Sample Size Considerations When Using the Synthesis Method for Non-inferiority Trials

Huei Wang, Amgen Inc and Rachel Wei, Amgen Inc

Abstract:

Non-inferiority (NI) trials are widely used in drug development. The choice of NI margin has important practical consequences, e.g. a smaller margin requires a larger sample size and a large margin may lead to false conclusion of drug effectiveness. In NI trials comparing test drug to active control, one may consider two margins (1) the margin based on that whole active control effect (M1) (2) the largest clinically acceptable difference of the test drug compared to the active control (M2). Showing the effect size of M1 would only provide assurance that the test drug has an effect greater than placebo. Fixed margin and synthesis approaches are the two conventional strategies to show NI to M2. For situations in some therapeutic areas, it is challenge what preservation rate for M2 should be chosen based on synthesis approach in order to sufficiently demonstrate the test drug effect over placebo and active control. In addition, it is mathematically possible that sample size required for 2nd stage (M2) is less than N required for 1st stage (M1). This poster presentation discusses the issues and potential solutions by using a real example for phase 3 trial planning.

Background:

- Conceptually, the non-inferiority (NI) study design provides two comparisons[1, 2, 3]:
    - a direct comparison of the test treatment (X) with the active standard treatment (S)
    - an indirect comparison of treatment X to placebo (P), based on what is known about the effect of the active comparator compared to placebo.
- For regulatory approval. it is required that treatment X is shown to preserve some fraction of the effect of the treatment S in addition to demonstrating treatment X superior to P via an indirect comparison[2,4].
- The requirement of two margins[2] are

- 1st step: select the margin of M_1 is to rule out loss of the entire assumed effect of the treatment S so we can conclude that the treatment X is superior to P. Based on the draft guideline, the fixed margin approach is preferable for the 1st stage.
- 2nd step: choose NI margin (called M_2 from clinical judgment) based on a specified portion of the control effect (M1) whose loss by treatment X must be ruled out. FDA thinks that the synthesis approach, appropriately conducted, can be considered in ruling out the clinical margin M_2.

Issues and Methods

- Based on FDA draft guideline2, a sponsor may design a non-inferiority trial based on the two-stage approach:

  1) Use the fixed margin approach to meet the requirement of 1st stage and may use the lower or upper bound of the two-sided 95% confidence interval comparing treatment S vs. P for margin M_1
  2) Use synthesis approach to ensure certain preservation, for example, 50% of the treatment S effect (i.e. clinical margin M_2) to meet the requirement of 2nd stage

- This two-stage approach seems a logical approach in order to meet regulatory approval since usually the sample size (N) is driven by 2nd stage.
- However, it is mathematically possible that N required for 2nd stage is less than N required for 1st stage and it will occur in real examples.

Assumptions:

- RR_{xs} and V_{xs} as the treatment effect and variance of treatment X relative to treatment S
- RR_{ps} and $V_{ps}$from the historical data of treatment P and S
- Power 1- β and preservation rate γ

The N1 and N2 required for 1st stage and 2nd stage, respectively, are shown below,

1st stage:

$$N \geq \frac{(\frac{p_x + p_s}{p_x p_s} - 2)(1.96 - Z_\beta)^2}{(rr_{xs} - [rr_{ps} + 1.96\sqrt{v_{ps}}])^2})$$

2nd stage: $$P(Z < \frac{1.96 * \sqrt{V_{xs} + (1-\gamma)^2 v_{ps}} - rr_{xs} + (1-\gamma) rr_{ps}}{\sqrt{v_{xs}}}) \leq 1-\beta$$

where $v_{xs} = ([p_x + p_s]/p_x*p_s) -2)/N_1$, rr_{xs}, $v_{xs,}$ $rr_{ps,}$ and v_{ps} are in log scale

Objective: explore the parameters for determining the relationship of Ns between the stages 1 vs. 2 and demonstrate by an real example

- RR_{xs}
- P_x and P_s (or V_{xs})
- Preservation rate
- Historical data (RR_{ps} and V_{ps})

