# Sense and sensibility of estimands for health technology assessment

Jason Wang, Celgene

Christine Fletcher, Amgen

Neil Hawkins, Glasgow University

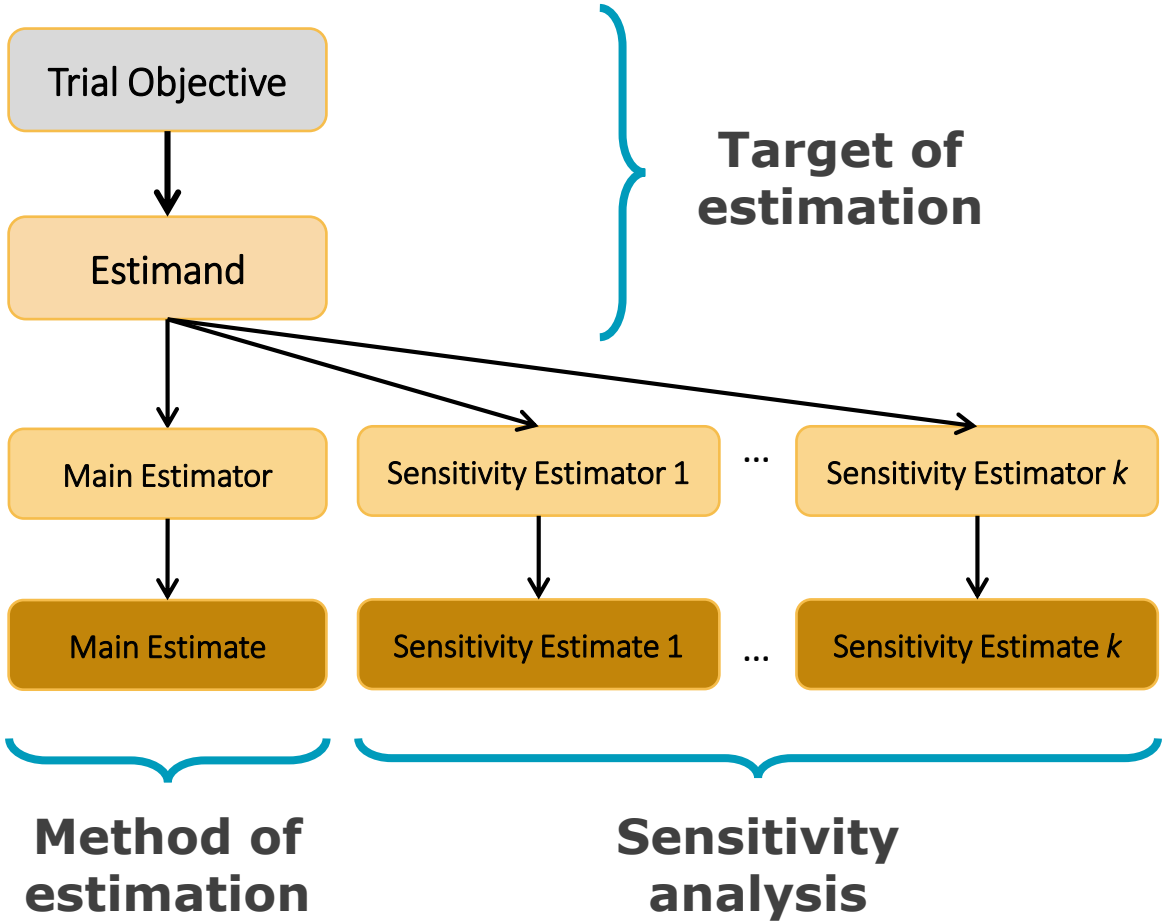
## Disclaimer

The views expressed are those of the presenters and should not be understood or quoted as being made on behalf of their affiliations.

