ESTIMANDS IN CLINICAL TRIALS WITH COMPLEX LIFE HISTORY PROCESSES

Richard Cook

Joint Work with Alexandra Bühler and Jerry Lawless

Department of Statistics and Actuarial Science
University of Waterloo

15th Annual Conference on Statsitical Issues in Clinical Trials

Advances in Time to Event Analyses

April 17, 2023

UNIVERSITY OF WATERLOO

Outline

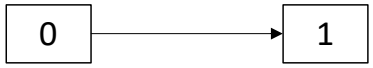
I. Multistate Processes in Clinical Trials

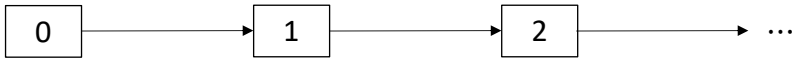
Bühler A, Cook RJ, Lawless JF (2023a). Multistate models as a framework for estimand specification in clinical trials of complex processes. *Statistics in Medicine*.

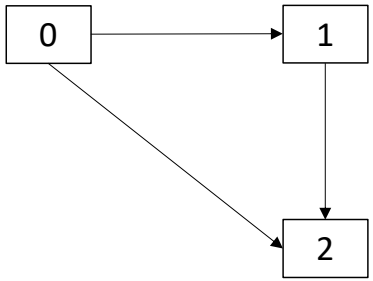
II. Joint Models Incorporating Intercurrent Events

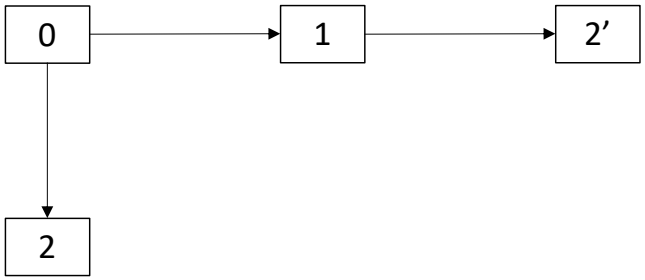
Bühler A, Cook RJ, Lawless JF (2023b). Generalized linear models for marginal features of complex life history processes. *Manuscript*.

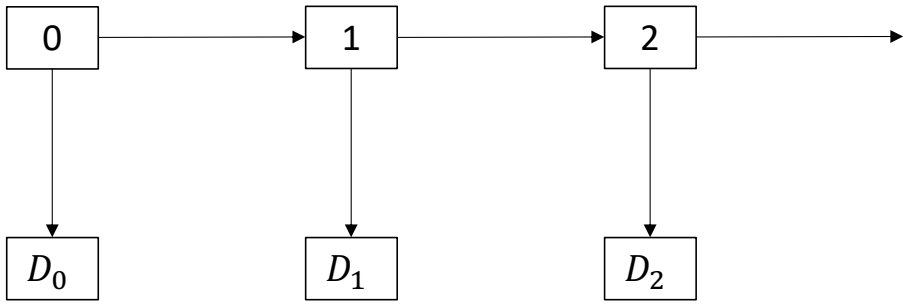
Some Canonical Multistate Processes











$\mathcal{S} = \{0, 1, \dots\}$ is state space with absorbing states $\mathcal{A}$

$\{Z(s), 0 < s\}$ is process

Intensity Functions

- $X = 1$ for treated and $X = 0$ for control and $\{V(s), 0 < s\}$ is a marker process
- $\mathcal{H}(t) = \{Z(s), V(s), 0 < s < t, X\}$

Intensity function of a $k \longrightarrow l$ transition Andersen et al. (1993)

$$\lim_{\Delta t \downarrow 0} \frac{P(Z(t + \Delta t^-) = l \mid Z(t^-) = k, \mathcal{H}(t))}{\Delta t} = \lambda_{kl}(t \mid \mathcal{H}(t)) \qquad k, l \in \mathcal{S}$$

Remarks

Intensity-based models play a critical role in understanding causal effects of life history processes

Aalen, Røysland and Gran (2012)

Causal effects are manifest over time so *attempts to understand causal mechanisms* should be based on intensity functions

Remarks

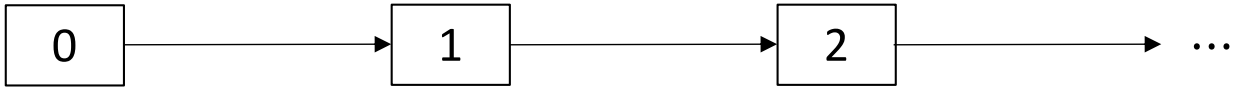
Intensity-based models play a critical role in understanding causal effects of life history processes — Aalen, Røysland and Gran (2012)

Causal effects are manifest over time so *attempts to understand causal mechanisms* should be based on intensity functions

Conditioning on the life history induces "collider bias" — Hernán (2010)

Clinical trials are not *primarily* designed to enhance understanding of causal mechanisms but rather to *test and estimate effects on marginal process features* and *facilitate regulatory decision making*

A Simple Case: Recurrent Events



If $\lambda_{k,k+1}(t \mid Z(t^-) = k, \mathcal{H}(t))$, $k = 0, 1, \ldots$ the rate function

$$\rho(t \mid X) = \lim_{\Delta t \downarrow 0} \frac{P(Z(t + \Delta t^-) - Z(t^-) = 1 \mid X)}{\Delta t}$$

is

$$\rho(t \mid X) = E_{\mathcal{H}(t)} \left\{ \lambda_{k,k+1}(t \mid Z(t^-) = k, \mathcal{H}(t)) \mid X \right\}$$

A working Poisson model $\rho(t \mid X) = \rho_0(t) e^{\gamma X}$ gives Lawless (1995)

$$\mu(t \mid X) = \mu_0(t) e^{\gamma X}$$

Andersen-Gill (1982) model offers a basis for causal inference. Lin et al. (2000)

For Markov processes

$$\mathbb{P}(t \mid X) = \prod_{(0,t]} \{\mathbb{I} + d\mathbf{\Lambda}(u \mid X)\}$$

Marginal features for multistate processes include

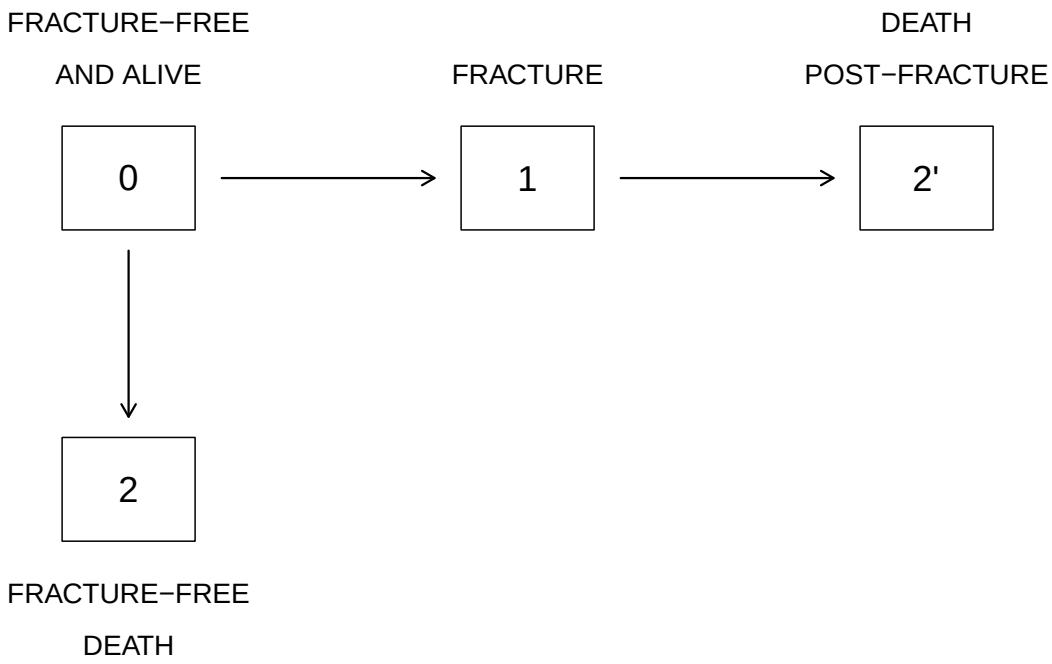
$$P(Z(t) = k \mid Z(0) = 0, X)$$

or

$$P(Z(t) \in \mathcal{A} \mid Z(0) = 0, X)$$

On their own, marginal features for complex processes are inadequate for assessing treatment effects.

