STATISTICAL CONSIDERATIONS IN NONINFERIORITY CLINICAL TRIALS

1. INTRODUCTION

Clinical trials are the core of evidence-based medical practice. They rely greatly on statistical methodologies to extract meaningful conclusions from complex sets of data. This project intends to provide a meticulous analysis of the role played by statistics in clinical trials, with a specific focus on noninferiority trials.

Clinical trials constitute the gold standard for assessing the efficacy and safety of medical interventions. Statistics strengthen these trials and guide researchers in designing study protocols and analysing data to derive scientifically valid conclusions.

In the context of clinical trials, statistics serve a dual purpose. First, they assist in the design of methodologically sound studies (calculations of sample sizes, power and significance). But they also decode the intricacies of obtained data, transforming information into insights capable of informing clinical practice.

Unlike traditional superiority trials, where a new treatment is evaluated against a placebo or an active control to establish superiority, noninferiority clinical trials introduce a different objective: their primary goal is to demonstrate that a new treatment is not significantly worse than an existing standard treatment.

This shift in focus requires meticulous consideration of statistical nuances, particularly in defining an acceptable margin of noninferiority, a predefined threshold beyond which the new treatment is deemed clinically acceptable.

The relevance of noninferiority trials is underlined by ethical and practical considerations in contemporary clinical research. When it is unethical to conduct a placebo-controlled trial, noninferiority trials offer a viable alternative. They become essential when it is necessary to establish the benefits of a new treatment over existing standards, taking into account factors such as safety, convenience, compliance and cost-effectiveness.

As the landscape of healthcare evolves, with innovations being more incremental than revolutionary, the focus shifts from providing outright superiority to establishing noninferiority. Noninferiority trials balance between statistical rigour and clinical relevance, and the determination of noninferiority margins requires a careful evaluation regarding these matters.

Throughout this project, we will delve into several statistical considerations inherent to noninferiority trials, including aspects such as sample size, noninferiority margin, population analysis or assay sensitivity. The aim is to improve the theoretical understanding of future researchers about the application of statistics in impactful noninferiority clinical research.

2. TYPES OF CLINICAL TRIALS: SUPERIORITY, EQUIVALENCE AND NONINFERIORITY

Clinical research relies on carefully planned trials, each designed to address specific scientific questions. In this section, we shall break down different types of clinical trials, with a focus on the unique statistical aspects that distinguish superiority, equivalence, and noninferiority trials.

Superiority trials

Superiority trials, essential in clinical research, aim to prove that a new intervention is better than a comparison (placebo or active control). The key statistical goal is to present strong evidence against the null hypothesis, stating that the effects of the new treatment are significantly better than the comparator.

A carefully determined sample size ensures that the trial is sensitive enough to detect meaningful differences, strengthening the reliability of the findings. The statistical process in superiority trials is relatively straightforward, involving significance testing with p-values and confidence intervals. This approach yields a clear outcome: either the new treatment is superior, or it is not.

Raquel Romero González
MSc in Pharmacological Research – Data Analysis

Equivalence trials

Equivalence trials, a more nuanced component of the trio, deal with the intricate task of showing that the effects of two treatments are not significantly different. Achieving this delicate balance demands statistical skill to establish equivalence within a predetermined margin of clinical significance.

The statistical complexity in equivalence trials arises from the bidirectional nature of hypothesis testing. In contrast to the clear-cut results of superiority trials, equivalence trials involve accepting a null hypothesis that suggests a predefined difference between treatments. Confidence intervals play a critical role in this process, capturing the allowable range within which the true difference is likely to be found.

Noninferiority trials

Noninferiority trials, the central focus of our analysis, revolve around a specific goal: to confirm that a new treatment is not significantly worse than a standard treatment by more than a predefined margin. This shift in approach introduces unique statistical challenges.

