Estimands in oncology

Steven Teerenstra[1,2]

[1]Radboud university Nijmegen medical center, NL

[2]Statistical assessor at Medicines Evaluation Board (MEB), NL

Disclaimer

Views expressed

- my own, based on my experience at MEB as statistical assessor
- not to be quoted or interpreted as those of MEB

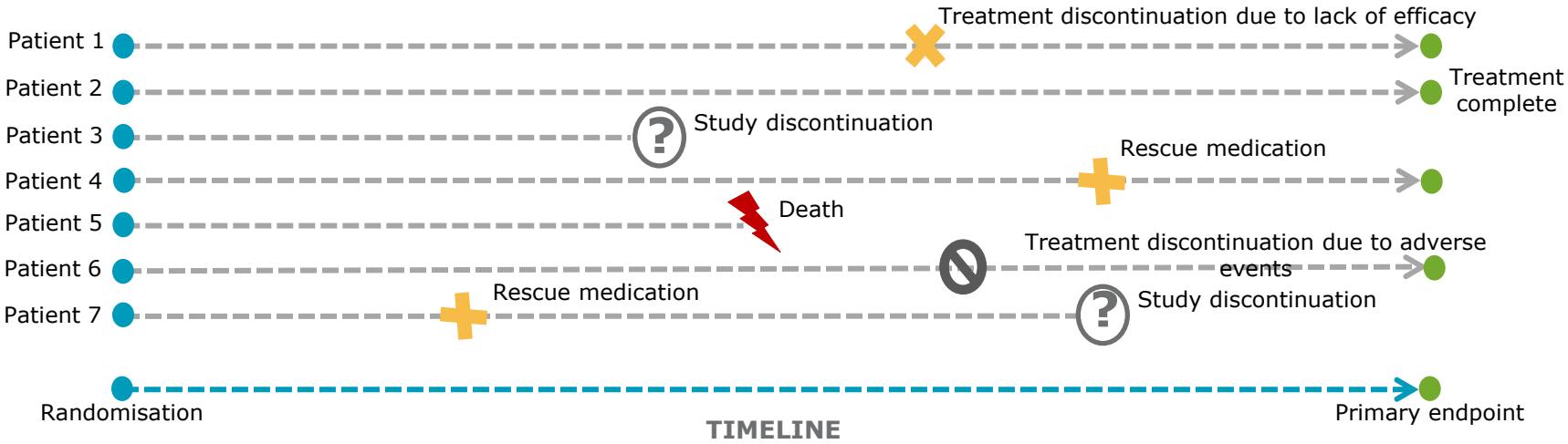
estimand

Estimand

Origin:

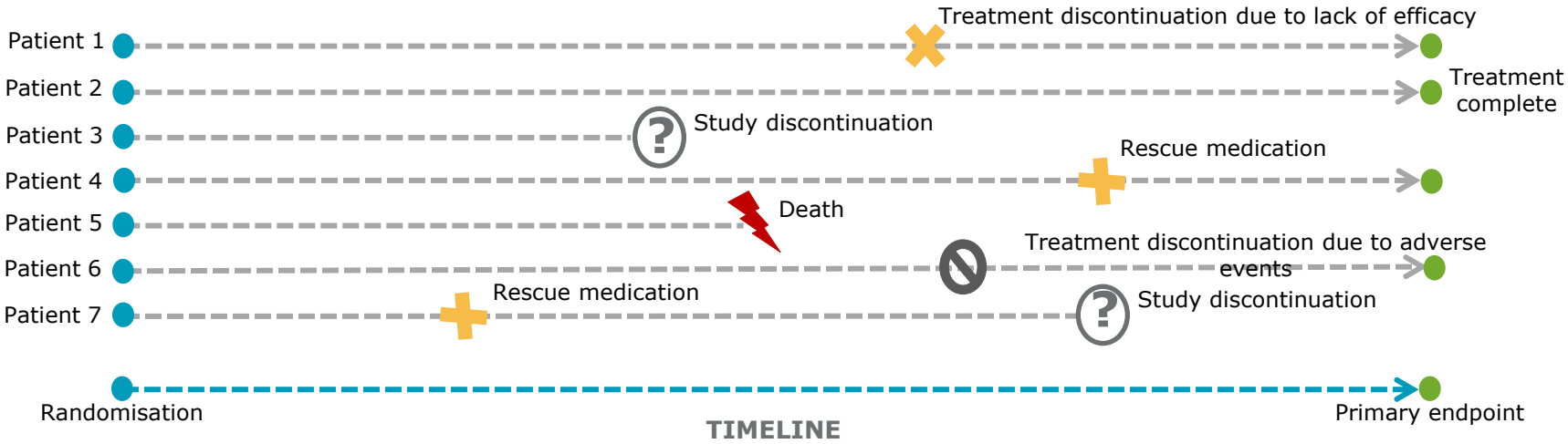
- **gerundive** from the Latin verb *estimare*
- *estimandum= (the property) that is to be estimated*