Results

Hypercalcemia of malignancy (HCM) trial

- Hypercalcemia of malignancy (HCM) has been reported to occur in 10% to 30% of patients with advanced cancer and is indicative of poor prognosis
- IV bisphosphonates (IV BPs) are standard care treatment for HCM
- To assess the efficacy and safety of a new treatment X vs. the standard care treatment S

Historical data

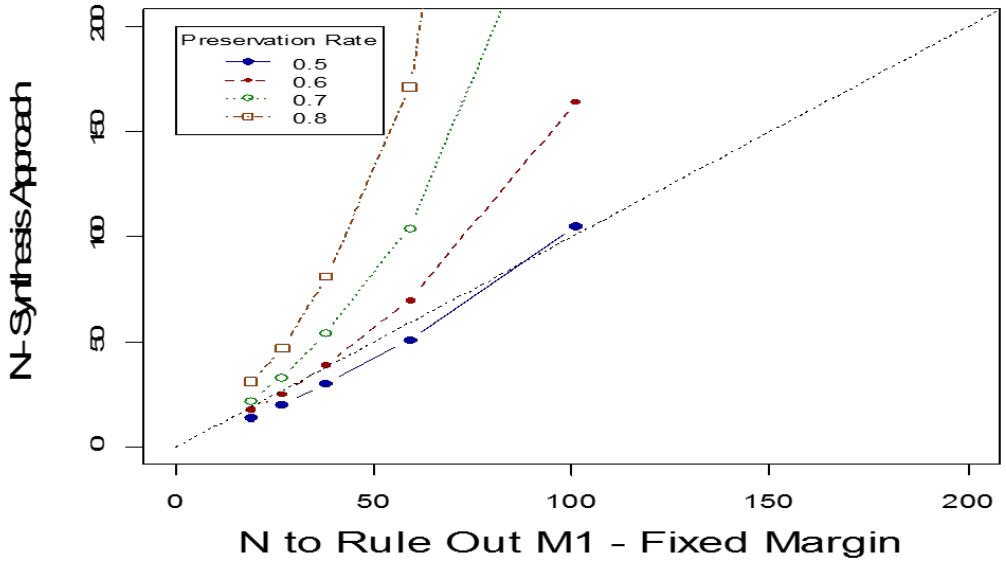
Response Rate		RR_{xs}	Response Rate	RR_{ps} (95% CI)
X (P_x)	S (P_s)	P_x/P_s	P (P_p)	P_p/P_s
80%	70%	1.14	22%	0.31 (0.13, 0.74)

Assumption for figure 1:

- Based on historical data - RR_{ps} with 95% CI: 0.31 [0.13, 0.74])

- Vary expected RR_{xs} by varying response rates P_x (70% to 90%) and fixing P_s of 70%
- Assume 90% power; N is required sample size per treatment group (1: 1 ratio of treatment X vs. treatment S) ; alpha=0.025 (one sided test)

Figure 1: Relationship of Ns between 1st vs. 2nd stages with various high P_x and P_s



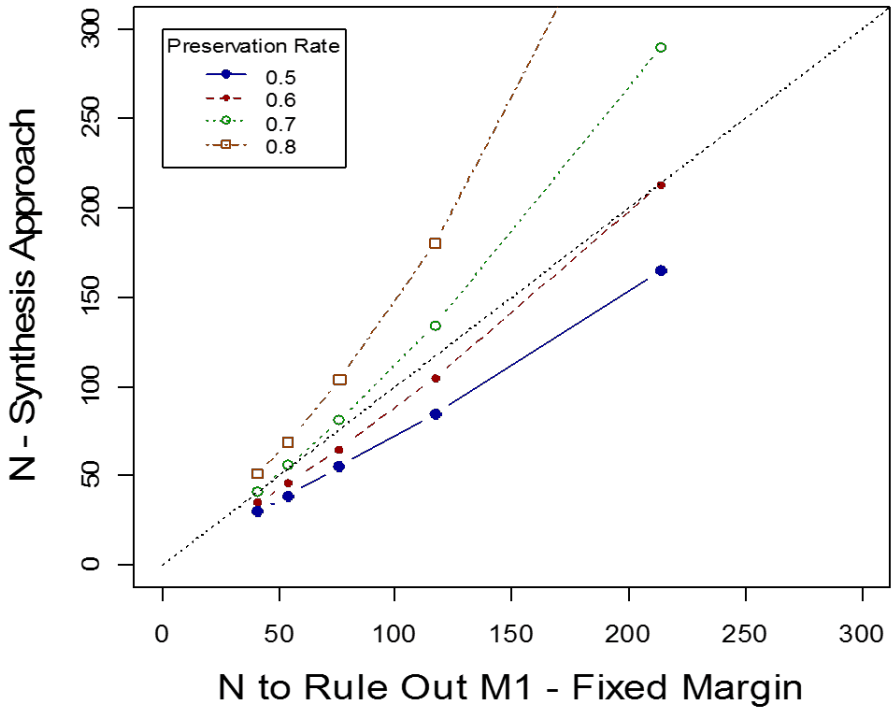
Observations on figure 1:

