# Some biases arising from patient withdrawal in RCTs and how to address them

Richard Cook
University of Waterloo

CANNecTIN Seminar
December 12, 2008

## OUTLINE

- DEPENDENTLY MISSING RESPONSE DATA
- RECURRENT EVENTS WITH DEPENDENT CENSORING
- PROTOCOL DRIVEN DEPENDENT CENSORING
- TAKE HOME MESSAGES

## AIMS OF THIS SEMINAR

- Raise awareness of the impact of dependently missing/incomplete data
- Consider ways of assessing whether this is an issue in studies
- Discuss ways of dealing with it in analysing data from trials
- To develop an understanding of features most likely to be affected by dependently missing/incomplete data.

## SOME BASIC PRINCIPLES REGARDING MISSING DATA

Incomplete data can arise from

- missed assessment
- drop-out
- protocol driven study withdrawal

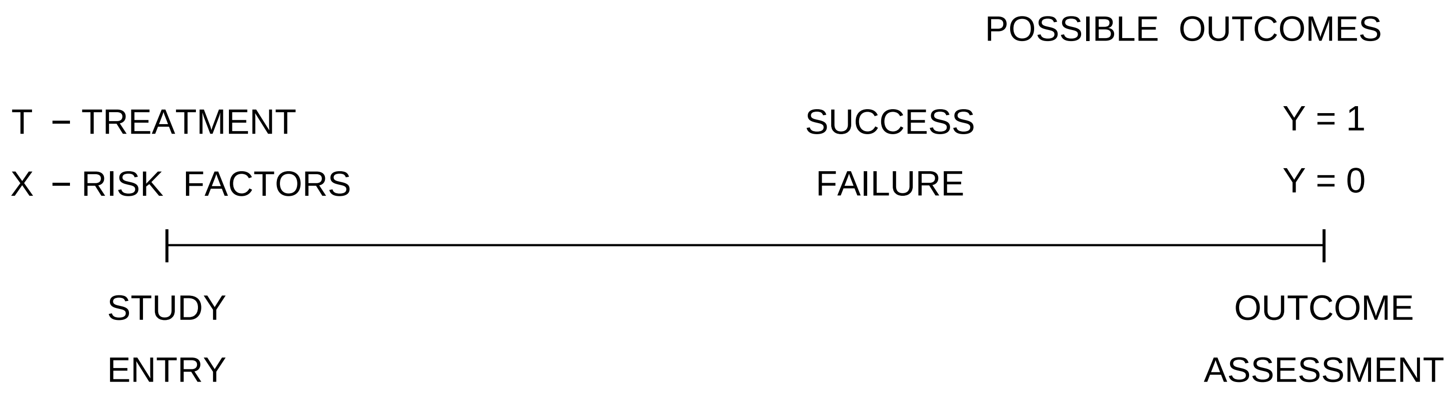
Standard analyses can give seriously biased estimates of means, event rates, and associated treatment effects

Careful thought and additional analyses are required to investigate the impact of incomplete data on inferences

## A SIMPLE FOLLOW-UP STUDY

With a binary response, interest often lies in

- the probability of success for treated patients: $P(Y = 1|T = 1)$
- the probability of success for control patients: $P(Y = 1|T = 0)$
- associated measures of treatment effect (ARR, RRR, OR, NNT)

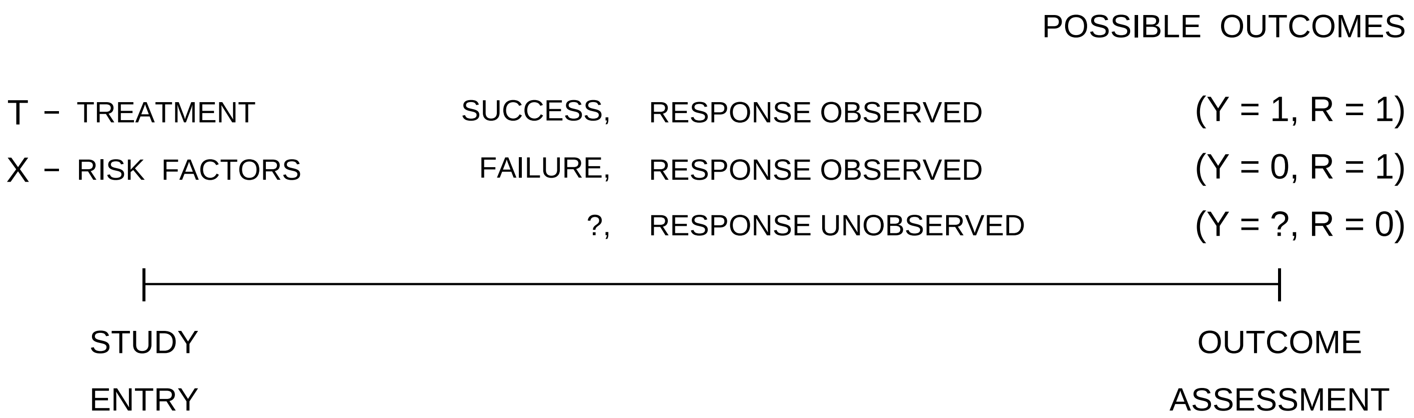


## PROBLEM

With incomplete data, $R = 1$ if response is observed, and $R = 0$ otherwise.

We then have three "outcomes"!

- $(Y = 1, R = 1)$ - success and response observed
- $(Y = 0, R = 1)$ - failure and response observed
- $(Y = ?, R = 0)$ - response *not observed*



## IMPLICATIONS OF MISSING DATA

- It is tempting to analyse *available data* in the standard way.
- In this case we are estimating the

  probability of success given treatment, *and that response was observed*.

- This is $P(Y = 1|T = 1, R = 1)$

## CENTRAL QUESTION

- How similar is the probability of success among those subjects observed and those subjects unobserved?
- Does $P(Y = 1|T, R = 1) = P(Y = 1|T, R = 0)$?
- Is sub-sample available at end of study representative of sample recruited?

## Smoker's Help-lines

- Smoker's wishing help to quit smoking call a "Help-line" available in many provinces
- Caller's receive counselling to help them quit
- Caller's are asked if they will participate in a study and consent to be contacted for a six month follow-up assessment
- Attempts are made to contact consenting participants six months later
- Not all people consenting people are contacted.
- How does this impact estimation of quit rates among callers to the help-lines?

## A SIMPLE ILLUSTRATIVE ONE-SAMPLE EXAMPLE

- Population is heterogenous
- Suppose a covariate $X$ explains this heterogeneity
- $X = 1$ for patient with a low response rate; $X = 0$ otherwise
- Suppose half of the patients have a low response rate, so $P(X = 1) = 0.5$

- OUTCOME
  - For patients with a low response rate : $P(Y = 1|X = 1) = 0.40$
  - For other patients : $P(Y = 1|X = 0) = 0.80$

- MISSING STATUS
  - For patients with low response rate: $P(R = 1|X = 1) = 0.50$
  - For other patients : $P(R = 1|X = 0) = 1.00$

In clinical trials, primary interest is in *marginal response rates*, $P(Y = 1)$

- In this example, the marginal response rate in the population is

$$P(Y = 1) = 0.60$$

- *Among those with an observed response*,

$$P(Y = 1|R = 1) = 0.66!$$

- This difference of 6% arises because there is a lower percentage of individuals with a low response rate available at study completion.

