{(a)}(t)$: Counting process for recurrent events (e.g., hospitalization)
  - No event after death
  - $\mathrm{d}N^{(a)}(t) = 0$ for $t > D^{(a)}$

Standard methods

Two broad-based approaches…

- **Conditional event rate**
  - Standard analysis treating death as censoring, e.g., proportional rates model (LWYY) (Lin et al. 2000)
  - Estimand: $E\left\{\mathrm{d}N^{(a)}(t) \mid D^{(a)} \geq t\right\}$
  - Lacks causal interpretation (condition on post-treatment status) (ICH 2020)
- **Cumulative frequency**
  - Mean number of events in presence of death as a competing risk (Ghosh and Lin 2000; Ghosh and Lin 2002; Mao and Lin 2016)
  - Estimand: $E\left\{N^{(a)}(t)\right\}$
  - Ignores length of exposure…

A new proposal

While-alive event rate

- **Estimand**

$$\ell^{(a)}(\tau) = \frac{E\left\{N^{(a)}(\tau)\right\}}{E\left(D^{(a)} \wedge \tau\right)} = \frac{\text{Mean \# of events by } \tau}{\text{Mean survival time by } \tau}$$

  - τ: prespecified time horizon; $b \wedge c := \min(b, c)$
  - Average event rate in $[0, \tau]$ per person-time alive
  - Proposed as a clinically interpretable measure to Committee for Medicinal Products for Human Use (CHMP) of European Medicines Agency (Akacha et al. 2018; CHMP 2020)
  - Also called “exposure-weighted” event rate

Early work

A series of **follow-up papers** (Schmidli, Roger, and Akacha 2023a, 2023b; Wei et al. 2023; Fritsch et al. 2023)…

- Gamma shared-frailty models: analytic expression of $\ell^{(a)}(\tau)$
- Inference under *minimal mortality*: Poisson/negative-binomial/LWYY regressions
- Method-of-moment estimator under a *fixed censoring point*

A similar idea is (independently) considered for mortality (Uno and Horiguchi 2023)

$$E\{N_D^{(a)}(\tau)\}/E(D^{(a)} \wedge \tau)$$

Gaps

- Generalization of estimand
- General nonparametric inference procedure

General estimands - definition

While-alive *loss* rate (Mao 2023)

$$\ell^{(a)}(\tau) = \frac{E\left\{\mathcal{L}\left(\mathcal{H}^{(a)}\right)(\tau)\right\}}{E\left(D^{(a)} \wedge \tau\right)}$$

- $\mathcal{H}^{(a)}(t) = \left\{N_D^{(a)}(u), N^{(a)}(u) : 0 \leq u \leq t\right\}$: total outcomes by t
- $\mathcal{L}\left(\mathcal{H}^{(a)}\right)(t)$: user-specified *loss function* satisfying **two conditions**
  - A function only of $\mathcal{H}^{(a)}(t)$
  - $\mathcal{L}\left(\mathcal{H}^{(a)}\right)(\mathrm{d}t) \equiv 0$ for $t > D^{(a)}$ (what does this mean?)
- Average *loss* rate in $[0, \tau]$ per person-time alive

General estimands - examples

- **Original while-alive event rate** (Schmidli, Roger, and Akacha 2023a)

$$\mathcal{L}\left(\mathcal{H}^{(a)}\right)(t) = N^{(a)}(t)$$

- **Per-person-time mortality rate** (Uno and Horiguchi 2023)

$$\mathcal{L}\left(\mathcal{H}^{(a)}\right)(t) = N_D^{(a)}(t)$$

- **More generally…**

$$\mathcal{L}\left(\mathcal{H}^{(a)}\right)(t) = \int_0^t \left\{ w^D_{N^{(a)}(u-)}(u)\mathrm{d}N_D^{(a)}(u) + w_{N^{(a)}(u-)}(u)\mathrm{d}N^{(a)}(u) \right\}$$

