Non-inferiority trials and the ICH E9 estimands framework

Oliver Keene, GSK
29 Jan 19

Disclosure

– I am a full time employee of GSK and hold shares in the company
– The views expressed in this presentation are personal and do not represent the views of GSK

Outline of talk

▶ What is done currently

- ITT and per protocol analysis

▶ New estimands framework for non-inferiority trials

- Revised role of per protocol analysis

▶ Population for non-inferiority comparison

- ITT or those meeting inclusion/exclusion?

▶ Strategy for intercurrent events

- Treatment policy vs. hypothetical vs. other approaches

▶ Non-inferiority margins

▶ Missing data
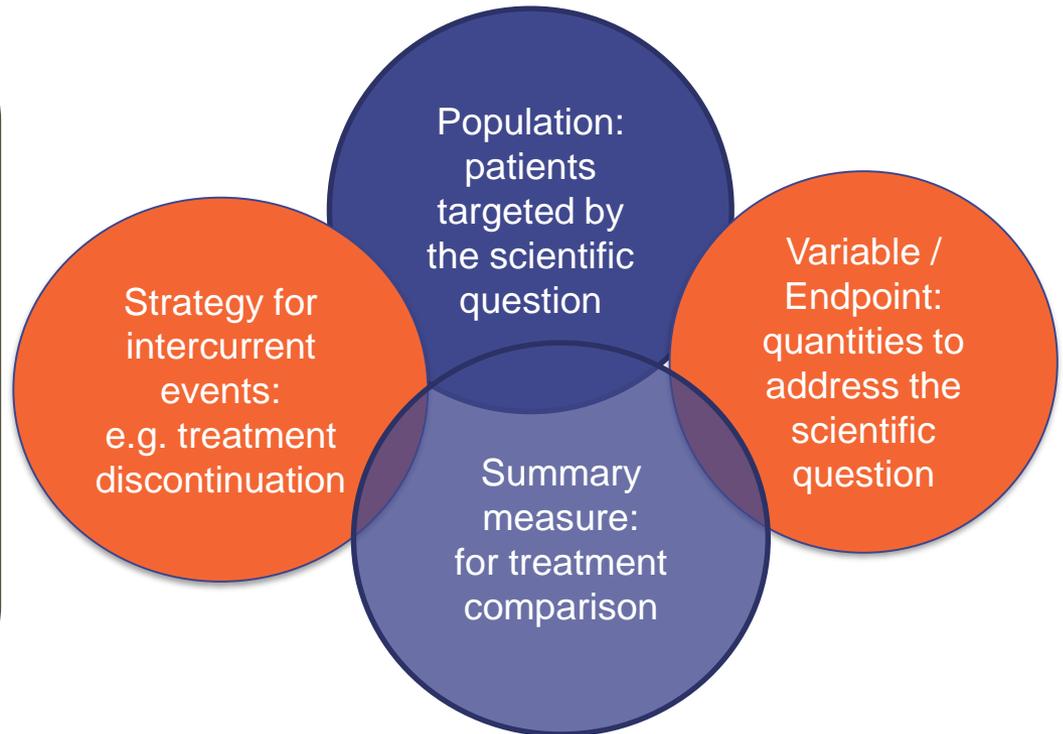
▶ Summary

Non-inferiority: pre-estimand approaches

– "Intent-to-treat" analysis
  – Include all patients randomised and all data following randomisation regardless of protocol deviation
  – Preserves randomisation
  – **Key concern: not necessarily conservative in an NI study**
    – can lead to an incorrect finding of non-inferiority

– Per Protocol analysis
  – Exclude patients who don't fulfil entry criteria
  – Exclude data / patients following an important protocol deviation
  – May be more sensitive to true differences between treatments
  – Potential bias from exclusion of data from randomised patients