- Based on the preservation rate of 50%, the N required for the 1st stage (ruling out of M1 by using Fixed Margin approach) is larger than the N required for the 2nd stage (preserving 50% of M1 effect by using Synthesis approach).
- Based on such historical data, the preservation rate >= 64% will ensure the N required for the 2nd stage larger than the one required for the 1st stage.

Assumption for figure 2:

- Vary expected RRxs by varying response rates Px (35% to 55%) and fixing Ps of 30%
- Other assumptions are same as figure 1

Figure 2: Relationship of Ns between 1st vs. 2nd stages with various low P_x and P_s



Observation on figure 2:

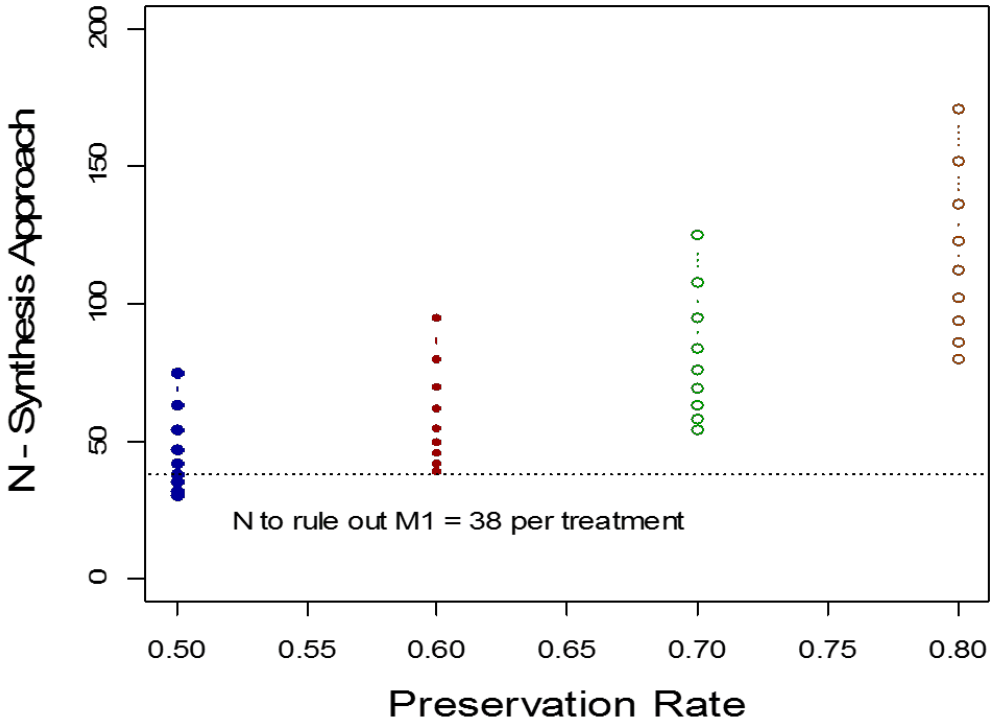
- Based on the preservation rate of 50%, the N required for the 1st stage (ruling out of M1 by using Fixed Margin approach) is larger than the N required for the 2nd stage (preserving 50% of M1 effect by using Synthesis approach).
- Based on such historical data, the preservation rate >= 69% will ensure the N required for the 2nd stage larger than the one required for the 1st stage.
- Small P_x and P_s require large N.

