UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

SAFETY ANALYSES: THE CINDERELLA OF BIOSTATISTICS?

Academic perspective

EFSPI Statistics Leaders Meeting, 23 July 2020

Tim Friede

Department of Medical Statistics

University Medical Center Göttingen, Germany

ACKNOWLEDGEMENTS

This presentation contains results from the following two working groups

- **Joint ATF/APF Project Group** "Analysis of adverse events at varying follow-up times in the context of benefit assessments" of GMDS and IBS-DR
  - Resulted in the following manuscript: Unkel et al (2019) Pharm Stat 18: 166–183.
- The "**Survival analysis for AdVerse events with VarYing follow-up times**" (SAVVY) project

HOW **NOT** TO ESTIMATE ADVERSE EVENT PROBABILITIES

- **Incidence proportion** (# patients with AE within time t / n)
  - Useful with identical follow-up times
  - Underestimates AE probability in the presence of censoring with varying follow-up times
- **1 - Kaplan-Meier** (censoring competing events)
  - Overestimates AE probability
  - 1-KM approximates a distribution function, i.e. assuming that eventually all patients experience the adverse event
- **Reference**: Unkel et al (2019), Section 4.1

HOW TO ESTIMATE ADVERSE EVENT PROBABILITIES

UNIVERSITÄTSMEDIZIN GÖTTINGEN : UMG

- **Aalen-Johansen estimator** (cumulative incidence function)

$$\widehat{\boldsymbol{P}}(\boldsymbol{T} \leq \boldsymbol{t}, \boldsymbol{AE}) = \sum_{\boldsymbol{u} \leq \boldsymbol{t}} \widehat{\boldsymbol{P}}(\boldsymbol{T} > \boldsymbol{u} -) \frac{\#\ \text{AE at u}}{\#\ \text{at risk,}\ \ \text{no AE prior to u}}$$

  - generalizes the KM estimator to multiple event types
  - computation straightforward
- **Nelson-Aalen estimator of the cumulative hazard**
  - cumulative nonparametric counterpart of commonly used incidence rate
  - As incident rate, does not estimate a probability (probability transform would need to be used, see below)
- **Reference**: Unkel et al (2019), Section 4.1

ILLUSTRATIVE (TOY) EXAMPLE

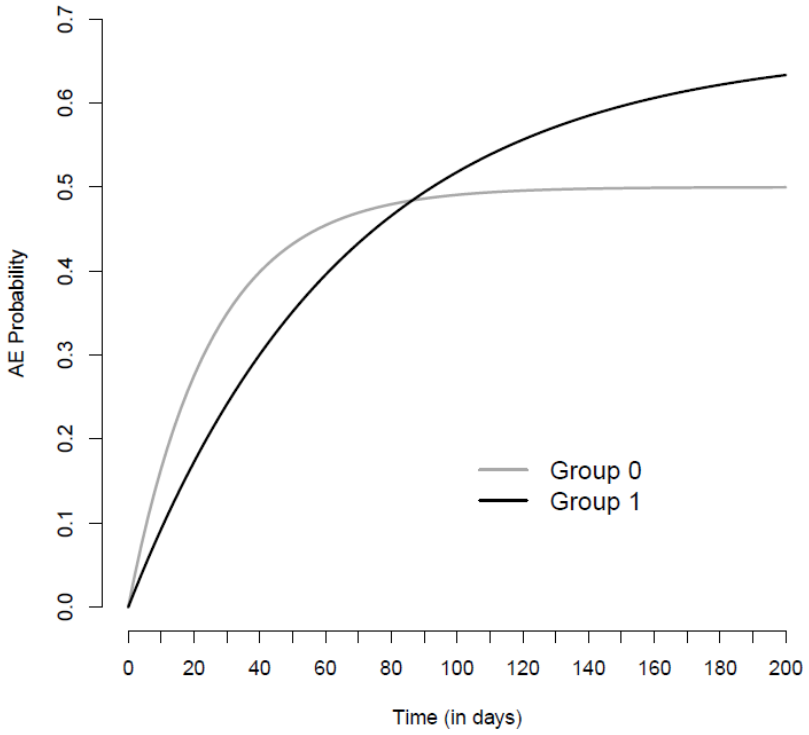


Figure 3 from Unkel et al (2019)

