EDITORIAL

EQUIVALENCE AND NON-INFERIORITY TRIALS IN A SNAPSHOT

Salima Kerai
Department of Paediatrics, Aga Khan University, Karachi-Pakistan

Randomized controlled trials conventionally entails superiority hypothesis where researcher is interested to know whether a new treatment is better than the standard of care. However, developing and testing novel techniques or therapies which are better than standard is not always feasible. There are circumstances when effective treatment exists and new treatment is not substantially better than existing treatment and conducting placebo controlled trials are unethical. Then goal of the investigator changes; if new treatment is equivalent or non-inferior in comparison to current standard of care with respect to a-priori set endpoints. The current paper aims to discuss few key principles of equivalence and non-inferiority trial design and some challenges to think about before designing or conducting these trials.

Keywords: Equivalence trials, Non-inferiority trials

J Ayub Med Coll Abbottabad 2017;29(3):371-2

Generally, when we think of a trial hypothesis of interest, we think of superiority hypothesis whereby the investigator intends to determine that a new (experimental) treatment is better than the control treatment (standard or placebo). In superiority trials researcher claims that a new treatment is different (two sided) or better (one sided) than the control arm and the null hypothesis is that there is no difference between the treatment arms. These trials are useful in testing new therapies (preventive, therapeutic or rehabilitative) in order to establish a standard of care or test therapies which are better than the standards in place.

Sometimes developing novel techniques or therapies which are better than the standard of care are more difficult or treatment effect differs no more than a specific amount. Then goal of the investigator changes; if the new treatment is equally safe as thecurrent standard of care. For this type of question, equivalence design is the most suited where the researcher's hypothesis is that difference between two treatments is equal to predefined 'x' which is called equivalence margin. In other words, researcher tends to prove that new treatment is no better or no worse than standard of care as far as difference between two treatments is within equivalence margin.

In a multicentre randomized controlled trial at district level in Uganda, investigators claimed that the safety and effectiveness of diagnosing and treatment of incomplete abortion with misoprostol was equivalent between midwives and physicians.[2] They randomly allocated women with first-trimester incomplete abortion to clinical assessment and treatment with misoprostol either by a physician or a midwife. The Primary outcome was complete abortion not needing surgical intervention within 14–28 days after initial treatment with predefined equivalence of -4% to 4%. The estimated risk difference for midwives versus physicians group was -0.8% (95% CI -2.9 to 1.4) which was falling within equivalence margin proving abortion with misoprostol by midwives is equally safe and effective when provided by physicians.

Besides, equivalence design requires a larger sample size in order to rule out large differences and to have a high probability of detecting a difference within a margin. In equivalence designs researchers' hypothesis are revered as compared to superiority hypothesis. For the equivalence design null hypothesis is that there is a difference between two groups and researcher claims that there is no difference or difference is within the equivalence margin (two sided). Correspondingly, type I and type II errors are also flipped. In equivalence design type I error is failure to reject the null when difference exists and type II error is wrongly rejecting null when there is no difference.

However, situations arise where investigators are unable to accept claims of equivalence. There are circumstances when effective treatment exists and new treatment is not substantially better than existing treatment and conducting placebo controlled trials are unethical. This issue can be addressed by non-inferiority hypothesis where investigator claims that new treatment is 'no worse' than existing treatment by more than a pre-specified amount which is called non-inferiority (NI) margin.

Similar to equivalence trials, we are detecting small differences in non-inferiority trials with reversed hypothesis and type I and type II errors. While hypothesis of non-inferiority trials are one sided as compared to equivalence hypothesis which are two sided. Therefore, lesser sample size is required in non-inferiority trials to reject the null. Because essentially in non-inferiority trials, we are showing that new treatment's response if worse, still sufficiently close to the established treatment response, so that new treatment is as good as or not worse than the established treatment.

Non-inferiority trials offer other advantages in experimenting therapies for endpoints like better safety profile or lower cost or ease of administration or compliance which are important from a clinical point of view. For example, in case of childhood pneumonia, amoxicillin is standard of care. Testing any new drug against placebo would be unethical in presence amoxicillin. In such case, the investigator can aim to test non-inferiority of new drug in comparison to amoxicillin using lower cost or safety as clinical endpoints.