- Rates are the same if
  - Variable X is not associated with missingness (e.g. $P(R|X) = P(R)$ )
  - Outcome (Y) and “missingness” (R) are independent

## APPROACH 1: ADOPT A MORE COMPLETE MODEL FOR RESPONSE PROCESS

- Control for $X$ in analysis (analysis of covariance, ANCOVA)
- This approach renders missingness unimportant
- Then $P(Y = 1|X, R = 1) = P(Y = 1|X)$
- But, we abandoned our original objective of estimating $P(Y = 1)$!
- With some work we can average over covariate distribution to obtain

$$E_X(P(Y = 1|X)) = P(Y = 1)$$

## APPROACH 2: MODEL THE MISSING DATA PROCESS

- Model $P(R = 1|X)$ via logistic regression, say
- Then construct an estimating equation

$$\sum_{i=1}^{m} \frac{R_i}{P(R_i|X_i)}(Y_i - P(Y_i = 1))$$

giving a *weighted estimate*

$$P(Y_i = 1) = \frac{\sum_{i=1}^{m} R_i Y_i / P(R_i = 1|X_i)}{\sum_{i=1}^{m} R_i / P(R_i = 1|X_i)}$$

- Numerator and denominator are weighted sums where each observed person's contribution is weighted since they *represent individuals in the original sample* for whom $R = 0$

## EXAMPLES OF RECURRENT EVENT PROCESSES

- Exacerbations in respiratory diseases such as asthma or cystic fibrosis
- Occurrence of seizures in neurology (e.g. epilepsy)
- Graft rejection episodes in transplant studies and total graft rejection
- Trials of cancer patients with bone metastases at risk of fractures and death

## TIMELINE DIAGRAM

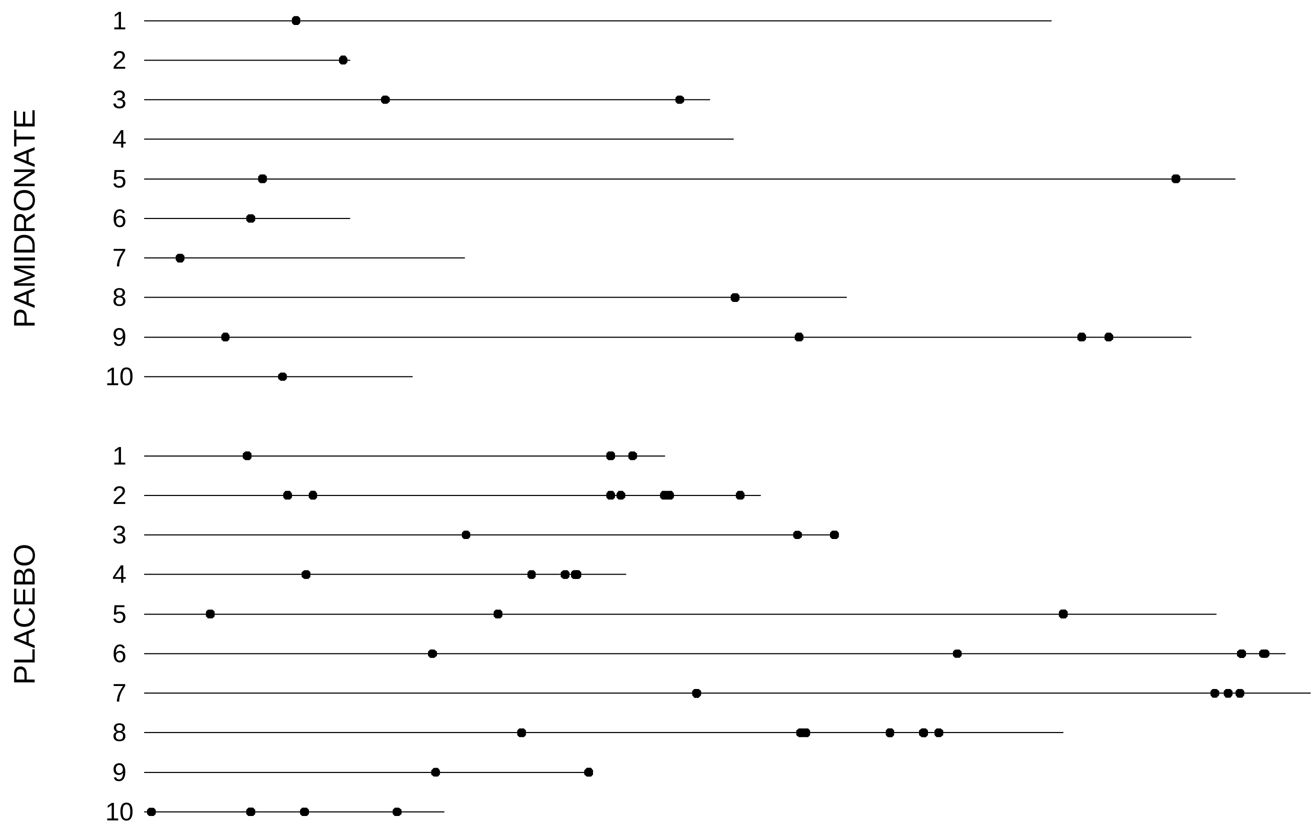
$T_k$ is the time of the $k$th event

## A TRIAL OF PATIENTS WITH SKELETAL METASTASES [1]

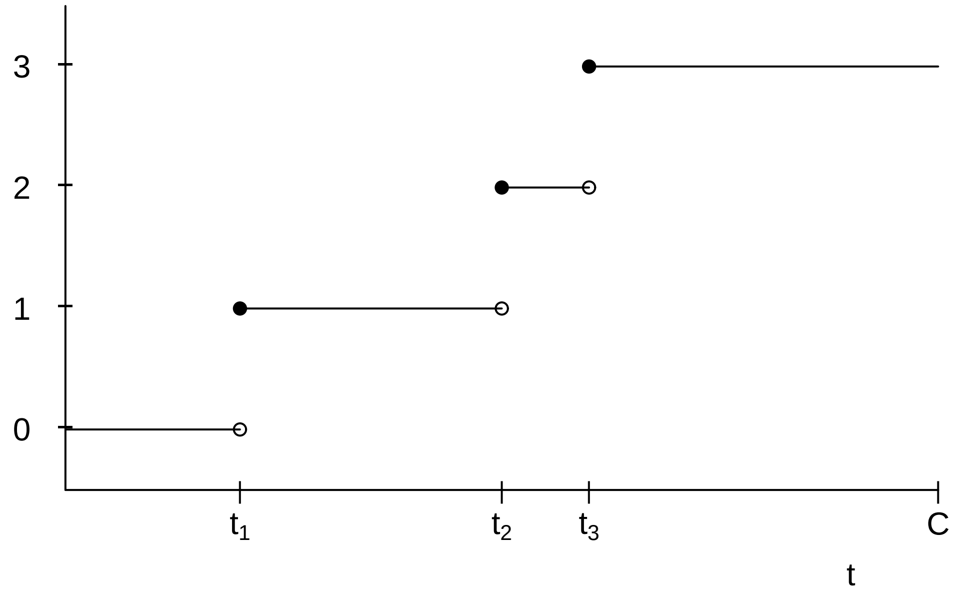
- An international multi-center randomized placebo-controlled trial of stage IV breast cancer patients with at least one $\geq 1$ cm lytic bone lesion (metastasis)
- Bone metastases compromise the integrity of skeletal structure and cause bone pain
- Aim of trial is to improve quality of life rather than affect survival
- Clinical event is a "skeletal event" (e.g. fracture) which arise from bone metastases
    - $185$ received pamidronate and $187$ received placebo
    - 24 months follow-up in extension phase

[1] Hortobagyi GN, Theriault RL, Lipton A, Porter L, Blayney D, Sinoff C, Wheeler H, Simeone JF, Seaman J, Knight RD, Heffernan M, Mellars K, and Reitsma DJ (1998). Long-term prevention of skeletal complications of metastatic breast cancer with pamidronate. *J. Clin. Oncol.* **16**, 2038–2044.

# TIMELINE DIAGRAMS FOR SELECTED PATIENTS



# COUNTING PROCESS $N(t)$ FOR A SINGLE SUBJECT




## NOTATION

- $\{N(s), 0 < s\}$ is event process where $N(s) = \sum_{k=1}^{\infty} I(T_k \leq s)$
- $H(s) = \{N(u), 0 < u < s\}$ is process history
- $dN(s) = 1$ if event at time $s$; $dN(s) = 0$ otherwise.

## MEAN AND RATE FUNCTION ESTIMATION

- Let $\{N_i(s), 0 < s\}$ be counting process for subject $i$
- $C_i$ is *random* right censoring time and $Y_i(s) = I(s \leq C_i)$
- $Y_i(s) = I(s \leq C_i)$
- Marginal mean and rate functions offer a natural basis for treatment comparisons

$$\mu(t) = E\{N(t)\} \qquad \text{and} \qquad d\mu(t) = \mu'(t)dt$$

## ESTIMATING FUNCTION

$$\sum_{i=1}^{m} I(C_i \geq t)\,\{dN_i(t) - d\mu(t)\} \tag{2.1}$$

$$d\widehat{\mu}(t) = \frac{d\bar{N}.(t)}{Y.(t)} \qquad \text{and} \qquad \widehat{\mu}(t) = \int_0^t d\widehat{\mu}(s)$$