- **In group 0**, AE and competing event hazard rates set to 0.02 events per day, eventually leading to an AE probability of 1/2
- **In group 1**, the AE and competing event hazards reduced by factor of 0.5 and 0.25, respectively
- Although AE hazard in group 1 lower compared to group 0, the cumulative AE probability in group 1 is eventually greater than in group 0
- **Conclusion: need to model all events**

COMPARING TREATMENT GROUPS

- **Risk difference, relative risk or odds ratio of incidence proportion potentially misleading**
  - comparing two quantities that both underestimate the probability of interest
- **Alternatives** include
  - Cox proportional (event-specific) hazards regression
  - Fine & Gray proportional subdistribution hazards model
- **Recommendation**
  - Model not only the AE, but also the competing event
- **Reference**: Unkel et al (2019), Section 4.2

SURVIVAL ANALYSIS FOR ADVERSE EVENTS WITH VARYING FOLLOW-UP TIMES (SAVVY)

Academic leads

- Jan Beyersmann (Ulm)
- Claudia Schmoor (Freiburg)
- Tim Friede (Göttingen)



All the hard work is done by

- Regina Stegherr (Ulm)

SAVVY: THE PROJECT GROUP

- **Steering Committee**
  - To ensure project runs smoothly and on target; to develop strategy for future activities
  - Members: Jan Beyersmann, Claudia Schmoor, Tim Friede, Valentine Jehl (Novartis), Friedhelm Leverkus (Pfizer), Kaspar Rufibach (Roche)
- **Participating organizations**
  - Providing data for empirical study; engaging in discussions on design and analysis of empirical study
  - Bayer, Boehringer Ingelheim, BMS, Janssen, Lilly, Merck, Novartis, Pfizer, Roche, University Medical Center Freiburg

STUDY CHARACTERISTICS

- 10 participating organizations contributing 17 randomized controlled trials including 186 adverse events (AEs)

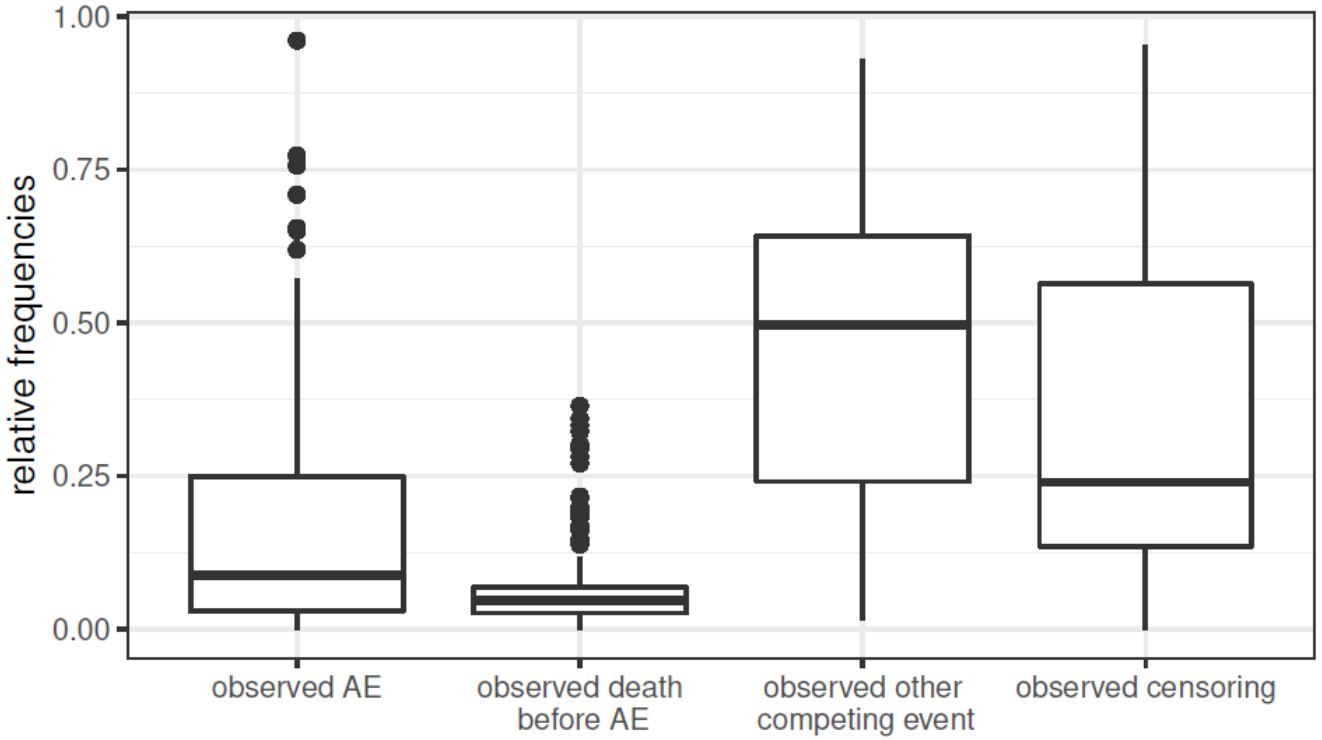
	Frequency (%)
AEs per study: median (Q1, Q3)	7.5 (3, 19)
Type of control: Placebo	8 (47)
Disease area: Oncology	12 (71)

