ulm university universität ulm

Rationale and first results from the SAVVY project

Regina Stegherr[1], on behalf of the SAVVY Study Group

[1]Institute of Statistics, Ulm University, Germany

27.08.2019

regina.stegherr@uni-ulm.de

Thanks to...

The SAVVY project group

- Academic leads: Jan Beyersmann (Ulm), Tim Friede (Göttingen) and Claudia Schmoor (Freiburg)
- Steering Committee: Valentine Jehl (Novartis), Friedhelm Leverkus (Pfizer), Kaspar Rufibach (Roche) and the academic leads
- Participating companies: Bayer, Boehringer Ingelheim, BMS, Janssen, Lilly, Merck, Novartis, Pfizer, Roche

Survival analysis for AdVerse events with VarYing follow-up times - SAVVY

- **Safety in terms of adverse events (AEs)** is a relevant aspect of risk-benefit assessent of therapies (Unkel et al., 2019).
- **Compare commonly** used (but biased) **estimators quantifying the AE probability** to estimators **accounting for competing events** in time-to-event studies and also compare safety comparisons between treatment groups.
- In analyses of AEs (of a certain kind), observation may be precluded by **death, progression** or some other competing event. Moreover, recording of AEs is limited to a restricted period of time (**censoring**) and varying follow-up times (Allignol et al., 2016).
- **Aim:** Investigate in an **empirical study** of several randomized controlled trials whether the use of different estimators for analyses of AEs leads to different conclusions about therapies' safety

Estimating AE probabilities: Commonly used but biased methods

- **Incidence proportion**: $\frac{\#\text{ AEs in } [0, \tau]}{\#\text{ patients}}$
  - Usually only calculated at the end of follow-up $\Rightarrow$ Assumes identical follow-up times in all patients
  - **Underestimation** of AE probability in presence of censoring
- **Incidence density**: $\mathrm{ID}(\tau) = \frac{\#\text{ AE in } [0, \tau]}{\text{patient-time at risk restricted by } \tau}$
  - Assumption of constant hazards
  - Estimator of hazard rate $\Rightarrow$ probability scale requires transformation: $1 - \exp\left(-\mathrm{ID}(\tau) \cdot \tau\right)$
  - Parametric version of 1-Kaplan-Meier
- **1-Kaplan-Meier**: competing risks censor follow-up time
  - **Overestimation** of AE probability in presence of competing events

Estimating AE probabilities: Commonly used but biased methods

- **Incidence proportion**: $\frac{\#\text{ AEs in } [0, \tau]}{\#\text{ patients}}$
  - Usually only calculated at the end of follow-up $\Rightarrow$ Assumes identical follow-up times in all patients
  - **Underestimation** of AE probability in presence of censoring
- **Incidence density**: $\mathrm{ID}(\tau) = \frac{\#\text{ AE in } [0, \tau]}{\text{patient-time at risk restricted by } \tau}$
  - Assumption of constant hazards
  - Estimator of hazard rate $\Rightarrow$ probability scale requires transformation: $1 - \exp\left(-\mathrm{ID}(\tau) \cdot \tau\right)$
  - Parametric version of 1-Kaplan-Meier
- **1-Kaplan-Meier**: competing risks censor follow-up time
  - **Overestimation** of AE probability in presence of competing events

Estimating AE probabilities: Commonly used but biased methods

- **Incidence proportion**: $\frac{\#\text{ AEs in } [0, \tau]}{\#\text{ patients}}$
  - Usually only calculated at the end of follow-up $\Rightarrow$ Assumes identical follow-up times in all patients
  - **Underestimation** of AE probability in presence of censoring
- **Incidence density**: $\mathrm{ID}(\tau) = \frac{\#\text{ AE in } [0, \tau]}{\text{patient-time at risk restricted by } \tau}$
  - Assumption of constant hazards
  - Estimator of hazard rate $\Rightarrow$ probability scale requires transformation: $1 - \exp\left(-\mathrm{ID}(\tau) \cdot \tau\right)$
  - Parametric version of 1-Kaplan-Meier
- **1-Kaplan-Meier**: competing risks censor follow-up time
  - **Overestimation** of AE probability in presence of competing events