Intercurrent events



e.g. treatment discontinuation, switching, rescue/additional medication, competing/terminal event, deviations from protocol (measurements), …

Intercurrent events



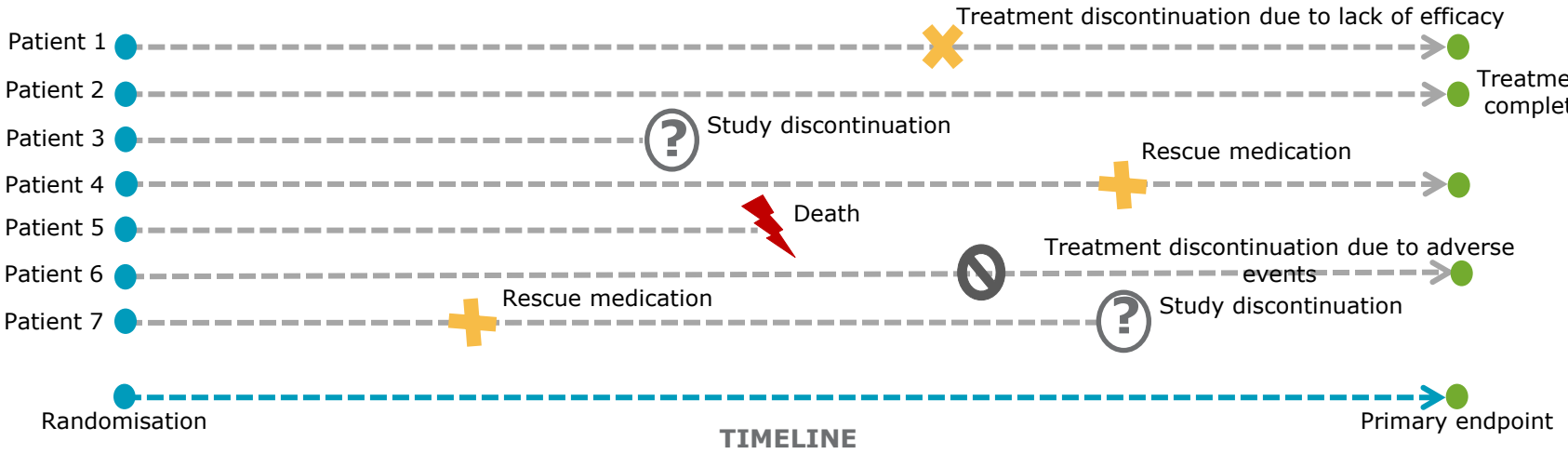
e.g. treatment discontinuation, switching, rescue/additional medication, competing/terminal event, deviations from protocol (measurements), …

often causing 'missing' information as to the situation when patients would have adhered to

* randomised treatment and
* (assessment as per) protocol

until end of trial

Intercurrent events



e.g. treatment discontinuation, switching, rescue/additional medication, competing/terminal event, deviations from protocol (measurements), …

often causing 'missing' information as to the situation when patients would have adhered to