ADVERSE EVENT CHARACTERISTICS

- Out of a total of n=186 adverse events

	Frequency (%)
Serious adverse event	12 (7)
Proportion of censored obs.: median (Q1, Q3)	0.18 (0.12, 0.55)
Frequency category	
Very rare	6 (3)
Rare	0 (0)
Uncommon	6 (3)
Common	86 (46)
Very common	88 (47)
Higher AE probability in experimental group	138 (74)

FREQUENCIES OF EVENTS



Stegherr et al (2020a) in preparation

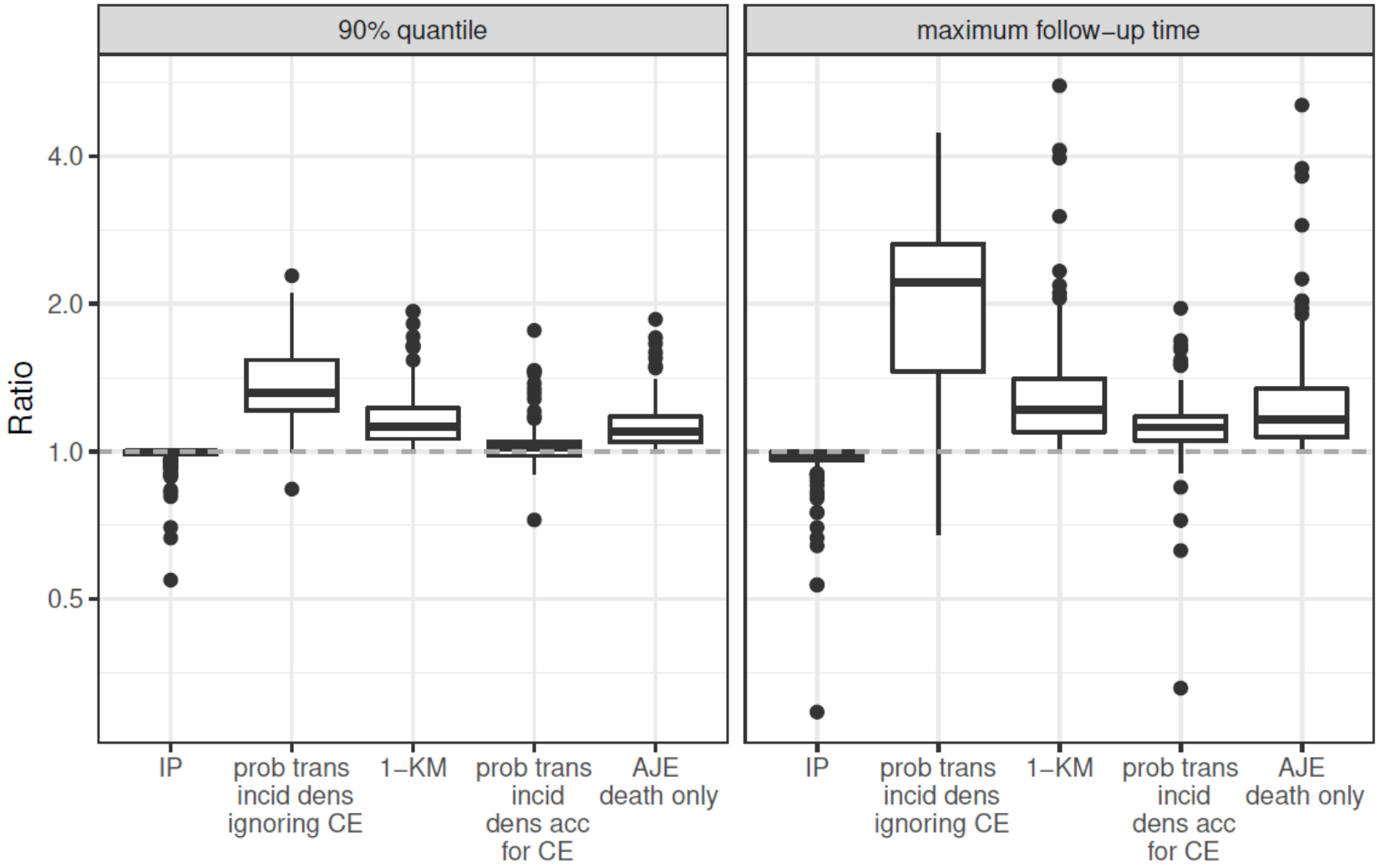
ESTIMATING AE PROBABILITIES

Table 1. Overview whether the estimators deal with the possible sources of bias.

	Accounts for censoring	Makes no constant hazard assumption	Accounts for CEs
Incidence proportion	No	Yes	Yes
Probability transform incidence density ignoring CEs	Yes	No (AE Hazard)	No
1-Kaplan-Meier	Yes	Yes	No
Probability transform incidence density accounting for CEs	Yes	No (AE and CE Hazard)	Yes
death only Aalen-Johansen estimator	Yes	Yes	Yes (Death only)
gold-standard Aalen-Johansen estimator	Yes	Yes	Yes

Stegherr et al (2020a) in preparation

BIAS IN ESTIMATING AE PROBABILITIES



Stegherr et al (2020a) in preparation

CONCLUSIONS FROM EMPIRICAL STUDY

AE probability

- Choice of estimator crucial
- Incidence Proportion similar to Aalen-Johansen estimator with ‘all events’ definition of competing events (resulting here in low censoring); differences in studies with substantial (especially late) censoring
- Kaplan-Meier estimator not appropriate as it censors competing events
- Ignoring competing events is more of a problem than falsely assuming constant hazards
- Death-only definition of competing events censors other competing events and therefore results in overestimation of AE probability

CONCLUSIONS FROM EMPIRICAL STUDY

Between-group comparisons (results not shown)

- Choice of estimator also crucial for group comparisons
- In most cases the results of the group comparisons at no tau and at max are equal, but be careful with situations characterized by many late AEs in one group
- Hazard ratios for the AE hazard from Cox analyses: Need to model additionally competing event hazards