At the core of noninferiority trials is the establishment of the noninferiority margin, a predetermined threshold that outlines the maximum acceptable degree of inferiority. Statistical analysis in these trials goes beyond simply rejecting the null hypothesis; it involves ensuring that the lower end of the confidence interval falls within the defined margin. The careful determination of this margin requires a thoughtful combination of clinical judgment, statistical reasoning, and historical evidence.

A) Superiority Trial   B) Equivalence Trial   C) Non-Inferiority Trial
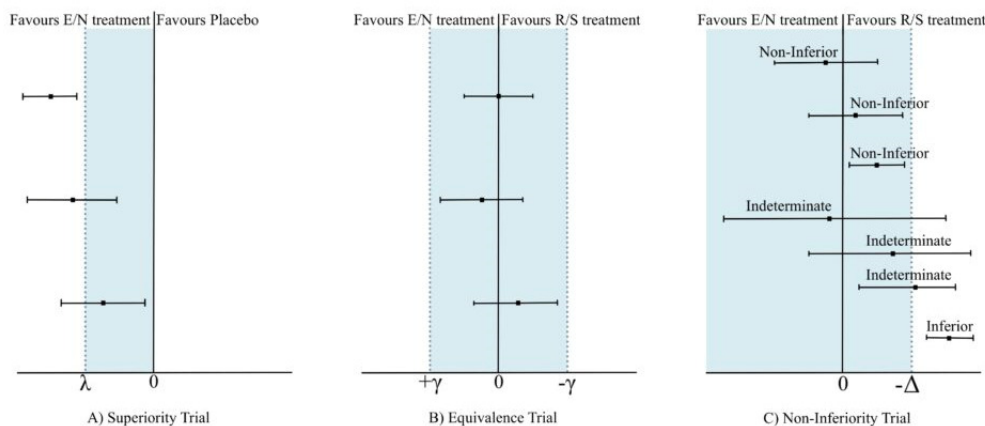
3. ESTABLISHMENT OF THE NONINFERIORITY MARGIN

In noninferiority trials, the noninferiority margin assumes a critical role, serving as a predefined threshold that marks the extent of clinical noninferiority. This section examines the noninferiority margin, analysing its function in establishing the limit for acceptable clinical noninferiority.

The noninferiority margin is a pre-established threshold delineating clinical noninferiority. Its determination involves a forward-looking approach, requiring foresight and a deep comprehension of the clinical context. Researchers are tasked with foreseeing the range within which the actual difference between experimental and standard treatments may lie without clinical significance.

In order to establish the noninferiority margin, a key point of the provided by the FDA is to ascertain the smallest plausible benefit offered by the standard strategy, indicated by the lower limit of the 95% confidence interval (CI) or the CI closest to no effect. While the lower bound of the 95% CI can serve as the noninferiority margin ($-\Delta$), several considerations influence this decision:

- Preservation of effect is important, ensuring the new treatment never is less effective than placebo. The noninferiority margin is linked to a percentage of the lower limit of the 95% CI.
- Assay sensitivity emphasizes the undeniable superiority of the standard treatment over placebo.
- The consideration of previous meta-analyses aids in the setting of the noninferiority margin.

Raquel Romero González
MSc in Pharmacological Research – Data Analysis

The prospect of a therapeutic intervention being potentially worse than a gold standard poses challenges in informing and enrolling participants, emphasising the need for careful consideration and communication in trial design. It is particularly important to be conservative regarding the noninferiority margin when considering drugs with important toxic effects.

In summary, the noninferiority margin is not just a number in statistics but a mix of clinical importance and accurate statistical measurement. Defining it beforehand requires a careful combination of foresight, clinical insight and statistical rigour.

4. TRIAL DESIGN AND ASSAY SENSITIVITY

Consistency is crucial when dealing with noninferiority clinical trials. To make reliable claims, it is necessary to be watchful of variations in how the trial is carried out, the people involved, and how the outcomes are measured. Keeping things standardized (following the same methods throughout the trial) is vital. Making false claims is a real risk, especially if there's no consistent approach.