* randomised treatment and
* (assessment as per) protocol

until end of trial

Relevance, definition, existence endpoint questionable

Oncology: intercurrent events->time-to-event endpoints

- No different from continuous, binary, rate outcomes in principle
- Deviation from protocol treatment
  - discontinuation of treatment
    - lack of efficacy: e.g. progression
    - AE
    - (substantial) resolution of disease
    - other
  - Especially treatment switching (often no rescue medication)
  - change in concomittant therapy (additional medication)

Oncology: intercurrent events->time-to-event endpoints

- No different from continuous, binary, rate outcomes in principle
- Deviation from protocol treatment
  - discontinuation of treatment
    - lack of efficacy: e.g. progression
    - AE
    - (substantial) resolution of disease
    - other
  - Especially treatment switching (often no rescue medication)
  - change in concomittant therapy (additional medication)
- Deviation from protocol assessment
  - undocumented (spontaneous) assessments
    - e.g. undocumented progression altering treatment of disease
  - (intermittent) missed assessments visits
  - leave the trial (permanently):
    - withdrawal by patient or by physician (for reasons as above)
    - Lost-to-followup
    - Terminal/competing event (e.g. death when looking at time to progression)

Handling intercurrent events partly different

Intercurrent events:

Continuous /binary/rate outcomes:

- currently often 'hidden' in missing data
- Discussions about non-missing at random

time-to-event outcomes:

- currently often 'hidden' in censored data
  e.g. PFS/DFS/EFS
  - censoring for start new anti-cancer therapy
  - Censoring for > 1 missed visits
  - Censored for undocumented progression
- Discussion about informative censoring

Current practice

- Set rules regarding
  - data collection wrt intercurrent events (in design)
  - handling of intercurrent design (in analysis)

- Example:
  - Censor control patients who switch to experimental treatment
  - No data collection (on e.g. new progressions) after switch

Caveat

- However: handling of intercurrent events in design and analysis
  - dictates which effect is estimated (which question is answered)
  - under which assumptions

- Example:
  handling of intercurrent event= censor when switched
  - One may think this address the ITT question,
    because it enables to include everyone starting the trial
  - Effect "when everyone would have adhered to treatment:
    no control patient would have switched to experimental treatment"
    (*Hypothetical estimand)*

  - Kaplan-Meier/Cox: unbiased for this effect under the assumption:

    "patients in control group that do not switch
    are representative for those that do switch"
    (censoring due to switching is non-informative wrt outcome)

Usual practice

- Handling of intercurrent events in design/analysis

Effect estimated / Question answered *(estimand)*

Aim of ICH-E9 addendum

- Handling of intercurrent events in design/analysis

Effect to be estimated / Question to be answered *(estimand)*

Aim of ICH-E9 addendum

- Handling of intercurrent events in design/analysis

Effect to be estimated / Question to be answered *(estimand)*

- Discussion with stakeholders (regulators, industry, HTA, patients, ..) about
  - **which** *estimand* is primary and **why**
  - **how** design and analysis will estimate it (*estimator)*
    - under which assumptions
  - *Sensitivity analyses* **per estimand**
    - i.e. different assumptions (different estimators) for the same question
  - Possibly: which secondary/other estimands
    - For other stake holders / broader view

Survival outcomes in oncology estimands

Examples

- PFS (DFS, EFS,..) primary endpoint
  - Censoring rules / data collection after censoring:
    - what is the implied estimand?
    - Which analyses address the same question (estimand) and which different estimands?

- OS primary endpoint
  - cross-over of control patients to experimental treatment
  - Switch to treatment that changes course of disease substantially e.g.
    - (completely) different mechanism of action
    - (completely) different efficacy (safety)

    - For example: stem cell transplantation
    - (underlying consideration: some stakeholders consider it 'unfair' to use an treatment policy estimand then)
  - Late survival / tolerators effects

PFS: EMA guideline

- EMA:
  - Appendix 1 to GL on evaluation of anticancer medicinal products in man (CHMP/EWP/205/95 rev.3)

- progression date= first evidence of objective progression
- Regardless of violations, discontinuation, change of therapy.
  - If, for a particular study, a different approach is considered to be more appropriate, a justification is expected and CHMP Scientific Advice agreement is recommended at the planning stage.

