Haben wir schon längst aus den Augen verloren, was wir in der Überlebenszeitanalyse schätzen sollten?

Jan Beyersmann

Institute of Statistics, Ulm University

- Estimands: 'intercurrent' or **post-randomization** events
- Time-to-event: censor post-randomization events?
- Examples:
  - MACE (and censor 'other death'.)
  - Leukemia: censor allogeneic stem cell transplantation.
  - 'Censoring rules' in RCTs
  - (Clinical Hold)
- Censoring independent, random, non-informative?
- Censoring and causality?

Nissen et al.: Muraglitazar and Adverse Events, JAMA 2005, (antidiabetic drug, not unlike Rosiglitazone)

- Authors use **incidence proportions** (IPs), but discuss that survival analysis would have been more powerful.
- That's not the point: IPs estimate $P(\textbf{observed } \mathrm{AE}) < P(\mathrm{AE})$

Editorial by Brophy: Selling Safety — Lessons from Muraglitazar, JAMA (2005)

- Editor uses **incidence densities** (IDs).
- IP and ID the two major workhorses in safety analyses, differing in the denominator only, but
  - IPs probabilities, but do not account for censoring.
  - IDs account for censoring, but are no probabilities: $33.37 > 1$.
  - (IDs estimate hazards under a constant hazard assumption — a restrictive assumption, but not a 'dubious concept'.)

Why hazards?

- Outcome: Jan's death
- Observation process: the audience is looking at me.
- I'm **at risk**: alive and under observation.
- Independent censoring: your presence
  - does not scare me too much (which might increase my hazard)
  - does not please me too much (which might decrease my hazard)
- If I die right now,
  - it'll happen with the same hazard as without you looking,
  - you'll observe it.
- So you can **estimate** my hazard (based on 100 Jans. . .)

This has little to do with me dying. . .

- Outcome: Jan breaks his right arm, too. (A possibly recurrent event in the presence of competing risk 'death'.)
- Observation process: the audience is looking at me.
- I'm at risk: alive **with right arm** and under observation.
- Independent censoring: your presence
  - does not scare me too much (which might **change** my arm-breaking hazard)
  - does not please me too much (which might **change** my arm-breaking hazard)
- If I break my right arm right now,
  - it'll happen with the same hazard as without you looking,
  - you'll observe it.
- So you can **estimate** it.
- Martingales: Right-arm-breaking counting process retains intensity, if we additionally condition on the censoring process.

Major AE-workhorses IP and ID: simplified situation

- No censoring. (For the time being.)
- (Time to 1st) AE or 'death' whatever comes first.
- 'Death' some event that precludes AE occurrence.
- IP
$$\frac{\#\text{ AE Patients}}{n} = \frac{\#\text{ AE}}{n} \to P(\text{AE})$$
or even
$$\frac{\#\text{ AE Patients in } [0,t]}{n} = \frac{\#\text{ AE in } [0,t]}{n} \to P(\text{AE in } [0,t])$$
- ID is $\#$ AE/Population time at risk and
$$\frac{\#\text{AE/Population time at risk}}{(\#\text{AE} + \#\text{Deaths before AE})/\text{Population time at risk}} = \frac{\#\text{AE}}{n} = \text{IP}$$
- Competing risks in a nutshell with **competing hazard**
$$\#\text{ Deaths before AE/Population time at risk}$$

Independent vs. informative censoring

- Incidence proportion IP
$$\text{IP} = \frac{\#\text{AE}}{n} = \frac{\#\text{AE/Population time at risk}}{(\#\text{AE} + \#\text{Deaths before AE})/\text{Population time at risk}}$$
- Incidence densities IDs
$$\frac{\#\text{ AE}}{\text{Population time at risk}} \quad \text{and} \quad \frac{\#\text{ Deaths before AE}}{\text{Population time at risk}}$$
- AE incidence density censors observed Deaths before AE:
  - **independent censoring**: allows for estimating AE hazard (under a constant hazard assumption)
  - **informative censoring**: does not allow for probability statements without competing Deaths before AE incidence density
- **Random censoring**: T and C independent. Does not hold in common pharmaceutical RCTs.

Event-driven trials

- Toy example: 2 patients put on trial at the same time, stop after 1 observed event.
- The data are **not independent** (censoring not random):
$$T_1 \wedge T_2, \quad \mathbf{1}(T_1 \leq T_2)$$
$$T_1 \wedge T_2, \quad \mathbf{1}(T_2 \leq T_1)$$
- General counting process & martingale machinery copes with, e.g., event-driven trials:
  - **Independent censoring** does not disturb the intensity of a counting process. (A probabilistic concept.)
  - Whether or not this is **informative**, depends on the target parameter. (A statistical concept.)
- Welcome to Babylon: The literature is a mess this side of counting processes. . .

Babylonian confusion on censoring

- It is not uncommon to use 'independent censoring' and 'non-informative censoring' interchangeably, e.g., Collet, *Modelling Survival Data in Medical Research*, CRC 2015, p3.
- E.g., O'Quigley, *Proportional Hazards Regression*, Springer, 2008, p122:

  [. . .] the assumption of independent censoring, sometimes referred to as non informative censoring, [. . .]

  $$P(T_i > x, C_i > x) = P(T_i > x)P(C_i > x)$$

- But this is random censoring!
- This torpedoes competing risks methodology (needed for MACE, AEs, . . .)

