



# Statistical Design and Analysis of a Non-Inferiority Clinical Trial

Áine Glynn<sup>1</sup>, Filip Kłosowski<sup>1</sup>

## Background

Non-inferiority (NI) trials are increasingly used in clinical research<sup>1,2</sup>. However, they are frequently poorly designed and interpreted. Common sources of confusion include the specification of NI margins, selection of appropriate active controls, and interpretation of statistical conclusions, which can lead to adverse consequences for manufacturers, clinicians, and the wider public.

*What is a Non-Inferiority Trial?*

- Non-Inferiority Trials:** Clinical studies designed to demonstrate a new treatment is not clinically worse than an active control by more than a pre-specified margin<sup>3</sup>.
- Non-Inferiority Margin ( $\Delta$ ):** Pre-specified & approved threshold the new treatment must meet to prove it preserves a clinically meaningful portion of the active control's effect<sup>2</sup>.
- NI Trials typically run like a randomised control trial but comparing the new treatment with an **active control**- an established standard of care used, instead of a placebo.
- NI must show the new treatment's estimated effect, along with its confidence interval, lies within the pre-specified NI margin.

## Aims & Objectives

- To justify the  $\Delta$  using historical evidence.
- To develop and assess appropriate NI trial design parameters.
- To develop a Shiny application for exploring sample size requirements.
- To analyse simulated data and evaluate NI conclusions.
- To examine the assumptions and limitations of NI designs.

## Case-study: Biomimics 3D stent

It is 2008 and we have just been hired as statisticians for BioMimics 3D Vascular Stent's pivotal NI trial:



The BioMimics 3D stent is a peripheral vascular stent implanted in the leg or arm to improve blood flow in patients with peripheral vascular disease (narrowing of the peripheral blood vessels). Unlike conventional straight stents, it features a three-dimensional helical design, intended to improve vascular performance and blood flow in affected vessels.

From a statistical perspective, this study is not a traditional two-arm NI trial, but a single-arm trial evaluated against a fixed Performance Goal (PG). In medical device development, particularly for implantable medical devices, two-arm non-inferiority trials are often impractical due to cost, extended timelines, and challenges with recruitment and blinding. In such settings, regulators may accept a single-arm design, provided that the device's observed performance, together with its confidence interval, exceeds a pre-specified PG within an agreed non-inferiority margin.

The PG and corresponding NI margins are derived from a random-effects meta-analysis of safety and effectiveness outcomes from existing, approved comparator devices. For safety, the margin is set at the upper bound of the 95% confidence interval, representing the maximum acceptable level of harm. For efficacy, the margin is set at the lower bound of the 95% confidence interval, representing the

minimum clinical benefit that must be preserved. Crossing either bound results in failure to demonstrate NI.

Our project involves working with BioMimics' lead scientist, to design, assess, and interpret this trial from a statistician's perspective. This includes:

- Defining and justifying appropriate safety and efficacy endpoints.
- Ensuring the trial is statistically powered and ethically justified.
- Selecting valid analysis methods.
- Correctly interpreting and communicating non-inferiority conclusions.

## Meta-Analysis & Evidence base

Prior to being available on the market all medical products must obtain approval from the relevant regulatory authority. This is done through a series of clinical studies where the medical product is monitored for safety and effectiveness.

Initially, a feasibility study is conducted. This is a small, single-arm prospective study for a narrowly defined patient population, designed to ensure preliminary safety and device performance. If successful, a pivotal clinical study is proposed. Our pivotal study is a single-arm objective performance goal study which requires pre-specified and approved primary efficacy and safety endpoints. These endpoints must be statistically justified, adequately powered, and grounded in clinical evidence of safety and benefit in the target population.

### Evidence-Based Safety Performance Goals (30 Days)

**Safety Endpoint:** Must be able to identify potential harm, including any adverse events (AEs), serious adverse events (SAEs), lab abnormalities, or physical changes<sup>3</sup>.

Endpoint	Pooled Estimate	95% CI	Recommended Performance Goal	Justifiable Range
30-Day Amputation	0	[0.0000, 1.0000]	1%	0-1%
30-Day Death	0	[0.0000, 1.0000]	1%	0-1%
30-Day Target Vessel Revascularisation (TVR)	0.0517	[0.0234, 0.1104]	11%	2-11%

### Evidence-Based Efficacy Performance Goals (12 Months)

**Efficacy:** The treatment's capacity to produce a desired, measurable, and beneficial therapeutic effect under ideal and strictly controlled conditions<sup>3</sup>.

Endpoint	Outcome	Pooled Estimate	95% CI	Recommended Performance Goal	Justifiable Range
Rutherford Classification Change (12 months)	Improved or No Change	0.9583	[0.8786, 0.9865]	87%	87-99%
Rutherford Classification Change (12 months)	Increase by One Class	0.0441	[0.0143, 0.1280]	1%	1-13%

To determine appropriate safety and efficiency endpoints for the Biomimics pivotal study we conducted a targeted literature review of clinical trials previously run involving peripheral vascular devices. This included a meta-analysis which we analysed & utilized at an individual participant data (IPD) level to justify our chosen endpoints<sup>4-6</sup>.