  - $w_m^D(u), w_m(u)$: weights for incident death or nonfatal event at u if patient has experienced m nonfatal events by then

General estimands - effect size

Survival-completed (SC) cumulative loss

$$L^{(a)}(\tau) = \ell^{(a)}(\tau)\tau$$

- Better graphics: $\ell^{(a)}(\tau) \approx 0/0$ unstable for $\tau \approx 0$
- Properties: $L^{(a)}(0) = 0$, $L^{(a)}(t) \uparrow$ with t, and $L^{(a)}(t) \geq E\{\mathcal{L}(\mathcal{H}^{(a)})(t)\}$

Measuring the treatment effect…

- **Risk (loss rate) ratio** (RR): $r(\tau) = \ell^{(1)}(\tau)/\ell^{(0)}(\tau)$
  - Treatment reduces average loss rate by $100\{1 - r(\tau)\}\%$
- **Absolute risk reduction** (ARR): $d(\tau) = \ell^{(1)}(\tau) - \ell^{(0)}(\tau)$
  - Treatment reduces average loss rate by $-d(\tau)$ (per person-time alive)

Nonparametric estimation

Observed data $\left\{\mathcal{H}\left(X^{(a)}\right), X^{(a)}\right\}$

- $X^{(a)} = D^{(a)} \wedge C^{(a)}$; $C^{(a)}$: independent censoring time
- Two samples: $\left\{\mathcal{H}_i\left(X_i^{(a)}\right), X_i^{(a)}\right\} (i = 1, \ldots, n_a; a = 1, 0)$

Estimating $\ell^{(a)}(\tau) = E\{\mathcal{L}(\mathcal{H}^{(a)})(\tau)\}/E(D^{(a)} \wedge \tau)$...

- Demonstrator (easy) $= \int_0^\tau S^{(a)}(t)\mathrm{d}t$ (restricted mean survival time; RMST) (Royston and Parmar 2011)
  - $S^{(a)}(t) = \mathrm{P}(D^{(a)} > t)$: plug in Kaplan–Meier estimator
- Numerator (moderate) $= \int_0^\tau S^{(a)}(t-)E\{\mathcal{L}(\mathcal{H}^{(a)})(\mathrm{d}t) \mid D^{(a)} \geq t\}$
  - Integrator estimated by $\sum_{i=1}^{n_a} I(X_i^{(a)} \geq t)\mathcal{L}(\mathcal{H}_i^{(a)})(\mathrm{d}t) / \sum_{i=1}^{n_a} I(X_i^{(a)} \geq t)$
- Robust variance estimator by delta method

Nonparametric testing

J-sample testing $(a = 0, 1, \ldots, J-1)$

$$H_0 : \ell^{(0)}(\tau) = \cdots = \ell^{(J-1)}(\tau)$$

- χ^2_{J-1} test on the $\log \hat{r}^{(a)}(\tau) = \log \hat{\ell}^{(a)}(\tau) - \log \hat{\ell}^{(0)}(\tau)$ $(a = 1, \ldots, J-1)$

Joint test of morbidity & mortality

$$H_0 : \ell^{(0)}(\tau) = \cdots = \ell^{(J-1)}(\tau), \quad \mu^{(0)}(\tau) = \cdots = \mu^{(J-1)}(\tau)$$

- $\mu^{(a)}(\tau) = E(D^{(a)} \wedge \tau)$: τ-RMST
- $\chi^2_{2(J-1)}$ test on the $\log \hat{r}^{(a)}(\tau)$ and $\log \hat{\mu}^{(a)}(\tau) - \log \hat{\mu}^{(0)}(\tau)$ $(a = 1, \ldots, J-1)$

R-package `WA` - usage

CRAN: https://cran.r-project.org/web/packages/WA (Mao 2021)

Main function

```
LRfit(id, time, status, trt, Dweight = 0, wH = NULL, wD = NULL)
```

