

On estimands and the analysis of adverse events in the presence of varying follow-up times within the benefit assessment of therapies

Steffen Unkel

Department of Medical Statistics
University Medical Center Göttingen, Germany

This is joint work with the members of the ATF/APF project group
“Analysis of adverse events in the presence of varying follow-up times
in the context of benefit assessments”.

Acknowledgement

- In August 2016, the aforementioned joint project group was established by...
 - 1 the **Working Group Therapeutic Research** (ATF) of the German Society for Medical Informatics, Biometrics and Epidemiology (GMDS) and
 - 2 the **Working Group Pharmaceutical Research** (APF) of the German Region of the International Biometric Society (IBS-DR).
- I am grateful to all members of the ATF/APF project group.

Setting the scene

- The analysis of **adverse events** (AEs) is a key component in the assessment of a drug's **safety profile**.
- Inappropriate analysis methods may result in misleading conclusions on a therapy's safety and consequently its **benefit-risk ratio**.
- A **variety of methods** are available for the analysis of AE data, but their **complexity** and the imposed **assumptions** differ.
- The **statistical analysis of AE data** is complicated by the fact that the **follow-up times can vary** between patients, treatment groups, and/or studies.

Examples

Table: Some examples from early benefit assessments with considerably different follow-up times. The dossier assessments can be obtained from <https://www.iqwig.de>.

Dossier evaluation	Intervention	Control	Ratio of follow-up times
Oncology	Median follow-up + safety follow-up		
A14-48 prostate	16.6 months + 28 days	4.6 months + 28 days	31%
A15-17 lung	336 + 28 days	105 + 28 days	37%
A15-33 melanoma	168 + 90 days	63 + 90 days	59%
A16-04 mantle cell lymphoma	14.4 months + 30 days	3.0 months + 30 days	26%
Hepatitis C	Planned follow-up + safety follow-up		
A14-44	8 - 12 weeks + 30 days	24, 28 or 48 weeks + 30 days	23% - 57%
A16-48	12 weeks + 30 days	24 weeks + 30 days	57%

Scenarios of typical AE follow-up periods in clinical trials

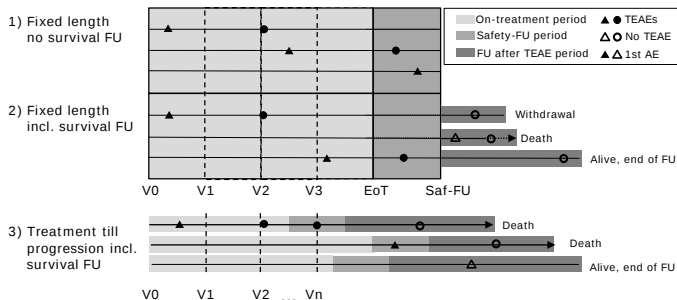


Figure: Description of different scenarios for typical AE follow-up (FU) in clinical trials (TEAEs: treatment emergent AEs (marked by bold symbols); EoT: end of treatment; Saf-FU: safety follow-up; V0: visit at the beginning of the trial; V1,...,Vn: visits during treatment). First occurrences of AEs are marked by triangles.

Main contributions of the ATF/APF project group

- We address the research gap in the analysis of AE data in the spirit of the current discussion on **clinical trial estimands**.
- We discuss which quantities should be estimated in the context of safety data, leading to the concept of **safety estimands**.
- Within the framework of estimands, we present **statistical methods for analysing AEs**. We discuss...
 - methods of estimation within one treatment group,
 - the comparison of AE occurrence between two treatment groups, and
 - methods for meta-analyses of AE data.
- We also give recommendations which **estimators** fit best to the described estimands.

Framework

- It is paramount to agree upon the **relevant target of estimation** defined by the following question:
 - what would happen to a specific patient when treated with a given drug as compared with another drug or to not being treated at all?
- Four different elements are required to describe the estimand of interest (draft addendum R1 to ICH E9 guideline):
 - 1 the **targeted population**,
 - 2 the **endpoint** (variable),
 - 3 the **intervention effect** that describes how **intercurrent events** such as treatment discontinuation, death, or switch to the other study treatment that potentially influence the endpoint are accounted for, and
 - 4 the **summary measure** that summarizes the comparison of the treatment groups under investigation.

Nomenclature of estimands in the regulatory context

- Currently, the following **classes of estimands** are discussed in the regulatory context:
 - 1 Treatment policy
 - 2 Composite
 - 3 Hypothetical
 - 4 Principal stratum
 - 5 While on treatment
- Our focus is on the **time to the occurrence of the first AE** of a specific type (usually the occurrence of a specific side effect).
- Relevant **intercurrent (post-randomization) events** are treatment discontinuation or switch, death, or other side effects that may prevent from the event of interest.

