Treatment effect measures for time-to-event endpoints - Estimands and beyond

Kaspar Rufibach (F. Hoffmann-La Roche)
Joint work with Mouna Akacha (Novartis)

Frankfurt, June 13th, 2017

Problem statement

- Estimand discussion: majority of examples based on **longitudinal data**.
  - Dapagliflozin: Change from baseline to 24 weeks for HbA1c.
- Many indications (not only oncology!) use **time-to-event endpoints**: time between two defined timepoints, e.g. randomization to death.
- Examples:
  - Overall survival (OS) or surrogates such as
  - Progression-free survival (PFS) or
  - Disease-free survival (DFS).
- Such endpoints can be interpreted as binary data (alive or death) assessed over time → estimand concept equally applies.
- **But is «complicated» estimand framework necessary for these endpoints «familiar» to all of us?**

DFS = DFS = DFS?

- Consensus in literature about ambiguous definition of time-to-event endpoints.
- Persistent over time: Peto et al. (1977), Altman et al. (1995), Mathoulin-Pelissier et al. (2008), Bellera et al. (2013).
- Half of articles in major clinical journals fail to provide clear definition of time-to-event endpoint (Mathoulin-Pelissier et al., 2008).
- “Most of these time-to-event endpoints currently lack standardised definition enabling a cross comparison of results from different clinical trials” (Bellera et al., 2013).
- DATECAN initiative: Definition for the Assessment of Time-to-event Endpoints in CANcer trials (Bellera et al., 2013).
- Work up recommendations for relevant indications. **«Estimand» not used as a concept**.

Example: DFS in adjuvant breast cancer

Table 1. Example of Inconsistent Definitions of Disease-Free Survival

Trial	Local/Regional Recurrence	Distant Metastasis	Death From Any Cause	Invasive Contralateral Breast Cancer	Second Primary Invasive Cancer (nonbreast)	Ipsilateral DCIS	Contralateral DCIS	Ipsilateral LCIS	Contralateral LCIS
BIG 1-98[4]	X	X	X	X	X				
MA-17[1]	X	X		X		X	X	X	X
ATAC[2]	X	X	X	X		X	X		
IES[3]	X	X	X	X					
ARNO[5]	X	X		X					

NOTE: Event-free survival used by ARNO.
Abbreviations: DCIS, ductal carcinoma in situ; LCIS, lobular carcinoma in situ; BIG, Breast International Group; MA, National Cancer Institute of Canada Clinical Trials Group MA-17; ATAC, Arimidex, Tamoxifen Alone, or in Combination; IES, Intergroup Exemestane 031; ARNO, Arimidex, Nolvadex 95 Study.

Hudis et al (2007).

Implications of heterogeneous definitions

- Trial **objective** and **effect quantification** not aligned.
- Trial can be «significant» or «non-significant».
  - PETACC-03, colon cancer, DFS as primary endpoint.
  - Primary definition in protocol: counted 2nd primary cancer other than colon as event → «non-significant», Van Cutsem et al. (2005).
  - Relapse-free survival (RPS): 2nd primary other than colon not event → «significant».
  - RPS in PETACC-03 was DFS in MOSAIC!
- Relative (hazard ratio) and absolute (milestone values, medians) effect quantifiers are affected by heterogeneity in definitions.
- Meta-analysis? Historical controls?

Estimand framework for time-to-event endpoints

Framework proposed by ICH E9 working group:

1. **Population**: *Characterized through in- and exclusion criteria*.
   Nothing specific to time-to-event endpoints.

2. **Variable**: *Quantities required to address the scientific question.*
   - Starting date, e.g. randomization or registration.
   - Event defining the endpoint.
   - Intercurrent events happening inbetween.

Estimand framework for time-to-event endpoints

3. **Intervention effect**: *How potential intercurrent events are reflected in scientific question.*

   For time-to-event endpoints intercurrent event can be

   – **censored**: death for time-to-progression (TTP),

   – **made part of a composite**: death for PFS,

   – **ignored**: start of new therapy for OS,

   – treated as a **competing risk**: time to CNS PD, non-CNS PD, death in NSCLC.