In an undergoing double blind randomized controlled trial in Karachi, Pakistan, investigators claimed that 'no treatment' (3 days of placebo) is non-inferior to 'treatment' (3 days of amoxicillin-WHO standard of treatment) in management of children with WHO defined fast breathing pneumonia.[1] Children 2–59 months of age with fast breathing, without any danger sign are randomly allocated to receive either three days of placebo or amoxicillin. Primary outcome is the difference in cumulative treatment failure between the two groups with a non-inferiority margin of 2.5%. For analysis of non-inferiority trials, researchers need to get a confidence interval (CI) for difference between two arms and note that lower or upper bound of the CI does not exceed NI margin (in predefined direction). There are times when the new treatment is statistically better than the control arm where CI is entirely above or below zero showing superiority. In such cases, investigators need to plan in advance that if non-inferiority is established then they can test for superiority.

The important thing is to carefully choose equivalence and non-inferiority margin, which is based on clinical and statistical significance and prior research experience. It should be carefully decided before commencing the trial as it gives scientific credibility to the trial. Equivalence and non-inferiority margin is the minimum difference investigator can tolerate with respect to specified endpoints in given direction. In order to avoid type I error (false claim of equivalence or non-inferiority when outcome is actually equivalent or inferior), it is important to avoid margin that is too large and has potential to adversely affect the participants. In addition, the sample size must be sufficient enough on the carefully selected margin, alpha and power to declare equivalence and non-inferiority. Of note, conservative margins require more sample size to detect differences.

In contrast to intention to treat analysis, equivalence and non-inferiority trials conventionally incorporate per-protocol analysis. Intention to treat analysis includes participants who were initially randomized to treatment assignment whether they adhere to the group or not or were lost to follow-up. It has potential to bias results towards null if there is lots of crossover among treatment groups. Per protocol analysis on the other hand includes participants who adheres to specific treatment group and excludes protocol deviation and violations. However, excluding these data points may also bias results in either direction, particularly in survival trials where one might discontinue study drug due to other fatal condition. It is therefore advised to analyse equivalence and non-inferiority trials by using both approaches. Moreover, control arm must have a standard of care, otherwise, it would be unethical if the comparator is less optimal choice of treatment.

Quality of trial depends upon what endpoint is chosen and how it is measured and how many lost to follow-ups are there. In order to ensure validity of outcome estimate, quality control measures need to be in place to ensure adherence to standard operating procedures, to minimize protocol deviations and violations and minimize lost to follow-up. It should be clearly stated how missing data would be handled if any present and method used to deal with missing data. Moreover, sensitivity analysis must be carried out to ensure robustness of primary outcome with and without missing data values. Specifically, in non-inferiority trials, there should be consistency between type I error rate and confidence interval margin. And hypothesis is one sided based on one sided alpha which is conventionally taken as 5% . Correspondingly, one sided 95% CI or two sided 90% CI should be calculated. The *p*-values are generally not required and if P value is calculated, then it should be reported for one sided test only.

Correct understanding of trial design is pre-requisite for conducting and interpreting trials with valid outcomes and to answer the research question in an appropriate way.

REFERENCES

1. Klingberg-Allvin M, Cleeva A, Atuhairwe S, Tumwesigye NM, Faxelid E, Byamuhisha J, *et al*. Comparison of treatment of incomplete abortion with misoprostol by physicians and midwives at district level in Uganda: a randomised controlled equivalence trial. Lancet 2015;385(9985):2392–8.
2. Jehan F, Nisar MI, Kerai S, Brown N, Balouch B, Hyder Z, *et al*. A double-blind community-based randomized trial of amoxicillin versus placebo for fast breathing pneumonia in children aged 2-59 months in Karachi, Pakistan (RETAPP). BMC Infect Dis 2016;16:13.

Address for Correspondence:
Salima Kerai, BScN, MSc Epidemiology and Biostatistics, Aga Khan University, Karachi-Pakistan
Cell: +92 331 210 1770
Email: Salima.kerai@aku.edu