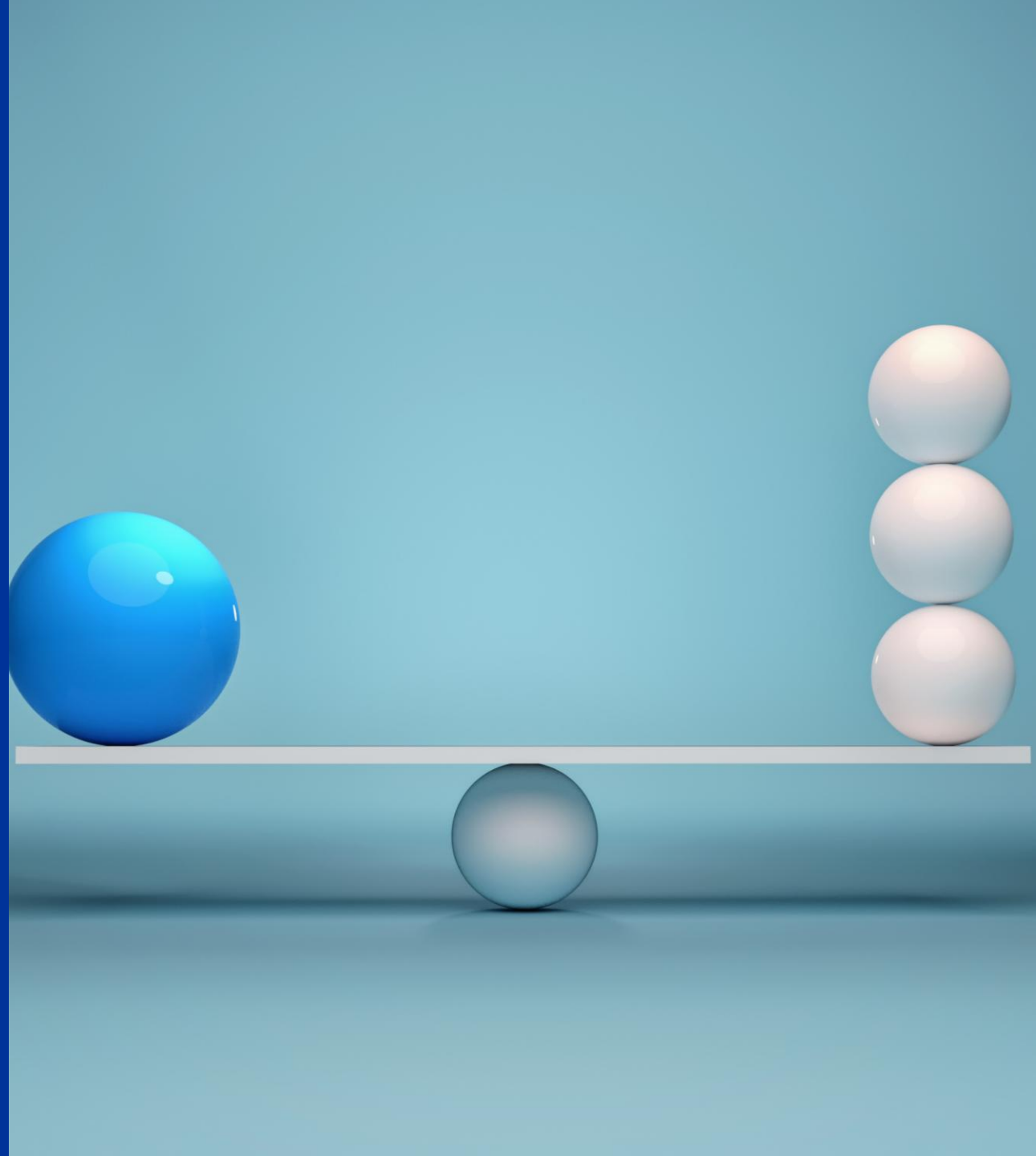
EMA
EUROPEAN MEDICINES AGENCY

Draft guideline on non-inferiority and equivalence comparisons in clinical trials