Aim is to evaluate the effect of bisphosphonates in reducing risk of fractures



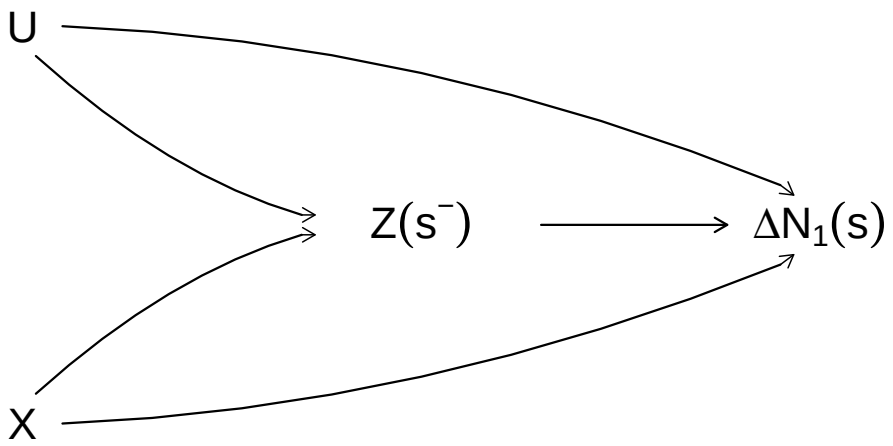
T_k is entry time to state k

$$N_k(s) = I(T_k \leq s), \;\; \Delta N_k(s) = N_k(s + \Delta s^-) - N_k(s^-)$$

Competing Risks

Cause-specific Cox regression

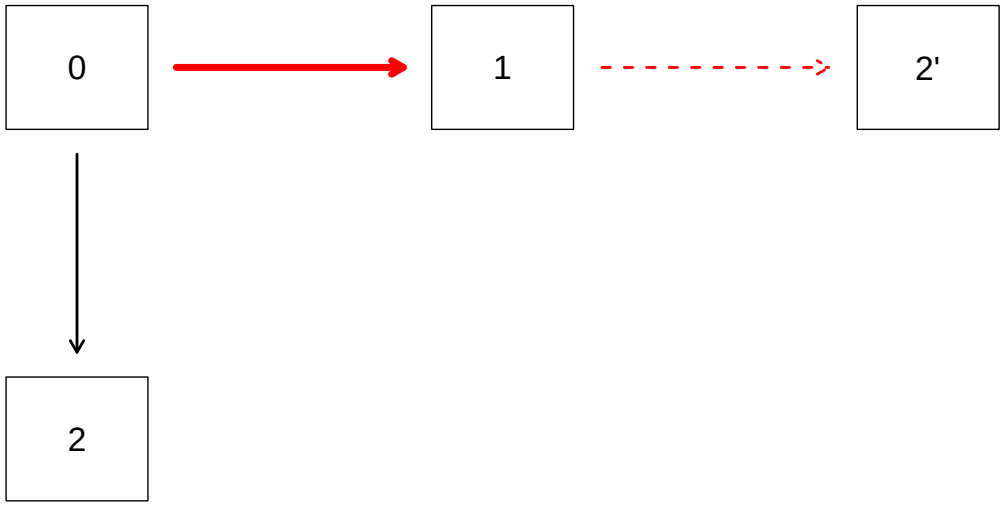
$$\lambda_1(s \mid X) = \lim_{\Delta s \downarrow 0} \frac{P(\Delta N_1(s) = 1 \mid Z(s^-) = 0,\ X)}{\Delta s} = \lambda_{10}(s)\, e^{\gamma_1 X}$$



$$U \not\perp X \mid Z(s^-) = 0$$

"Collider bias" from conditioning on $Z(s^-) = 0$

Marginal Models based on State Occupancy



Cumulative Incidence Function

Let $F_k(t) = P(T_k < t \mid Z(0) = 0) = P(N_k(t) = 1)$

General transformation models Fine and Gray (1999); Scheike et al. (2008)

$$g(F_1(t)) = \alpha_1(t) + x\,\beta_1 \tag{1}$$

This *fails to distinguish states* Putter et al. (2020)

0 – alive and event-free 2 – event-free death

Cumulative Incidence Function Regression

Fine and Gray (1999)

Let $Y_i^\dagger(t) = \mathbb{I}\left(T_{i1} \geqslant t\right)$

$$\sum_{i=1}^{n} w_i(t) Y_i^\dagger(t) \left(dN_{i1}(t) - \exp\left(\beta X_i\right) d\Gamma(t)\right) = 0$$

$$\sum_{i=1}^{n} \int_0^\infty w_i(t) Y_i^\dagger(t) \left(dN_{i1}(t) - \exp\left(\beta X_i\right) d\Gamma(t)\right) X_i = 0$$

and

$$w_i(t) = w_i^G(t) = \frac{\mathbb{I}\left(C_i > \min\left(T_i, t\right)\right)}{G_i\left(\min\left(T_i, t\right) \mid X_i\right)} \quad \text{with} \quad T_i = \min(T_{i1}, T_{i2})$$

The estimand $\beta^\star$ solves

Bühler et al. (2023b)

$$\mathbb{E}\left(U\left(\beta, G^*\right)\right) = \int_0^\infty \left\{ s^{(1)}(t) - \frac{s^{(1)}(t,\beta)}{s^{(0)}(t,\beta)} s^{(0)}(t) \right\} dt = 0 \ ,$$

$$s^{(r)}(t,\beta) = \mathbb{E}\left(w_i^\star(t) Y_i^\dagger(t) X_i^r \exp(\beta X_i) \right) \qquad s^{(r)}(t) = \mathbb{E}\left(w_i^\star(t) Y_i^\dagger(t) X_i^r dN_{i1}(t) \right)$$