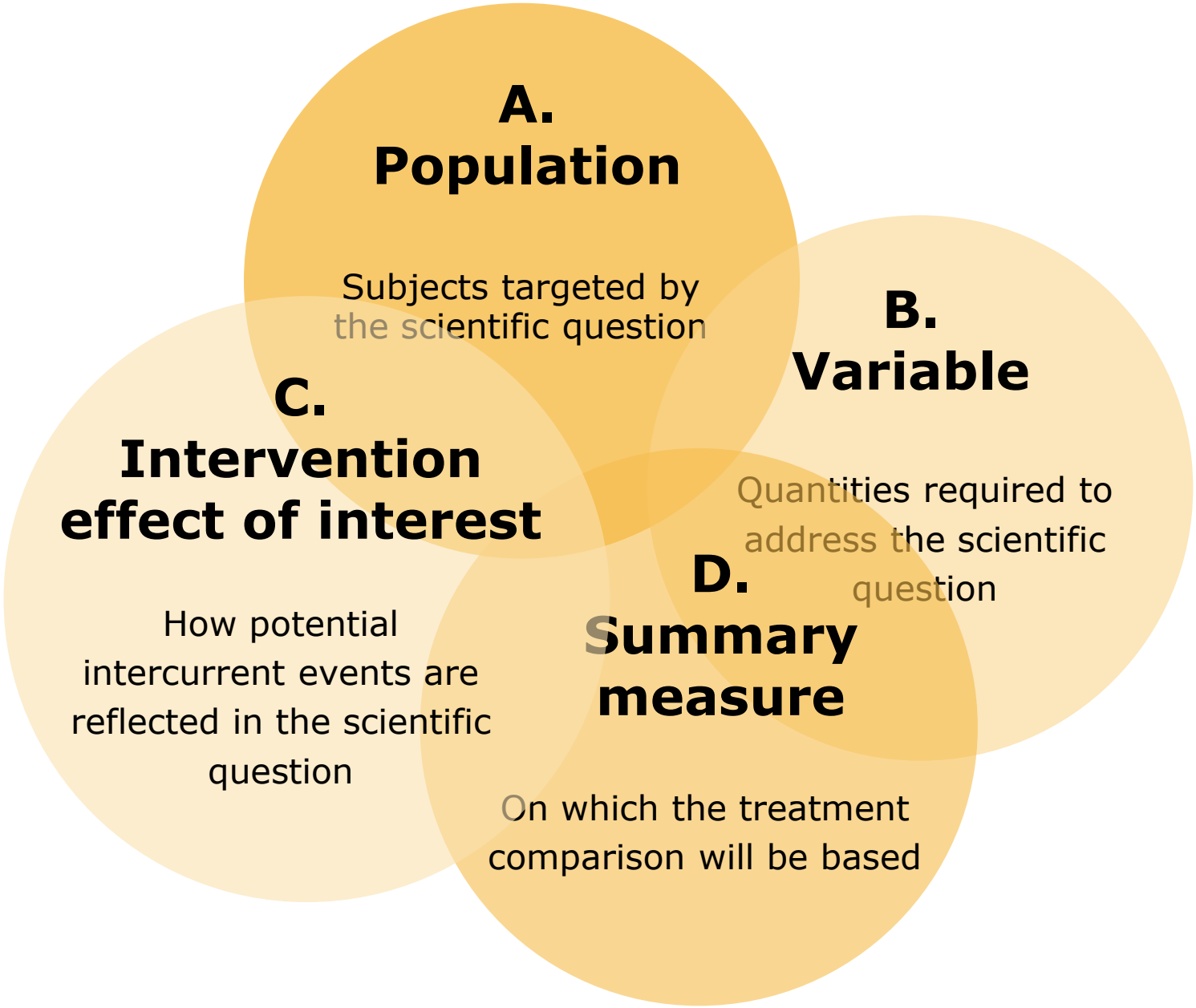
# Outlines

- Recent discussion on estimands include some considerations in HTA aspects, but has not fully covered the wide range of aspects in HTA.
- Why considering estimands is important for HTA and what  different from regulatory aspects
- Estimands that make sense for HTA.
- Which effect (when treatment changes)?
- Which population (when dropout/missing occurs)?
- Estimation vs hypothesis testing.