Estimating AE probabilities: Alternative, underused approaches

- **Probability transform of the incidence density accounting for competing events** (parametric version of Aalen-Johansen):

$$\frac{\mathrm{ID}(\tau)}{\mathrm{ID}(\tau) + \overline{\mathrm{ID}}(\tau)} \left(1 - \exp(-\tau \cdot [\mathrm{ID}(\tau) + \overline{\mathrm{ID}}(\tau)])\right)$$

$$\text{with } \overline{\mathrm{ID}}(\tau) = \frac{\#\text{ competing event in } [0, \tau]}{\text{patient-time at risk restricted by } \tau}$$

  - Assumption of **constant hazards** for both (AE and competing event) hazards

- **Aalen-Johansen estimator**:

$$\mathrm{CIF}(\tau) = \sum_{u \in (0,\tau]} \prod_{v \in (0,u)} \left(1 - \Delta\hat{\Lambda}(v) - \Delta\hat{\overline{\Lambda}}(v)\right) \Delta\hat{\Lambda}(u)$$

  - **Gold-standard**: accounts for censoring and competing events and is not restricted to constant hazards (non-parametric)
  - Generalizes the Kaplan-Meier estimator to multiple event types

Estimating AE probabilities: Alternative, underused approaches

- **Probability transform of the incidence density accounting for competing events** (parametric version of Aalen-Johansen):

$$\frac{\mathrm{ID}(\tau)}{\mathrm{ID}(\tau)+\overline{\mathrm{ID}}(\tau)}\left(1-\exp(-\tau\cdot[\mathrm{ID}(\tau)+\overline{\mathrm{ID}}(\tau)])\right)$$

with $\overline{\mathrm{ID}}(\tau)=\dfrac{\#\text{ competing event in }[0,\tau]}{\text{patient-time at risk restricted by }\tau}$

  - Assumption of **constant hazards** for both (AE and competing event) hazards

- **Aalen-Johansen estimator**:

$$\mathrm{CIF}(\tau)=\sum_{u\in(0,\tau]}\prod_{v\in(0,u)}\left(1-\Delta\hat{\Lambda}(v)-\Delta\hat{\overline{\Lambda}}(v)\right)\Delta\hat{\Lambda}(u)$$

  - **Gold-standard**: accounts for censoring and competing events and is not restricted to constant hazards (non-parametric)
  - Generalizes the Kaplan-Meier estimator to multiple event types

Definition of the competing event

- **Death only**: death without prior AE, i.e., events after which an AE can definitely not occur any more
- **All events**: death, loss to follow up, withdrawal of consent, treatment discontinuation, and progression, i.e., competing events after which an AE in principle still could occur, but not is not observed due to premature end of follow-up

Group comparisons and follow-up times

- **Risk difference** or **relative risk** of incidence proportions may be misleading
  - Comparing two quantities that both underestimate the AE probability
  - Comparing two quantities **evaluated at different follow-up times**, i.e., largest observed event time in treatment group τ_A may be greater/smaller than largest observed event time in comparison group τ_B(Incidence proportion only calculated at the end of follow-up)
- Evaluate estimators at $\tau = \min(\tau_A, \tau_B)$ (referred to as P100)
- As estimators (e.g. Kaplan-Meier) at the end of follow-up may have **larger variability** due to small numbers still at risk (Pocock et al. 2002):
  - Evaluate estimators at **earlier time point when more patients are still at risk**
  - Evaluate estimators at $\tilde{\tau} = \min(\tilde{\tau}_A, \tilde{\tau}_B)$, with $\tilde{\tau}_A(p)$ and $\tilde{\tau}_B(p)$ defined as event time when $p \cdot 100\%$ of all patients in group A and group B, respectively, are still at risk , e.g., $p = 0.9$ (P90), $p = 0.6$ (P60) and $p = 0.3$ (P30)