The HTA perspective in Germany

- The idea of estimands is also relevant to address the needs of **health technology assessment (HTA) bodies**.
- The **IQWiG (2017) methods paper** (Version 5.0) describes the methods used in the assessment of dossiers.
- From this framework, estimands can be derived for the HTA process in Germany (early benefit assessment is done according to **AMNOG**).
- The HTA bodies are most interested in the **treatment policy estimand**.

Flow chart

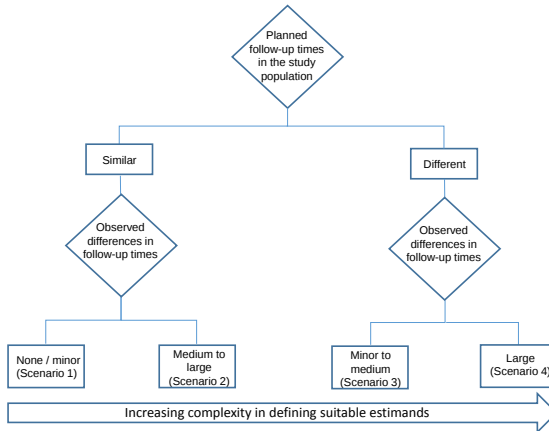


Figure: Flow chart displaying four different scenarios across indications for the consideration of safety estimands in an HTA system.

Methods of estimation within one treatment group

- **Crude "rate"**: $\hat{P}(\text{AE}) = a/n$, where a denotes the number of patients observed to experience at least one AE of a specific type and n is the total number of study patients.
- **Incidence proportion**: $\hat{P}(\text{AE in } [0, t]) = \frac{\sum_{u \leq t} a_u}{n}$, where a_u denotes the number of patients observed to experience at least one AE of a specific type at time u .
- In the presence of censoring, both the crude rate and the incidence proportion **underestimate** AE probabilities.
- One minus the Kaplan-Meier estimator for estimating $P(\text{AE in } [0, t])$, **censoring time to AE by both the end of follow-up and by competing events that preclude AE occurrence** (such as death without a prior AE), **overestimates** AE probabilities.

Methods of estimation within one treatment group (2)

- It is the **Aalen-Johansen estimator** that generalizes the Kaplan-Meier estimator to multiple event types.
- The nonparametric Aalen-Johansen estimator of the cumulative AE incidence function is

$$\hat{P}(T \leq t, \text{AE}) = \sum_{u \leq t} \hat{P}(T > u-) \cdot \frac{a_u}{n_u},$$

where T is the time until occurrence of an AE or of a competing event, $\hat{P}(T > u-)$ denotes the estimate of the probability of not experiencing an AE or the competing event just prior to time u and n_u is the number of patients at risk of observing an AE or a competing event just prior to u .

Methods of estimation within one treatment group (3)

- **Nelson-Aalen estimator** of the cumulative AE hazard $\int_0^t \alpha_{AE}(u) du$:

$$\sum_{u \leq t} \frac{a_u}{n_u} .$$

- It is the cumulative nonparametric counterpart of the commonly used **incidence rate** (or density) of AEs:

$$IR_{AE} = \frac{a}{\sum t_i} ,$$

where t_i is the time at risk for patient i and $\sum t_i$ denotes the population time (person-years) at risk.

- The incidence rate is an estimator of $\alpha_{AE}(t)$ under a constant hazard assumption, $\alpha_{AE}(t) = \alpha_{AE}$ for all t .

Methods of estimation within one treatment group (4)

- Translating incidence rates into probability statements requires incorporating competing events (CEs).
- IR_{AE} and the incidence rate of the CE, $IR_{CE} = c / \sum t_i$, can be used to obtain a **parametric counterpart** of the Aalen-Johansen estimator.
- If constant event-specific hazards are assumed, the cumulative incidence function of the event type AE is explicitly given as

$$\begin{aligned} P(T \leq t, AE) &= \int_0^t \alpha_{AE} \cdot \exp(-(\alpha_{AE} + \alpha_{CE})s) \, ds \\ &= \frac{\alpha_{AE}}{\alpha_{AE} + \alpha_{CE}} (1 - \exp(-(\alpha_{AE} + \alpha_{CE})t)) \quad , \end{aligned}$$

where α_{CE} denotes the competing event hazard.