Key changes in the draft guideline with potential impact on future practice and received public comments

August 2026

Susanne Urach, Jonathan Bergman, Florian Lasch



Background

- **Draft** updated guideline published in October 2025
- The guideline replaces
  - the 'Guideline on the choice of the non-inferiority margin' (EMEA/CPMP/EWP/2158/99) and
  - The 'Points to consider on switching between superiority and non-inferiority' (CPMP/EWP/482/99)
- The public consultation closed on the 31st of May 2026 - comments **not** yet implemented


EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

1 24 October 2025
2 EMA/301654/2025
3 Committee for Medicinal Products for Human Use (CHMP)/
4 Methodology Working Party (MWP)

5 Guideline on non-inferiority and equivalence comparisons
6 in clinical trials
7 Draft

# of comments: 482
General: 77
Specific: 405
# of commentors: 19
(7 industry, 3 regulatory, 9 others)

EMA and FDA Non-Inferiority Guidance Release History

How to calculate the margin is not described in this document, and there was little published experience on how to do this.

YEAR (FIRST PUBLISHED)	AGENCY	GUIDELINE TITLE
2000	ICH	Choice of Control Group in Clinical Trials (ICH-E10)
2000	EMA	Points to consider on switching between superiority and non-inferiority
2005	EMA	Guideline on the choice of the non-inferiority margin
2016 (2010)	FDA	(Draft) Non-Inferiority Clinical Trials to Establish Effectiveness Guidance
2024	EMA	Concept Paper for NI & Equivalence Comparisons
2025	EMA	Draft Guideline on NI & Equivalence Comparisons

EMA and FDA Non-Inferiority Guidance Release History

Delta can be defined as the lower bound of t minus r that ensures that the lower bound of the indirect confidence interval of t minus p will be above zero.

YEAR (FIRST PUBLISHED)	AGENCY	GUIDELINE TITLE
2000	ICH	Choice of Control Group and Related Issues in Clinical Trials (ICH-E10)
2000	EMA	Points to consider on switching between superiority and non-inferiority
2005	EMA	Guideline on the choice of the non-inferiority margin
2016 (2010)	FDA	(Draft) Non-Inferiority Clinical Trials to Establish Effectiveness Guidance
2024	EMA	Concept Paper for NI & Equivalence Comparisons
2025	EMA	Draft Guideline on NI & Equivalence Comparisons

EMA and FDA Non-Inferiority Guidance Release History

FDA provided more details on how to choose the margin (differing between the fixed margin and synthesis approach)

YEAR (FIRST PUBLISHED)	AGENCY	GUIDELINE TITLE
2000	ICH	Choice of Control Group and Related Issues in Clinical Trials (ICH-E10)
2000	EMA	Points to consider on switching between superiority and non-inferiority
2005	EMA	Guideline on the choice of the non-inferiority margin
2016 (2010)	FDA	(Draft) Non-Inferiority Clinical Trials to Establish Effectiveness Guidance
2024	EMA	Concept Paper for NI & Equivalence Comparisons
2025	EMA	Draft Guideline on NI & Equivalence Comparisons