- $\widehat{\mu}(t)$ is the Nelson-Aalen (NA) estimate

## DEPENDENT WITHDRAWAL

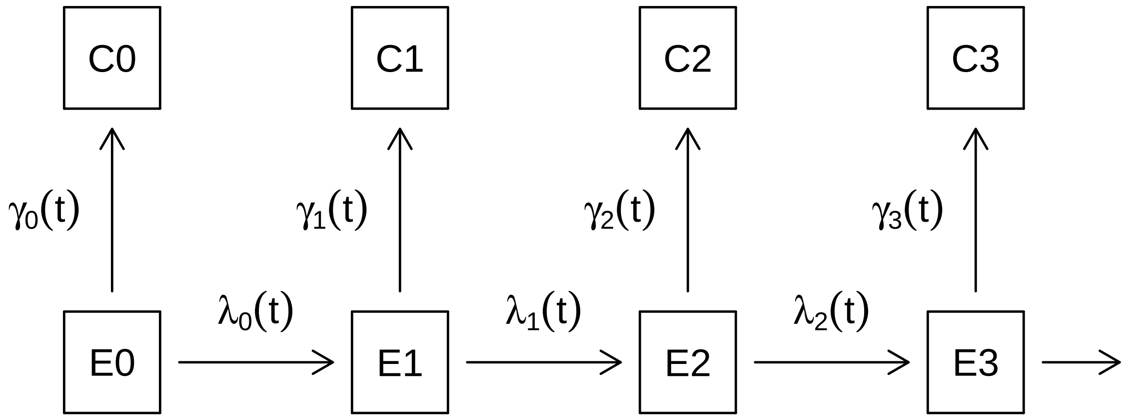
$$\sum_{i=1}^{m} I(C_i \geq t)\,\{dN_i(t) - d\mu(t)\} = 0$$

- Validity of (2.1) requires $C_i \perp \{N_i(s), 0 < s\}$ *so*

$$E\{dN_i(t)|C_i \geq t\} = E\{dN_i(t)\} = d\mu(t)$$

- This means that the decision to withdraw a patient from a trial cannot depend on their past responses (or future!)
- We say that "censoring is completely independent of the event process"
- Is this reasonable in the current study?
- How plausible is this more generally in clinical trials?

## ASSESSING DEPENDENT WITHDRAWAL




- Censoring rates denoted by $\gamma_k(t)$

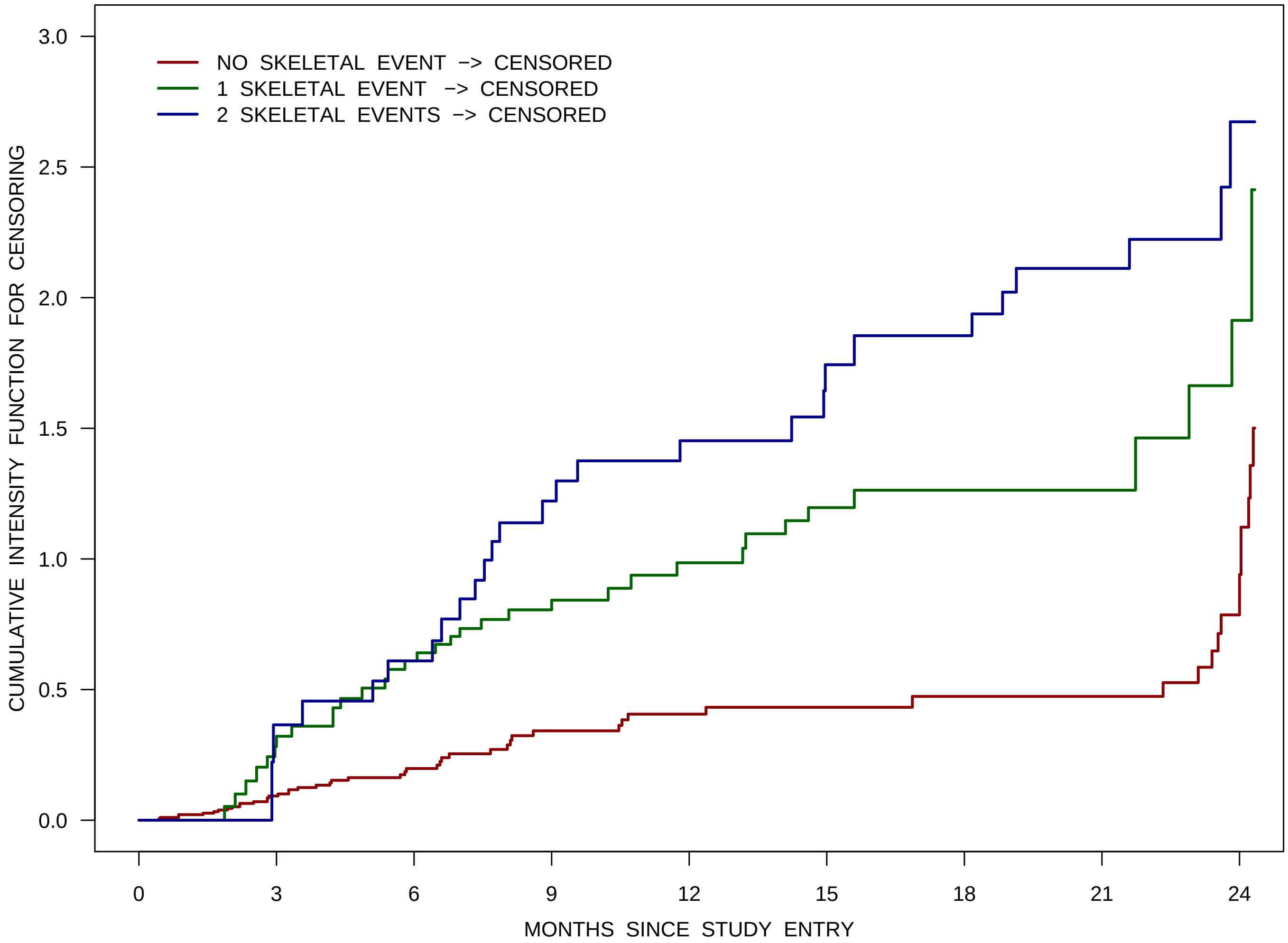
  If $\gamma_k(t) = \gamma(t)$, censoring is completely independent

  – Otherwise, censoring is event-dependent

- Event rates are denoted by $\lambda_k(t)$

  If $\lambda_{k+1}(t) > \lambda_k(t)$ then risk of events increases with each event

# CUMULATIVE CENSORING RATES [PLACEBO]



## HOW TO PROCEED WITH EVENT-DEPENDENT CENSORING?

As in the simple example of Section 1 we have two options.

If we are interested in estimating the expected number of events, we can

A. model the censoring process and adjust (2.1) by the inclusion of

"inverse probability of censoring weights"

B. model the process $\{N_i(s), 0 < s\}$ more fully and then "marginalize" to get $E\{N_i(t)\}$

## A. USING INVERSE PROBABILITY OF CENSORING WEIGHTS (IPCW)

$$\sum_{i=1}^{m} U_i(t) = \sum_{i=1}^{m} \frac{I(C_i \geq t)}{G_i(t)} \{dN_i(t) - d\mu(t)\} = 0 \tag{2.2}$$

- $G_i(t) = \Pr\left(C_i \geq t \middle| H_i(t)\right)$.
- Replace $G_i(t)$ in (2.2) with estimate $\widehat{G}_i(t)$ to give

$$d\widehat{\mu}(t) = \frac{\sum_{i=1}^{m} I(C_i \geq t) dN_i(t) / \widehat{G}_i(t)}{\sum_{i=1}^{m} I(C_i \geq t) / \widehat{G}_i(t)}$$

- $\widehat{\mu}(t) = \int_0^t d\widehat{\mu}(s)$ is the weighted Nelson-Aalen estimate

## A MODEL FOR THE CENSORING PROCESS

If $d\Lambda^c\left(s|H_i(s)\right)$ is the censoring intensity, let

$$G_i(t) == \exp\left\{- \int_0^t d\Lambda^c\left(s|H_i(s)\right) ds\right\} \tag{2.3}$$

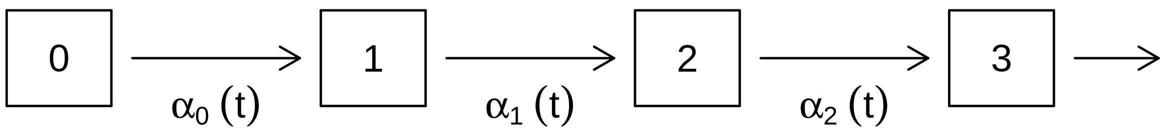
With event-dependent censoring, consider Markov models with

$$d\Lambda^c(t|H_i(t)) \; = \; d\Lambda^c(t|N_i(t^-) = j) = d\Lambda^c_j(t)$$

This means that censoring depends on the cumulative number of events

This is easily estimated using survival analysis software handling time-dependent stratification.