Comparison of treatment groups

- When **comparing two treatment groups** with respect to AE occurrence, measures such as risk difference, relative risk, or odds ratio of crude rates are often suggested.
- If such measures are used in the presence of censoring and are based on biased one-sample estimators, the result of such a comparison will be biased too.
- Furthermore, the direction of the bias is uncertain.
- In a parametric approach, the ratio of two incidence rates is an appropriate estimator of the hazard ratio under a constant hazard assumption.

Comparison of treatment groups (2)

- Semiparametric **Cox proportional hazards model**:

$$\alpha_{AE}(t|Z) = \alpha_{AE;0}(t) \exp(\beta_{AE}^T Z),$$

where $\alpha_{AE;0}(t)$ is an unspecified baseline AE hazard, β_{AE} is the vector of regression coefficients and Z a vector of covariates including treatment group.

- If the only covariate is treatment group, $Z \in \{0, 1\}$, then the ratio of the incidence rates estimates the hazard ratio $\exp(\beta_{AE})$ under the assumption: $\alpha_{AE;0}(t) \equiv \text{constant}$.
- The analysis remains incomplete without consideration of the CE hazard, e.g. via a second Cox model:

$$\alpha_{CE}(t|Z) = \alpha_{CE;0}(t) \exp(\beta_{CE}^T Z) .$$

Comparison of treatment groups (3)

- Let us illustrate this using a **simple example**, assuming constant hazards.
- Consider a treatment that modifies the AE hazard by a factor of 0.5 and the CE hazard by a factor of 0.25.
- As $t \rightarrow \infty$, $P(\text{AE} \mid \text{group 1})$ becomes

$$\frac{0.5 \cdot \alpha_{\text{AE}_0}}{0.5 \cdot \alpha_{\text{AE}_0} + 0.25 \cdot \alpha_{\text{CE}_0}} > P(\text{AE} \mid \text{group 0}) ,$$

where α_{AE_0} and α_{CE_0} denote the AE hazard and CE hazard in group 0, respectively.

Comparison of treatment groups (4)

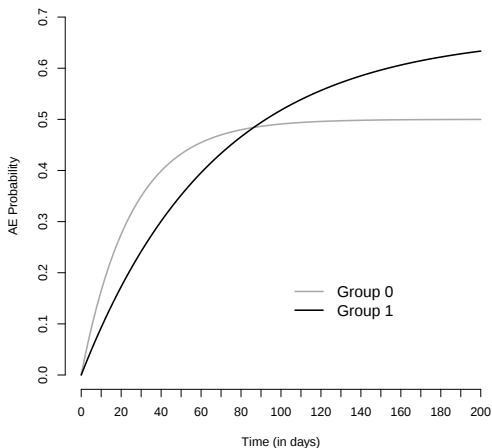


Figure: Cumulative AE probabilities for two groups and constant hazards.

Comparison of treatment groups (5)

- The **Fine and Gray approach** interprets one minus the cumulative incidence function as a survival function and fits a Cox model to the so-called **sub-distribution hazard**.
- The approach is useful in that a subdistribution hazard ratio greater (smaller) than one translates into an increase (decrease) of the cumulative incidence function.
- However, the subdistribution hazard, $\lambda(t|Z)$, is difficult to interpret, because it can be expressed as

$$\lambda(t|Z) = \frac{P(T > t|Z)}{1 - P(T \leq t, AE|Z)} \cdot \alpha_{AE}(t|Z) .$$

- Alternative: group comparisons based on confidence bands of the cumulative incidence functions.

Meta-analyses of adverse event data

- When data from more than one study are available one may **naively pool** the data across the studies (results might be biased due to **Simpson's paradox**).
- ICH E9: “any statistical procedures used to combine data across trials should be described in detail” and “attention should be paid [...] to the proper modelling of the various sources of variation”.
- **Random-effects meta-analysis** allows for **variation in both baseline outcomes and treatment effects** across studies.
- **Bayesian random-effects meta-analysis** has been suggested for the scenarios of a **few studies** or **rare events** (CIOMS Working group X).

Meta-analyses of adverse event data (2)

Example:

- Neal et al. (2017) report an integrated analysis of two large randomized placebo trials assessing the efficacy and safety of canagliflozin in patients with type 2 diabetes and an elevated risk of cardiovascular disease.
- Patients were followed up for varying lengths of time.
- In a stratified Cox regression, a beneficial effect of canagliflozin versus placebo on the primary outcome time to death from cardiovascular causes, nonfatal myocardial infarction or nonfatal stroke was demonstrated ($HR = 0.86$; 95% confidence interval (CI): $[0.75, 0.97]$).
- We consider here the AE “low trauma fracture”.