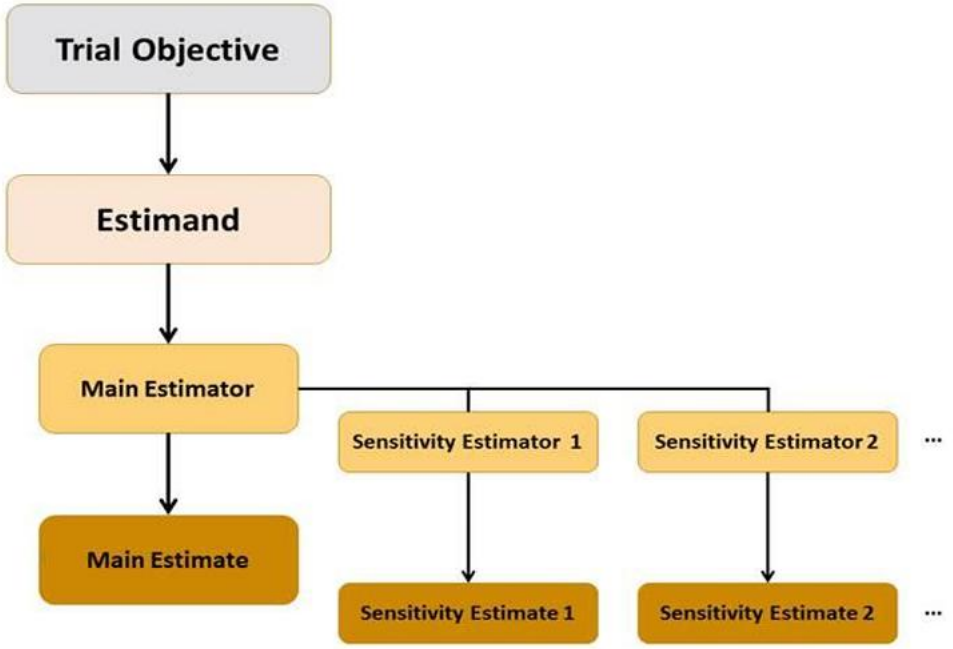
EMA and FDA Non-Inferiority Guidance Release History

Consolidation of previous guidance documents to align more closely with modern regulatory science, estimand framework, and statistical expectations

YEAR (FIRST PUBLISHED)	AGENCY	GUIDELINE TITLE
2000	ICH	Choice of Control Group and Related Issues in Clinical Trials (ICH-E10)
2000	EMA	Points to consider on switching between superiority and non-inferiority
2005	EMA	Guideline on the choice of the non-inferiority margin
2016 (2010)	FDA	(Draft) Non-Inferiority Clinical Trials to Establish Effectiveness Guidance
2024	EMA	Concept Paper for NI & Equivalence Comparisons
2025	EMA	Draft Guideline on NI & Equivalence Comparisons

Contents of the Draft GL

Contents of the Draft GL



Objectives of non-inferiority and equivalence comparisons

- **Clear objectives:**
  - Absolute efficacy: The new treatment is better than no treatment/placebo, but an active comparator is used. The analysis is essentially an indirect comparison between the test treatment and a putative placebo arm.
  - Relative efficacy: The new treatment is not worse than a clinically acceptable amount.
  - Therapeutic equivalence
  - Non-inferior safety
- **Why not always aim for showing relative efficacy?**

  - Showing absolute benefit risk often enough for marketing authorization.
  - The objective influences the non-inferiority margin and thereby affects sample size and study feasibility.

Contents of the Draft GL

EMA

Estimands

- Usually, one primary estimand is expected.
- Supplementary estimands are important to complement the primary estimand (e.g. to understand the role of additional / rescue medication) and need pre-specification.

Recommendation by type of non-inferiority trial:

- Demonstrating absolute or relative efficacy
  - The scientific question of interest should drive the choice of population, endpoint, ICE strategy: e.g. the intended indication, most relevant clinical endpoint
  - The constancy assumption might be violated by design -> preserving a defined fraction of the comparator's effect as safeguard
- Demonstrating therapeutic equivalence
  - 'Sensitive estimand' (population, endpoint, intercurrent event strategies)

Estimands

- Each objective (e.g. an additional objective to test the superiority of the new treatment over the active comparator) might require a dedicated estimand that could differ from each other.
- Concerning ICEs no general guidance is possible for NI comparisons:
  - Exception: For ICE ‘switching to active comparator’, treatment policy is not acceptable
  - A higher / lower rate of the intercurrent event in one or both treatment arms of the new trial can bias the comparison in favour of the new treatment
  - The frequency and pattern of relevant intercurrent events can be informative about relevant differences between the compared treatments
- Some guidelines only include a discussion on the preferred estimand for a superiority comparison, which should not be understood as the default preference of the same estimand for a non-inferiority setting.
- The efficacy of the active comparator should be established in the sought indication.