Limiting values under intensity-based processes

Suppose $0 - k$ intensities are

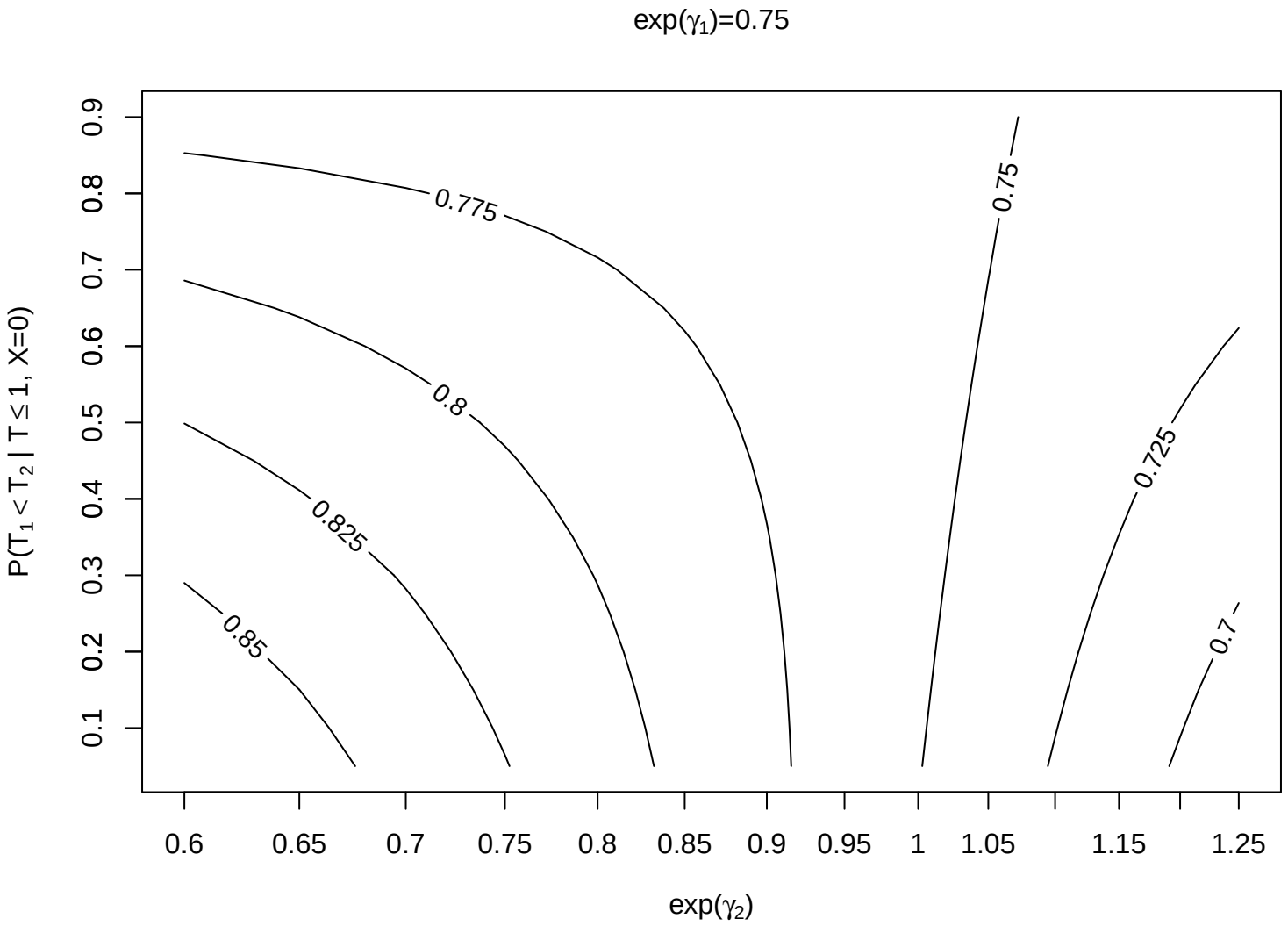
$$\lambda_{0k}(t \mid X) = \lambda_k \exp(\gamma_k X), \ k = 1, 2 .$$

Set $\exp(\gamma_1) = 0.75$

For given $\gamma = (\gamma_1, \gamma_2)'$, determine λ_k so that

- $P(T \leqslant 1 \mid X = 0) = 0.6$
- $P(T_1 < T_2 \mid T \leqslant 1, X = 0) \in (0, 1)$

Limiting Value of FG Estimator under Cause-Specific Hazards



Remarks

Analyses based on many *particular* marginal features are inadequate

Supplementary analyses are needed for a full understanding Scharfstein (2019)

– based on other marginal features
– using intensity-based or partially conditional models

Aalen-Johansen estimates of state occupancy probabilities are robust to violations of the Markov assumption Aalen et al (2001)

Utility-based Framework

Cook et al. (2003)

Marginal process feature:

$$U(\tau|X) = \sum_{k\in\mathcal{S}} \int_0^\tau \underbrace{U_k(u)}_{\substack{\text{state } k \text{ utility} \\ \text{at time } u}} \cdot P(Z(u) = k|X)du \ , \qquad U_k(u) = U_k \in [0,1] \ .$$

Model-free estimands:

$$\beta = U(\tau|X=1) - U(\tau|X=0) \quad \text{or} \quad \beta = \frac{U(\tau|X=1)}{U(\tau|X=0)}$$

Restricted Mean Survival Time (RMST)

McCaw et al. (2019)

Quality Adjusted Life Years

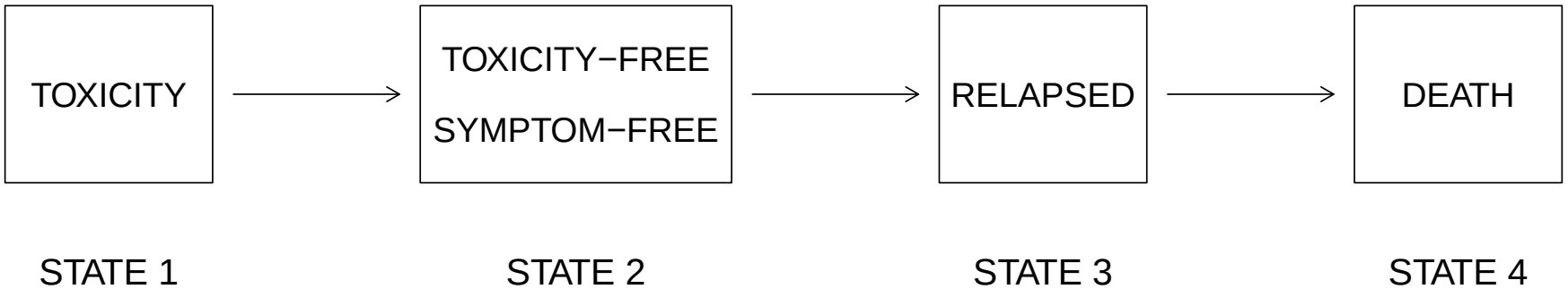
Gelber et al. (1989)

A Therapeutic Breast Cancer Trial — Glasziou et al. (1990)

A randomized trial designed to describe the quality adjusted survival over 84 months among breast cancer patients undergoing short or long duration chemotherapy

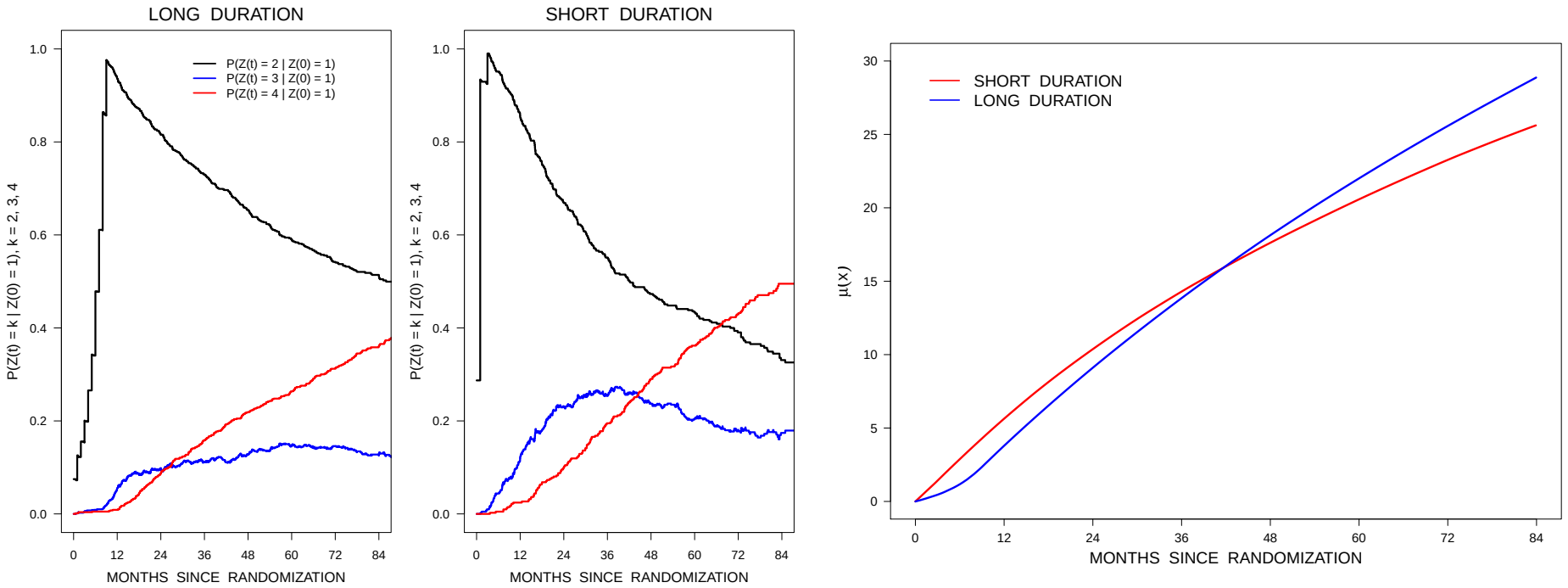
Total times spent in each transient state below are of interest