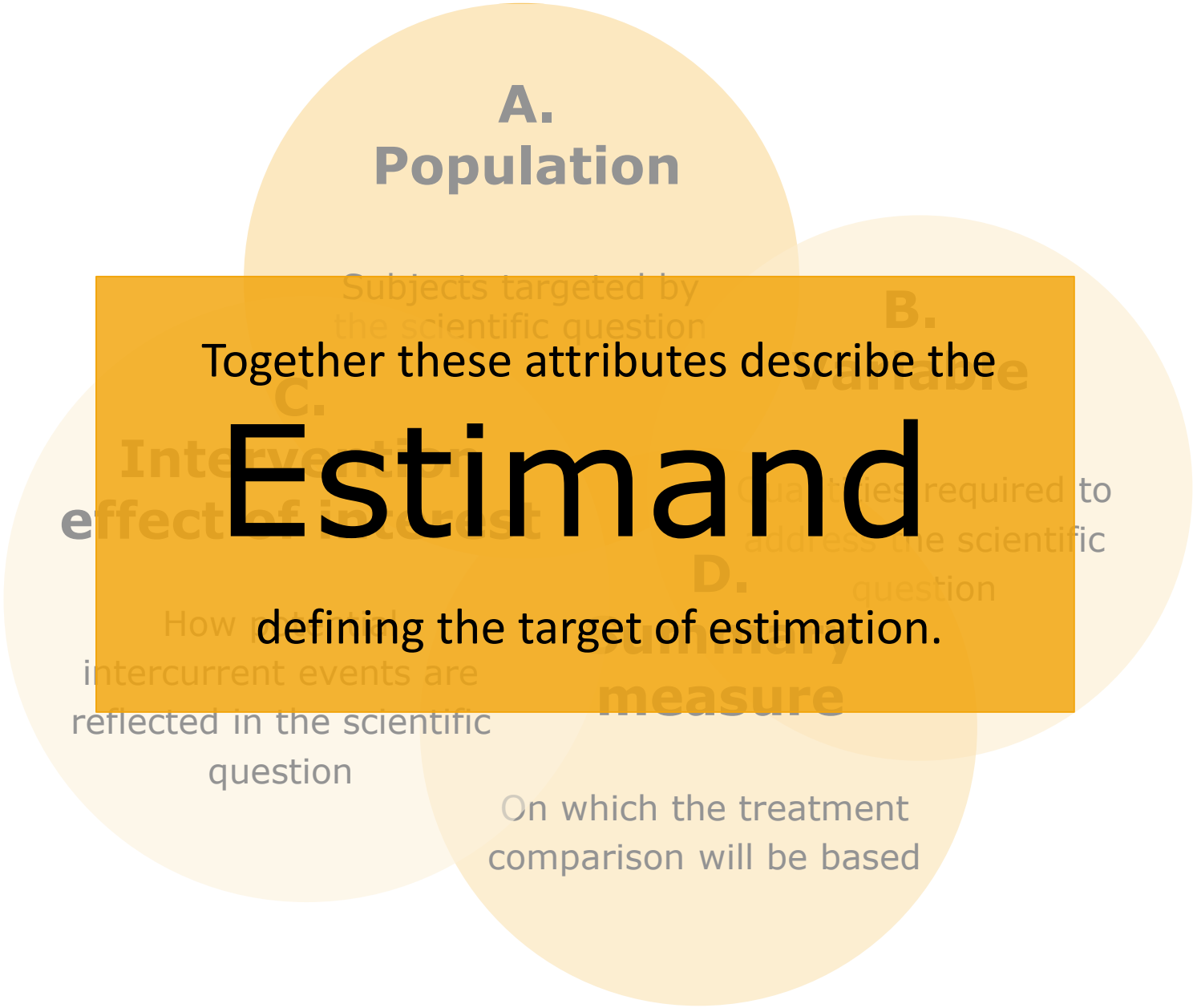
# Estimands: a new framework



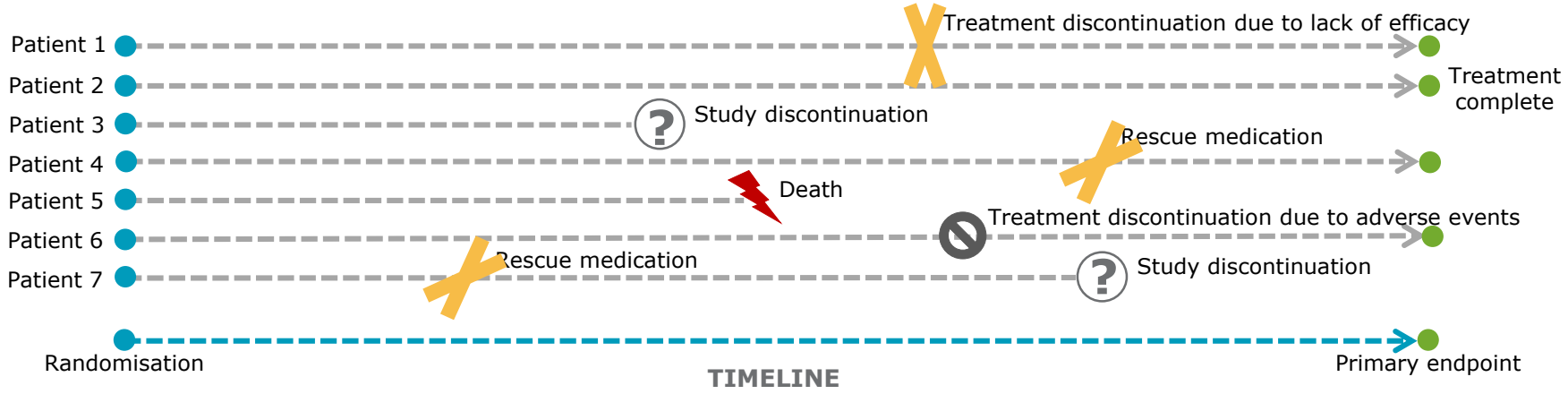
# Description of an estimand



# Description of an estimand



# Intercurrent events




- Events may occur that make the relevance, the definition, or even the existence of the primary variable questionable.
- Such events may include: death, treatment discontinuation due to adverse events or lack of efficacy, use of other medicines affecting the outcome, whether specified or prohibited by the protocol.

# The benefits of considering the use of estimands in HTA:

- It helps to clarify the needs of HTA at trial design stages or earlier, along with the development of value dossier.
- Coordination of estimands for multiple purposes can facilitate using planned trial analyses for HTA.
- It helps to identify statistical problems in estimation, e.g., confounding bias, at early stages.
- It is important for regulatory and HTA parallel consultation, e.g., to confirm appropriate estimands for HTA purposes.

# Common issues/different views on estimands

- Common issues about estimands for regulatory and HTA purposes:
  - Response-dependent treatment changes may make the direct comparison between randomized treatments less useful
  - Response dependent dropout/missing compromises causal estimation for the whole population
- However, for HTA we consider them in different aspects
  - Some response-dependent treatment changes reflect clinical practice.
  - It may still be possible to estimate useful causal effects despite response dependent dropout/missing
  - The analytic perspective and context varies between payers/decision-makers

# How RCT evidence is used for HTA?

- HTA approaches to using evidences from randomized controlled trials (RCT) vary, but some common features are:
  - Using treatment effect estimates to predict long term clinical and economical outcomes.
  - Evaluating outcomes when the test treatment is available, possibly as a part of a treatment policy, compared to current clinical practice.
  - To estimate the impact of treatment on a variety of outcomes representing how patients feel, function and survive, and resource use.
  - With a focus on causal estimation, rather than inference, e.g. hypothesis tests

# RCT evidence for HTA: examples

- Oncology trials
  - Some registration trials use progression free survival (PFS) and some use overall survival (OS) as the primary endpoint. Switching from the control to the test treatment upon disease progression is sometimes allowed.
  - OS is the key endpoint for HTA, but PFS may also be used as a surrogate of OS (e.g., the ASCO value framework).
  - Cost-effectiveness analysis (CEA) typically considers both PFS and OS.
  - For CEA, switching to the test treatment from the control arm will confound causal estimates, as it unlikely to reflect current clinical practice. Switching from the test treatment to other options may not confound causal estimates if it reflects clinical practice.