# B. MODELING THE EVENT PROCESS: WORKING MARKOV MODELS




## STEPS IN ESTIMATION

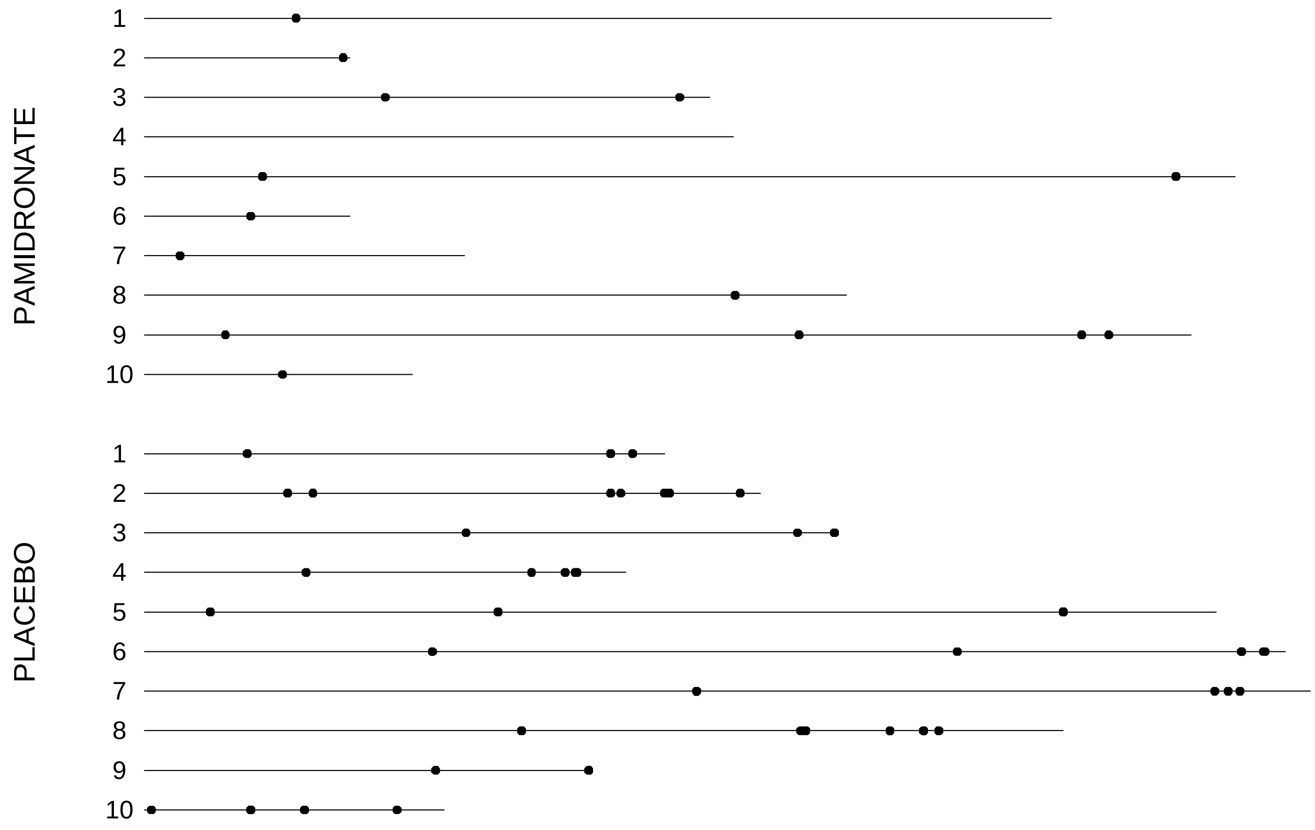
- Estimate "transition intensities" $\alpha_k(u)$
- Compute $P_{jk}(s,t) = P(Y(t) = k | Y(s) = j)$ is the transition probability matrix under Markov model
- Estimates are consistent for $P(0,t)$ in non-Markov models [2] [3] [4]
- We obtain a robust estimate of the mean function based on

$$\widehat{\mu}(t) = \sum_{k=1}^{\infty} k \widehat{P}_{0k}(0,t)$$

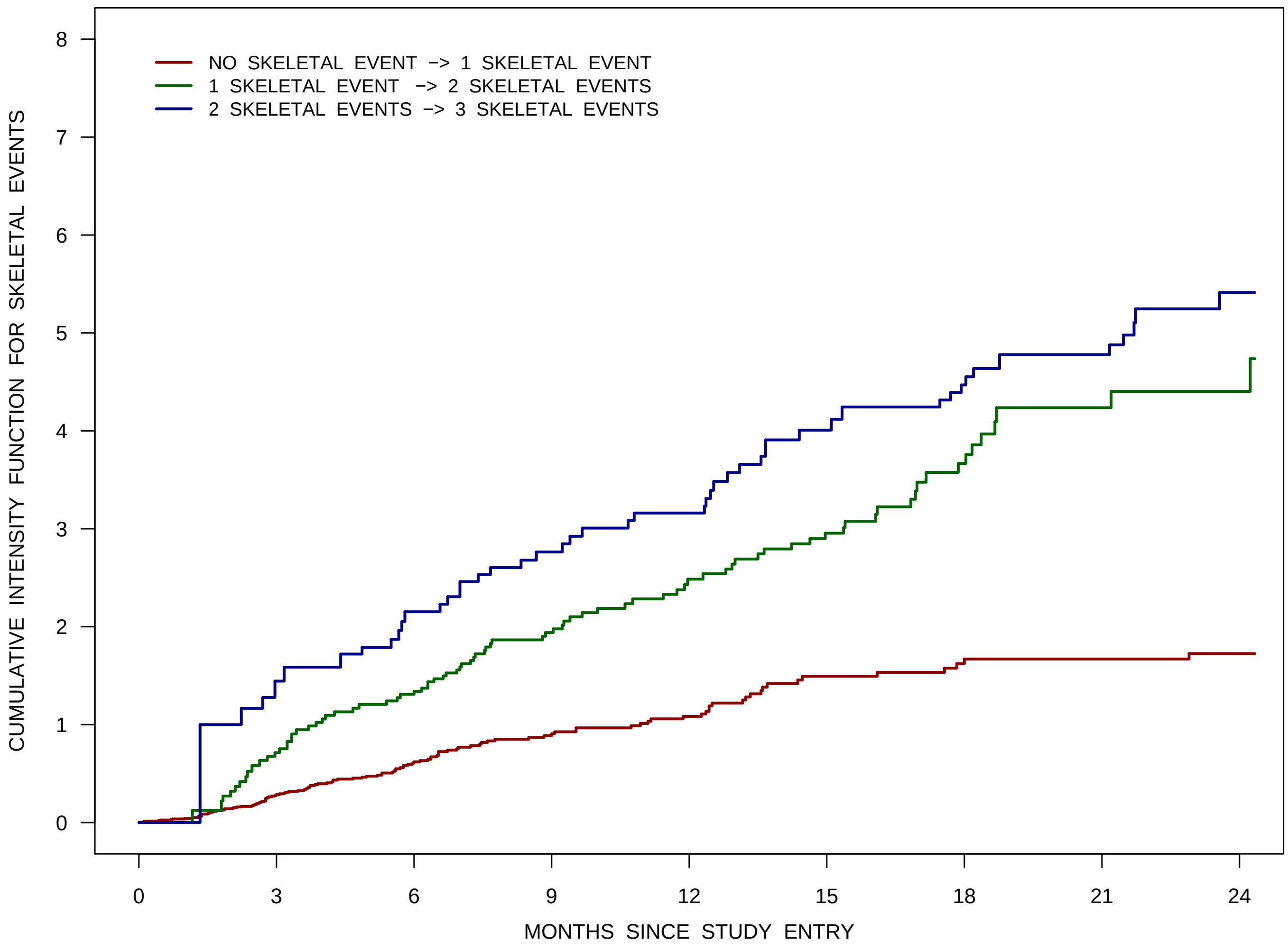
- A *partially conditional* model protects against event-dependent censoring

[2] Aalen et al. (2001). Biometrics
[3] Datta and Satten (2001). Statistics and Probability Letters
[4] Glidden (2002). Biometrics.