Meta-analyses of adverse event data (3)

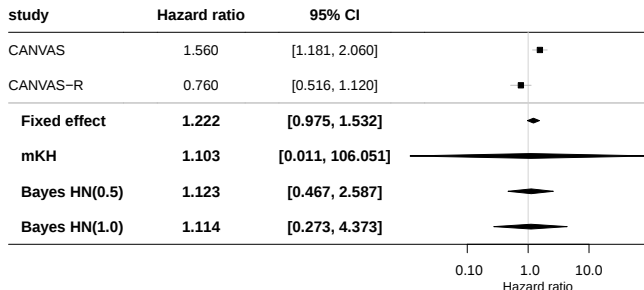


Figure: Forest plot of hazard ratios for low trauma fractures as observed in CANVAS and CANVAS-R with 95% CIs and four combined hazard ratios from a fixed-effect meta-analysis, modified Knapp-Hartung (mKH) meta-analysis and Bayesian random-effects meta-analysis with two half-normal (HN) priors for the heterogeneity parameter τ .

Estimators for estimands

- The **treatment policy estimand** focuses on the comparison of treatment groups with respect to AE occurrence until death or end of follow-up.
- It includes all AEs until death or end of study and requires the collection of AE data after treatment discontinuation.
- The AE hazards of the treatment groups can be compared by calculating the hazard ratio in a Cox regression model where for patients without an AE, the time to AE is censored by death or by end of follow-up.
- Estimation of AE probabilities within treatment groups: Aalen-Johansen estimator.

Estimators for estimands (2)

- Treatment groups can also be compared with respect to the AE probabilities by
 - estimating the difference between AE probabilities at a specified time point,
 - fitting a Fine and Gray model to the AE data,
 - calculating the odds ratio from a proportional odds cumulative incidence model.
- The **while on treatment estimand** includes AEs until discontinuation of treatment and requires the collection of AE data up to this event.
- Treatment groups can be compared with the same methods as used for the treatment policy estimand, now treating discontinuation of treatment before AE and death without prior AE as CEs.

Summary and further work

Concluding remarks

- We formulated a framework based on safety estimands within which we proposed statistical methods for analysing AE data, including methods for evidence synthesis.
- For the described estimands, we have also given recommendations which estimators should be used.
- In particular, we would like to advocate the use of time-to-event methodology for the analysis of AE data.
- Estimands of primary interest may differ between drug approval agencies and the HTA bodies in certain instances.
- Further work, some of it under way, can be done or is required in several areas, some of which have already been mentioned.

Empirical investigation

Survival analysis for Adverse events with VarYing follow-up times (SAVVY)

- The aim of the SAVVY study is to investigate in a large number of RCTs whether the different analyses of AEs lead to different decisions when comparing safety between groups.
- Data analyses are done within the company using R or SAS code provided by the project collaborators. Release of **individual patient data is not required**.
- Only aggregated data summarizing the results will be shared plus some information on the studies.
- Nine companies currently participate in the SAVVY study. Contact: tim.friede@med.uni-goettingen.de

References



MAIN PAPER



Open Access



On estimands and the analysis of adverse events in the presence of varying follow-up times within the benefit assessment of therapies

Steffen Unkel, Marjan Amiri, Norbert Benda, Jan Beyersmann, Dietrich Knoerzer, Katrin Kupas, Frank Langer, Friedhelm Leverkus, Anja Loos, Claudia Ose, Tanja Proctor, Claudia Schmoor, Carsten Schwenke, Guido Skipka, Kristina Unnebrink, Florian Voss, Tim Friede

First published: 20 November 2018 | <https://doi.org/10.1002/pst.1915>

References (2)



Special issue: Analysis of adverse event data (2016):

Four papers along with an editorial, *Pharmaceutical Statistics*, Vol. 15 No. 4, pp. 287-379.



European Medicines Agency (EMA) (2017):

ICH E9(R1) addendum to the guideline on statistical principles for clinical trials on estimands and sensitivity analysis in clinical trials, step 2b, EMA/CHMP/ICH/436221/2017.



IQWiG (2017):

Allgemeine Methoden, Version 5.0, Institute for Quality and Efficiency in Health Care, Cologne, Germany.



Council for International Organisations of Medical Sciences (CIOMS) Working group X (2016):

Evidence synthesis and meta-analysis for drug safety: report of CIOMS Working Group X, Geneva, Switzerland.



Neal, B. et al. (2017):

Canagliflozin and cardiovascular and renal events in type 2 diabetes, *The New England Journal of Medicine*, Vol. 377, pp. 644-657.