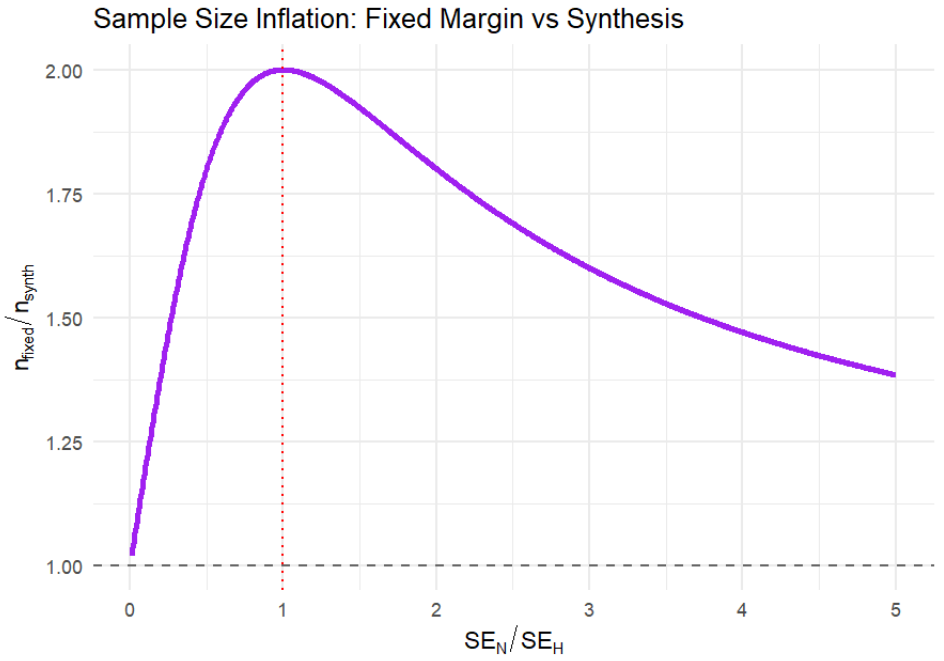
Contents of the Draft GL

Selecting the margin

- Selected based upon a combination of statistical reasoning and clinical judgement
- **Statistical justification**: ensures superiority over (a putative) placebo (demonstrating absolute efficacy). Clinical judgement needed to decide on safeguard due to violations of the constancy assumption.
- **Clinical justification**: ensures that the test product is not inferior to the comparator by more than a clinically acceptable amount (demonstrating relative efficacy, non-inferior safety, equivalence comparisons).
- Clinical justification of the margin **not** required for absolute efficacy objective only
- The margin chosen for demonstrating relative efficacy should logically also be small enough to demonstrate absolute efficacy as well.
- An **advantage in another important aspect** should **not** be used for justifying a larger than usually acceptable non-inferiority margin and claiming relative efficacy

Fixed margin vs synthesis approach

- Only the **fixed margin approach** is possible for showing **relative efficacy, non-inferior safety or equivalence comparisons.**
- Two possible approaches for showing **absolute efficacy:**
  1. **Fixed margin (e.g. 95%-95% method)**
  2. **Synthesis approach**
- An **estimate and confidence interval for the comparison against putative placebo** should be reported.
- The fixed margin approach is the synthesis approach with an inflated variance (see annex of GL). Same assumptions apply.



Discussion points

- Two-stage framework: Stand-alone interpretation of a trial (without external data) as step 1 at least?
- Studies including a non-inferiority comparison should always be assessed holistically, and the context provided by the relative efficacy versus the active control should not be disregarded by limiting the study objective solely to the demonstration of absolute efficacy.
- Focused primarily on clinical trials intended to assess the efficacy of a treatment: Removing non-inferior safety objective from the guideline?
- Guidance on estimating NI margin where there are no placebo-controlled trials of the comparator
- Guidance for when it is not possible to reconstruct the estimand of a historical study
- Do additional estimand(s) also have to meet pre-specified NI margins to conclude a study has demonstrated non-inferiority? Should the margin not depend on the chosen estimand?