E.g., more from O'Quigley, *Proportional Hazards Regression*, Springer, 2008, p124, f, on censoring by a competing risk

We will need make some assumptions, most often that of independent censoring [of the competing risk] [. . .] in order to make progress

- In the counting process world, independent censoring by a competing risk is not an assumption, but a theorem.
- But here it says: We will need to assume that, e.g., time-to-CV-death and time-to-non-CV-death are independent.
- Are they?
- And what is time-to-CV-death if one dies from other causes?
- Not the good concept of 'independent censoring', and time-to-CV-death not a well defined random variable.

Two scans from Kleinbaum and Klein, *Survival Analysis — A Self-Learning Text*, Springer 2012

- Kleinbaum and Klein have yet another suggestion: Independent censoring is random censoring within strata.
- And censoring is non-informative when T provides no information on C. (Somewhat circular. And what is T with MACE or AEs or some other post-randomization event other than all-causes death?)

And now some more circular ‘definitions’. . .

From Ibrahim et al., *Bayesian Survival Analysis*, Springer 2001, p15

Throughout the book, we will also assume that censoring is noninformative in the sense that inferences do not depend on the censoring process.

- So, the assumption is that censoring is such that inference based on censored data is o.k..

From Moore, *Applied Survival Analysis Using R*, Springer 2016, p3

[. . .] one cause of random censoring is patient dropout. If the dropout occurs truly at random [. . .]

- So, censoring because of dropout is random censoring provided that dropout is random.

From Hosmer et al., *Applied Survival Analysis*, Wiley, 2008, p6

Incomplete observation of a survival time due to the end of the study or follow-up is considered a right censored observation because the process by which subjects entered the study is random at the subject level.

- With bold face:

  *Incomplete observation of a survival time due to the end of the study or follow-up is considered a right **censored** observation **because** the process by which subjects entered the study is **random** at the subject level.*

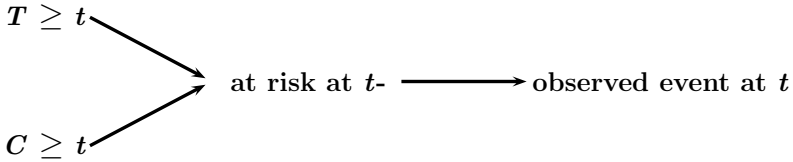
- Following this argument, a censored observation is randomly censored by definition.

A first summary

- Survival and event history analysis
  - based on hazards,
  - does complex event patterns and complex incomplete data mechanisms,
  - is way beyond survival functions and Kaplan-Meier.
- Censoring post-randomization events
  - independent: valid hazard inference.
  - informative: requires competing risks methodology for probability inference.
- Counting process machinery makes this rigorous.
- Alas, the literature this side of counting processes is a nightmare...
- We need to sort this out before moving forward in the estimand debate.
- Now, what about censoring and causality?

Random censoring and causal interpretation of KM

- A causal directed acyclic graph:



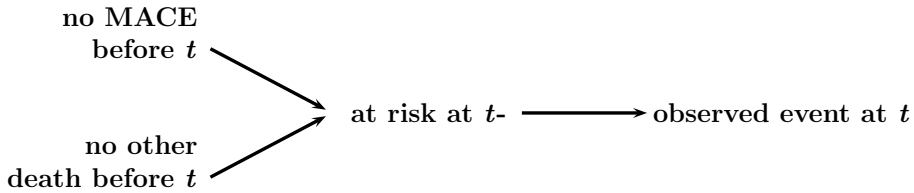
- Estimand $P(T > t \,|\, \text{do(no censoring)})$ — the survival function of (potential) outcomes in a world w/o censoring.
- Well known for Kaplan-Meier (e.g., Gill 1980)

$$\#\text{ observed survivors} > t = \text{KM}(t) \cdot \text{Censoring-KM}(t)$$

- Simple example of
  - **I**nverse **P**robability of **C**ensoring **W**eighting (or propensity score for 'censoring treatment')
  - g-computation aka **truncated** product rule

A common misinterpretation of KM

- This



not the proper causal graph for a post-randomization event like 'other death': there are common causal 'parents'.

- You must not use Kaplan-Meier censoring the other post-randomization event - still a common mistake (Schumacher et al., J Clin Epi 2016).
- Need more complex DAG/multistate model/g-computation (e.g., Aalen et al, 2008).

Some references

- Aalen, Borgan, Gjessing: *Survival and Event History Analysis*, Springer 2008
- Andersen: Censored data. In *Encyclopedia of Biostatistics*, Wiley, 2005
- Allignol, Beyersmann, Schmoor (2016) Statistical issues in the analysis of adverse events in time-to-event data. *Pharm Stat*, 15(4):297–305.
- Unkel et al. (2018): On estimands and the analysis of adverse events in the presence of varying follow-up times within the benefit assessment of therapies. *Submitted.*