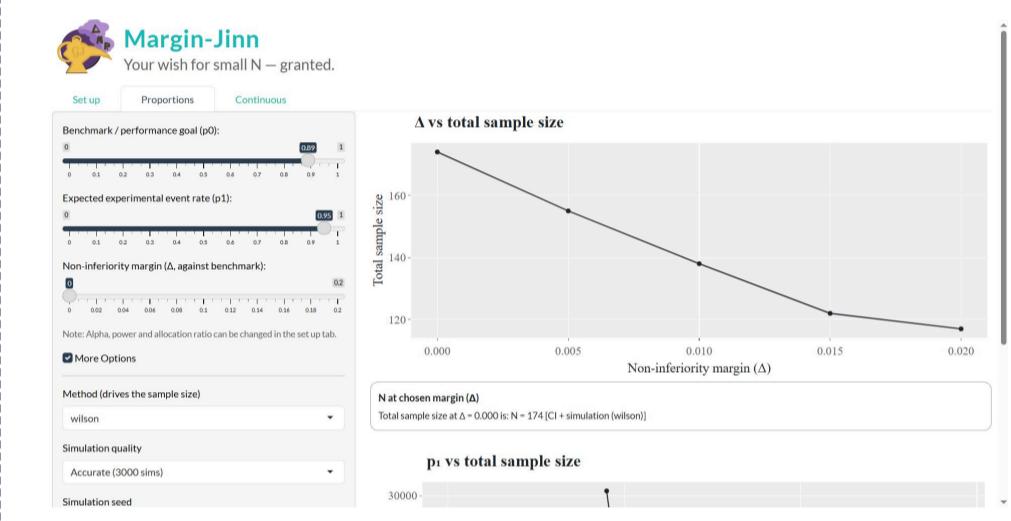
## Sample Size Determination

With performance goals fixed, sample size was calculated for freedom from 30-day TVR using a single-arm performance-goal framework. NI was assessed using a one-sided CI-based decision rule ( $\alpha = 0.025$ , power = 90%), with  $p_0 = 0.89$  derived from

historical evidence and expected performance  $p_1 = 0.95$  based on the pooled estimate. A margin of  $\Delta = 0$  was used, reflecting no allowance beyond the pre-specified performance goal. Sample size was determined via Monte Carlo simulation using Wilson confidence intervals, chosen to avoid overly conservative intervals when success rates are high. Under these assumptions, the required total sample size was  $n = 219$ .

## Margin-Jinn

Margin-Jinn is an interactive Shiny application developed as part of this project to support statistical planning for NI trials. The app allows users to explore how sample size requirements vary under different design assumptions, including power, significance level, effect size, and choice of confidence-interval method.



Multiple approaches to interval estimation for proportions (e.g. Clopper-Pearson, Wilson, Agresti-Coull) are supported, enabling evaluation of how methodological choices influence NI conclusions, particularly in single-arm trials with small sample sizes and low event rates.

In this study, Margin-Jinn is applied to the BioMimics 3D case study to justify and critically assess sample size decisions under realistic design scenarios.

GitHub: <https://github.com/FilipMKgit/Margin-Jinn>

## Next Steps

- Extend the BioMimics 3D case study through simulation of NI trial data under the defined Performance Goals.
- Finalise a Statistical Analysis Plan (SAP) specifying endpoints, estimators, confidence-interval methods, and NI decision rules prior to data unblinding.
- Analyse simulated trial outcomes to assess NI conclusions under alternative assumptions.
- Conduct sensitivity and tipping-point analyses to evaluate the robustness of conclusions to small changes in assumptions or observed event counts.

## References

- Sandie A et al. Non-inferiority test for a continuous variable with a flexible margin in an active controlled trial: An application to the Stratall ANRS 12110 / ESTHER trial. *Trials* 2022;23:202.
- Cuzick J, Sasieni P. Interpreting the results of noninferiority trials — a review. *British Journal of Cancer* 2022;127:1755–1759.
- U.S. Food and Drug Administration. Non-inferiority clinical trials to establish effectiveness: Guidance for industry. Silver Spring, MD, USA: U.S. Department of Health & Human Services, 2016;
- Werk M et al. Inhibition of restenosis in femoropopliteal arteries: Paclitaxel-coated versus uncoated balloon — the femoral paclitaxel randomized pilot trial. *Circulation* 2008;118:1358–1365.
- Tepe G et al. Local delivery of paclitaxel to inhibit restenosis during angioplasty of the leg. *The New England Journal of Medicine* 2008;358:689–699.
- 6.