- => no censoring for violations, discontinuation, change therapy
  => treatment policy estimand
  is preferred by EMA guideline

PFS: FDA guideline

- FDA:
  - GfI: Clin. trial endpoints for the approval of cancer drugs and biologics

- progression date = earliest time of observing any progression
  - without prior missing assessments and censoring at the date when the last radiological assessment determined a lack of progression.
  - so: when missing visits: censoring

- "For instance, for the primary analysis, patients going off-study for undocumented clinical progression, change of cancer treatment, or decreasing performance status can be censored at the last adequate tumor assessment. The secondary sensitivity analysis would include these dropouts as progression events."

PFS: FDA guideline

- censoring when going off study for
  - missed visits
  - undocumented clinical progression,
  - change of cancer treatment, or
  - decreasing performance status
- going of study: no PFS data collection implied after censoring
- *hypothetical estimand*
  "effect if no patient would switch to another treatment and stay in the study until documented clinical progression (measured at prespecified assessments) "

- (Kaplan-Meier/Cox regression)
  unbiased for this effect under the assumption:

  "patients that do *not* switch/undocu.PD/PS↓/miss visits are representative for those that do switch/undocu.PD/.."

PFS: sensitivity analyses

- Often many 'sensitivity' analyses presented

Scenario 1: If disease progression was documented between scheduled visits, the date of the next scheduled visit was used as the date of progression.
Scenario 3: If the patient started new antineoplastic therapy with or without subsequent progression/death event, the patient was censored at the date of the last disease assessment before the start of antineoplastic therapy.
Scenario 4: If the patient started new antineoplastic therapy (with or without subsequent progression/death event), the patient was treated as progressed at the date of last disease assessment before the start of antineoplastic therapy.
Scenario 5: If death or progressive disease occurred after more than 1 missed visit, the patient was censored at the last disease assessment before the missed visits.
Scenario 6: If death or progressive disease occurred after more than 1 missed visit, the patient was censored at the last disease assessment before the missed visits. If the patient started new antineoplastic therapy with or without subsequent progression/death event, then the patient was censored at the date of the last disease assessment before the start of antineoplastic therapy.

PFS: sensitivity analyses

- Often many 'sensitivity' analyses presented

Scenario 1: If disease progression was documented between scheduled visits, the date of the next scheduled visit was used as the date of progression.
Scenario 3: If the patient started new antineoplastic therapy with or without subsequent progression/death event, the patient was censored at the date of the last disease assessment before the start of antineoplastic therapy.
Scenario 4: If the patient started new antineoplastic therapy (with or without subsequent progression/death event), the patient was treated as progressed at the date of last disease assessment before the start of antineoplastic therapy.
Scenario 5: If death or progressive disease occurred after more than 1 missed visit, the patient was censored at the last disease assessment before the missed visits.
Scenario 6: If death or progressive disease occurred after more than 1 missed visit, the patient was censored at the last disease assessment before the missed visits. If the patient started new antineoplastic therapy with or without subsequent progression/death event, then the patient was censored at the date of the last disease assessment before the start of antineoplastic therapy.

- May give a <u>false sense of consistency</u> as to the primary estimand (question at hand)
  if these sensitivity analyses give *similar answers*
  but actually *address different* estimands (questions)
- May give a <u>false sense of inconsistency</u> if these do *not give similar answers* because they are *addressing different* estimands (questions)

PFS

Important:

- Clarify which estimand / question is addressed
- Especially for sensititivity analyses
- Present per estimand (question)
  - the primary analysis (estimator)
  - The sensitivity analyses (estimators) explaining
    - that the same estimand (question) is addressed
    - which assumptions are varied

OS

- Many intercurrent events:

- always subsequent therapy:
  - may argue whether this is intercurrent or not
    - If started after protocol specified treatment regimen
  - Some acknowledgement of part of difference in survival due to difference in subsequent therapies
  - *hypothetical estimands* (e.g. "what would be the effect if no subsequent therapy", "what would be the effect is all arms had similar subsequent subsequent therapy") in principle may be interesting for some stakeholders, but difficult to estimate

- Often a treatment policy estimand (implicity) chosen
  - so intercurrent events and subsequent therapy no issue

OS

Situations where not only trt policy estimand could be used:

- switch to experimental treatment (“cross-over”)
- switch to therapies that substantially change course of disease
- Early / late survival effects

Switch to experimental therapy

- Two estimands (implicitly) used
  - Treatment policy effect
  - Effect if no cross-over would have been available (hypothetical estimand)
    - Latter often proposed by Sponsor for 'pure efficacy of randomised trt'

- Issues:
  - if hypothetical estimand *post-hoc proposed* when treatment policy estimand fails
  - Estimators for this hypothetical estimand typically require strong assumptions
    - IPCW: no unmeasured confounders
    - RPSFT: Main assumption: (relative) treatment effect when switching is the same as time of randomisation
    - Two-stage: switch at designated 'secondary baseline' (e.g. progression) and no unmeasured confounders at this secondary baseline

Switch to experimental therapy

Important issues:

- Preplan (and discuss) which estimand is primary and why
- Preplan to collect *extra* data to motivate that assumptions of estimators are plausible
- Avoid switch to experimental therapy (?!)

OS (switch to disease altering therapy)

Example: leukaemia

- EU trial, but different standard of care in regions
  - west-EU: stem cell transplantation part of standard of care
  - east-EU: stem cell transplantation much much less available

- Preplanned analysis:
  - OS including transplant
  - No censoring for transplant
    => treatment policy estimand

- Treatment policy OS:
  - not clinically relevant difference,
  - not statistically significant

OS (switch to disease altering therapy)

- Post hoc proposed: hypothetical estimand:
  - "OS when no transplant would have been available"
  - Censoring for transplant
    - with inverse probability weighting on certain characteristics to 'replace' a patient having transplant by a patient not having transplant with those characteristics similar (IPCW).
  - No unmeasured confounders assumption:
    - prognosis, switch, and time of switch are predictable by (time dependent) characteristics used in IPCW

- But: data-collection (timing & type of variables) often not preplanned and not (fully) convincing that 'no unmeasured confounders' assumption is plausible

- Even if clin. relevant & stat.sign. effect for hypothetical estimand: ……may be relevant for east-EU, but not for west-EU (?!)

