Estimands and Their Role in Clinical Trials

Frank Bretz (Novartis)

EFSPI/PSI – UK – September 28, 2015

Acknowledgements

- ICH E9(R1) Expert Working Group
- M Akacha, D Ohlssen, H Schmidli, G Rosenkranz (Novartis)

NOVARTIS

Aims of this talk

- Discuss the need for a new framework
- Illustrate the choice of estimands with real and hypothetical examples

Aims of this talk

- Discuss the need for a new framework
- Illustrate the choice of estimands with real and hypothetical examples

Background – Measure of treatment effect

- Clinical trials aim to draw inference about the benefit of a medicinal product
- Trialist selects an appropriate measure of treatment effect
  - For example in a 2-arm diabetes study:

    'Difference in change in HbA1c from baseline to 24 weeks based on all randomized patients'
- Selection often does not account for post-randomization events, e.g. dropout
- Post-randomization events introduce ambiguity to the chosen measure of treatment effect

→ Lack of clarity leads to challenges in communication

Example to illustrate the problem – Dapagliflozin

- Primary endpoint: Change in HbA1c from baseline to 24 weeks
- Analysis set: modified intention to treat
- Data after initiation of rescue medication was considered as missing
- Primary analysis: ANCOVA using LOCF
- Comment by FDA statistical reviewer:

> "While FDA has implicitly endorsed LOCF imputation for diabetes trials in the past, there is now more awareness in the statistical community of the limitations of this approach. [...] My preferred analysis simply uses the observed values of patients who were rescued."

NOVARTIS

Dapagliflozin – Sponsor's interest versus regulatory interest

Sponsor:

What was done?

- Remove data after initiation of rescue medication

Implied 'measure treatment benefit':

- Attempt to establish the treatment effect of the initially randomized treatments had no patient received rescue medication

FDA:

- Include all data regardless of initiation of rescue medication

- Compare treatment policies: 'dapagliflozin plus rescue' versus 'control plus rescue'

Dapagliflozin – Sponsor's interest versus regulatory interest

Implied objectives / scientific questions of interest differ for both parties.

This is hidden behind the method of estimation / handling of 'missing data'.

Need to avoid such 'miscommunications'.

patient received rescue medication

NOVARTIS

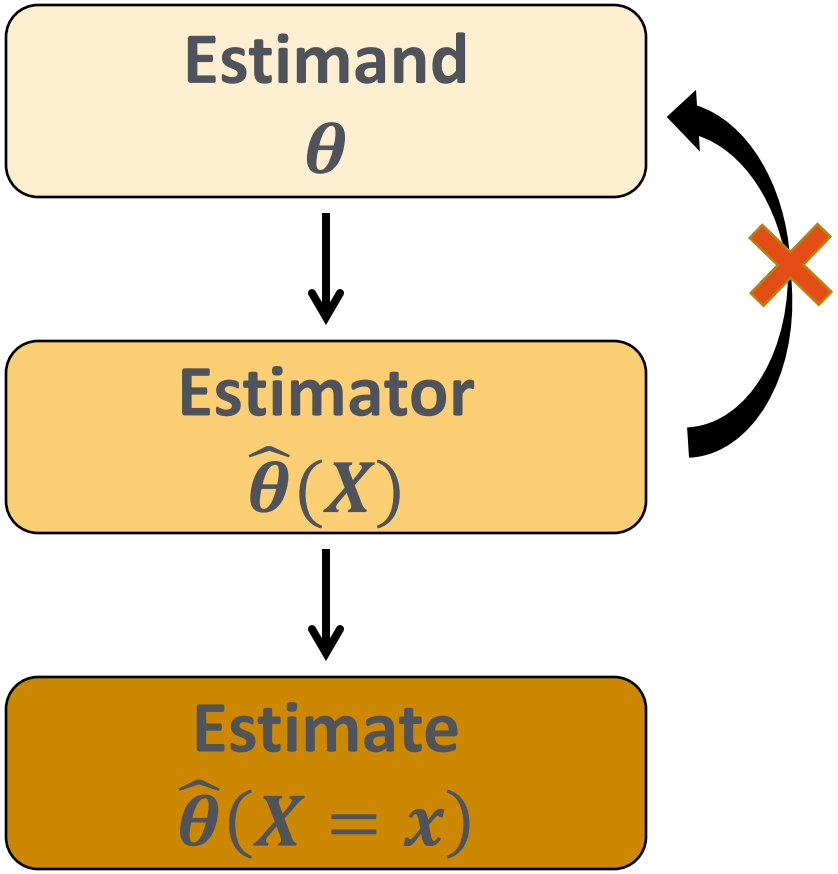
Distinguish ‘target of estimation’ and ‘method of estimation’