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

Any questions?

susanne.urach@ages.at

Follow us

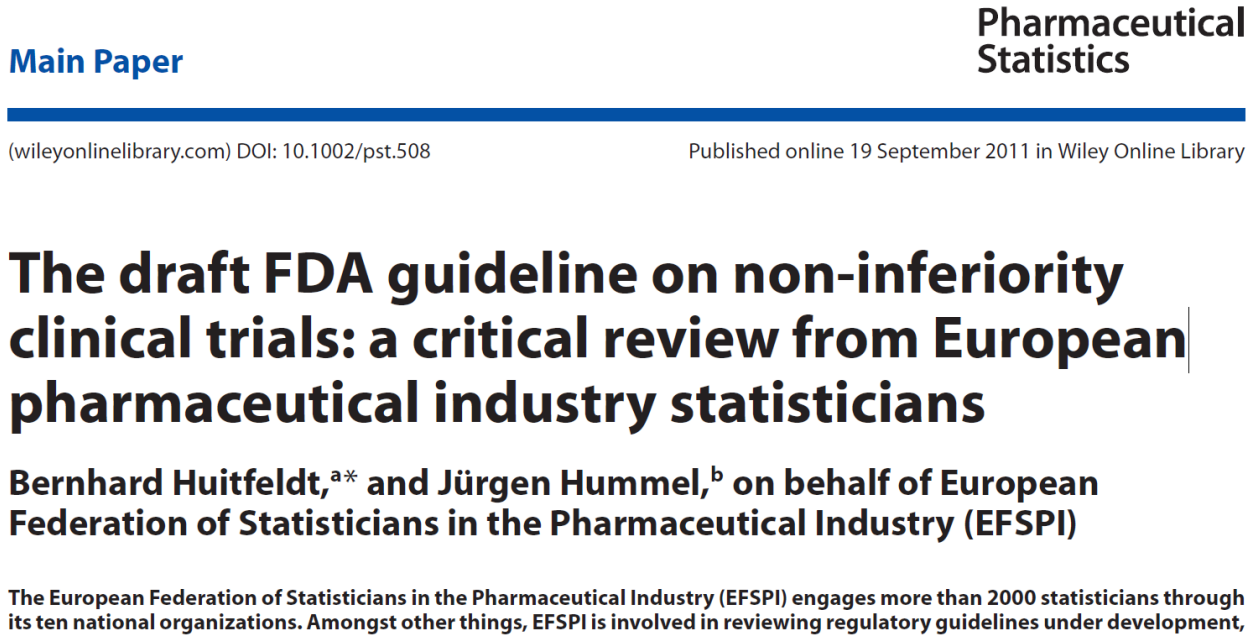
Summary of novelties

- Clear objectives: absolute efficacy – relative efficacy – equivalence - NI safety
- Aligned methodological approach to the objective: objective – estimand – main analysis
- Stronger role of pre-specification and impact assessments (both before and after the trial) for assay sensitivity/constancy assumption
- Clinical margin justification not needed for absolute efficacy objective only
- Additional approach for showing absolute efficacy: synthesis approach
- For relative efficacy: an advantage in another important aspect does not justify planning for a larger non-inferiority margin than usually acceptable
- Areas where it is difficult to justify a non-inferiority margin of any size, e.g. superiority using an increased significance level not mentioned anymore
- Analysis: no default "ITT and PP analyses": sufficiently conservative main analysis
- When switching from non-inferiority to superiority, the type-1 error is not controlled if there are secondary endpoints unless the tests are included in a multiple testing procedure.

Addressed some critical comments on FDA guidance

Critical comments

- Single standard for proving efficacy
- Fraction of effect to be preserved
- Statement of primary hypothesis
- Synthesis method versus fixed margin method



Main Paper

Pharmaceutical Statistics

(wileyonlinelibrary.com) DOI: 10.1002/pst.508

Published online 19 September 2011 in Wiley Online Library

The draft FDA guideline on non-inferiority clinical trials: a critical review from European pharmaceutical industry statisticians

Bernhard Huitfeldt,[a]* and Jürgen Hummel,[b] on behalf of European Federation of Statisticians in the Pharmaceutical Industry (EFSPI)

The European Federation of Statisticians in the Pharmaceutical Industry (EFSPI) engages more than 2000 statisticians through its ten national organizations. Amongst other things, EFSPI is involved in reviewing regulatory guidelines under development,

BUT: Focuses on absolute efficacy not relative efficacy
EMA: It is not appropriate to define the non-inferiority margin as a proportion of the difference between active comparator and placebo.