- `id`: vector of patient IDs
- `time`: vector of times
- `status`: vector of event types (`1`= recurrent event; `2`= death; `0`= censoring)
- `trt`: vector of (binary or multiclass) treatment groups
- `Dweight`: weight for death relative to each recurrent event
- `wH` and `wD`: user-supplied R-functions of `(m, t)` implementing $w_m(t)$ and $w_m^D(t)$ (override `Dweight`)

Summarize and plot results by `summary()` and `plot()`

R-package WA - code example (i)

First install the package if it hasn't been installed…

```
# install package from CRAN
install.packages("WA")
```

Load the package and the HF-ACTION dataset…

```
# load the package
library(WA)
# load the HF-ACTION dataset
dat <- hfaction_cpx12
dat[1:16,] # what the data look like
```

```
           id       time status trt
1  HFACT00001 0.60506502      1   0
2  HFACT00001 1.04859685      0   0
3  HFACT00002 0.06297057      1   0
4  HFACT00002 0.35865845      1   0
5  HFACT00002 0.39698836      1   0
6  HFACT00002 3.83299110      0   0
7  HFACT00007 0.29021218      1   1
8  HFACT00007 1.80424367      1   1
9  HFACT00007 2.42573580      1   1
10 HFACT00007 2.68583162      1   1
11 HFACT00007 2.91307324      2   1
12 HFACT00008 0.01916496      1   0
13 HFACT00008 0.02737851      2   0
14 HFACT00011 0.06570842      0   0
15 HFACT00019 3.66598220      1   0
16 HFACT00019 4.23271732      0   0
```

R-package WA - code example (ii)

Unweighted (while-alive) hospitalization rate

```
1  obj <- LRfit(dat$id, dat$time, dat$status, dat$trt)
2  obj
```

```
Call:
LRfit(id = dat$id, time = dat$time, status = dat$status, trt = dat$trt)
    N Rec. event Death Med. Follow-up
0 377        747    75       2.496920
1 364        644    49       2.536619
```

Summarize inferential results for restriction time $\tau = 3.5$ years...

```
1  summary(obj, tau = 3.5)
```

```
Call:
LRfit(id = dat$id, time = dat$time, status = dat$status, trt = dat$trt)

Analysis of log loss rate (LR) by tau = 3.5:
               Estimate   Std.Err Z value  Pr(>|z|)
Ref (Group 0) -0.223940  0.054308 -4.1235 3.731e-05 ***
Group 1 vs 0  -0.186019  0.084590 -2.1991   0.02787 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Test of group difference in while-alive LR
X-squared = 4.835905, df = 1, p = 0.02787301

Point and interval estimates for the LR ratio:
             LR ratio 95% lower CL 95% higher CL
Group 1 vs 0 0.8302577     0.703412     0.9799773
```

R-package WA - code example (iii)