– Confidence in conclusion is increased if non-inferiority shown according to both analyses
  EMA points to consider on switching between superiority and non-inferiority:
  "In a non-inferiority trial, the full analysis set and the PP analysis set have equal importance and their use should lead to similar conclusions for a robust interpretation."
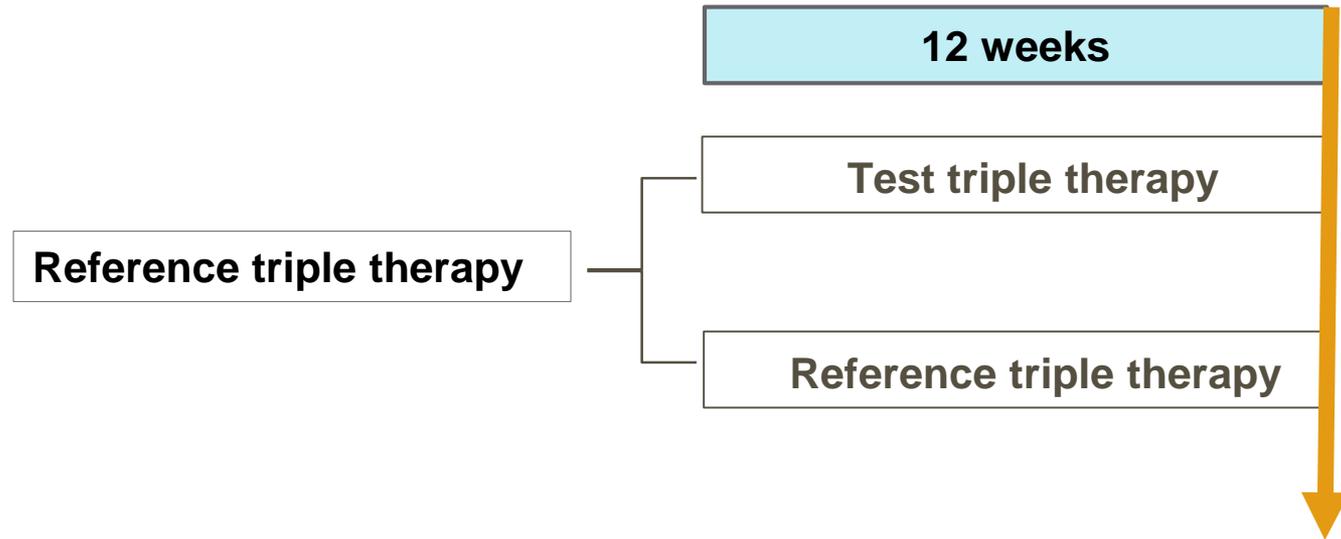
Estimands framework: ICH E9 (R1)

An **estimand** reflects what is to be estimated to address the scientific question of interest

Estimand involves:
- **Population of interest**
- **Variable or endpoint**
- **Summary measure of intervention effect**
- **Strategy for intercurrent events**

Population: patients targeted by the scientific question

Strategy for intercurrent events: e.g. treatment discontinuation

Variable / Endpoint: quantities to address the scientific question

Summary measure: for treatment comparison

Example non-inferiority trial in COPD

12 weeks

Reference triple therapy

Test triple therapy

Reference triple therapy

Variable: Weighted mean FEV_1 over 24h period at 12 weeks

Summary measure: Difference between treatments in mean change from baseline
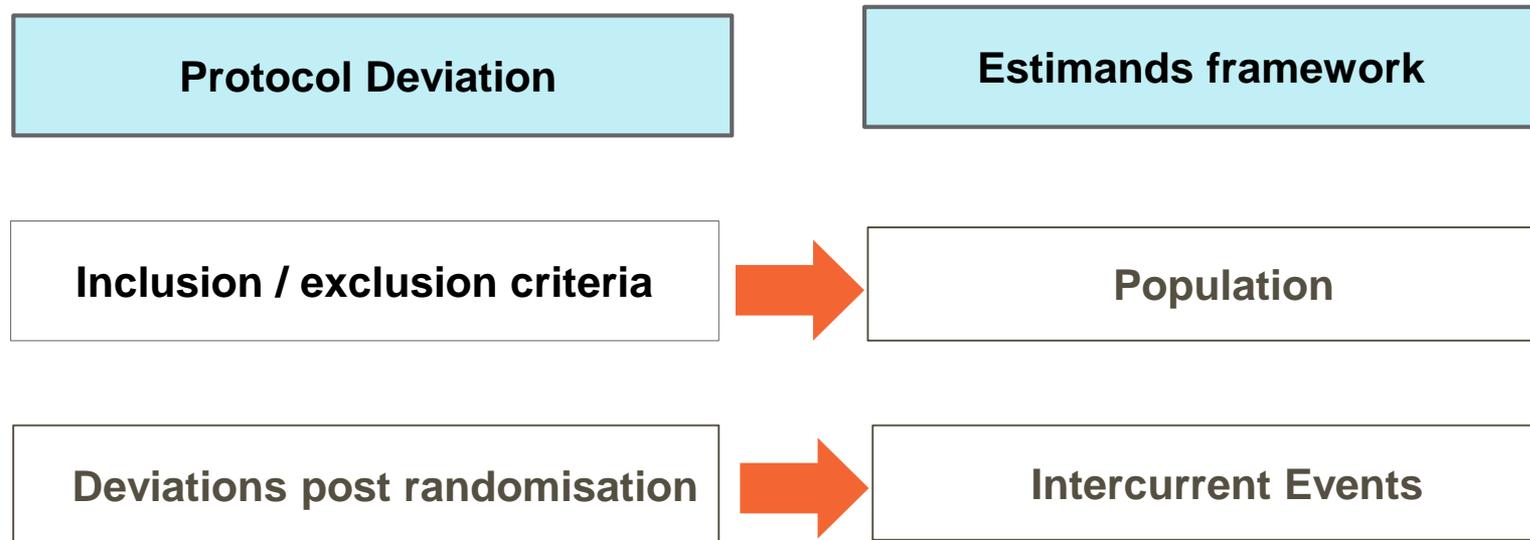
Triple therapy: LABA / LAMA / ICS combination
LABA = long-acting bronchodilator
LAMA = long-acting muscarinic
ICS   = Inhaled corticosteroid