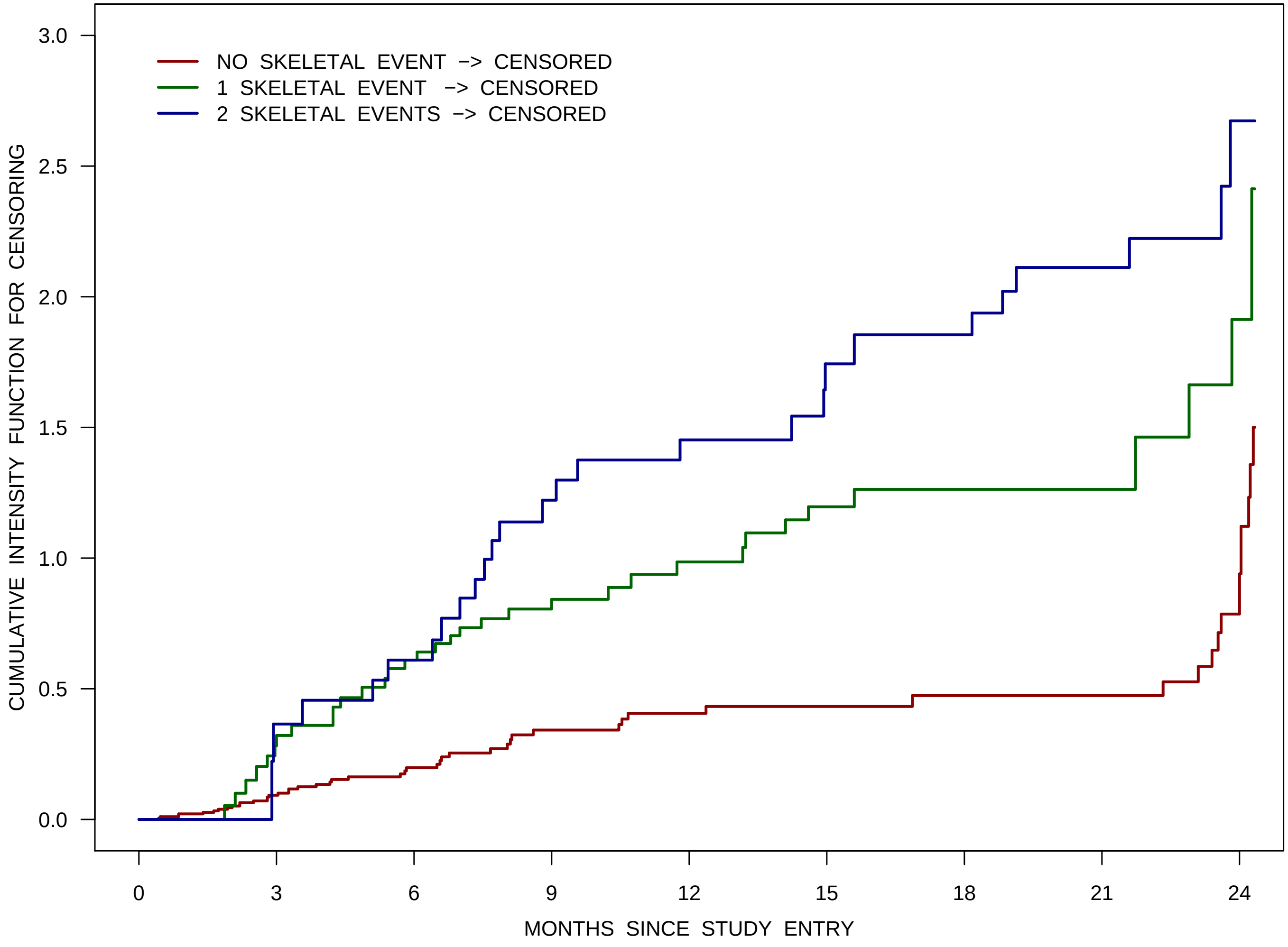
# EVENT PLOTS



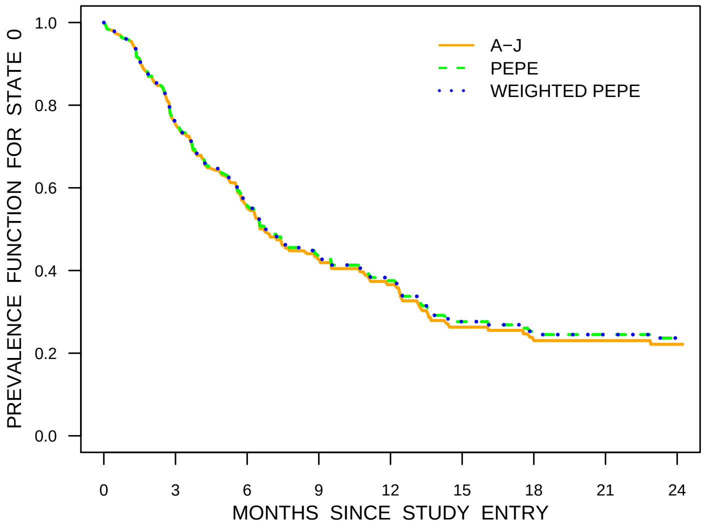
# CUMULATIVE EVENT INTENSITIES [PLACEBO]

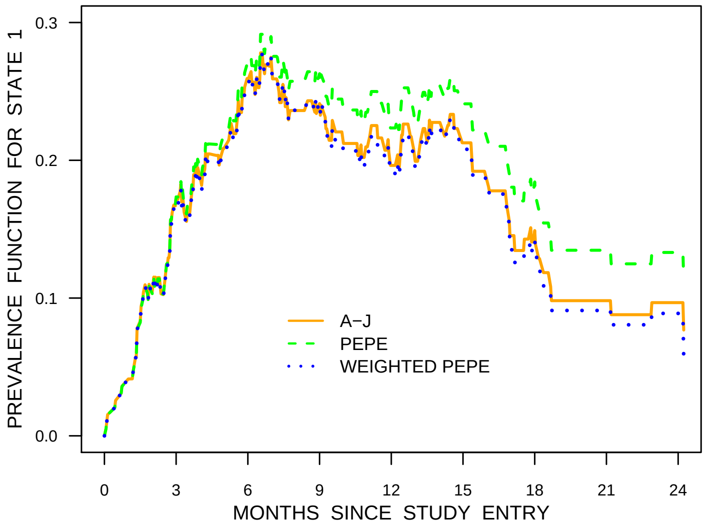


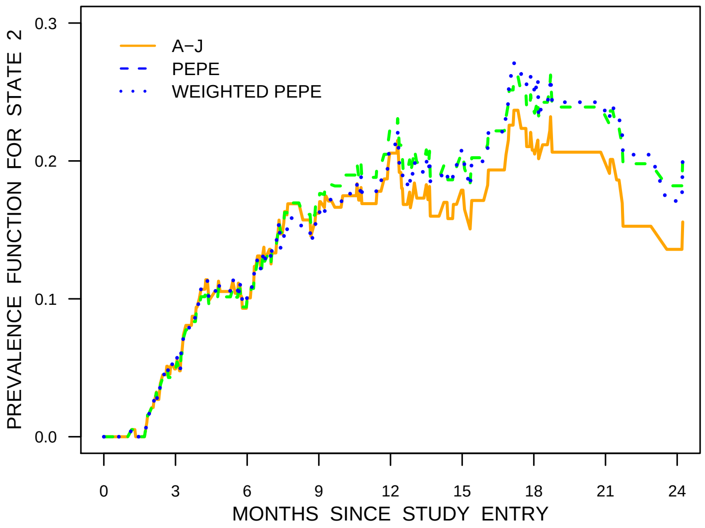
# CUMULATIVE INTENSITIES FOR CENSORING [PLACEBO]