Assumption for figure 3:

- Fix M_1 as 0.74 and vary the historical RR_{ps} : 0.3 to 0.7
- Assume expected RR_{xs} as 1.14 (i.e. 80% for P_x and 70% for P_s)

- Assume 90% power; N is required sample size per treatment group (1: 1 ratio of treatment X vs. treatment S); alpha=0.025 (one sided test)

Figure 3: Plot of N for Synthesis approach with various historical RR_{ps}



Observation on figure 3:

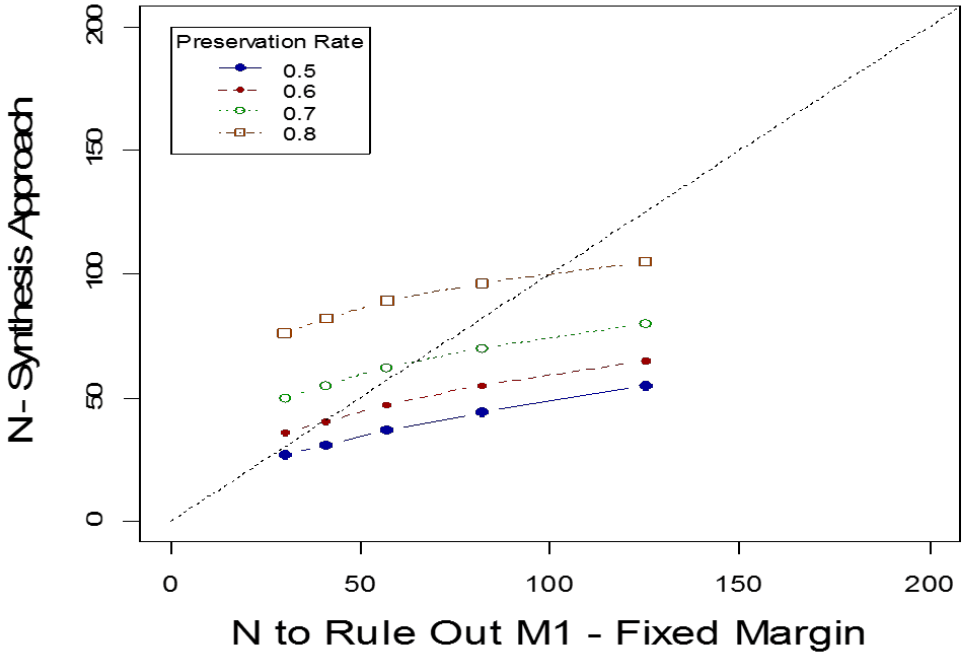
- Based on various historical RR_{ps} and SE_{ps} (by fixing M_1), the N required for the 2nd stage (preserving 50% of M_1 effect by using Synthesis approach) may not larger than the N required for the 1st stage (ruling out of M_1 by using Fixed Margin approach) for a specific preservation rate.
- Large RR_{ps} requires large N.

Assumption for figure 4:

- Fix RR_{ps} = 0.31; vary SE_{ps} by varying M_1 (0.7 to 0.9)

- Other assumptions are same as figure 3

Figure 4: Plot of N for Synthesis approach with various historical SE_{ps}



Observation on figure 4:

- Based on M_1 and various historical SE_{ps} (by fixing RR_{ps}), the N required for the 2nd stage (preserving 50% of M_1 effect by using Synthesis approach) may not larger than the N required for the 1st stage (ruling out of M_1 by using Fixed Margin approach) for a specific preservation rate.

- The large M1 is, the difference of N between 1st and 2nd stages increases.

Conclusions and Discussion

- The sample size required by 1st stage is crucial to provide assurance that the treatment X has an effect greater than placebo.

- Two approaches (synthesis and fixed margin methods) have commonly been used for the 2nd stage design. A synthesis method with an appropriately chosen value of preservation

rate is always more efficient than a fixed-margin approach that achieves the same control of the type 1 error rate[6].