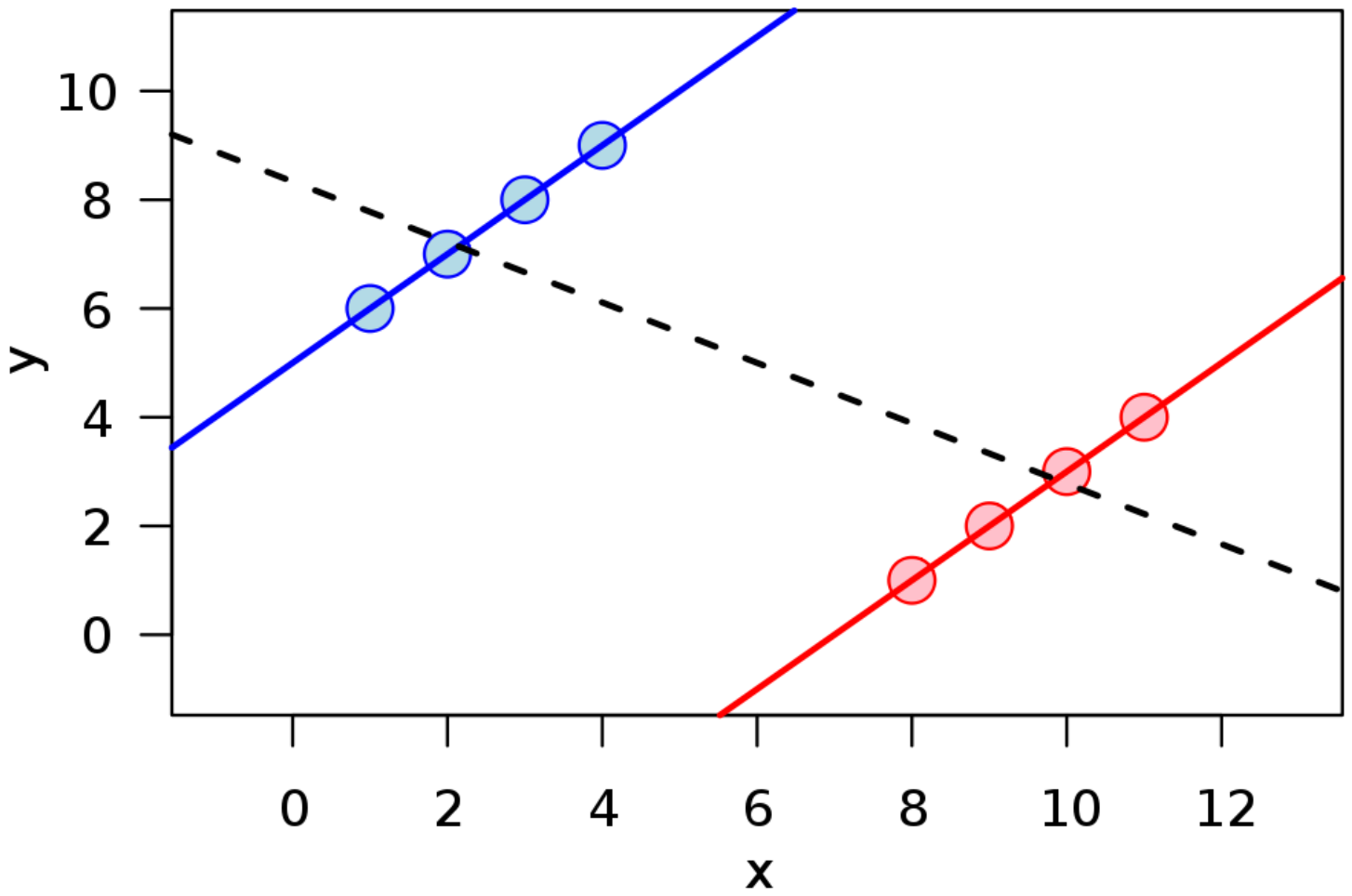
DATA MONITORING COMMITTEES

- **Monitoring of adverse events**
  - Even in studies with fixed follow-up per patient, follow-up times vary at interim (e.g. safety review by DMC)
  - Incidence proportions commonly used in this setting, although inappropriate
- **Presentation of adverse events**
  - (Very) long tables and listings the (unfortunate) standard
  - Visual displays accounting appropriately for competing events (e.g. cumulative incidence functions) recommended
- In summary, there is some **space for improvement** …

DO NOT POOL NAIVELY, BUT USE META-ANALYSIS TECHNIQUES

- **Not uncommon to naïvely pool data across studies for AE analyses**, e.g. by “simply combin[ing] the numerator events and the denominators for the selected studies.” (FDA, 2005)
- Results might be biased due to Simpsons’s paradox (McEntegart, 2000; Rücker and Schumacher, 2008; Chuang-Stein and Beltangady, 2011)
- ICH E9 states “any statistical procedures used to combine data across trials should be described in detail” and that “attention should be paid [...] to the proper modelling of the various sources of variation.”

SIMPSON‘S PARADOX



http://en.wikipedia.org/wiki/Simpson%27s_paradox

DO NOT POOL NAIVELY, BUT USE META-ANALYSIS TECHNIQUES