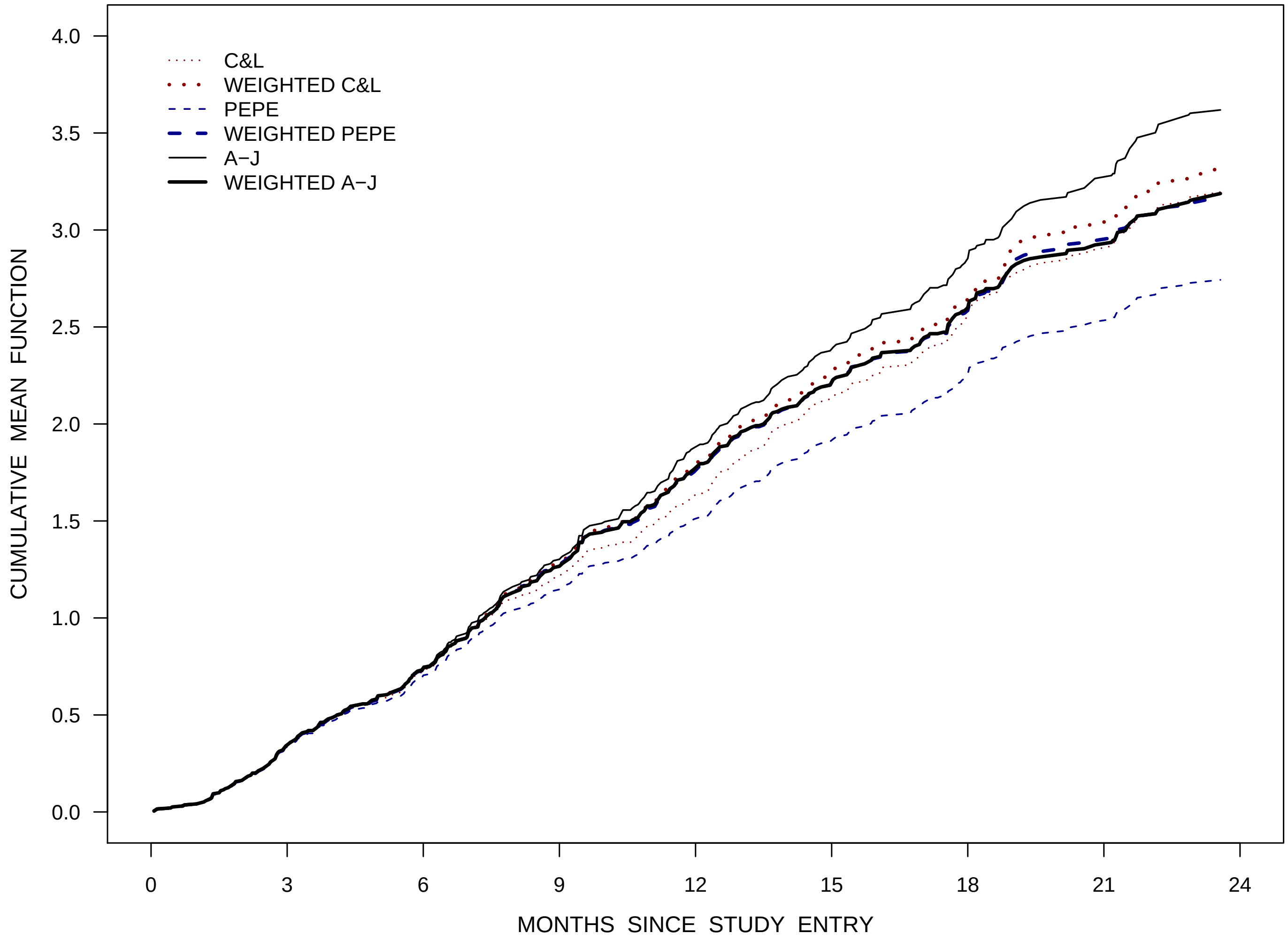
# ESTIMATES OF STATE OCCUPANCY PROBABILITIES [PLACEBO]



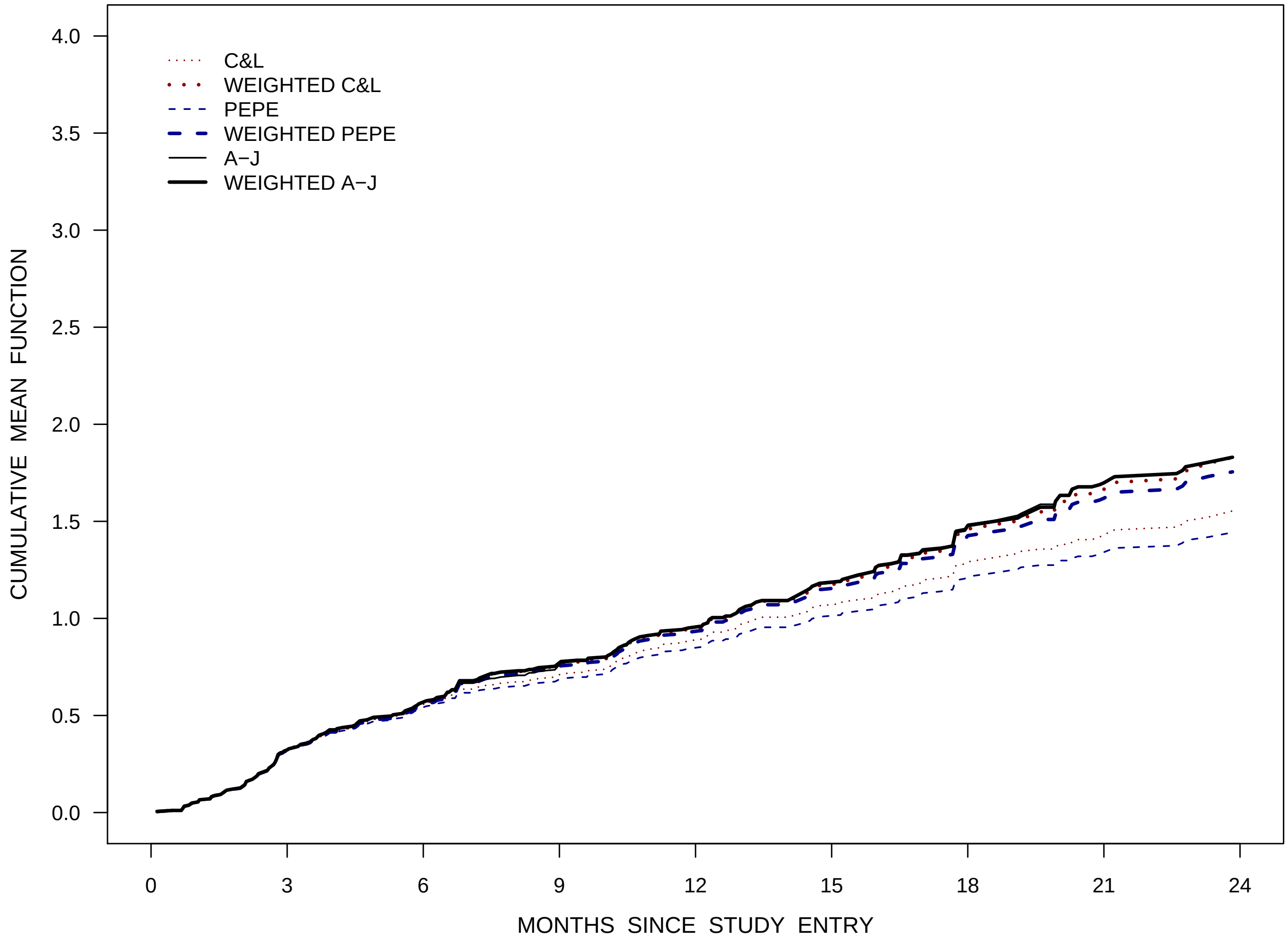



# MEAN FUNCTION ESTIMATES [PLACEBO]



# MEAN FUNCTION ESTIMATES [PAMIDRONATE]



## ASSESSING THE TREATMENT EFFECT

- The methods of inverse weighting we've discussed this far can be adapted for regression analyses
- An unweighted analysis is carried out on the data from Hortobagyi et al. (1996) we obtain