EMA

Testing absolute efficacy

Fixed margin approach	**Synthesis approach**
Involves pre-specifying a margin based on historical studies of the active comparator.	Combines data from historical trials with data from the current trial to derive an estimate and confidence interval for the comparison of the new product to a putative placebo.
Formally **tests a non-inferiority objective**, misinterpretation as demonstrating relative efficacy $\mu_T - \mu_{C_{new}} < -M$	Formally **tests a superiority objective** which directly corresponds to demonstrating absolute efficacy $\mu_T - \mu_P < 0$
Estimate and confidence interval compares **new treatment to comparator**	Estimate and confidence interval compares **new treatment to putative placebo**

Why does the fixed margin approach even work for showing absolute efficacy?

Testing absolute efficacy

Fixed margin approach	**Synthesis approach**
Involves pre-specifying a margin based on historical studies of the active comparator.	Combines data from historical trials with data from the current trial to derive an estimate and confidence interval for the comparison of the new product to a putative placebo.
Formally **tests a non-inferiority objective**, misinterpretation as demonstrating relative efficacy $\mu_T - \mu_{C_{new}} < -M$	Formally **tests a superiority objective** which directly corresponds to demonstrating absolute efficacy $\mu_T - \mu_P < 0$
Estimate and confidence interval compares **new treatment to comparator**	Estimate and confidence interval compares **new treatment to putative placebo**

Why does the fixed margin approach even work for showing absolute efficacy?

Assuming the whole lower bound of the 95% confidence interval for the effect estimate is used as M:

$$M = \overline{c_{hist}} - \bar{p} - 1.96\, SE(\overline{c_{hist}} - \bar{p})$$

$\bar{t} - \overline{c_{new}} + \overline{c_{hist}} - \bar{p} - 1.96\left(SE(\bar{t} - \overline{c_{new}}) + SE(\overline{c_{hist}} - \bar{p}) \right) > 0$	Lower $\bar{t} - \overline{c_{new}} + \overline{c_{hist}} - \bar{p} - 1.96\sqrt{SE(\bar{t} - \overline{c_{new}})^2 + SE(\overline{c_{hist}} - \bar{p})^2} > 0$