OS

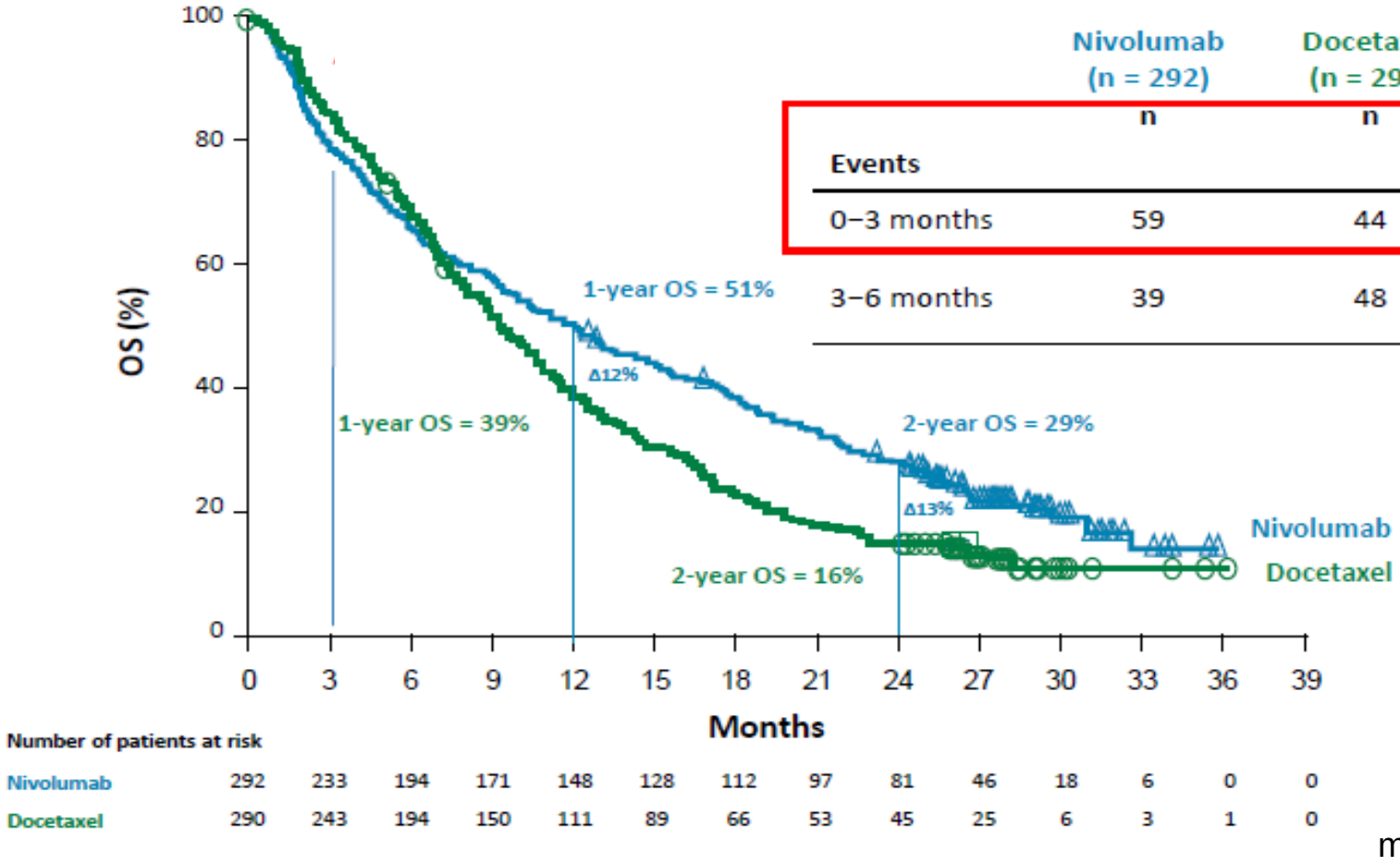
Important issue:

- A priori discussion which estimand primary,
  - for which stakeholder

- Preplan data-collection/design accordingly
  - Here:
    A priori data collection to make ‘no unmeasured confounders’ assumption needed by IPCW stronger