  $\widehat{\beta} = -0.617$, s.e.($\widehat{\beta}$)= 0.095

  RR$= \exp(-0.617) = 0.540$, p$<$0.0001

- A weighted analysis gives a slightly smaller estimate

  $\widehat{\beta} = -0.584$, s.e.($\widehat{\beta}$)= 0.182

  RR$= \exp(-0.584) = 0.558$, p=0.0013

## AN INTERESTING RESPIRATORY TRIAL

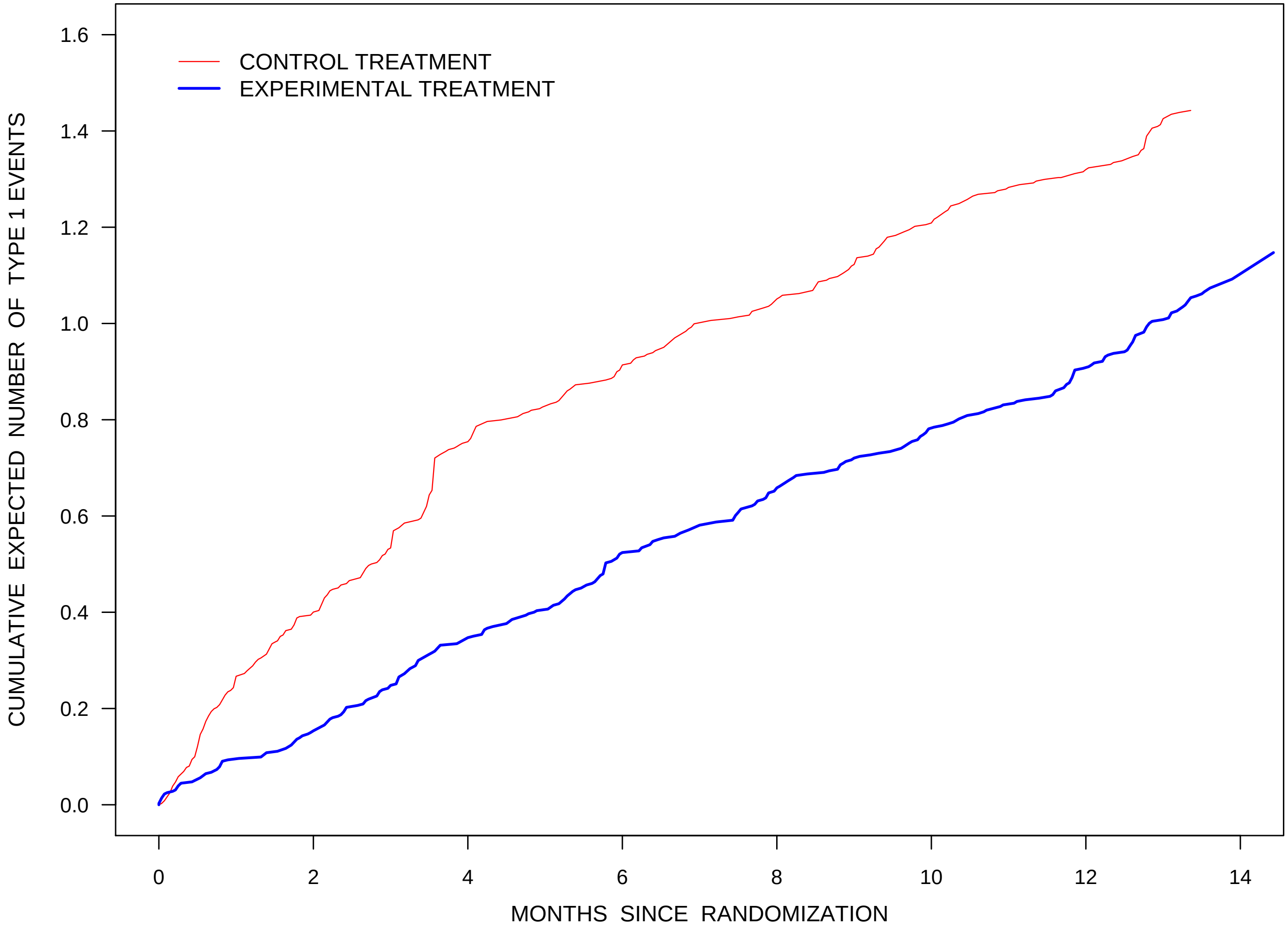
- multicenter international randomized trial of patients with COPD

- 358 patients randomized to experimental treatment
- 361 patients randomized to control

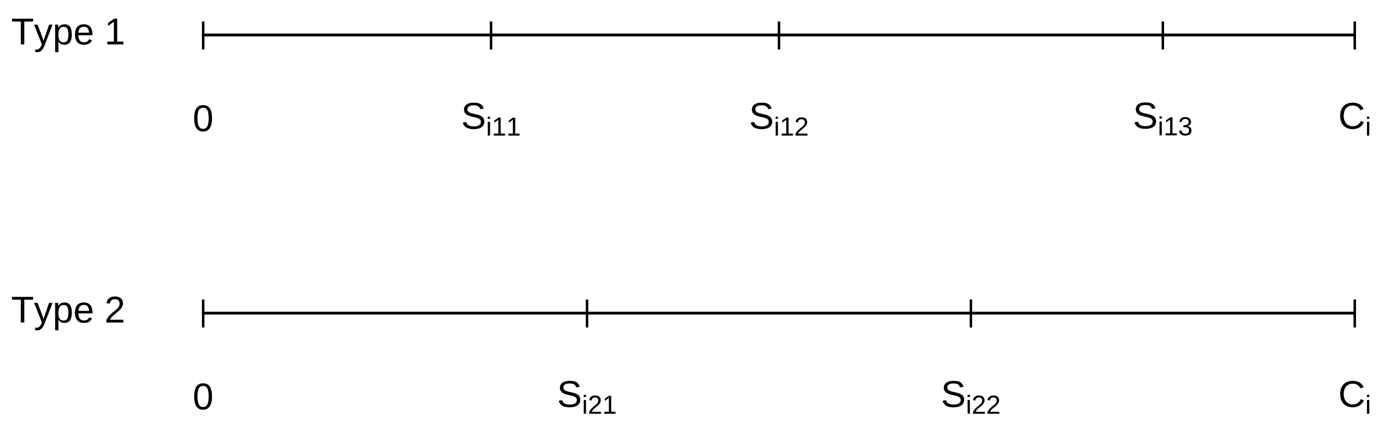
- Follow-up scheduled for 12 months

- Recurrent events (exacerbations) were recorded as secondary endpoints and classified by type
  - **TYPE 1:** moderately serious
  - **TYPE 2:** serious/very serious

# EXPECTED NUMBER OF TYPE 1 EVENTS



# BIVARIATE RECURRENT EVENT DATA

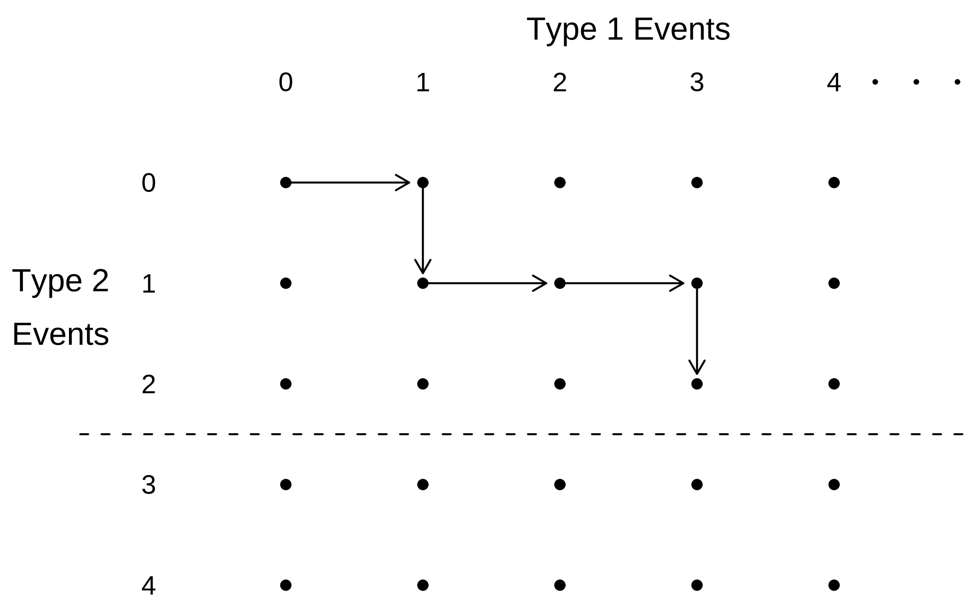



- $\{N_{ij}(s), 0 \leq s\}$ records events of type $j$ experienced by individual $i$
- $dN_{ij}(s) = 1$ if a type $j$ event occurs at time $s$ ; $dN_{ij}(s) = 0$ otherwise
- bivariate counting process is $\{N_i(s), 0 \leq s\}$ where $N_i(s) = (N_{i1}(s), N_{i2}(s))'$
- $C_i$ is right censoring time