Differences in how the trial is conducted could introduce bias, making it hard to trust claims of noninferiority. If the patients included in the trial have different characteristics or get different standards of care, it might affect the understanding of the treatment effects. Different measurement methodologies could create confusion and affect the reliability of the study.

The principle of standardization emerges as a safeguard against these challenges. Adhering to consistent methodologies and protocols increases study reliability and validity. Commitment to uniformity extends across the trial, including participant selection criteria, treatment administration, data collection methods, and the application of statistical analyses.

Noninferiority trials are characterised by the importance of precise statistical methodologies, and the adherence to them becomes an ethical responsibility, as it ensures that noninferiority claims are held by robust evidence, inspiring confidence in the validity of trial outcomes and their subsequent implications for clinical practice.

Assay sensitivity refers to a trial's ability to distinguish between effective and ineffective treatments. In the context of clinical trials, it is crucial for evaluating the efficacy of a new treatment compared to a standard or control. Unlike superiority trials, noninferiority trials face challenges in verifying assay sensitivity, particularly as a placebo arm is absent. They frequently resort to historical evidence as a guide: historical control groups, drawn from previous studies or established databases, serve as a reference for anticipated outcomes. However, this reliance introduces complexities, as shifts in patient populations, standard care or study methodologies over time may affect the comparability of historical data.

The design of noninferiority trials calls for a nuanced strategy to address the challenges associated with assay sensitivity. Opting for a comparison with an active control with established efficacy underlines the importance of meticulous trial planning to ensure the selected control aligns with the therapeutic context and has the sensitivity to detect treatment effects.

5. SAMPLE SIZE CONSIDERATIONS

Deciding on the sample size is a crucial point that demands careful thought. This section explores the details of sample size considerations, highlighting the sensitivity in this decision-making process. Factors such as anticipated treatment effects, assumptions about treatment effectiveness, and the nuanced interplay with the noninferiority margin are thoroughly examined. The sample size required for a noninferiority clinical trial may be calculated using this formula:

$$n = \frac{(r + 1)(Z_{1-\beta} + Z_{1-\alpha})^2 \sigma^2}{r((\mu_A - \mu_B) - d_{NI})^2}$$

Population variance (σ^2)

Allocation ratio (r)   Type II error (β)   Type I error (α)

Difference between mean of treatments A and B

Non-inferiority Limit (d_{NI})

Raquel Romero González
MSc in Pharmacological Research – Data Analysis

Expected treatment effect

One of the main aspects to consider when calculating sample size is the estimation of treatment effects. Predicting a treatment effect that is too modest may result in an underpowered study, while an overestimation may lead to an inflated sample size. Achieving precision calls for a careful assessment of the expected treatment effects, based on a deep understanding of the therapeutic concepts.

The estimated effectiveness, quantified through parameters like response rates or mean differences, serves as the foundation for these calculations. However, the uncertainty of these assumptions adds complexity, requiring a meticulous exploration to align statistical power with precision.

Impact of the noninferiority margin

The consideration of the noninferiority margin significantly influences decisions about sample size, as it has an impact on statistical power. As the margin broadens, the required sample size tends to increase, impacting the trial's ability to detect genuine differences. A broader margin, while providing a laxer standard for noninferiority, requires a larger sample size to keep statistical robustness.

Caution against inflated margins

Inflated margins pose the risk of compromising statistical power, jeopardising the trial's ability to distinguish real clinical differences from chance. As we have previously discussed, setting the noninferiority margin requires precision and a balance between statistical reasoning, historical evidence and clinical understanding.

6. INTENTION-TO-TREAT VS. PER-PROTOCOL ANALYSIS

The selection of an analysis population is an important decision. There are two main approaches to this choice: intention-to-treat (ITT) analysis and per-protocol analysis. Each strategy can carry inherent biases, so this section will discuss each of them, together with their advantages and drawbacks.