- 413 randomized to short duration chemotherapy ($X = 0$)
- 816 randomized to long duration chemotherapy ($X = 1$)



Mean cumulative quality of life

Cook and Lawless (2018)



With utilities $u_1 = 0.1$, $u_2 = 0.5$, $u_3 = 0.1$ and $u_4 = 0$,

$\widehat{U}(84 \mid X = 0) = 25.61$ (S.E. $= 0.73$) and $\widehat{U}(84 \mid X = 1) = 28.87$ (S.E. $= 0.47$)

II. Models Incorporating Intercurrent Events Bühler et al. (2023a)

Intercurrent events are defined as

"events occurring after treatment initiation that affect either the interpretation or the existence of the measurements associated with the clinical question of interest" ICH E9(R1)

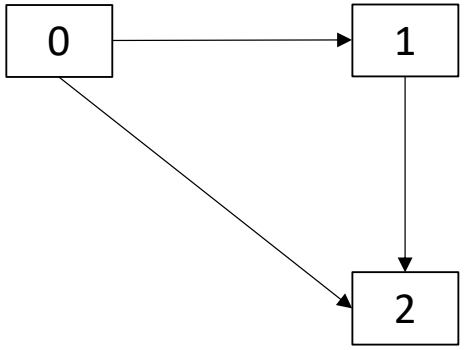
Type 1 Intercurrent Event: *precludes observation or occurrence* of event of interest

- Loss to follow-up
- Death

Type 2 Intercurrent Event: *changes the interpretation* of event of interest

- Rescue treatment
- crossover or treatment discontinuation

1. An estimand should target a *marginal process feature* with clear *scientific relevance*.
2. Features and estimands should be *interpretable* in the "*real world*" rather than in any hypothetical world.
   - Target of inference should be an element of the observable process!
3. Estimands should not be sensitive to *uncheckable assumptions*, and consistent with *observed data* and *scientific background*.
   - Models on which an estimand is based should be assessed using available data.



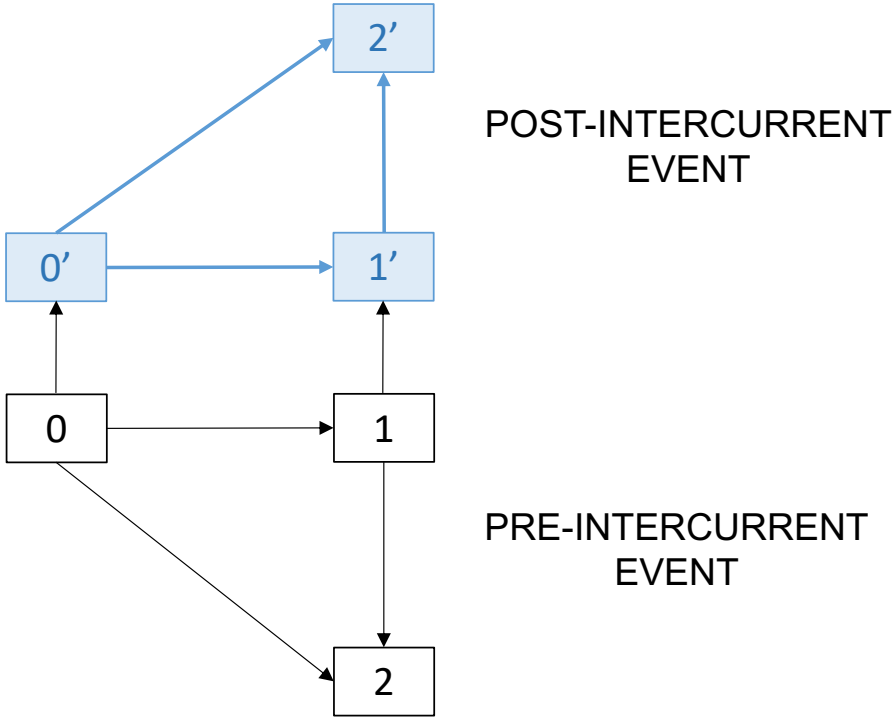
A. Illness-Death Process

$\{Z(s), 0 < s\}$

$\mathcal{H}(t) = \{Z(s), 0 < s < t, X\}$

$\lambda_{kl}(t \mid \mathcal{H}(t))$

$$q_{kl}(t \mid X) = \lim_{\Delta t \downarrow 0} \frac{P(Z(t + \Delta t^-) = l \mid Z(t^-) = k, X)}{\Delta t}$$



B. Illness-Death-IE Joint Process

$\{Z^\circ(s), 0 < s\}$

$\mathcal{H}^\circ(t) = \{Z^\circ(s), 0 < s < t, X\}$

$\lambda^\circ_{kl}(t \mid \mathcal{H}^\circ(t))$

$$q^\circ_{kl}(t \mid X) = \lim_{\Delta t \downarrow 0} \frac{P(Z^\circ(t + \Delta t^-) = l \mid Z^\circ(t^-) = k, X)}{\Delta t}$$

$$Y_c(t) = I(t \leq C)$$

$$Y_k(t) = I(Z(t^-) = k)$$

$$\frac{Y(t)Y_c(t)}{P(C > \min(T,t) \mid \mathcal{H}^\circ(t))} \{dN_1(t) - d\Gamma(t \mid X)\}$$

IPCW renders $C \perp \mathcal{H}(\cdot) \mid X$

A Type II Intercurrent Event: Rescue treatment

$Y_E(t) = I(t \leq E)$ where E is time of *introduction of rescue treatment*.

$$\frac{Y_E(t)Y_k(t)}{P(E \geq \min(T,t) \mid \mathcal{H}^\circ(t))} \{dN_1(t) - d\Gamma(t \mid X)\}$$

IPW renders $E \perp \mathcal{H}^\circ(\cdot) \mid X$

This hypothetical scenario *has little bearing on clinical care*.

Remarks

- Inverse probability of censoring weighting deals with unrepresentativeness due to selective attrition

- When censoring at intercurrent events IPCW creates a pseudo-sample not representative of any real-world setting

- Intention-to-treat analyses preferred approach incorporating IE into response process. Scharfstein (2019)

References

Hortobagyi, G. N., Theriault, R. L., Porter, L., Blayney, D., Lipton, A., Sinoff, C., Wheeler, H., Simeone, J. F., Seaman, J., and Knight, R. D. (1996). Efficacy of pamidronate in reducing skeletal complications in patients with breast cancer and lytic bone metastases. *New England Journal of Medicine*, **335**(34): 1785–1791.

Andersen, P. K., Borgan, O., Gill, R. D., and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer-Verlag, New York, NY.

Cook, R. J. and Lawless, J. F. (2018). Multistate Models for the Analysis of Life History Data. *CRC Press*.

Hernán, M. A. (2010). The hazards of hazard ratios. *Epidemiology*, **21**(1): 13–15.

Bühler, A., Cook, R. J., and Lawless, J. F. (2023a). Multistate models as a framework for estimand specification in clinical trials of complex processes. *Statistics in Medicine*.

Bühler, A., Cook, R. J., and Lawless, J. F. (2023b). Generalized Linear Models for Marginal Features in Complex Life History Processes. *Manuscript*.

Cook, R. J., Lawless, J. F., and Lee, K.-A. (2003). Cumulative processes related to event histories. *SORT*, **27**(1): 13–30.

Aalen O. O., Røysland K, Gran JM (2012a). Causality, mediation and time: a dynamic viewpoint. *Journal of the Royal Statistical Society: Series A*, **175**(4): 831–861.