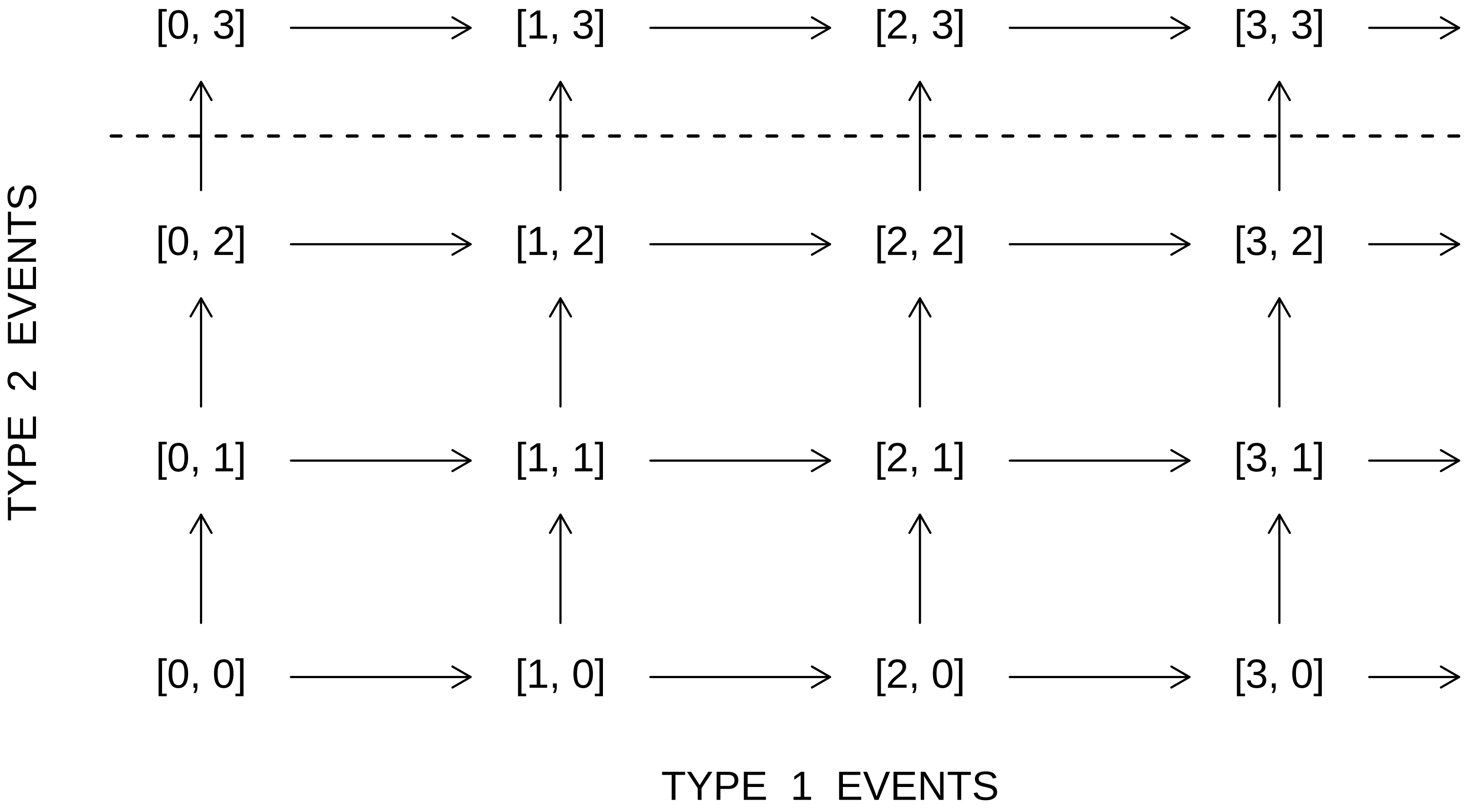
## RESPIRATORY TRIAL FEATURES DEPENDENT CENSORING

- Trial involves withdrawal of patients from trial when they have had **two** type 2 events
- What is the impact on this analysis of type 1 events?

**Marginal analysis is invalid if event types are associated!**



# MULTISTATE ANALYSIS



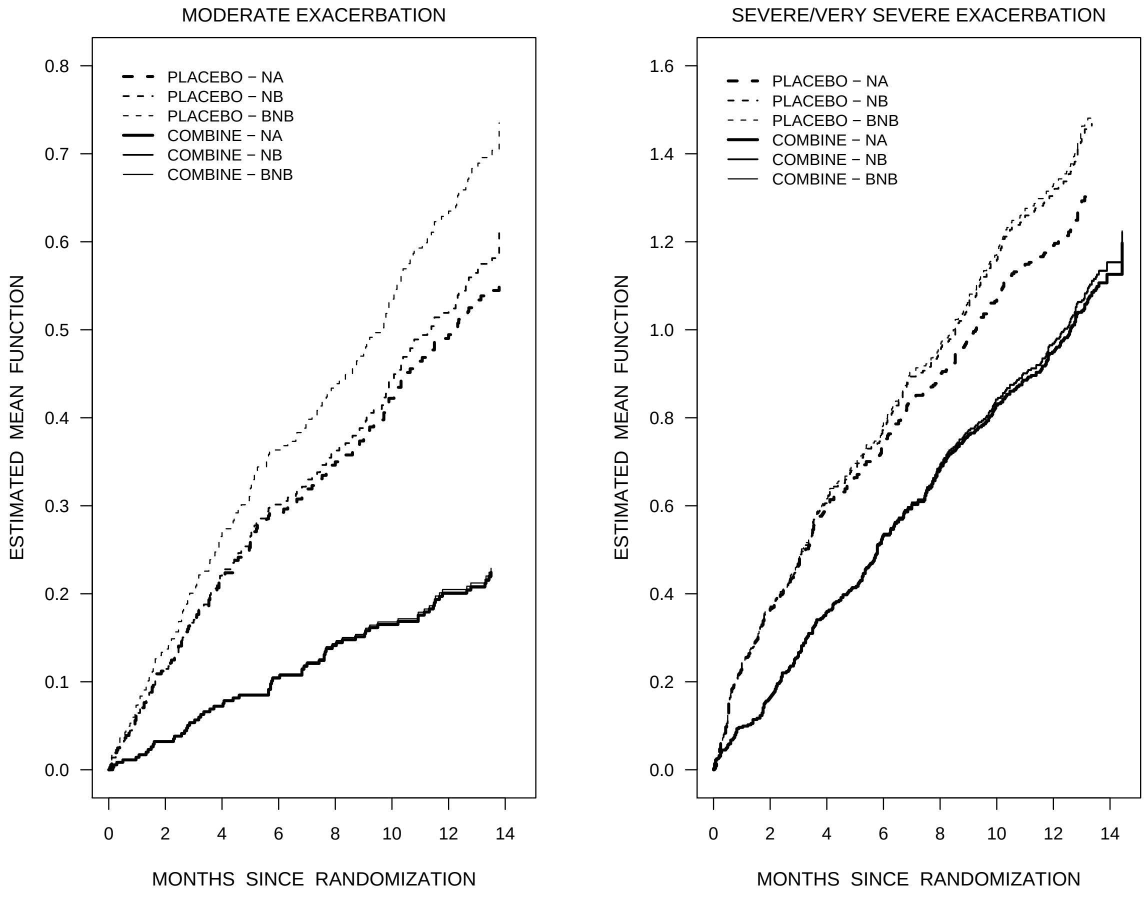
## WHAT CAN BE ESTIMATED HERE?

- Let $P_{r_1 r_2}(t) = P(N_1(t) = r_1, N_2(t) = n_2)$
- Then
$$\mu_1(t) = \sum_{r_1=0}^{\infty} \sum_{r_2=0}^{\infty} r_1 P_{r_1 r_2}(t) = \sum_{r_1=0}^{\infty} r_1 P(N_1(t) = r_1) \ .$$
- We need to estimate joint probabilities $P_{r_1,r_2}(t)$ consistently, but this is is inestimable nonparametrically for $r_2 > k$ so $\mu_1(t)$ is nonparametrically inestimable.
- We can nonparametrically estimate

$$P(N_1(t) = r_1, N_2(t) = K)$$
$$P(N_1(t) = r_1 | N_2(t) \leq K)$$
$$E\{N_1(t) | N_2(t) \leq K\}$$

- More fully specified models which characterize the event process are useful.

# ESTIMATING EXPECTED NO. EXACERBATIONS WITH DEPENDENT CENSORING



## GENERAL REMARKS

- We have discussed
  - missing data where the dependence is on a baseline covariate (part I)
  - event-dependent censoring with recurrent event analyses (part II)
- Dependent censoring can also arise when people drop out of a study for reasons related to a response
- In survival analysis (not recurrent event analysis), dependent censoring can have a significant impact
  - In this case, dependence is induced by related time-varying covariates we do not wish to control for
- Can also arise in multistate analyses for more complex disease processes

## TAKE HOME MESSAGES

- Be aware of possible effects of an association between the withdrawal/censoring process in trials and the responses of interest
- Although this was not discussed, similar issues arise in observational studies
- The issues are similar whether dealing with drop-out in longitudinal studies with regularly scheduled assessments or study withdrawal when time to event analyses are planned (although the models for dealing with them are different)
- Survival analysis techniques can be used to assess whether it is cause for concern in a particular study
- Estimates of marginal features like proportions responding, or the probability of surviving 1 year, are typically more affected than estimates of treatment effects

- Inverse probability of censoring weighted approaches can be used to address these concerns
- The “price” is the need to model the censoring process, which is often not of direct interest