Non-inferiority trials in new estimands framework

Protocol Deviation	Estimands framework

Inclusion / exclusion criteria	→	Population

Deviations post randomisation	→	Intercurrent Events

ICH E9 Addendum:

Analysis of the per-protocol data set does not achieve the goal of estimating the effect in adherent subjects because it does not compare similar subjects on different treatments

The role of such an analysis is therefore limited to investigating whether the extent of protocol violations and deviations compromises confidence in the trial results.

Discussion point - population

- Primary population
    - a) all randomised or
    - b) those who fulfil the inclusion/exclusion criteria (i.e. excluding subjects with pre-randomisation protocol violations)?

    - All randomised strictly preserves randomisation but bias in excluding those with pre-randomisation protocol violations appears minimal / negligible

Discussion point – intercurrent events

Intercurrent events (= Post randomisation protocol deviations)

- Treatment discontinuation due to lack of efficacy or AE

- Treatment discontinuation for other reason e.g. subject choice

- Use of oral steroids for a COPD exacerbation or pneumonia (potentially affects FEV_1 measurement)

- Use of a prohibited medication

- Subject unblinded for safety concern

Strategies for each intercurrent event in the non-inferiority setting? Main choices:

- Treatment policy (closest to ITT) or

- Hypothetical (closest to Per protocol) or

- Some events treatment policy, other hypothetical

Example estimands

| Strategy	Description
Treatment Policy	Difference in mean FEV1 change from baseline at 12 weeks in the population represented by the study inclusion/exclusion criteria comparing *all* patients assigned to test LAMA/LABA/ICS to *all* patients assigned to reference LAMA/LABA/ICS *regardless of discontinuation of assigned treatment, regardless of the addition or switches to alternative medications, regardless of use of oral steroids or treatment unblinding*
Hypothetical	Difference in mean FEV1 change from baseline at 12 weeks in the population represented by patients *who fulfill the study inclusion/exclusion criteria* comparing patients assigned to test LAMA/LABA/ICS to patients assigned to reference LAMA/LABA/ICS *in a hypothetical scenario where patients do not discontinue assigned treatment, have no addition or switches to alternative medications, no use of oral steroids*

Treatment policy vs. hypothetical

▶ Reflects same debate as for superiority studies

- Extra issue is concern that treatment policy blunts true differences between treatments

- Potential bias towards conclusion of NI

Possible approaches:

▶ Co-primary estimands: one estimand using treatment policy and one using hypothetical?

- Patient / physician may prefer one primary estimate of treatment effect

▶ One primary estimand using treatment policy, a supportive estimand using hypothetical?

- Primary of hypothetical only?

▶ If different strategies for intercurrent events for non-inferiority and superiority testing

- lead to different point estimates and confidence intervals
- Complicates shift from non-inferiority to superiority testing