Hortobagyi GN, Theriault RL, Porter L et al. (1996). Efficacy of Pamidronate in reducing skeletal complications in patients with breast cancer and lytic bone metastases. *New England Journal of Medicine*, **335**: 1875 – 1792.

Putter H, Schumacher M, van Houwelingen HC (2020). On the relation between the cause-specific hazard and the subdistribution rate for competing risks data: The Fine-Gray model revisited. *Biometrical Journal*, **62**: 790–807.

McCaw ZR, Yin G and Wei LJ (2019). Using the restricted mean survival time difference as an alternative to the hazard ratio for analyzing clinical cardiovascular studies. *Circulation*, **140** (17), 1366–1368.

Gelber RD, Gelman RS and Goldhirsch A (1989). A quality-of-life-oriented endpoint for comparing therapies. *Biometrics* **45**: 781–795.

Glasziou PP, Simes RJ and Gelber RD (1990). Quality adjusted survival analysis. *Statistics in Medicine* **9**: 1259–1276.

Cook RJ, Lawless JF and Lee K-A (2003). Cumulative processes related to event histories. *SORT*, **27**(1): 13–30.

Andersen, Per Kragh, and Richard D. Gill. (1982). Cox's regression model for counting processes: a large sample study. *The annals of statistics*, 1100-1120.

Ich, C. H. M. P. (2019). E9 (R1) Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. In *Proceedings of the international conference on harmonisation of technical requirements for registration of pharmaceuticals for human use*.

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

Fine, J. P., and Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American statistical association*, **94**(446): 496-509.

Scheike, T. H., Zhang, M. J., and Gerds, T. A. (2008). Predicting cumulative incidence probability by direct binomial regression. *Biometrika*, **95**(1): 205-220.

Aalen, O. O., Borgan, Ø., and Fekjær, H. (2001). Covariate adjustment of event histories estimated from Markov chains: the additive approach. *Biometrics*, **57**(4): 993-1001.

Scharfstein, D. O. (2019). A constructive critique of the draft ICH E9 addendum. *Clinical Trials*, **16**(4): 375-380.

Lawless, J. F., and Cook, R. J. (2019). A new perspective on loss to follow-up in failure time and life history studies. *Statistics in Medicine*, **38**(23): 4583-4610.

Lawless, J.F. (1995). The analysis of recurrent events for multiple subjects. *Applied Statistics*, **44**: 487-498.

Effect of Radiosurgery Alone vs Radiosurgery With Whole Brain Radiation Therapy on Cognitive Function in Patients With 1 to 3 Brain Metastases

A Randomized Clinical Trial

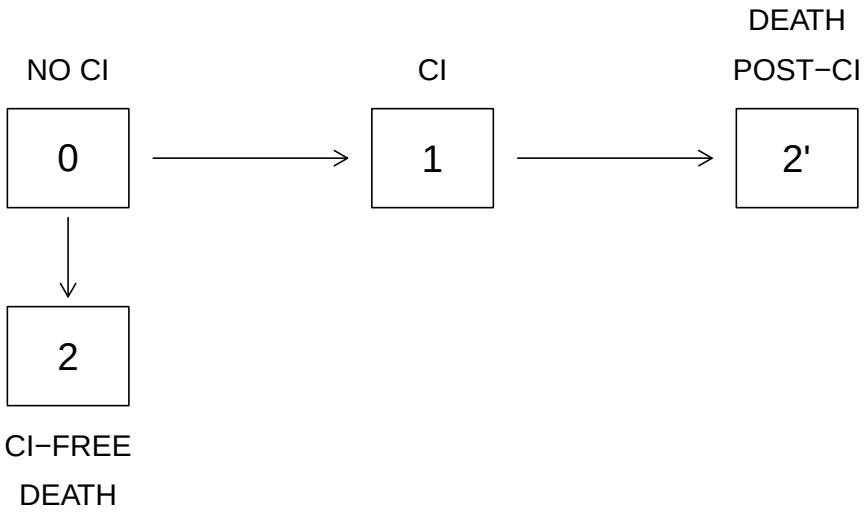
Paul D. Brown, MD; Kurt Jaeckle, MD; Karla V. Ballman, PhD; Elana Farace, PhD; Jane H. Cerhan, PhD; S. Keith Anderson, MS; Xiomara W. Carrero, BS; Fred G. Barker II, MD; Richard Deming, MD; Stuart H. Burri, MD; Cynthia Ménard, MD; Caroline Chung, MD; Volker W. Stieber, MD; Bruce E. Pollock, MD; Evanthia Galanis, MD; Jan C. Buckner, MD; Anthony L. Asher, MD

IMPORTANCE Whole brain radiotherapy (WBRT) significantly improves tumor control in the brain after stereotactic radiosurgery (SRS), yet because of its association with cognitive decline, its role in the treatment of patients with brain metastases remains controversial.

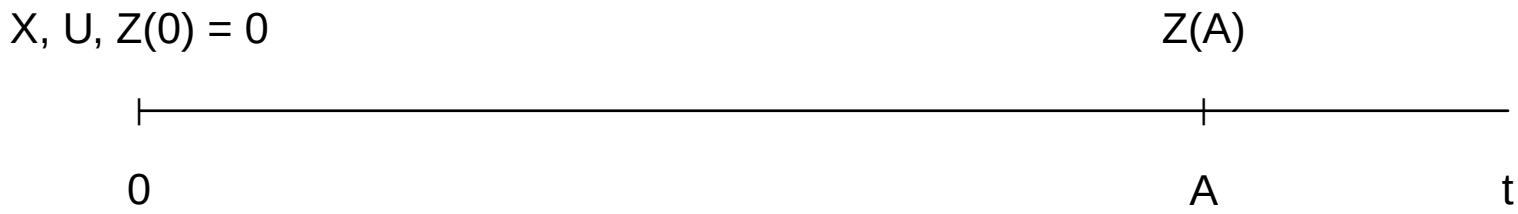
OBJECTIVE To determine whether there is less cognitive deterioration at 3 months after SRS alone vs SRS plus WBRT.

Brown PD, Jaeckle K, Ballman KV, et al. (2016). Effect of radiosurgery alone vs. radiosurgery with whole brain radiation therapy on cognitive function in patients with 1 to 3 brain metastases: a randomized clinical trial. *Journal of the American Medical Association*, **316**(4): 401–409.

A Data Generating Process



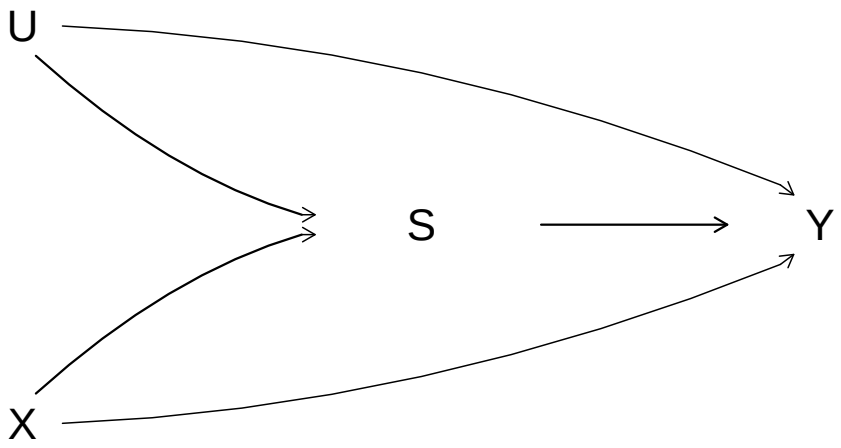
- $\{Z(t),\ 0 < t\}$ is the stochastic process
- $X = 1$ for treated and 0 for control
- U are prognostic variables
- $\mathcal{H}(t) = \{Z(s),\ 0 < s < t,\ X, U\}$