Early / Late survival effect

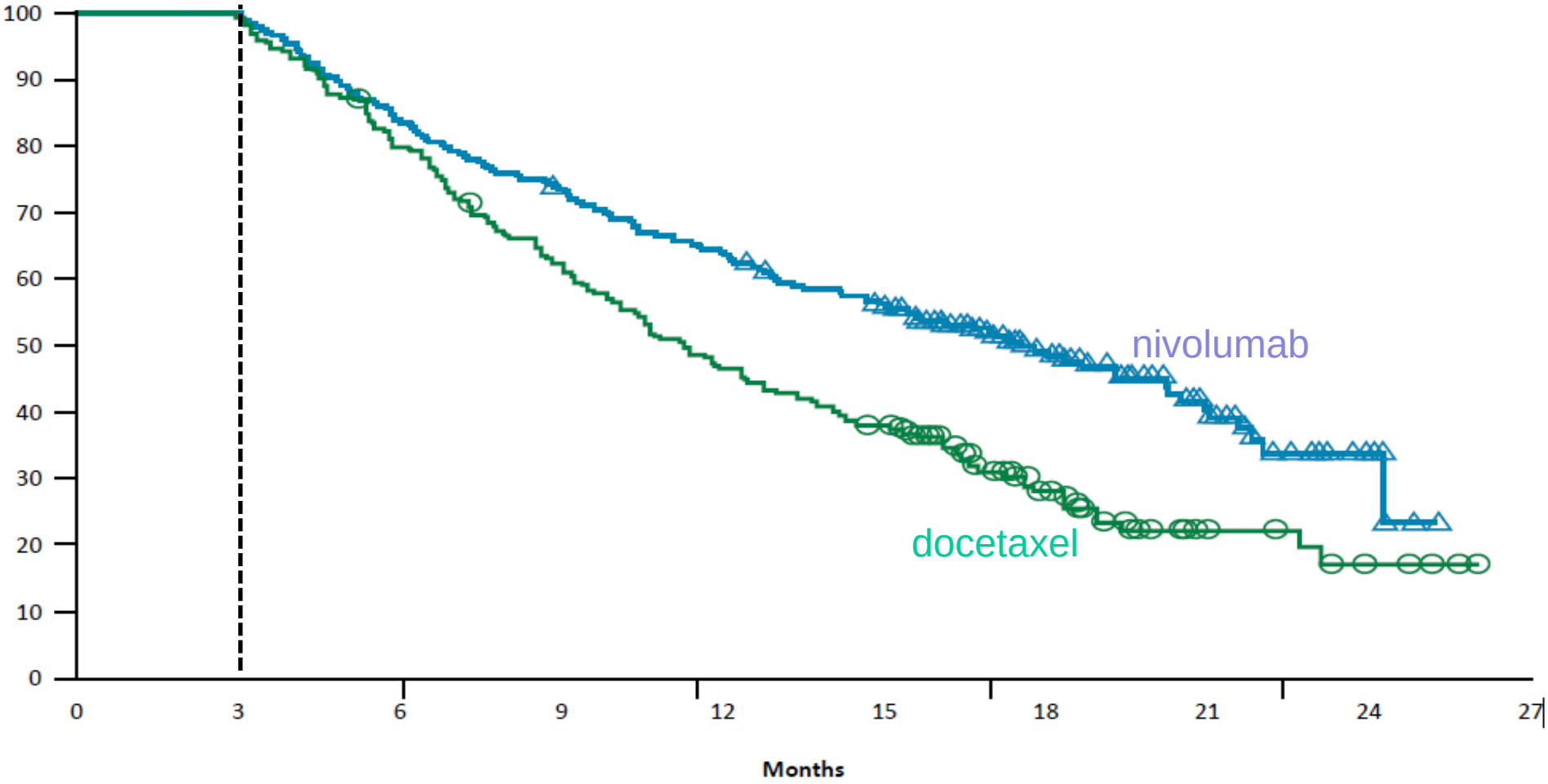
- Based on
  - *Fattori predittivi di efficacia ed interpretazione della risposta all' immunoterapia* by Matteo Brighenti, March 13-14 March, 2017
  - Borghaei, et al. Presented at the American Society of Clinical Oncology 2016 Annual Meeting; June 3–7, 2016; Chicago, IL, USA. Abstract 9025.