Estimand framework helps distinguishing between

- target of estimation (estimand)
- method of estimation (estimator)

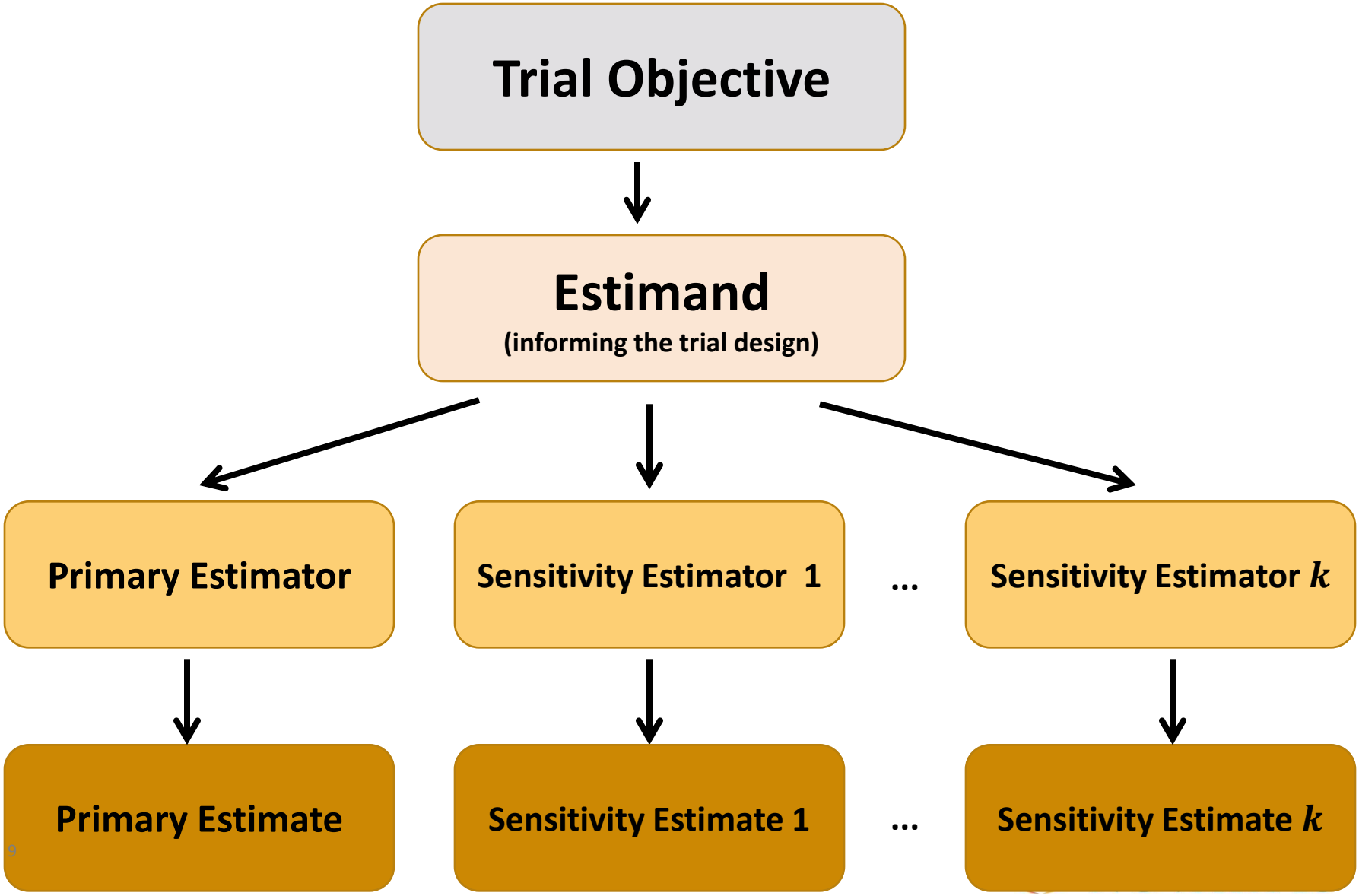
Especially in the context of ‘missing data’ the estimand and method of estimation are often confused

However, estimand framework applies to a broader setting than missing data



NOVARTIS

Structured framework to bridge trial objectives with statistical inference



Estimand – A proposed definition

An estimand reflects what is to be estimated to address the scientific question of interest posed by a trial.