$S = I(\text{survival to follow-up time } A)$

Y indicates cognitive impairment

Aim to evaluate intervention effect on cognitive impairment at A



$X \perp U$ due to randomization

$X \not\perp U \mid S = 1$

Remarks

- Rubin (2006) gives rationale for SACE as a causal estimand
- *Principal strata are latent* so identifiability and estimation of the SACE requires additional modeling assumptions Egleston et al. (2007)

Questions

Is the *interpretation of the SACE aligned with scientific aims?*

> *If we cannot translate our causal question into a target trial, the question is not well-defined* Hernán (ISCB, 2020)

Should we *re-evaluate the specification of study objectives in presence of mortality?*

Remarks on Counterfactuals

> *"Many counterfactual analyses are based, explicitly or implicitly, on an attitude that I term fatalism. This considers the various potential responses $Y_i(u)$, when treatment i is applied to unit u, as predetermined attributes of unit u, waiting only to be uncovered by suitable experimentation."*
>
> Dawid (JASA, 2000a)

In their discussion of his paper, Robins and Greenland (2000) and Rubin (2006) *reaffirm the need for counterfactuals*, however in Dawid's rejoinder . . .,

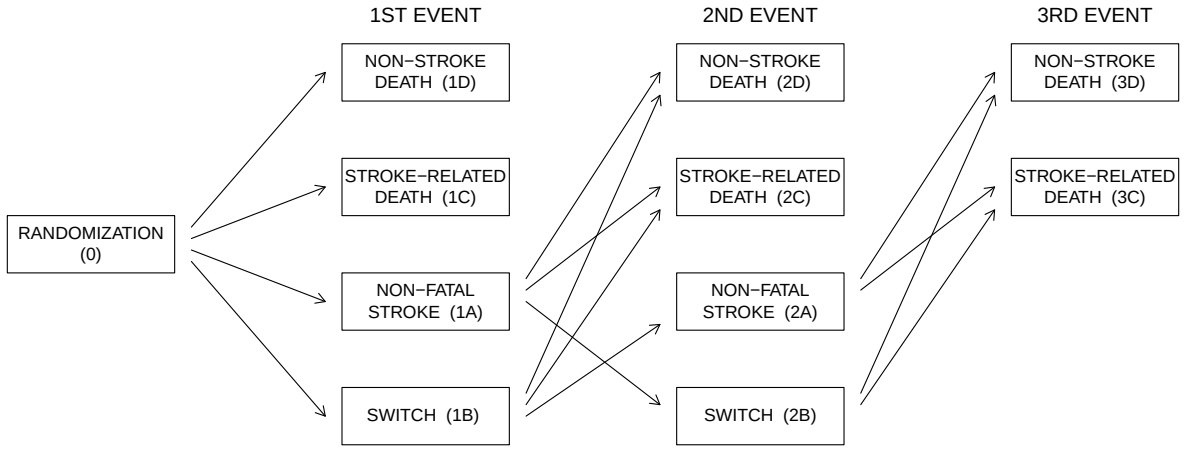
> *"... my own attitude is that, as there is no difficulty in determining an empirically meaningful probability structure for the observable (Y, Z) given treatment – even though this is defined over an unusual space, where Z automatically takes the value "undefined" whenever $Y = 0$ ... The real problem is how to define a sensible utility measure on this outcome space."*
>
> Dawid (2000b)

Carotid endarterectomy vs. medical care in stroke prevention

Barnett et al. (1998) report on a multicenter clinical trial designed to evaluate the effect of carotid endarterectomy vs. medical therapy.

Endpoints include

- any ipsilateral stroke
- any stroke
- stroke or stroke-related death
- any stroke or death



Dynamic Path Analyses

Aalen et al. (2012a, 2012b) point out the *importance and appeal of a stochastic intensity-based point of view* in the study of casual mechanisms

Regulatory Decision Making Among Regulators

Randomized clinical trials are typically designed to estimate and test treatment effects defined on marginal quantities

Some desirable features of analyses include

- interpretable estimands
- robustness to misspecification
- validity
- reproducibility
- facilitates decision making

Guiding Principles for Defining Estimands

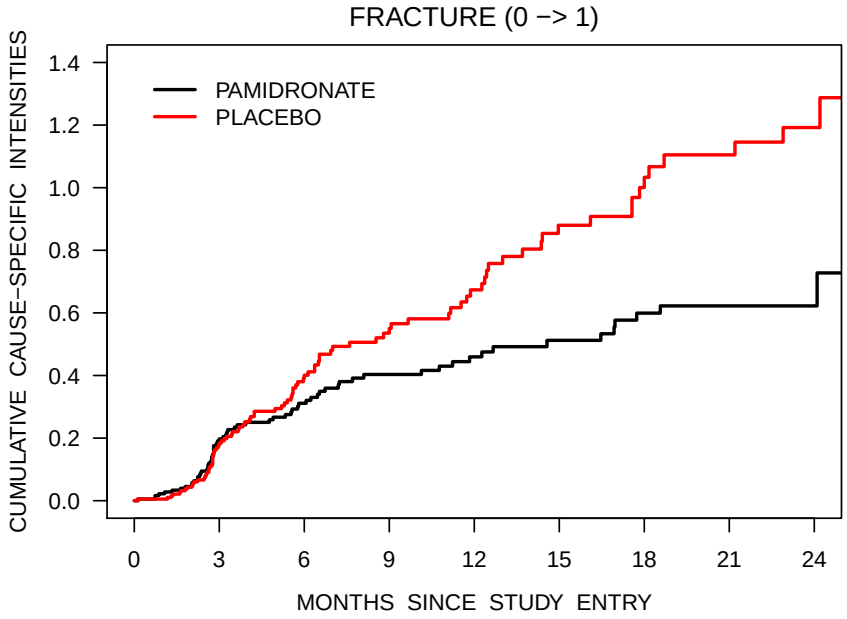
Process feature is any functional of the set of intensities Andersen and Keiding (2012)

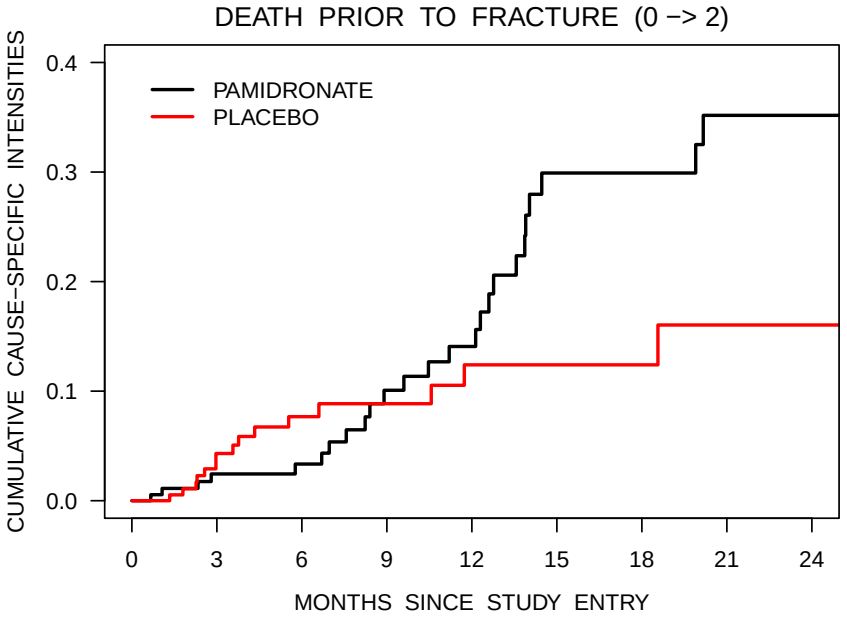
Estimand β = one-dimensional measure in the difference of a process feature between the treatment groups $X = 1$ and $X = 0$.