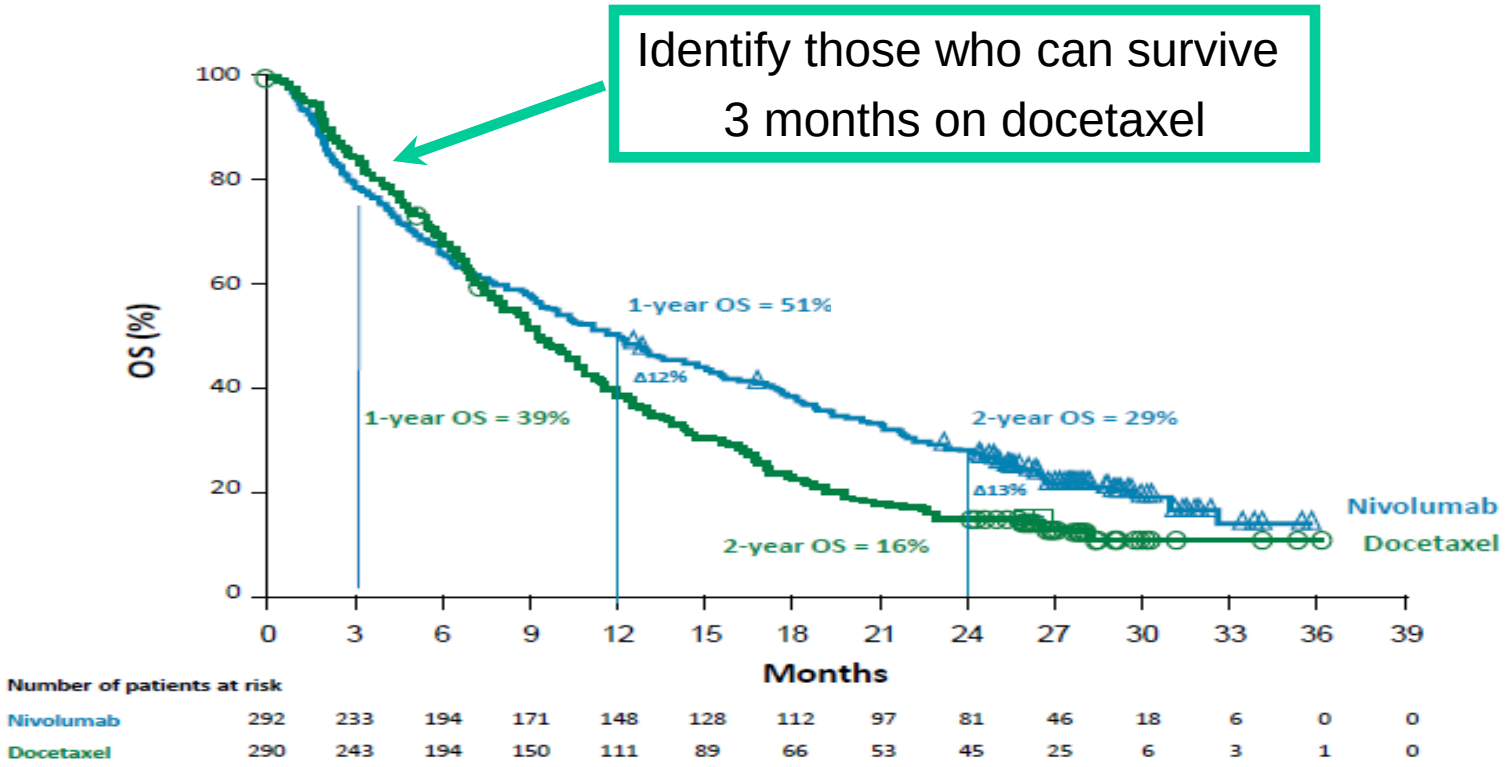
Events	Nivolumab (n = 292) n	Docetaxel (n = 290) n	Difference n
0–3 months	59	44	15
3–6 months	39	48	[9]

February 18, 2016 database lock; min.follow-up: 2 years

- What is the effect for those able to survive at least 3 months <u>on nivolumab</u> (‘late effect’, ‘tolerator effect’)
- *Principal stratum estimand*
- Interesting for patients



- estimator 1: Landmark analysis with 3-month landmark
- But groups based on selection on post-baseline covariate
  => those able to survive 3 months on nivolumab may be prognostically different from those able to survive 3 month on docetaxel
  => non-randomised comparison => bias



- estimator 2 based on identification of those who can survive at least 3 months on nivolumab via predictive modeling
- 'convincing' predictive model for nivolumab survival at 3 months possible? (e.g. because only 44 deaths available for modeling)
- principal strata estimand often difficult because of selecting the stratum (without bias)
- Can also look at those who tolerate docetaxel at least 3 months

Estimands:

- different way to look at questions that were relevant before
  - what question are we trying to answer,
  - under which assumptions and
  - how to translate this in good design and analysis

- Reframing good practice we (at least partially) did before

- (R)enforcing clarity in discussions (planning & interpretation)

- (Re)empowering stake holders

Thank you for your attention