If you want joint test (χ^2_2) with mortality …

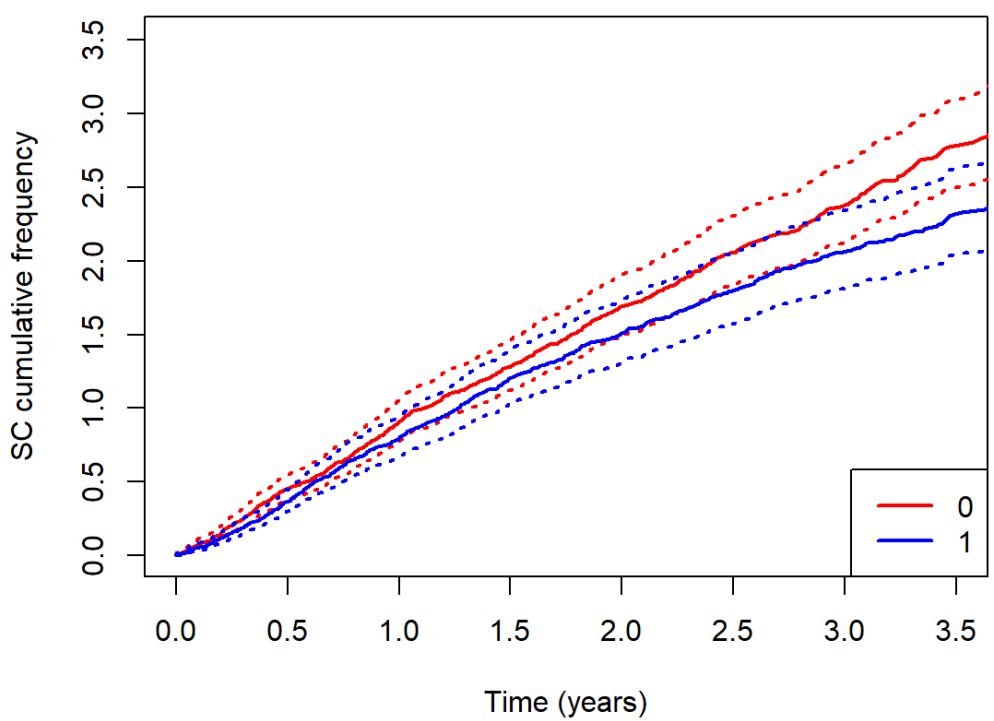
```
summary(obj, tau = 3.5, joint.test = TRUE)
```

```
Call:
LRfit(id = dat$id, time = dat$time, status = dat$status, trt = dat$trt)

Analysis of log loss rate (LR) by tau = 3.5:
               Estimate   Std.Err Z value  Pr(>|z|)
Ref (Group 0) -0.223940  0.054308 -4.1235 3.731e-05 ***
Group 1 vs 0  -0.186019  0.084590 -2.1991   0.02787 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Test of group difference in while-alive LR
X-squared = 4.835905, df = 1, p = 0.02787301

Point and interval estimates for the LR ratio:
             LR ratio 95% lower CL 95% higher CL
Group 1 vs 0 0.8302577     0.703412     0.9799773


Analysis of log RMST (restricted mean survival time) by tau = 3.5:
              Estimate  Std.Err Z value  Pr(>|z|)
Ref (Group 0) 1.107544 0.016064 68.9443 < 2.2e-16 ***
Group 1 vs 0  0.056689 0.020349  2.7858  0.005339 **
```

R-package WA - code example (iv)

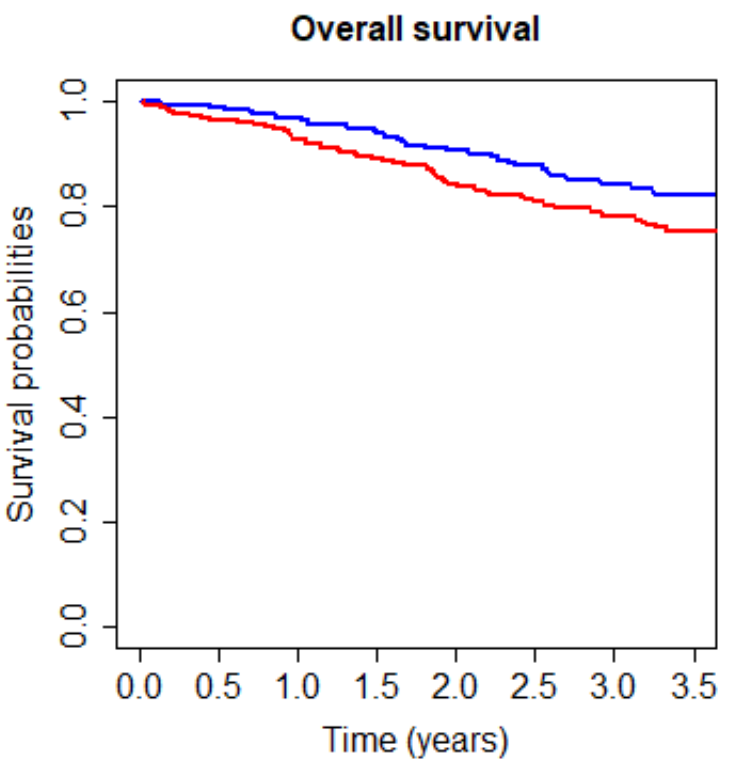
Graphics …

```
1  # plot the estimated survival-completed cumulative loss
2  # by group, with 95% confidence intervals
3  plot(obj, conf = TRUE, xlab = "Time (years)", xlim = c(0, 3.5),
4       ylab = "SC cumulative frequency")
```