Empirical Study

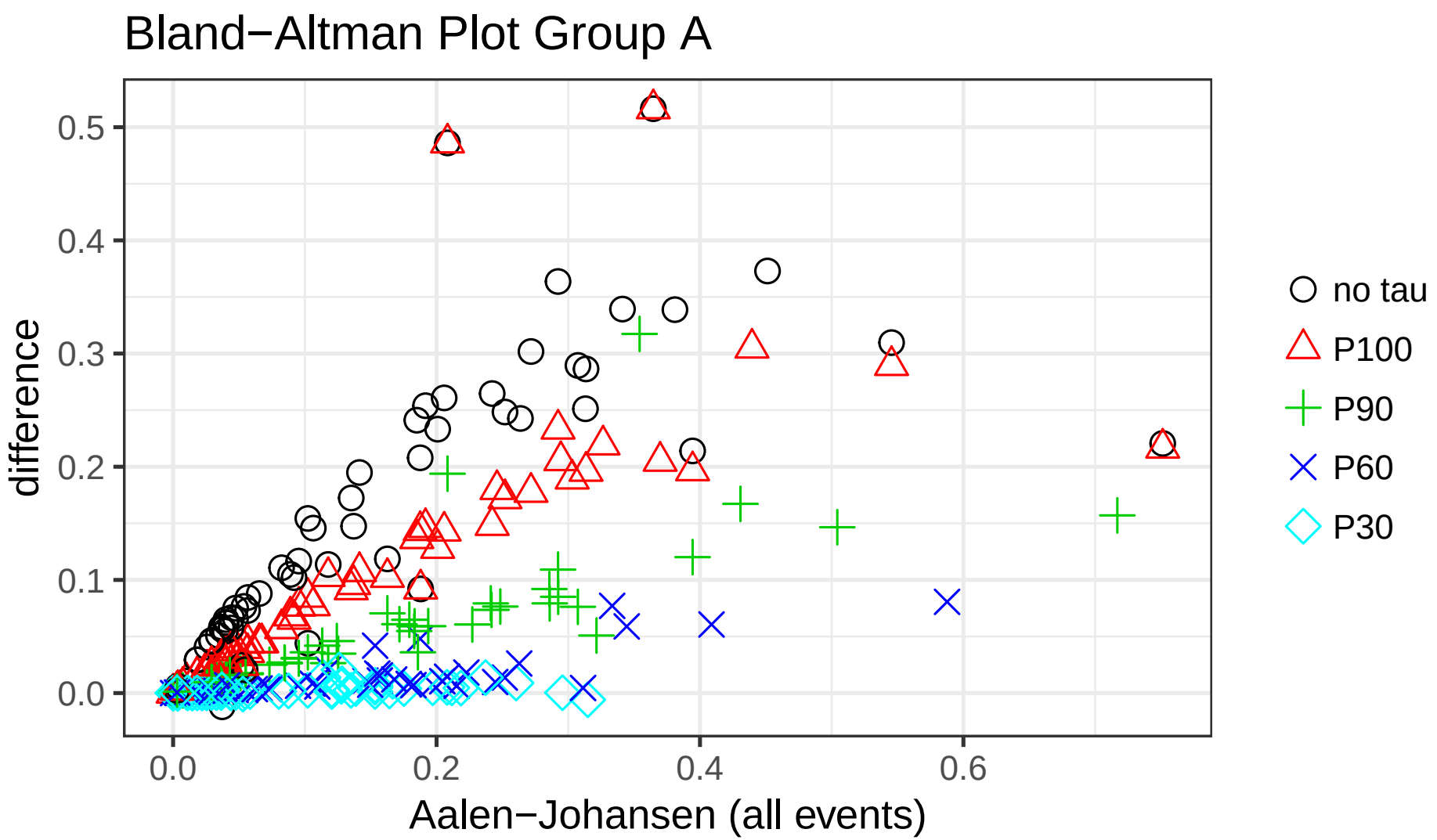
- Only **aggregated** data shared with the project collaborators: Trial level analyses ran within the sponsor company / organization using SAS (and R) code provided by the project collaborators $\Rightarrow$ **no release of individual patient data was required**.
- **Pilot study** to **develop SAS macros**, to **assess feasibility** of macros and output data structure, to **check output dataset** whether they contain all necessary information and to **train meta-analysis** and obtain early results
- Participants of the pilot study: 3 companies providing 5 studies and a total of 62 AEs (range 2-51 per study)

