Falling in the margin

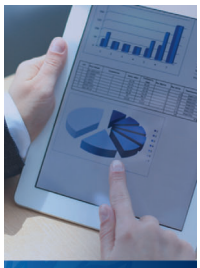
Middleton, Lee Jonathan

DOI:
10.1111/1471-0528.16702

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Middleton, LJ 2021, 'Falling in the margin: Randomised controlled trials with a non-inferiority design', *BJOG: An International Journal of Obstetrics and Gynaecology*. https://doi.org/10.1111/1471-0528.16702

BJOG An International Journal of Obstetrics and Gynaecology
Royal College of Obstetricians & Gynaecologists



BJOG RESEARCH METHODS GUIDES
Series edited by Karla Hemming and Natalie AM Cooper

Falling in the margin

Randomised controlled trials with a non-inferiority design

BJOG PERSPECTIVES

LEE JONATHAN MIDDLETON, UNIVERSITY OF BIRMINGHAM, BIRMINGHAM, UK

Learning points

- Sometimes new treatments with known advantages, such as being better tolerated, easier to administer or more accessible, might be good alternatives to the standard treatment, despite potentially being less effective.
- Non-inferiority trials allow the assessment of whether these new treatments are less effective, but only by an acceptable degree.
- Determining non-inferiority should be based on comparing confidence intervals of treatment differences with a suitable predetermined margin, known as the non-inferiority margin.

In a randomised controlled trial (RCT) we are usually trying to answer the question of which treatment is best. Sometimes, though, we may be considering a treatment that we do not believe will perform better than an existing option but has other benefits: it could be cheaper, easier to deliver or more acceptable to patients. For example, the need for urodynamic testing compared with simple office evaluation before stress incontinence surgery was evaluated in an RCT (Nager et al. *N Engl J Med* 2012;366:1987–97). The authors believed that simple office-based evaluation would be less invasive than urodynamic testing and that it may also reduce the risk of urinary tract infections. It was also expected that simple office-based evaluation was unlikely to reduce treatment success.

Issues with non-inferiority designs

The key consideration is how much worse in terms of outcome is going to be acceptable. There is no consensus about how this should be evaluated, but intuitively the difference should be small and should take into account the nature of the outcome, balanced with the potential benefits. One option is to ask clinicians and patients what level of difference would cause them to switch treatments. This level is termed a non-inferiority margin and determines the size of the trial.

Confidence intervals are used to determine whether a new treatment is 'not unacceptably worse' than a standard treatment. If the upper end of the confidence interval (CI) for the treatment effect is smaller than the non-inferiority margin, then we can conclude non-inferiority (scenario B in Figure 1).
A number of other design issues should be considered in this setting and may be why some consider non-inferiority trials to have more weaknesses than superiority trials. These include: (i) ensuring that the efficacy of the standard comparator has been proven on a similar population; (ii) selecting an appropriate analysis population in light of a reverse hypothesis (i.e. in trying to demonstrate similarity, per-protocol populations could be considered alongside intention-to-treat populations); and (iii) taking care with trial conduct, as carelessness (such as poorly executed randomisation) may lend itself to the conclusion of similarity. Further reading on these issues is recommended (see below).

Example

The primary outcome in the study described above was treatment success (i.e. symptoms that have substantially improved) at 12 months. The non-inferiority margin was set at 11 percentage points. The study found that in patients tested with urodynamics the treatment success rate was 76.9% (203/264), and for patients undergoing simple office evaluation the treatment success rate was 77.2% (200/259). The percentage difference between the two groups was –0.3% (95% CI –7.5% to 6.9%). This confidence interval tells us that office-based evaluation might be as much as 6.9% worse (upper end of CI) or as much as 7.5% better (lower end of CI) than evaluation with urodynamics. The results therefore indicated that office evaluation is certainly less than 11% worse. The authors declared non-inferiority, recommending that urodynamics need not be performed.

Useful resources

- General discussion: Mulla et al. *JAMA* 2012;308:2605–2611.
- CONSORT reporting: Piaggio et al. *JAMA* 2012;308:2594–2604.

Disclosure of interests

None declared. Completed disclosure of interests form available to view online as supporting information.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article. ■

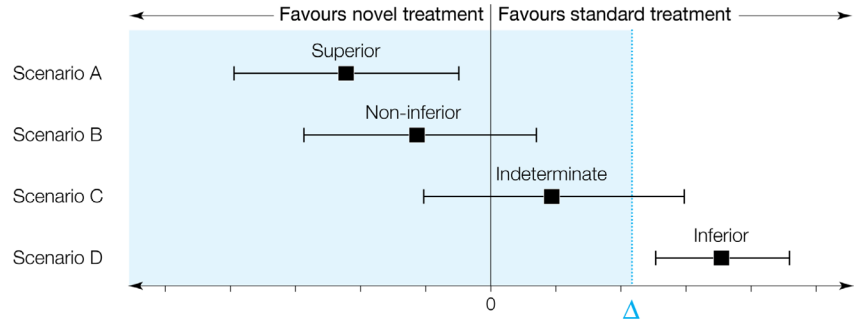


Figure 1. Possible outcome scenarios in a clinical trial. The triangle indicates the non-inferiority margin. Black lines indicate confidence intervals around treatment effect estimates (square boxes), which were risk differences in this example.