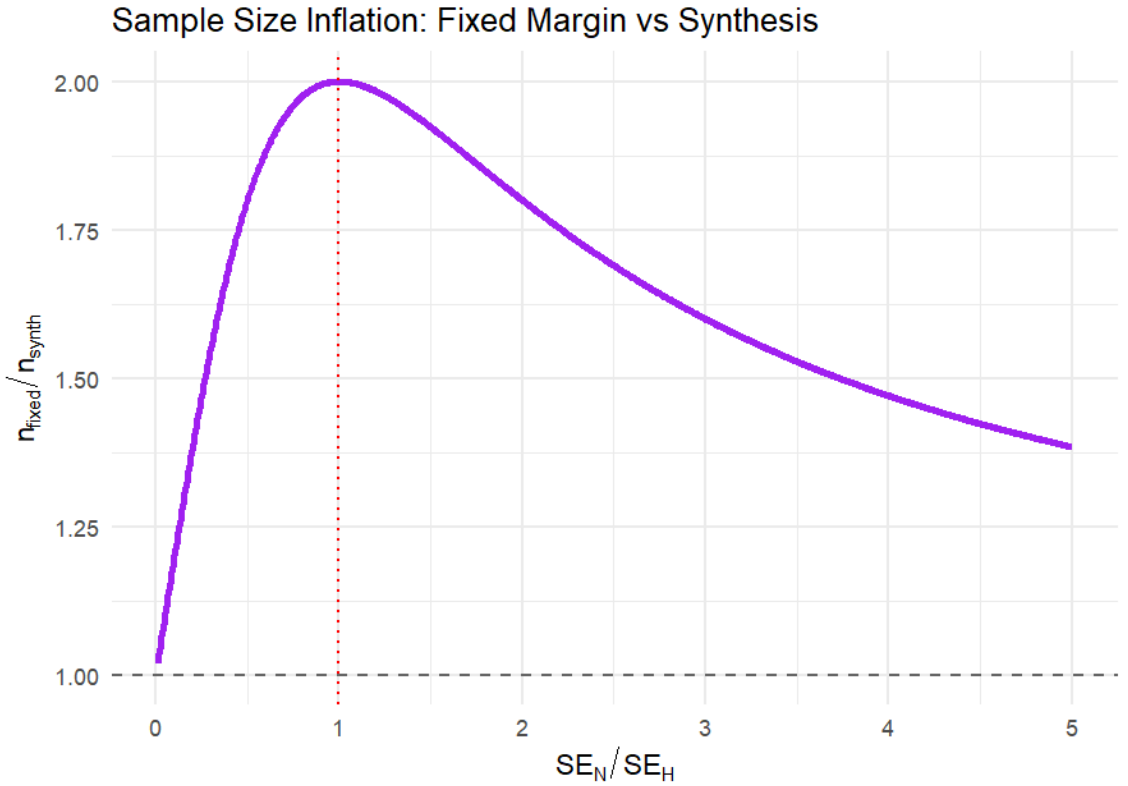
triangle inequality

$$a + b \geq \sqrt{a^2 + b^2}$$

EMA

Fixed margin vs synthesis approach

Fixed margin approach	Synthesis approach
Studies for calculating the effect of the comparator against placebo must be **pre-specified**	
Can be made more **conservative** by **preserving a defined fraction** of the comparator's effect	
Use external data for showing absolute efficacy	
more conservative	Substantial increase in power possible
In addition, delivers a stand-alone relative comparison (even when not showing relative efficacy)	Potential use of population-adjusted indirect comparison methods or Bayesian methodologies to relax the constancy assumption



The most efficient way to control the type 1 error rate is always by choosing no variance inflation at all, i.e. the synthesis approach with a discounting factor. (Snapinn and Jiang, 2008)

Assay sensitivity and trial quality

- Assay sensitivity is the ability of a clinical trial to distinguish an effective treatment from a less effective or ineffective treatment
- The constancy assumption assumes that the effect of the comparator treatment (over a putative placebo) in the new trial is the same as the effect over placebo in the old trial.
- Stronger role of pre-specification and impact assessments (both before and after the trial) in the protocol / CSR
- The pre-defined non-inferiority margin may no longer be appropriate at all if the comparator performs differently from what was assumed when defining the non-inferiority margin.

Assay sensitivity and trial quality - Comments

- Distinctions and similarities between assay sensitivity and the constancy assumption
- Examples where violations of the constancy assumption are acceptable?
- Meaningful differences to historical results for the active comparator could be quite subjective.
- Explain what should be checked regarding the constancy assumption

Estimands -Comments

- Constancy assumption can only be made for the same estimand
- Guidance for when it is not possible to reconstruct the estimand of a historical study
- Do additional estimand(s) also have to meet pre-specified NI margins to conclude a study has demonstrated non-inferiority?
- Comparator treatment not the best but still indicated/not indicated
- The change to a more sensitve population is very risky.
- How is 'sensitive endpoint' defined and when does it make sense?
- Guidance for intercurrent events that risk the treatment policy estimand to go towards the null

Synthesis vs fixed margin approach

- For showing superiority compared to putative placebo, the fixed margin approach is the synthesis approach with an inflated variance (see annex of GL). Same assumptions apply.
- Lower bounds of 95% CIs for no safeguard (q=0)