- Usually regulatory agencies in the US would accept 50% or greater preservation rate in the 2nd stage design[5]. However, for situations in some therapeutic areas, it is challenge to choose the preservation rate based on synthesis approach in order to sufficiently demonstrate the treatment X effect over treatment S. We show in our case that the sample size required for 2nd stage (e.g. 50% preservation rate) is not sufficient to show treatment X effect greater than placebo by using Synthesis method in the 2nd stage. Some Observations are discussed below:

  - One may increase the preservation rate to have the sample size required in 2nd stage at least as large as the one in 1st stage. For example, one may increase the preservation rate to 64% in our case rather than using 50%. This suggestion seems sufficient by "any effect" criterion[5]; however, this may not sufficient for regulatory approval since the preservation of effect (by regulatory) requires treatment X effect above some threshold for clinical importance. The regulatory may not agree whether such threshold (for example 64% of preservation rate) means clinically important.

  - When the historical data is promising, the claim of non-inferiority becomes easier to achieve. However, the corresponding margin using Synthesis method with 50% or 60% preservation rate may be wide and not clinically meaningful. It may not easy for physicians to choose clinically meaningful margin. For example, in our case, the 50% of preservation rate is corresponding to 15% of fixed margin; when the preservation rate increases to 67%, the fixed margin reduces to 10%. The 10% margin has been used in the non-inferiority HCM trial to compare zoledronic acid with pamidronate[7].

  - What happens if one uses Synthesis method in the cases of promising historical data of S vs. P present? For example, if the historical data is very promising (eg RR point estimate as 0.1 or 0.2), N required for 2nd stage may be much smaller

than N for 1st stage; even though with high preservation rate, it may not make the sample size required in 2nd stage at least as large as the one in 1st stage.

- One may choose to discount historical data: rather than use the lower bound of 95% confidence interval from the historical data for the sample size calculation for 1st stage, one may use the lower bound of 90% CI specially in the cases of promising historical data

- Regulatory agencies may ask using fixed margin method for 2nd stage rather than synthesis method. For examples, the sample sizes needed per group (1:1 ratio) are 122 and 349 for synthesis and fixed margin methods, respectively, assuming 50% preservation rate (i.e. 15% of fixed margin), 90% power, alpha=0.025 (one sided) and same assumption as the historical data table above. When the fixed margin decreases to 10%, the sample size needed per group increase to 197 and 519 for synthesis and fixed margin methods, respectively.

- In the cases of small effect size of historical trial comparing treatment S with placebo, the size of historical data may be too small to allow possibility of powered a study with preservation rate of 50% or above.

The authors like to thank Chunlei Ke and Qi Jiang's helpful discussions on this poster.

References:

[1] ICH E-10. International Conference on Harmonization: Choice of Control Group in Clinical Trials, January 2001

[2] FDA Guidance for Industry: Non-Inferiority Clinical Trials, March 2010. Available at:
http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM202140.pdf

[3] Committee for Medicinal Products for Human Use (CHMP): Guideline on the choice of the non-inferiority margin, July 2005. Available at:

http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003636.pdf.

[4] Snapinn, Steve and Jiang, Qi. Indirect comparisons in the comparative efficacy and non-inferiority settings. Pharmaceutialcal Statistics, 2011, 10 420-426.

[5] Snapinn, Steve and Jiang, Qi. Preservation of effect and the regulatory approval of new treatment on the basis of non-inferiority trials. Statistics in Medicaine. 2008, 27:382-391

[6] Snapinn, Steve and Jiang, Qi. Controlling the type 1 error rate in non-inferiority trials. Statistics in Medicine 2008: 27: 371 – 381

[7] A. Lortholary, J. Hon, E. Abdi, G. Mills, H.D. Menssen, F. Yunus, R. Bell, J. Body, E. Quebe-Fehling, and J. Seaman: Zoledronic Acid Is Superior to Pamidronate in the Treatment of Hypercalcemia of Malignancy: A Pooled Analysis of Two Randomized, Controlled Clinical Trials. Journal of Clinical Oncology, Vol 19, No 2, 2001: pp 558-567.