Some results of the Pilot study

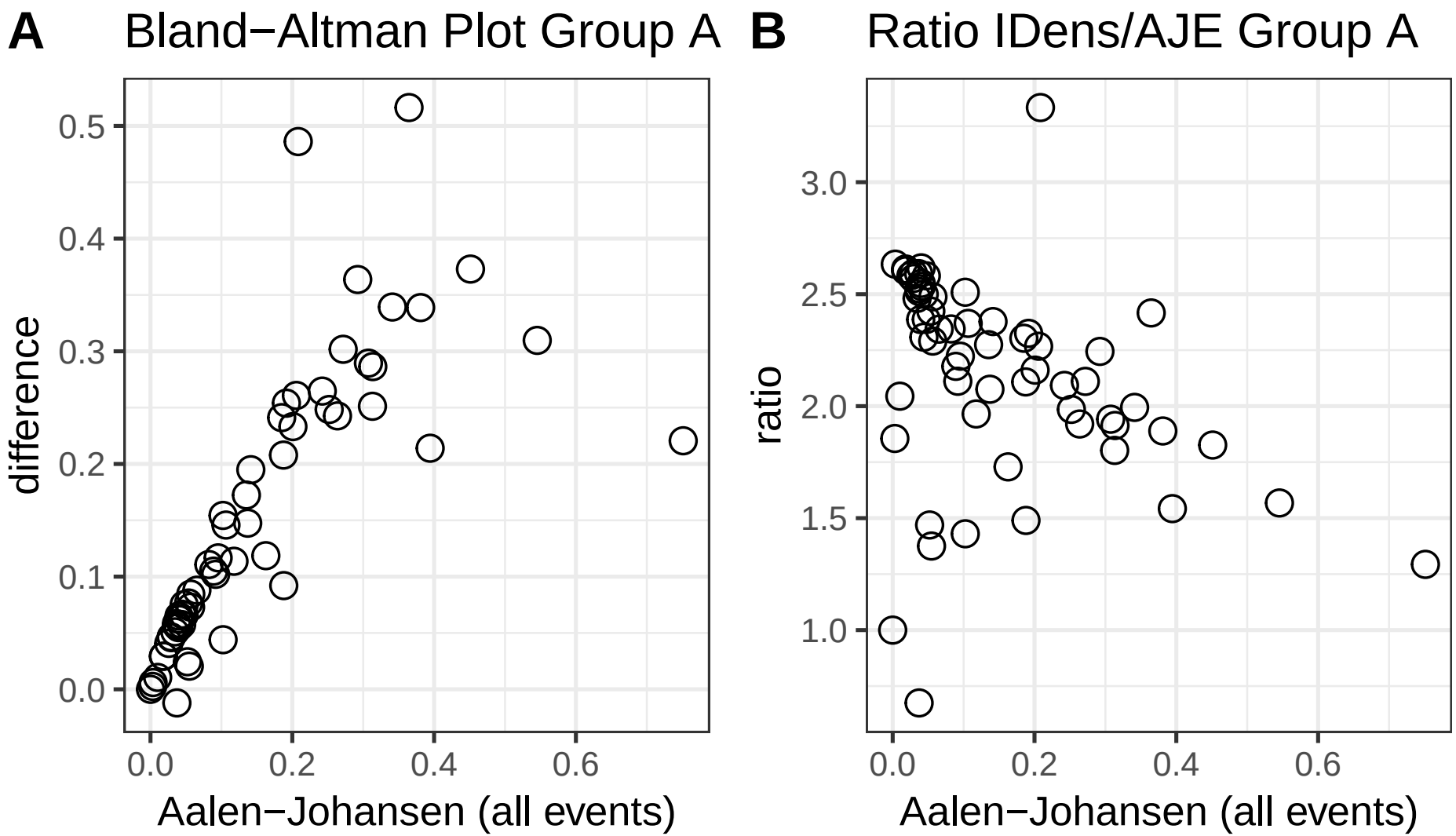
In the following, **some selected results** are presented to illustrate the kind of results the SAVVY project can generate