- FM: $\bar{t} - \overline{c_{new}} + \overline{c_{hist}} - \bar{p} - 1.96\left(SE(\bar{t} - \overline{c_{new}}) + SE(\overline{c_{hist}} - \bar{p}) \right) > 0$
- Syn: $\bar{t} - \overline{c_{new}} + \overline{c_{hist}} - \bar{p} - 1.96\sqrt{SE(\bar{t} - \overline{c_{new}})^2 + SE(\overline{c_{hist}} - \bar{p})^2} > 0$

  - Fixed margin approach more conservative
  - $r \to 0$: SE_N/SE_H small (new trial very precise compared to historical): both methods similar
  - $r \to \infty$: SE_N/SE_H large (historical trial very precise compared to new): both methods similar
  - Maximum inflation occurs when $SE_N/SE_H=1$
  - Notable increase in power possible using synthesis approach.

triangle inequality

$$a + b \geq \sqrt{a^2 + b^2}$$

Statistical considerations

- **Analysis sets**

  Per-protocol sets are discouraged (‘The default use of per-protocol sets should generally be avoided.’)

  All randomised patients should typically be included.

  Use the estimand framework to handle treatment discontinuations and other post-baseline events.

- **Sample size re-estimation**

  Should be avoided because it increases the type-1 error (even if it is blinded)

- **Missing data**

  Should be handled conservatively:

    Equivalence: A bias towards similarity is not ok.

    Non-inferiority: A bias towards similarity can be ok, but only if it disadvantages the test treatment.

  Two imputation methods are discussed:

    Reference-based imputation (discouraged for primary analyses).

    Imputation using off-treatment information (recommended).

Switching between objectives or changing the objective

Replaces Points to consider on switching between superiority and non-inferiority (CPMP/EWP/482/99)

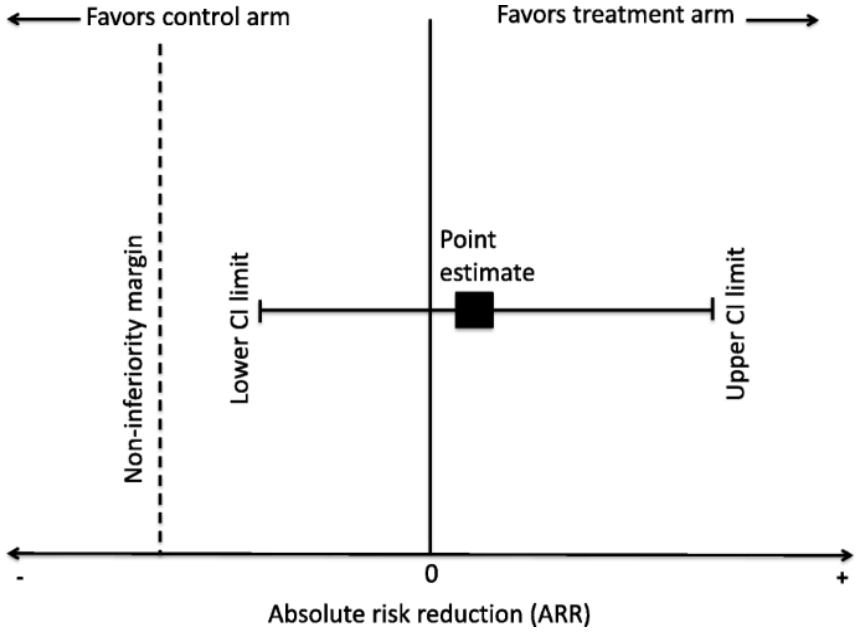
A single trial may aim to demonstrate:

- **Absolute efficacy** (vs placebo)
- **Relative efficacy** (vs active comparator)
- **Superiority** (vs active comparator)

When including multiple objectives:

- All objectives must be **pre-specified in the protocol**
- Must include a **multiplicity control strategy**
- Must ensure proper **Type I error control**
- May require **different estimands**
- Trial design must be suitable for **each objective**

The same confidence interval is used for multiple tests by moving the NI margin:



Changing the Objective Mid-Trial is strongly discouraged