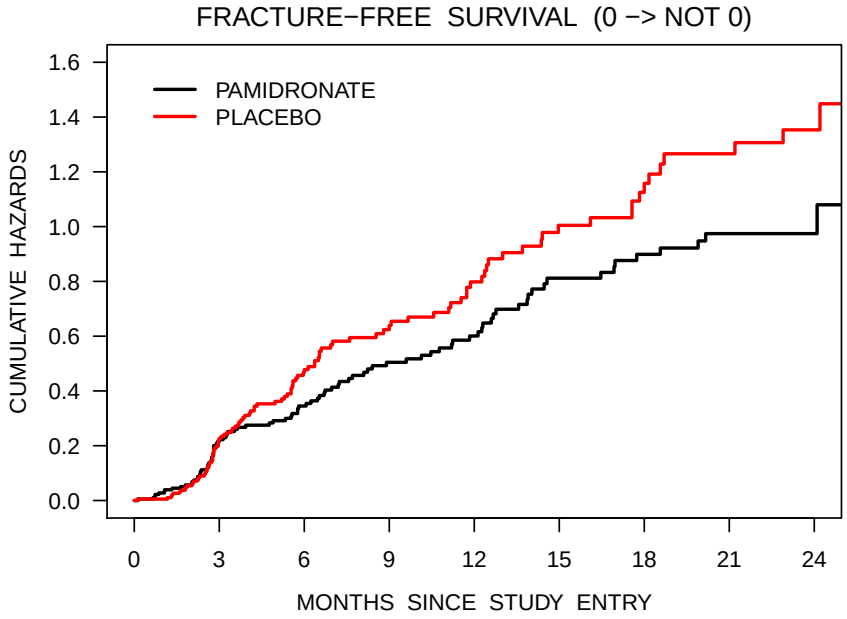
- ***Conditional (or dynamic) features*** have a *dynamic causal* interpretation Aalen (2012)
  - conditioning on $\mathcal{H}(t)$ induces time-dependent confounding, i.e., $X \not\perp\!\!\!\perp V \mid \mathcal{H}(t)$ Hernán (2010); Aalen et al. (2015)
  - crucial to a full understanding of the disease process
  - however: not suited for "simple" causal inference based on randomization

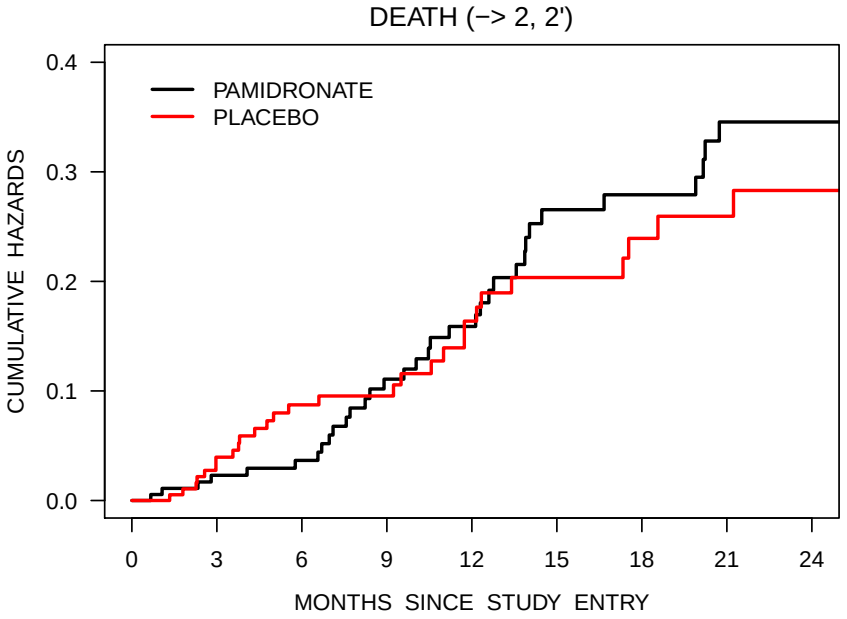
- ***Marginal features*** do not condition on $\mathcal{H}(t)$ and have a *descriptive causal* interpretation!