- Use of **meta-analysis techniques** encouraged because
  - **variation in baseline** (control group) outcomes across the various studies
  - Random-effects meta-analysis in addition allows **for variation in treatment effects** across studies (so-called between-trial heterogeneity)
- In the context of **safety analyses**, a number of **specific problems** arise (see, e.g. Berlin et al, 2013) including
  - Considerably varying follow-up time
  - Small number of studies included in meta-analysis
- **Reference**: Unkel et al (2019), Section 4.4

SOME REFERENCES

- Unkel S, Amiri M, Benda N, Beyersmann J, Knoerzer D, Kupas K, Langer F, Leverkus F, Loos A, Ose C, Proctor T, Schmoor C, Schwenke C, Skipka G, Unnebrink K, Voss F, Friede T (2019) On estimands and the analysis of adverse events in the presence of varying follow-up times within the benefit assessment of therapies. Pharmaceutical Statistics 18: 166–183.
- Allignol A, Beyersmann J, Schmoor C (2016) Statistical issues in the analysis of adverse events in time-to-event data. Pharmaceutical Statistics 15: 297–305.
- Beyersmann J, Allignol A, Schumacher M (2011) Competing Risks and Multistate Models with R. Springer: New York.
- SAVVY Statistical analysis plan: https://arxiv.org/abs/1912.00263