- Choice between TTP and PFS = **choice of estimand**.

Estimand framework for time-to-event endpoints

4. **Summary measure**: *On which the treatment comparison will be based*.
   - Two-fold: **hypothesis test** on null hypothesis of interest and **quantification of effect**.
   - Logrank test: (virtually) always valid. Highest power for proportional hazards (PH), but non-PH does not invalidate it. Weighted?
   - If logrank rejects → efficacy established. Quantify effect using any relevant and meaningful measure.
   - **Hazard ratio**: most meaningful for PH, but has also interpretation for non-PH (geometric mean of piecewise HRs, weighted proportional to #events per period).
   - **Milestone** survival, difference of medians → vertical / horizontal «slices» only, massive information reduction.
   - **Restricted-mean survival**.
   - **Concordance probability**: generalization of P(X < Y) known from Wilcoxon test (Schemper et al., 2009).
   - **Cause-specific hazards** (incl. ratios) or **cumulative incidence**

Estimand framework for time-to-event endpoints

- **Heterogeneity** in endpoint definitions for endpoints.
- Harmonization underway (DATECAN). Discussion could / will gain if embedded in **estimand framework**!
- Language and definitions need to be improved. Just saying «DFS is the time between randomization and the earliest of relapse, 2nd tumor, or death» typically not be sufficient.
- Discussion on 4 estimand components needs to involve **clinicians**, **statisticians**, **regulators**, **payers**.

Example: PFS in lymphoma

	Gallium NCT01332968	Goya NCT01287741
Indication	Indolent lymphoma (FL)	Aggressive lymphoma (DLBCL)
Comparison	Gazyva + CT vs. Rituximab + CT	
N	1200	1400
Endpoint	PFS: event = PD/death, censored at last tumor assessment.	
Outcome	HR = 0.66 @ efficacy interim	HR = 0.92 @ final analysis

- Cheson et al (2007): «…in studies in which failure to respond without progression is considered an indication for another therapy, such patients **should be censored** at that point for the progression analysis.»
- Fleming et al (2009): «…patients **should not be censored** at the time other treatments are initiated when analyzing the PFS end point.»
- Potential intercurrent events:
  - Failure to respond to induction therapy, but no PD (IC1).

Example: PFS in lymphoma

- **Indolent** lymphoma: NALT typically initiated after PD.
- **Aggressive** lymphoma: failure to achieve complete response → very bad prognosis. Investigators tempted to initiate NALT if patient «only» achieves partial or stable response, even before PD.

Nr	Estimand	Description	IC1	IC2	Comment
1	Pure PFS based on PD/death only	PFS irrespective of whether a patient responds to initial treatment or starts NALT.	-	-	This would correspond to the definition recommended in Fleming et al (2009). Risk of investigators starting NALT before PD in aggressive.
2	PFS stopping at NALT	PFS irrespective whether a patient responds to initial treatment, as long as a patient stays on initially assigned therapy.	-	0	
3	PFS stopping at end of induction non-response	PFS up to a end of induction non-response, irrespective of initiation of NALT.	0	-	This would correspond to the definition recommended in Cheson et al (2007).
4	PFS stopping at NALT or end of induction non-response	PFS up to a end of induction non-response or initiation of NALT.	0	0	
5	EFS	Event-free survival: Count NALT (= IC2) as event, in addition to PD and death.	-	1	Similar to PFS, may be useful in evaluation of highly toxic therapies, only acceptable for regulators if NALT is supported by some ``objective'' evaluation of treatment failure in the absence of PD.

Example: PFS in lymphoma - conclusions

- DLBCL and FL: two related diseases, often same treatment → defining relevant estimand still
  - might **not be straightforward**,
  - needs **careful assessment** on what studied therapy should achieve.
- Irrespective of chosen estimand: we recommend
  - trial developers identify **all** potential intercurrent events upfront,
  - indicate for each whether ignored, censored, made part of composite, or competing risk.

Although current estimand focus is on longitudinal data it will come to your data type (whatever it is) as well!

Doing now what patients need next