Intention-to-treat analysis

Intention-to-treat (ITT) analysis takes a practical approach, as it analyses participants based on their initially assigned treatment, regardless of posterior deviations or non-compliance. This strategy mimics the real-world clinical practice, as it considers unpredictabilities encountered during the trial course.

In current clinical research, there is increasing endorsement for ITT analysis, as it ensures that the comparison between groups reflects the real situations seen in clinical practice. This means that ITT analysis increases external validity by including complexities such as patient compliance, dosing mistakes and waiting times.

However, adhering to real-world dynamics comes with inconveniences. ITT analysis can introduce biases that dilute the estimated treatment effects. The inclusion of patients that did not follow through with the intended treatment can mitigate the true efficacy of the intervention.

Per-protocol analysis

Opposed to ITT analysis, we can find per-protocol analysis. This approach focuses strictly on those participants that have adhered to trial protocol without major deviations. It intends to isolate the "pure" treatment effect by excluding those who deviate, providing insight into the efficacy of the treatment under optimal conditions.

Per-protocol analysis provides greater internal validity than ITT, as it is able to detect the true potential of the intervention, without disturbances from deviations, drop-outs or non-compliance.

Nevertheless, there is a toll on external validity, as only a selective sample of participants are analysed, potentially overestimating the real efficacy of the treatment. In the end, this approach introduces bias by narrowing the focus to a sample that matches idealised conditions.

Raquel Romero González
MSc in Pharmacological Research – Data Analysis

Comprehensive analysis strategies

Considering the contrast between these two approaches, it is possible to come up with an integration of both ITT and per-protocol analyses. This approach would aim to combine the strengths of each, while mitigating their bias. Presenting exploratory results from both analyses allows a deeper understanding of the real effect of the studied intervention or treatment. Also, analysing different outcomes with specific population analyses can improve the understanding of real effect while keeping caution in safety analysis.

This strategy seeks to increase the reliability of the conclusions and provide a complete perspective that addresses both real world circumstances and internal validity.

7. SWITCHING BETWEEN SUPERIORITY AND NONINFERIORITY

When approaching noninferiority trials, the doubt might arise on whether it is possible to interpret the results of noninferiority as superiority, and vice versa. There is consensus on attributing credibility to a noninferiority trial transitioning to superiority. However, it is important to keep caution when following the reverse approach, as previous planning in the trial protocol is necessary to maintain the credibility of the conclusions.
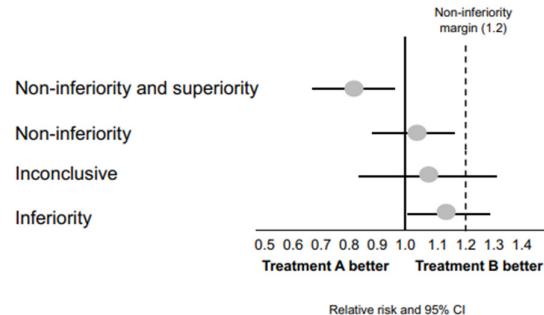
The foundation of a reliable transition is on the predefined hypothesis. If the aim is to prove that a new treatment is not significantly worse than the current standard, data could not only establish noninferiority but also suggest superiority. Therefore, a credible change to a superiority perspective would not incur in statistical detriment.