The choice of an estimand involves:

- **Population of interest**
- **Endpoint of interest**
- **Measure of intervention effect**

Estimand – A proposed definition
Population of interest

A estimand reflects what is to be estimated to address the scientific question of interest posed by a trial.

The choice of an estimand involves:

- **Population of interest**
- Endpoint of interest
- Measure of intervention effect

- Population for which we are assessing the scientific question of interest (parameterized through the estimand)
- Characterized through the inclusion / exclusion criteria

Estimand – A proposed definition
Endpoint of interest

A estimand reflects what is to be estimated to address the scientific question of interest posed by a trial.

The choice of an estimand involves:

- Population of interest
- **Endpoint of interest**
- Measure of intervention effect

- Characterized through measurement and time point/period of interest

Estimand – A proposed definition
Measure of intervention effect

A estimand reflects what is to be estimated to address the scientific question of interest posed by a trial.

The choice of an estimand involves:

- Population of interest
- Endpoint of interest
- **Measure of intervention effect**

- Taking into account the impact of post-randomization events, e.g.
  - non-compliance,
  - discontinuation of study,
  - discontinuation of intervention,
  - treatment switching,
  - rescue medication,
  - death etc.

Aims of this talk

- Discuss the need for a new framework
- Illustrate the choice of estimands with real and hypothetical examples
  - Dapagliflozin – Diabetes trial (above)
  - Toy example – Diabetes trial
    - See talk by Plamen Kozlovski for a clinician's perspective on estimands in diabetes trials

Example – Diabetes trial

- Randomized, 2-arm (drug A and drug B) diabetes trial in patients with type 2 diabetes mellitus (T2DM)
- Endpoint is the change of HbA1c levels to baseline after 24 weeks of randomization
- HbA1c levels are measured at baseline and at 4, 8, 12, 16, 24 weeks
- For ethical reasons, patients are switched to rescue medication once their HbA1c values are above a certain threshold
- Regardless of switching to rescue medication all (!) patients are followed up for the whole study duration, i.e.
  - there are no missing observations in this study
  - patients never discontinue their study medication, unless they start rescue medication

Three potential estimands of interest

Differ only in their 'measure of intervention effect'

	Estimand 1	Estimand 2	Estimand 3
Population	Intended post-approval population of T2DM patients	Intended post-approval population of T2DM patients	Intended post-approval population of T2DM patients
Endpoint	Change of HbA1c level to baseline after 24 weeks of randomization	Change of HbA1c level to baseline after 24 weeks of randomization	Change of HbA1c level to baseline after 24 weeks of randomization
Measure of intervention effect			

Three potential estimands of interest

Differ only in their 'measure of intervention effect'

	Estimand 1	Estimand 2	Estimand 3
Population	Intended post-approval population of T2DM patients	Intended post-approval population of T2DM patients	Intended post-approval population of T2DM patients
Endpoint	Change of HbA1c level to baseline after 24 weeks of randomization	Change of HbA1c level to baseline after 24 weeks of randomization	Change of HbA1c level to baseline after 24 weeks of randomization
Measure of intervention effect	Effect regardless of what treatment was actually received, i.e. • effect of treatment policies 'drug A until start of rescue followed by rescue' versus 'drug B until start of rescue followed by rescue'.		

Three potential estimands of interest

Differ only in their 'measure of intervention effect'

	Estimand 1	Estimand 2	Estimand 3
Population	Intended post-approval population of T2DM patients	Intended post-approval population of T2DM patients	Intended post-approval population of T2DM patients
Endpoint	Change of HbA1c level to baseline after 24 weeks of randomization	Change of HbA1c level to baseline after 24 weeks of randomization	Change of HbA1c level to baseline after 24 weeks of randomization
Measure of intervention effect	Effect regardless of what treatment was actually received, i.e. • effect of treatment policies 'drug A until start of rescue followed by rescue' versus 'drug B until start of rescue followed by rescue'.	Effect of the initially randomized treatments assuming that the treatment effect disappears and no rescue effect occurs after meeting rescue criteria, i.e. • effect of 'drug A until intake of rescue followed by a disappearing drug A effect' versus 'drug B until intake of rescue followed by a disappearing drug B effect'.	