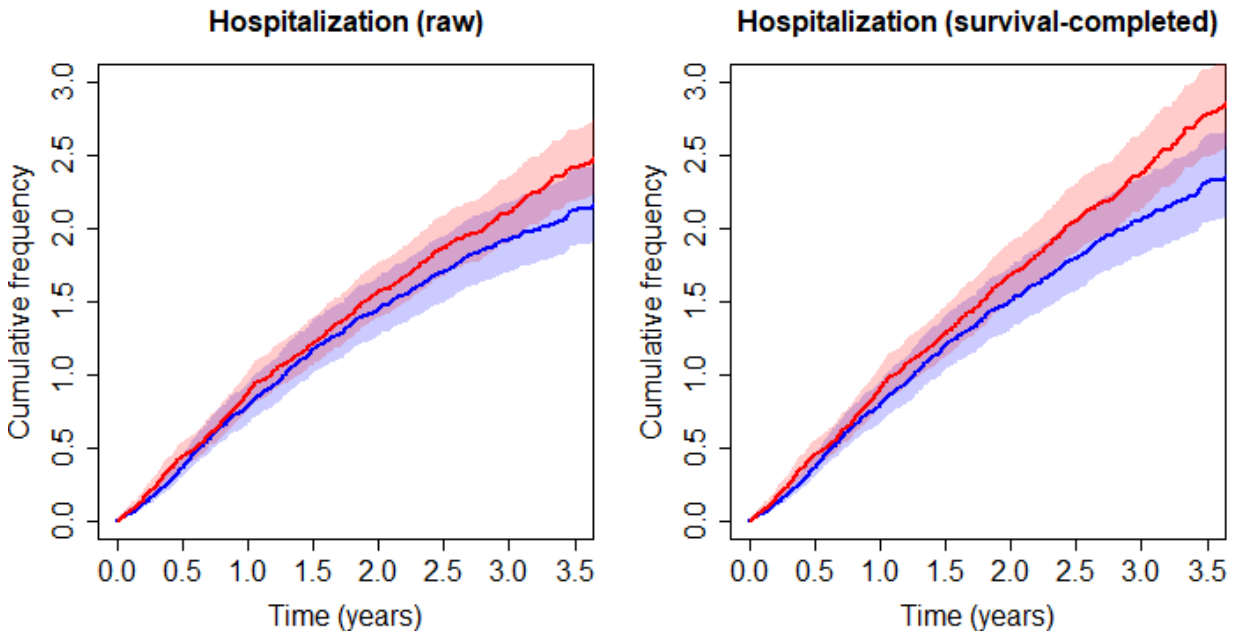
Data example - HF-ACTION (i)

	Exercise training	Usual care
Death rate	13.5%	19.9%
Avg # hospitalization (SD)	1.8 (2.1)	2.0 (2.1)



Blue: exercise training; Red: usual care

Data example - HF-ACTION (ii)



Blue: exercise training; Red: usual care

- Raw (left) shrinks treatment difference
  - Trained surviving longer $\rightarrow$ hospitalized more
- SC (right) corrects this by using event rate while alive

Data example - HF-ACTION (iii)

For $\tau = 3.5$ years...