Margin for non-inferiority
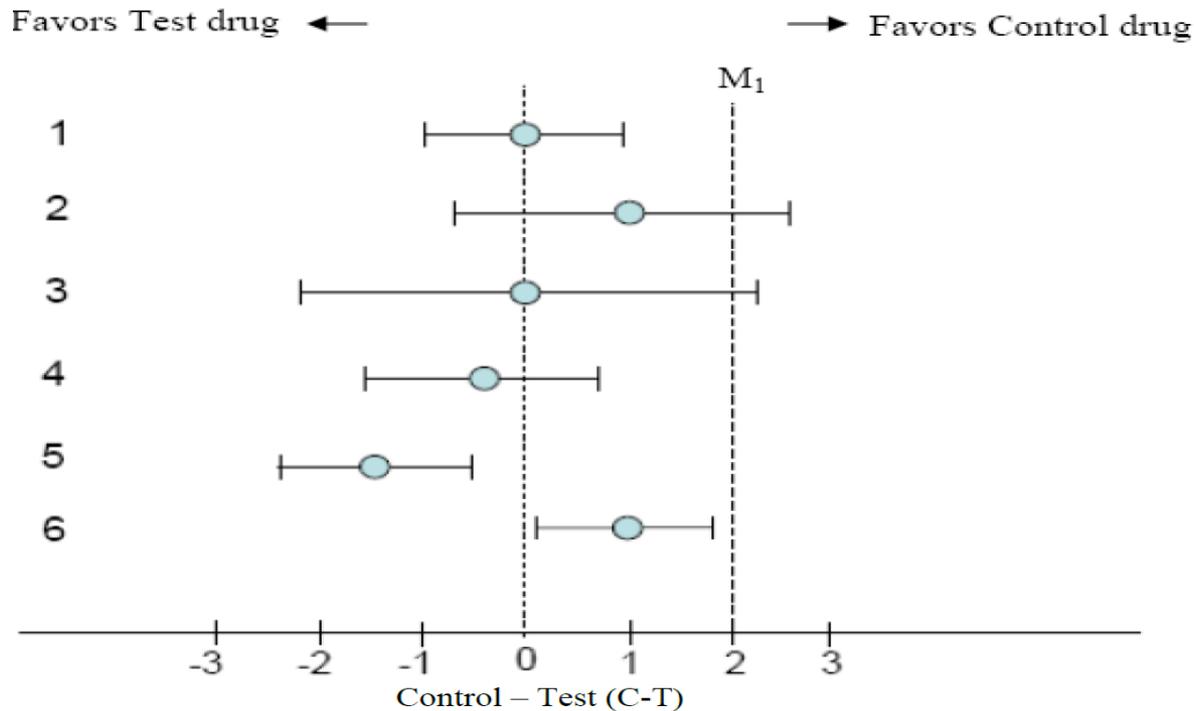
- Margin for non-inferiority typically based on previous superiority studies
- Superiority studies were based on treatment policy strategy for intercurrent events
- How to determine margin if hypothetical strategy is used in the NI study?
  - Should it be the same as for a treatment policy strategy?

Other strategies for intercurrent events

▶ Principal strata

- Seeks to compare treatment effect in strata of patients who would tolerate both treatment (or at least the reference treatment)
- Requires assumptions regarding whether patients would have completed treatment on the opposite treatment to which they were assigned

▶ While on-treatment

- May be similar to hypothetical in analysis
- Depends whether clinical outcome of interest is clinical status at end of study or summary of experience while receiving treatment

▶ Composite

- Not proposed here as change in endpoint relative to superiority study

Missing data in non-inferiority trials

Period 1 Observed On-treatment	Period 2 Observed Off-treatment	Period 3 Unobserved (Off-treatment)

▶ Treatment policy requires follow-up after randomised treatment discontinuation

▶ Typically will still be missing data due to patients withdrawing consent

▶ **Reverse concern to superiority trials: missing data may show treatment effect, exclusion may bias to no effect**

▶ Possible option:

- Impute missing outcomes under MAR within treatment groups
- Reduce imputed outcomes for test group by NI margin (Koch 2008)
- Alternative is to use tipping point approach, how large does penalty for missing have to be in test to change conclusion of non-inferiority

Estimands based on intercurrent events

- ICH E9 addendum refers to comparison of intercurrent events between treatments

- *Estimands could be constructed to directly address those intercurrent events which can lead to the attenuation of differences between treatment arms e.g. use of rescue medications … In this situation, the estimand might target a measure of treatment effect with high sensitivity to detect differences between treatments, if they exist*

- Required to describe frequency of intercurrent events

  - May be useful in settings such as pain to formally compare amount of rescue with associated non-inferiority margin

  - In general for events such as discontinuation of randomised treatment, outcome is proportion of subjects and will be unable to set non-inferiority margins on these proportions

Summary: estimands in non-inferiority trials

- Estimand framework has led to a reappraisal of use of Per Protocol Population

  - Subjects who do not fulfil inclusion/exclusion criteria may be excluded based on the population of interest

  - Protocol violations after randomisation are intercurrent events

  - Role of a Per Protocol Population substantially diminished or eliminated entirely.

- Strategy for intercurrent events not established yet

  - Demonstration of non-inferiority may be required under both treatment policy and a hypothetical strategy for key intercurrent events

- Possible increased focus on comparing frequency of important intercurrent events across treatments