Nelson-Aalen Estimates of Cumulative Intensities







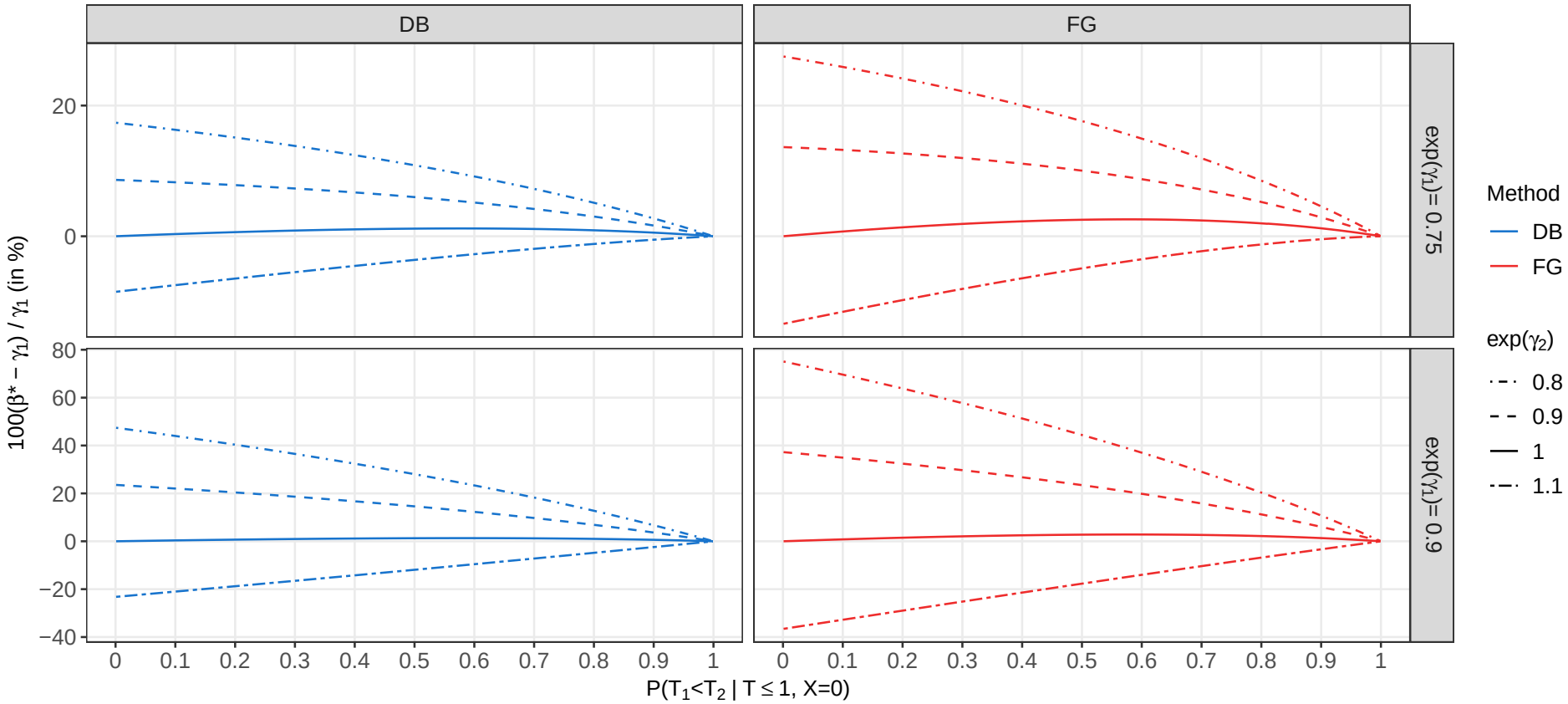


Remarks

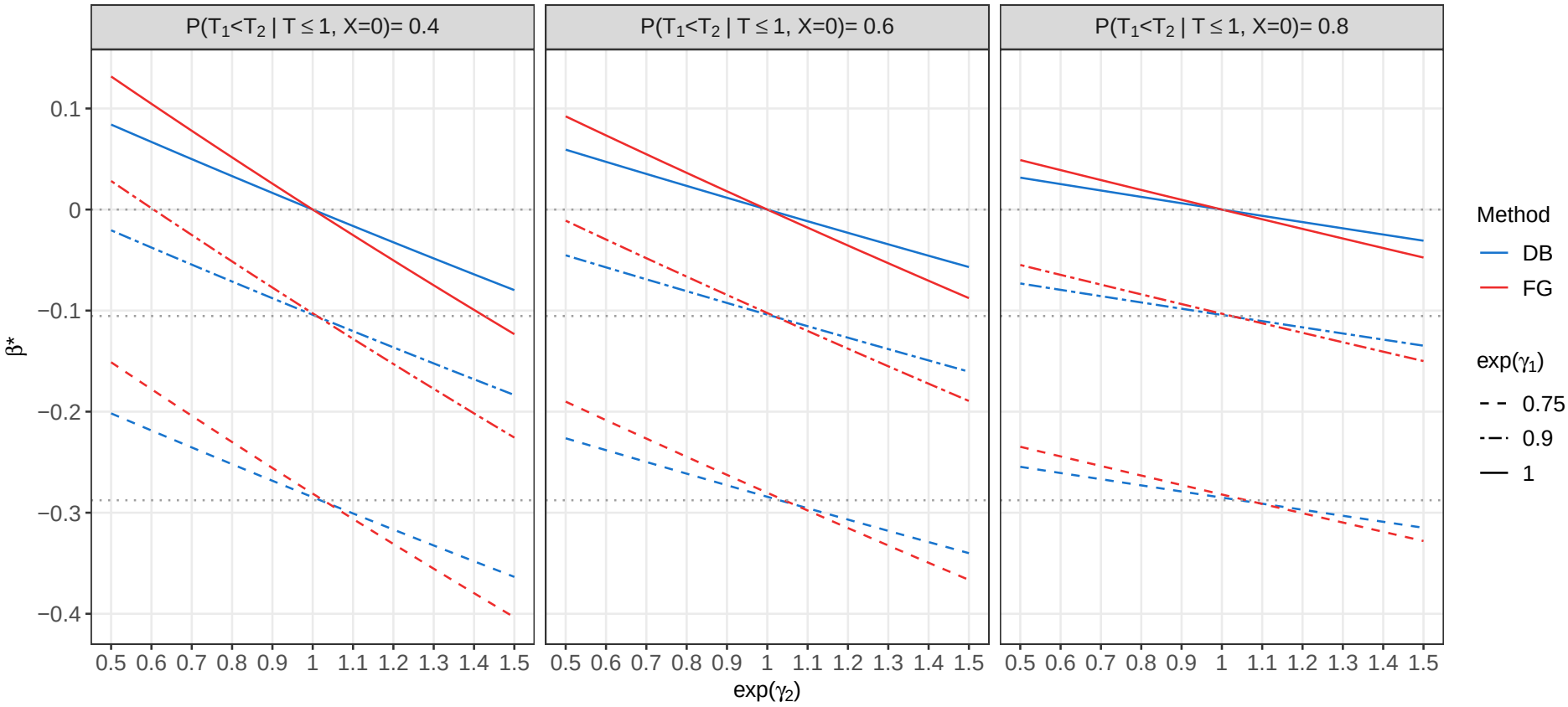
Multistate models provide a powerful framework for

1. characterizing disease processes and the introduction of interventions, loss to followup or other complications arising in the conduct of randomized clinical trials.
2. the analysis and interpretation of marginal process features used to define causal effects.
3. extensive secondary analysis of treatment effects and evaluation of overall response to treatment.

Asymptotic % Relative Difference: $100(\beta^\star - \gamma_1)/\gamma_1$



$\beta^\star$ vs. e^{γ_2} for different values of e^{γ_1} and $P(T_1 < T_2 | T \leq \tau, X = 0)$



Independence Assumption 1: For $(k,l) \in \{(0,1),(0,2),(1,2)\}$

$$\lim_{\Delta t \downarrow 0} \frac{P(Z^\circ(t+\Delta t^-) = l \mid Z^\circ(t^-) = k, \mathcal{H}^\circ(t))}{\Delta t} = \lambda_{kl}(t \mid \mathcal{H}(t)) \tag{X}$$

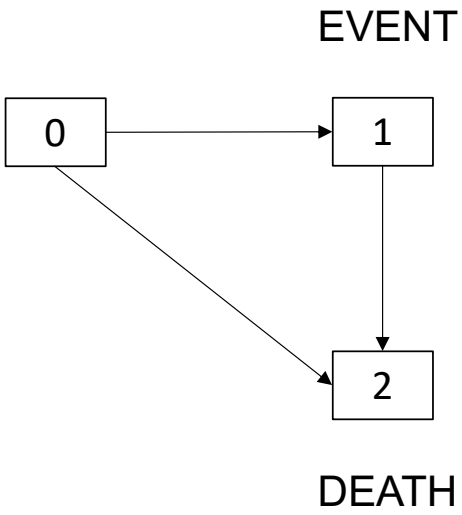
Independence Assumption 2: For $(k,l) \in \{(0,1),(0,2),(1,2)\}$

$$\lambda^\circ_{k'l'}(t \mid \mathcal{H}^\circ(t)) = \lambda_{kl}(t \mid \mathcal{H}(t)) \tag{Y}$$

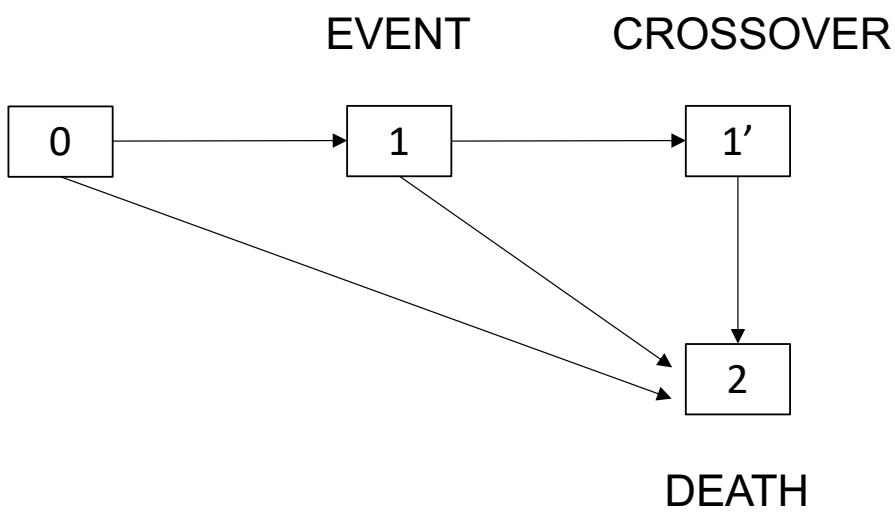
If (X) is violated we can model $0 \to 0'$ and $1 \to 1'$ to construct IPCW.

Note (Y) is not checkable without auxiliary data.

Crossover



EXPERIMENTAL ARM (X=1)



CONTROL ARM (X=0)

Event-free survival $P(Z^{\circ}(t) = 0 \mid X)$

Cumulative Incidence $g(P(Z^{\circ}(t) \geq 1 \mid X)) = \alpha_1(t) + \beta_1 X$

Survival? Intention to treat

$$P(Z(t) \neq 2 \mid X)$$

Remarks on potential outcomes

- Potential outcomes have played a central role in the development of causal inference theory and methods.
- The potential outcome framework can lead to specification of target estimands of dubious scientific relevance.
- Consider assessing the effect of a new intervention versus standard care on an outcome that can only be measured in individuals who are alive.
- "principal strata" are not identifiable (observable) from the available data Lipkovich et al. (2022).
- The survivor average causal effect does not convey the effect of treatment on all of those randomized, subset of the population may be small; see also Hernán and Scharfstein (2018)

Preferable strategies for dealing with intercurrent events include

- ignoring their occurrence in an intention-to-treat analysis targetting the effect of a "treatment policy" (i.e. the effect of prescribing one treatment versus the other at the time of study entry).
- incorporating into a composite endpoint.