TABLE Analysis of hospitalization and survival by $\tau = 3.5$ years in the HF-ACTION study

	Hospitalization				Survival	
	Raw frequency		SC frequency		RMST (years)	
	Est (95% CI)	*p*-value	Est (95% CI)	*p*-value	Est (95% CI)	*p*-value
Usual	2.43 (2.19, 2.68)	—–	2.80 (2.52, 3.11)	—–	3.03 (2.93, 3.12)	—–
Training	2.13 (1.88, 2.41)	—–	2.33 (2.05, 2.64)	—–	3.20 (3.13, 3.28)	—–
Ratio	0.88 (0.75, 1.03)	0.12	0.83 (0.70, 0.98)	0.03	1.06 (1.02, 1.10)	0.01

Note: SC, survival-completed (while-alive); Est, estimate; CI, confidence interval; Ratio, training vs. usual.

- In the first 3.5 years, exercise training on average reduces hospitalizations per person-year alive by 1 - 0.83 =17% (2%–30%; *p*-value 0.03)
- Joint test with RMST: χ^2_2=9.88, *p*-value 0.007

Data example - HF-ACTION (iv)

Weighted composites: $w_m(t) \equiv 1$ and $w_m^D(t) \equiv 1, 2, 3...$

	$w_m^D(t) \equiv 1$		$w_m^D(t) \equiv 2$		$w_m^D(t) \equiv 3$	
	Est (95% CI)	$p$val	Est (95% CI)	$p$val	Est (95% CI)	$p$val
Usual	3.08 (2.77, 3.43)	—	3.37 (3.02, 3.76)	—	3.65 (3.26, 4.09)	—
Training	2.52 (2.22, 2.86)	—	2.72 (2.38, 3.09)	—	2.91 (2.54, 3.33)	—
Ratio	0.82 (0.69, 0.97)	0.017	0.81 (0.68, 0.96)	0.013	0.80 (0.67, 0.95)	0.012

- In the first 3.5 years, exercise training on average reduces death and hospitalizations per person-year alive by 18%, 19%, and 20% under weights 1:1, 2:1, and 3:1, respectively

Summary

General while-alive loss rate: $\ell^{(a)}(\tau) = E\{\mathcal{L}(\mathcal{H}^{(a)})(\tau)\}/E(D^{(a)} \wedge \tau)$

- Summarizes loss profile while adjusting for length of exposure

Risk (loss rate) ratio: $r(\tau) = \ell^{(1)}(\tau)/\ell^{(0)}(\tau)$

- Treated experience $100r(\tau)\%$ as much loss as control does

R-package `WA`: https://cran.r-project.org/web/packages/WA

Open question: What if treatment delays nonfatal events without necessarily reducing their number by time τ

- Use a decreasing $w_m(t)$ to reward late occurrence?

Acknowledgments

This research is supported by NIH-NHLBI grant **R01HL149875**

> *Novel Statistical Methods for Complex Time-to-Event Data in Cardiovascular Clinical Trials*

HF-ACTION study data are provided by BioLINCC depository of NHLBI

References

Akacha, M., B. Binkowitz, F. Bretz, A. Fritsch, P. Hougaard, A. Jahn-Eimermacher, and et al. 2018. “Request for CHMP Qualification Opinion: Clinically Interpretable Treatment Effect Measures Based on Recurrent Event Endpoints That Allow for Efficient Statistical Analyses.” http://www.ema.europa.eu/documents/other/qualification-opinion-treatment-effect-measures-when-using-recurrent-event-endpoints-applicants_en.pdf.

CHMP. 2020. “Qualification Opinion of Clinically Interpretable Treatment Effect Measures Based on Recurrent Event Endpoints That Allow for Efficient Statistical Analyses.” https://www.ema.europa.eu/en/documents/other/qualification-opinion-clinically-interpretable-treatment-effect-measures-based-recurrent-event_en.pdf.