# RCT evidence for HTA: examples

- Type 2 Diabetes:
  - Although HbA1c is often used as the primary endpoint in RCTs for registration, HTA needs to assess clinical outcomes such as diabetic events.
  - CEAs are often based on two sets of models
    - Models for long term HbA1c changes, using treatment effect from RCTs
    - Models for the relationship between diabetic events and HbA1c
  - Long term effects of treatment policies are needed.
- Multiple sclerosis
  - CEAs typically use multi-state models to model relapsing-remitting and staged disease progression.
  - May use multiple treatments to delay the progression/relapse, hence CEA needs effects of long term treatment policy

# Treatment or policy estimand?

- Changes to subsequent treatments, including switching between randomized treatments, are common.
- Adjust for them or not, and if yes, to what?
  - In general we should always adjust artificial changes (or no change) to that of common practice to construct two scenarios: one with and another without the test treatment
  - E.g., if patients on the test treatment A may be switched (following clinical practices) to either B and C, while only B is licensed in the target population, then those switched to C should be adjusted to as if they had switched to B.

# Estimation vs hypothesis testing

- In HTA the primary goal is to estimate the effectiveness of a policy including the test treatment, compared with that without it.
- In contrast, the primary analysis of RCTS is often to test the primary hypothesis of no treatment effect difference, compared with a selected comparator, e.g., $H_0$: $\mu_1 - \mu_0 = 0$.
- In many cases, the hypothesis $H_0$ can be tested (i.e., with appropriate type I/II error control) with the ITT approach, but $\mu_1 - \mu_0$ may not be estimable.
- HTA evaluates multiple outcomes quantitatively, wouldn't make decisions based on a single significant hypothesis test only. Hence controlling familywise alpha level is not important.
- However, uncertainties are important and should be considered in the context of HTA decision making

# Strategies dealing with intercurrent events

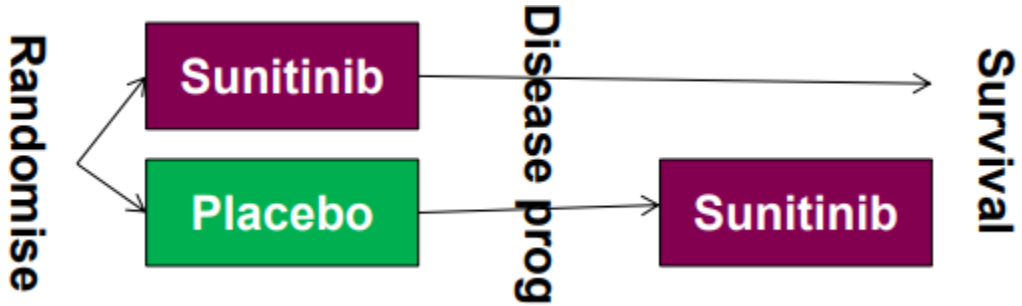
| Strategy | Summary | Treatment | Outcome | Population | Randomization based estimator |
|---|---|---|---|---|---|
| **Treatment policy** | Treatment changes as an integrated policy | Varying | Original | Whole | Yes |
| **Composite endpoint** | Intercurrent events as a part of endpoints | As randomized | Modified | Whole | Yes |
| **Hypothetical** | Assumes everyone stays on the randomized treatment | As randomized | Original | Whole | Varying |
| **Principal stratum** | The effects within a principal stratum. | As randomized | Original | Varying | Varying |
| **While on treatment** | Only counts effects while on the randomized treatments. | As randomized | Modified | Whole | Yes |

# Strategies dealing with intercurrent events

| Strategy | Pros | Cons |
|---|---|---|
| **Treatment policy** | • HTA typically evaluate efficiency of a policy rather than a treatment. | • Hard/impossible to replicate feasible policies for multiple payers in a single trial.<br>• Adjusting results to a policy different from the one in trial breaks the randomization |
| **Composite** | • A good composite endpoint (e.g., QALY) may fit to HTA well<br>• Meaningful for BRA and HTA if it reflects benefit-risk balancing | • Estimation of QALYs often requires long term extrapolation<br>• Not relevant to all decision-makers<br>• Reductive |
| **Hypothetical** | • Comparisons between randomized treatments | • The hypothetical scenario is meaningful only under some situations |
| **Principal stratum** | • Effects in principal strata are causal<br>• Randomization may still be useful | • Principal strata may not be identifiable<br>• They may not be of interest. |
| **While on treatment** | • May construct meaningful estimands, e.g., together with the composite approach | • May only measure short-term effects for some patients.<br>• May not be meaningful |

# Example: Treatment switching and overall survival (OS) as estimand

- Example: The Sunitinib trial
  - Patients on pcb can switch to sunitinib
  - Patients on Sunitinib can't switch