- Example comparison: **Probability transform of incidence density** ("parametric version of 1-KM" treating competing events as censored) and **Aalen-Johansen estimator** of AE probability in **treatment group**
- Only the "all events" definition of the competing events will be displayed

Pilot study: Assessment of bias



Pilot study: Assessment of bias (at end of follow-up)



Probability transform of incidence density overestimates AE probability compared to Aalen-Johansen estimator

Pilot study: Formal assessment of bias (maximum follow-up time)

Random effects meta-analysis

- Aggregated analysis datasets provide estimates (on appropriate scales) with bootstrapped estimates of the standard error
- Average ratio of probability transform of incidence density and Aalen-Johansen estimator:
  - On log scale: 0.773 (95% CI [0.728; 0.819]) with substantial between-study heterogeneity $\mathrm{SD} = 0.163$, $\mathrm{I}^2 = 90.22\%$
  - exp(0.773)=2.167 (95% CI [2.071; 2.268])
- **On average, the AE probability estimated with the probability transform of the incidence density is about twice the AE probability estimated with the Aalen-Johansen estimator.**
- For shorter follow-up times the estimated average ratio is smaller due to fewer competing events (results not shown)

Pilot study: Frequency categories

- at maximum follow-up time (no tau)

	Aalen-Johansen (all events)				
	very rare	rare	uncommon	common	very common
very rare	1	0	0	0	0
rare	0	0	0	0	0
uncommon	0	0	2	0	0
common	0	0	0	14	0
very common	0	0	0	15	30

- at P30

	Aalen-Johansen (all events)				
	very rare	rare	uncommon	common	very common
very rare	2	0	0	0	0
rare	0	0	0	0	0
uncommon	0	0	4	0	0
common	0	0	0	32	0
very common	0	0	0	0	24

SAVVY: Summary

- Choice of estimator of AE probability crucial (also for group comparisons)
- Ignoring competing events more of a problem than falsely assuming constant hazards (simulation result not shown)
- Under no censoring incidence proportion and Aalen-Johansen estimator are the same

SAVVY: Discussion and next steps

- Differences regarding group comparisons in detail and comparisons of hazard ratios
- Meta-regression to characterize the effect of, e.g., percentage of competing events on the average ratio
- Manuscript of the **statistical analysis plan** close to submission
- Manuscript focusing on the methodology and variances of the estimators and considering the comparisons in several **simulation** scenarios (close to submission)
- **Main study**: A total of 17 randomized controlled trials (1-3 trials per contributing organisation) including 186 AE (3-51 per study)

References

- Allignol, A., Beyersmann, J. and Schmoor, C. (2016). Statistical issues in the analysis of adverse events in time-to-event data. *Pharmaceutical Statistics* **15**, 297–305.
- Pocock, S. J., Clayton, T. C. and Altman, D. G. (2002). Survival plots of time-to-event outcomes in clinical trials: good practice and pitfalls. *The Lancet* **359**, 1686–1689.
- 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. and 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**, 165–183.

SAVVY: All planned comparisons

- The quantities marked with $\star$ are calculated in both groups.

Aimed quantity	Benchmark estimator	Compared estimator
AE probability $\star$	Aalen-Johansen	Incidence Proportion
	Aalen-Johansen	Probability Transform Incidence Density
	Aalen-Johansen	1-Kaplan-Meier
	Aalen-Johansen	Probability Transform Incidence Density accounting for CE hard
Composite endpoint $\star$	1-Kaplan-Meier	Incidence Proportion
Hazard Ratio	Cox	Ratio Incidence densities
	Cox	Ratio Nelson-Aalen estimators