Three potential estimands of interest

Differ only in their 'measure of intervention effect'

	Estimand 1	Estimand 2	Estimand 3
Population	Intended post-approval population of T2DM patients	Intended post-approval population of T2DM patients	Intended post-approval population of T2DM patients
Endpoint	Change of HbA1c level to baseline after 24 weeks of randomization	Change of HbA1c level to baseline after 24 weeks of randomization	Change of HbA1c level to baseline after 24 weeks of randomization
Measure of intervention effect	Effect regardless of what treatment was actually received, i.e. • effect of treatment policies 'drug A until start of rescue followed by rescue' versus 'drug B until start of rescue followed by rescue'.	Effect of the initially randomized treatments assuming that the treatment effect disappears and no rescue effect occurs after meeting rescue criteria, i.e. • effect of 'drug A until intake of rescue followed by a disappearing drug A effect' versus 'drug B until intake of rescue followed by a disappearing drug B effect'.	Effect of the initially randomized treatments had all patients remained on their randomized treatment throughout the study, i.e. • effect assuming patients did not receive rescue medication.

Three potential estimands of interest
Primary Analyses

	Estimand 1	Estimand 2	Estimand 3
Analysis Variable	• Change from baseline to week 24 in HbA1c • All HbA1c values are used, regardless of treatment	• Change from baseline to week 24 in HbA1c • HbA1c values after intake of rescue medication are set to missing	• Change from baseline to week 24 in HbA1c • HbA1c values after intake of rescue medication are set to missing
Primary Statistical Model	ANCOVA model		

Three potential estimands of interest
Primary Analyses

	Estimand 1	Estimand 2	Estimand 3
Analysis Variable	• Change from baseline to week 24 in HbA1c • All HbA1c values are used, regardless of treatment	• Change from baseline to week 24 in HbA1c • HbA1c values after intake of rescue medication are set to missing	• Change from baseline to week 24 in HbA1c • HbA1c values after intake of rescue medication are set to missing
Primary Statistical Model	ANCOVA model	Missing data will be multiply imputed under a 'missing not at random' assumption: • borrow information from placebo arm patients. For every completed data set fit an ANCOVA model. Overall inference is obtained by applying Rubin's rules on the estimates obtained from every imputed/completed data set.	

Three potential estimands of interest
Primary Analyses

	Estimand 1	Estimand 2	Estimand 3
Analysis Variable	• Change from baseline to week 24 in HbA1c • All HbA1c values are used, regardless of treatment	• Change from baseline to week 24 in HbA1c • HbA1c values after intake of rescue medication are set to missing	• Change from baseline to week 24 in HbA1c • HbA1c values after intake of rescue medication are set to missing
Primary Statistical Model	ANCOVA model	Missing data will be multiply imputed under a 'missing not at random' assumption: • borrow information from placebo arm patients. For every completed data set fit an ANCOVA model. Overall inference is obtained by applying Rubin's rules on the estimates obtained from every imputed/completed data set.	Missing data will be multiply imputed under a 'missing at random' assumption: • borrow information from patients in the same treatment arm. For every completed data set fit an ANCOVA model. Overall inference is obtained by applying Rubin's rules on the estimates obtained from every imputed/completed data set.

Three potential estimands of interest
Sensitivity analyses

	Estimand 1	Estimand 2	Estimand 3
Sensitivity Analyses	• Include/exclude covariates • Include/exclude outliers • Relax the normality assumption	• Include/exclude covariates • Include/exclude outliers • Relax the normality assumption • Modify the 'missing not at random' assumption	• Include/exclude covariates • Include/exclude outliers • Relax the normality assumption • Modify the 'missing at random' assumption

Summary

- Estimand reflects the ‘scientific question of interest’
- Offers a framework to formulate a clear and interpretable trial objective, which in turn provides a framework for targeted trial design and conduct
- Enables early discussions with clinicians and regulators to harmonize trial objectives
- Discussions of estimation and sensitivity analysis follow once an estimand was chosen

Potential impact of ICH addendum on our work

- Will likely have implications on trial design, protocol language, trial conduct and statistical analyses
- Identification of estimand(s) at the design stage requires informed discussion with all stakeholder - clinical teams, regulatory agencies, payers, patients
- Certain estimands may require innovative designs and endpoints - thus also new statistical methodologies and potentially new/updated clinical guidances
- Discussions within ICH suggest that estimands which account for treatment adherence may be of value so that
  - a ‘strict ITT estimand’ (= effect regardless of treatment adherence) may not always be the primary choice
  - data collection after study treatment discontinuation may not always be necessary.

NOVARTIS