Opposed to this, adopting a noninferiority perspective after establishing a superiority methodology without previously planning this in the trial protocol requires caution, as this approach is susceptible to data dredging and selective reporting. For example, in the figure below, it would not be correct to claim noninferiority for the drugs marked in green and red, as that would be just one of many possibilities.
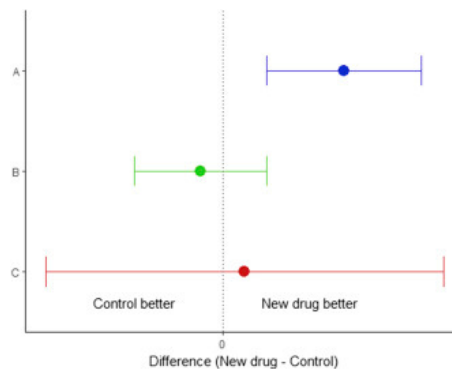
The main consideration in this regard is to meticulously plan the trial protocol. Predefined hypotheses prevent unwarranted shifts in perspective and ensure the trustworthiness of the results. A well-designed trial protocol should specify the conditions under which a noninferiority trial can legitimately be interpreted as proving (or hinting at) superiority. In the same manner, the trial protocol should establish whether a superiority trial could be turned into a noninferiority or equivalence trial in the case of not reaching statistical significance in the primary outcome.

In conclusion, changing between superiority and noninferiority approaches is acceptable if this shift is specified in the trial methodology, avoiding statistical pitfalls.

Raquel Romero González
MSc in Pharmacological Research – Data Analysis

8. CONCLUSIONS

Statistics are a guiding principle in clinical trials, as they shape the trajectory from study design to conclusive interpretations. This project has carefully explored statistical considerations in clinical research, with a focus on noninferiority trials.

Unlike traditional superiority trails, noninferiority trials shift towards establishing that a new treatment is not significantly (or unacceptably) worse than the existing standard. This change in perspective calls for an exploration of statistical matters, being the most important of them the determination of the noninferiority margin.

The noninferiority margin is a key threshold that delineates acceptable clinical noninferiority. Its establishment demands a careful balance between clinical insight and statistics, taking into account historical research regarding the standard treatment.

Even though consistency is fundamental in all clinical trials, the delicacies of statistics in noninferiority trials make it even more important in the case that concerns us. The clinical implications of wrongly accepting as noninferior a new treatment that actually is carries very important consequences for patients. Therefore, it is of utmost importance to carefully standard the methodology of noninferiority trials as a safeguard against false claims.

Sample size considerations unwind the balance between statistical power and the noninferiority margin, cautioning against inflated margins that jeopardise the trial's ability to detect genuine differences.

The tension between intention-to-treat and per-protocol analyses highlighted the need for a comprehensive strategy, combining the strengths of each to ensure robust and reliable conclusions and try an avoid the inherent bias of each of the separate strategies.

Finally, this project delved on the careful considerations that must be taken when interpreting noninferiority as superiority and vice versa, emphasising the need of predefined protocols to maintain result credibility.

In conclusion, this project illustrates the intricacies of statistics in the less well-known noninferiority clinical trials. All of the covered aspects must be carefully crafted in order to come up with valid and trustworthy trial conclusions. As clinical research evolves, the spotlight on noninferiority trials becomes brighter, demanding further commitment to precision, consistency and methodological rigour with the ultimate goal of achieving meaningful medical advances.

9. FURTHER INFORMATION

For those readers that wish to delve deeper into statistical considerations in noninferiority clinical trials, these articles might be of use:

- Non-inferiority statistics and equivalence studies.
  Doi: https://doi.org/10.1016%2Fj.bjae.2019.03.004
- Non-inferiority trials: understanding the concepts
  Link: Non-Inferiority Trials: Understanding the Concepts - Tutorials and Fundamentals (cochrane.org)
- Statistical issues and recommendations for noninferiority trials in oncology: a systematic review.
  Doi: https://doi.org/10.1158/1078-0432.CCR-11-1653
- Understanding noninferiority trials
  Doi: https://doi.org/10.3345%2Fkjp.2012.55.11.403
- Methodological aspects of superiority, equivalence, and non-inferiority trials
  Doi: https://doi.org/10.1007/s11739-020-02450-9
- Practical guide to sample size calculations: non-inferiority and equivalence trials
  Doi: https://doi.org/10.1002/pst.1716

Raquel Romero González
MSc in Pharmacological Research – Data Analysis