- Approaches to deal with treatment switchover

1. No adjustment: treatment policy approach: not applicable here
2. Adjusting to no switching: Hypothetical approach
   - RPAFTM: randomization based assuming no heterogeneity.
   - IPW: not randomization based assuming no unobserved confounders.
   - Sensitivity analysis: evaluating sensitivities to
     - Switching based on expected individual response (essential heterogeneity)
     - Unobserved confounders

# Example: Treatment switching and progression free survival (PFS) as estimand

- Disease progression is an intercurrent event.
- Two strategies lead to PFS
  - While on treatment: No switching before event/censoring happens
  - Composite endpoint: Combining death and progression
- Pros: Randomization based, hence estimation is easy.
- Cons:
  - PFS may not be a good OS predictor.  E.g., the ASCO framework counts PFS evidence as 70% of OS in the scoring algorithm.
  - CEA will need both PFS and OS.

# Principal stratum

## A typical principal stratum: the compliers stratum

- Compliers: those who take treatment A if given A, and take treatment C if given C.
- It depends on counterfactual scenarios so not identifiable.
- Effects with the stratum is called local average treatment effect (LATE).
- It may not be of interest: we are interested in those comply to the active treatment (as treated).
- But in some situations, **compliers = as treated**.

## Hypothetical example: switching A -> C due to safety events

- Assuming A has additional toxicity, hence if one can tolerate A, he will tolerate C.
- Assume A is 1 unit more effective than C.
- 100 patients are randomized to A (R=A) and 50 stay on A and 50 switch to C. 100 patients are randomized to C (R=C) and stay on C.
- Can we estimate the average treatment effect (ATE) of A in the whole population?
- Can we estimate the effect of A among those staying on A?

# Principal stratification and principal stratum

Example: (continued)

- Under some conditions (including our example) we can estimate the effect among treated without knowing the complier stratum.
- Assume the mean treatment effects in the R=A and R=C groups are 0.5 and 0.
- Intuitively since among R=A the exposure to A is 50% so the effect per exposed is 0.5 / 50% =1; we recover the true treatment effect
- This is the instrumental variable estimator for compliers
  (mean(eff.|R=A) –mean (eff.|R=C))/(rate(A|R=A)-rate(A|R=C))
- When this estimator is valid?
  - Constant treatment effect
  - Heterogeneous effects and switching only depends on baselines
- The IV estimator is generally not valid when switching depends on individuals responses (essential heterogeneity). But in this case, it is also valid as long as those tolerate A also tolerate C (monotonicity).

# Trial design/analysis and data collection

- Estimands for HTA should be taken into account at the design stage
- Good planning ensures key HTA estimands are estimable
- ISPOR recommendations provides view from HTA/CEA aspects

VALUE IN HEALTH 18 (2015) 161–172

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval


## Cost-Effectiveness Analysis Alongside Clinical Trials II—An ISPOR Good Research Practices Task Force Report

Scott D. Ramsey, MD, PhD[1,2,*], Richard J. Willke, PhD[3], Henry Glick, PhD[4], Shelby D. Reed, PhD, RPh[5], Federico Augustovski, MD, MSc, PhD[6], Bengt Jonsson, PhD[7], Andrew Briggs, MSc, DPhil[8], Sean D. Sullivan, PhD[1,2]

# Key recommendations

1. For generalizability/external validation:
   - Trial should follow common clinical practice as much as possible
   - Treatments should be compared as a part of treatment policies relevant to the decision under consideration
   - Appropriate endpoints and follow-up to capture long term effectiveness
2. For HTA specific estimands
   - Patient level resource data (e.g., duration and frequency of hospitalization). Distinguish protocol driven ones from those will occur in practice as much as possible.
   - Patient level preference based data (e.g., QoL data).
3. For analysis:
   - ITT principal (if 1. is followed, this would be the treatment policy approach).
   - Uncertainty assessment is important.

# Summary

- Consideration on estimands is important for HTA, even at an early stage of drug development.
- Estimands should either reflect effectiveness of treatments used in clinical practice, or can be extrapolated to estimate it.
- Knowing how estimands are used in HTA is key in the consideration
- Construction of meaningful estimands appropriate for HTA, using the 5 strategies.