Fritsch, Arno, Patrick Schlömer, Franco Mendolia, Tobias Mütze, and Antje Jahn-Eimermacher. 2023. “Efficiency Comparison of Analysis Methods for Recurrent Event and Time-to-First Event Endpoints in the Presence of Terminal EventsApplication to Clinical Trials in Chronic Heart Failure.” *Statistics in Biopharmaceutical Research* 15 (2): 268–79. https://doi.org/10.1080/19466315.2021.1945488.

Ghosh, Debashis, and D. Y. Lin. 2000. "Nonparametric Analysis of Recurrent Events and Death." *Biometrics* 56 (2): 554–62. https://doi.org/10.1111/j.0006-341x.2000.00554.x.

Ghosh, Debashis, and Danyu Y Lin. 2002. "Marginal Regression Models for Recurrent and Terminal Events." *Statistica Sinica*, 663–88.

ICH. 2020. "ICH E9 (R1) Addendum on Estimands and Sensitivity Analysis in Clinical Trials tothe Guideline on Statistical Principles for Clinical Trials, Step 5." London: European Medicines Evaluation Agency.

Lin, D. Y., L. J. Wei, I. Yang, and Z. Ying. 2000. "Semiparametric Regression for the Mean and Rate Functions of Recurrent Events." *Journal of the Royal Statistical Society Series B: Statistical Methodology* 62 (4): 711–30. https://doi.org/10.1111/1467-9868.00259.

Mao, Lu. 2021. "WA: While-Alive Loss Rate for Recurrent Event in the Presence of Death." https://CRAN.R-project.org/package=WA.

———. 2023. "Nonparametric Inference of General While-Alive Estimands for Recurrent Events." *Biometrics* 79 (3): 1749–60. https://doi.org/10.1111/biom.13709.

Mao, Lu, and D. Y. Lin. 2016. "Semiparametric Regression for the Weighted Composite Endpoint of Recurrent and Terminal Events." *Biostatistics* 17 (2): 390–403. https://doi.org/10.1093/biostatistics/kxv050.

O'Connor, Christopher M., David J. Whellan, Kerry L. Lee, Steven J. Keteyian, Lawton

S. Cooper, Stephen J. Ellis, Eric S. Leifer, et al. 2009. “Efficacy and Safety of Exercise Training in Patients With Chronic Heart Failure.” *JAMA* 301 (14): 1439. https://doi.org/10.1001/jama.2009.454.

Royston, Patrick, and Mahesh K. B. Parmar. 2011. “The Use of Restricted Mean Survival Time to Estimate the Treatment Effect in Randomized Clinical Trials When the Proportional Hazards Assumption Is in Doubt.” *Statistics in Medicine* 30 (19): 2409–21. https://doi.org/10.1002/sim.4274.

Schmidli, Heinz, James H. Roger, and Mouna Akacha. 2023a. “Estimands for Recurrent Event Endpoints in the Presence of a Terminal Event.” *Statistics in Biopharmaceutical Research* 15 (2): 238–48. https://doi.org/10.1080/19466315.2021.1895883.

———. 2023b. “Rejoinder to Commentaries on “Estimands for Recurrent Event Endpoints in the Presence of a Terminal Event”.” *Statistics in Biopharmaceutical Research* 15 (2): 255–56. https://doi.org/10.1080/19466315.2023.2166098.

Uno, Hajime, and Miki Horiguchi. 2023. “Ratio and Difference of Average Hazard with Survival Weight: New Measures to Quantify Survival Benefit of New Therapy.” *Statistics in Medicine* 42 (7): 936–52. https://doi.org/10.1002/sim.9651.

Wei, Jiawei, Tobias Mütze, Antje Jahn-Eimermacher, and James Roger. 2023. “Properties of Two While-Alive Estimands for Recurrent Events and Their Potential Estimators.” *Statistics in Biopharmaceutical Research* 15 (2): 257–67.