Session on “Statistical Issues in Non-Inferiority Trials”
SCT 37th Annual Meeting (2016) Montreal, Quebec, Canada - May 15-18, 2016

Group-sequential three-arm noninferiority clinical trial designs

Toshimitsu Hamasaki, PhD, PStat®
National Cerebral and Cardiovascular Center

joint work with

Toshimitsu Ochiai, PhD
Shionogi Co. & LTD.

Scott R Evans, PhD
Harvard T.H. Chan School of Public Health

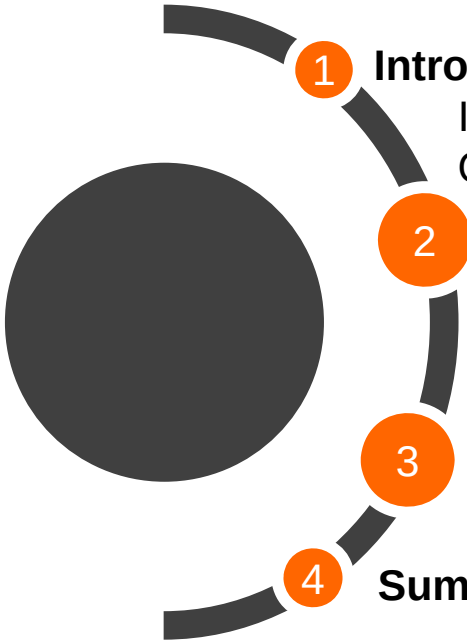
Koko Asakura, PhD
National Cerebral and Cardiovascular Center

Mayumi Fukuda, MD, MPH
National Cerebral and Cardiovascular Center

Haruko Yamamoto, MD, PhD
National Cerebral and Cardiovascular Center

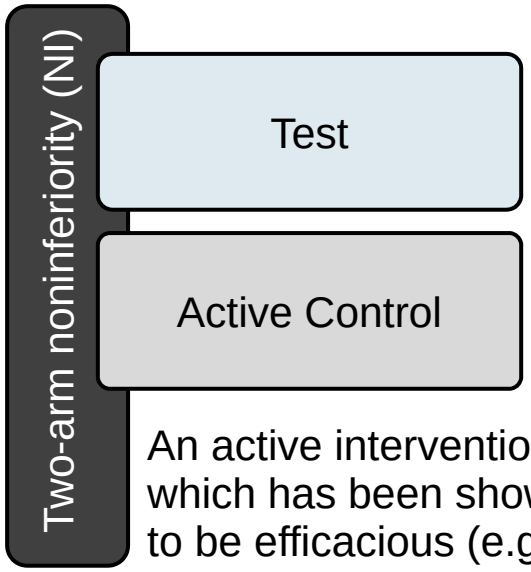
National Cerebral and Cardiovascular Center

Presentation outline



1. **Introduction**
   - Issues in two-arm and three-arm NI trials
   - Objectives
2. **Group-sequential designs for three–arm NI trials**
   - Notation and statistical settings
   - Fixed margin and fraction approach
   - Two decision-making frameworks
3. **An illustration:** Rotigotine trial
4. **Summary and findings, and possible extensions**

Issues in noninferiority trials: constancy and assay sensitivity



An active intervention which has been shown to be efficacious (e.g., superior to placebo) in a historical trial may considered as the active control in a NI trial

Constancy: assumption that the active control effect over placebo has not changed over time.

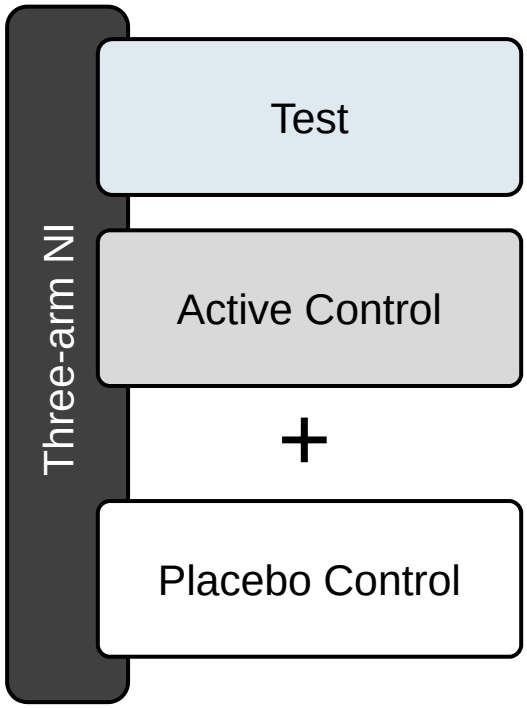
This is not testable in a trial without a concurrent placebo group.

Assay Sensitivity (AS): ability for the trial to be able to detect differences between strategies if they truly exist.

Many factors can affect AS: poor disease diagnosis, endpoint selection and timing, poor adherence, loss to follow-up, prior therapy, inclusion of subgroups, and use of concomitant therapies.

CHMP (2005) Guideline on the choice of the non-inferiority margin/D'Agostino RB et al (2003) *Stat Med* **22**:169–186/Evans SR, Follmann D (2015) *Stat Biopharm Res* **7**:331–336/FDA (2010) Guidance for industry non-inferiority trials. Rockville, MD, USA/ICH (2000) E10: Choice of control group and related issues in clinical trials.

Three-arm NI Trial as a gold standard design



Provide the opportunity of establishing the validity of the assay sensitivity via a comparison of the placebo with the active control intervention within the trial.

Provide challenges:

Ethical issue there may be ethical constraints to **using a placebo**

Difficulty there is the added difficulty of evaluating **two distinct co-objectives**: evaluation of (i) the superiority of the active control intervention to placebo (AS) and (ii) the NI of the test intervention to the active control intervention.

Feasibility it may result in a trial with too large and impractical of a sample size to conduct due to the two co-objectives

Objectives

To discuss group-sequential designs (GSDs) for three-arm NI clinical trials

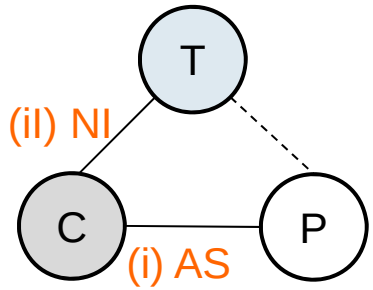
GSDs offer the possibility to stop a trial early when evidence is overwhelming and thus offers efficiency- potentially fewer trial participants and minimizing the amount of time that participants receive a placebo, compared to fixed-sample designs.

To extend two existing approaches for evaluating AS and NI into a GS setting

Fixed margin approach (Koch and Röhmel, 2004; Hida and Tango, 2011, 2013)

Fraction approach (Pigeot et al, 2003)

To discuss a three-arm NI trial that has two co-primary objectives: AS and NI

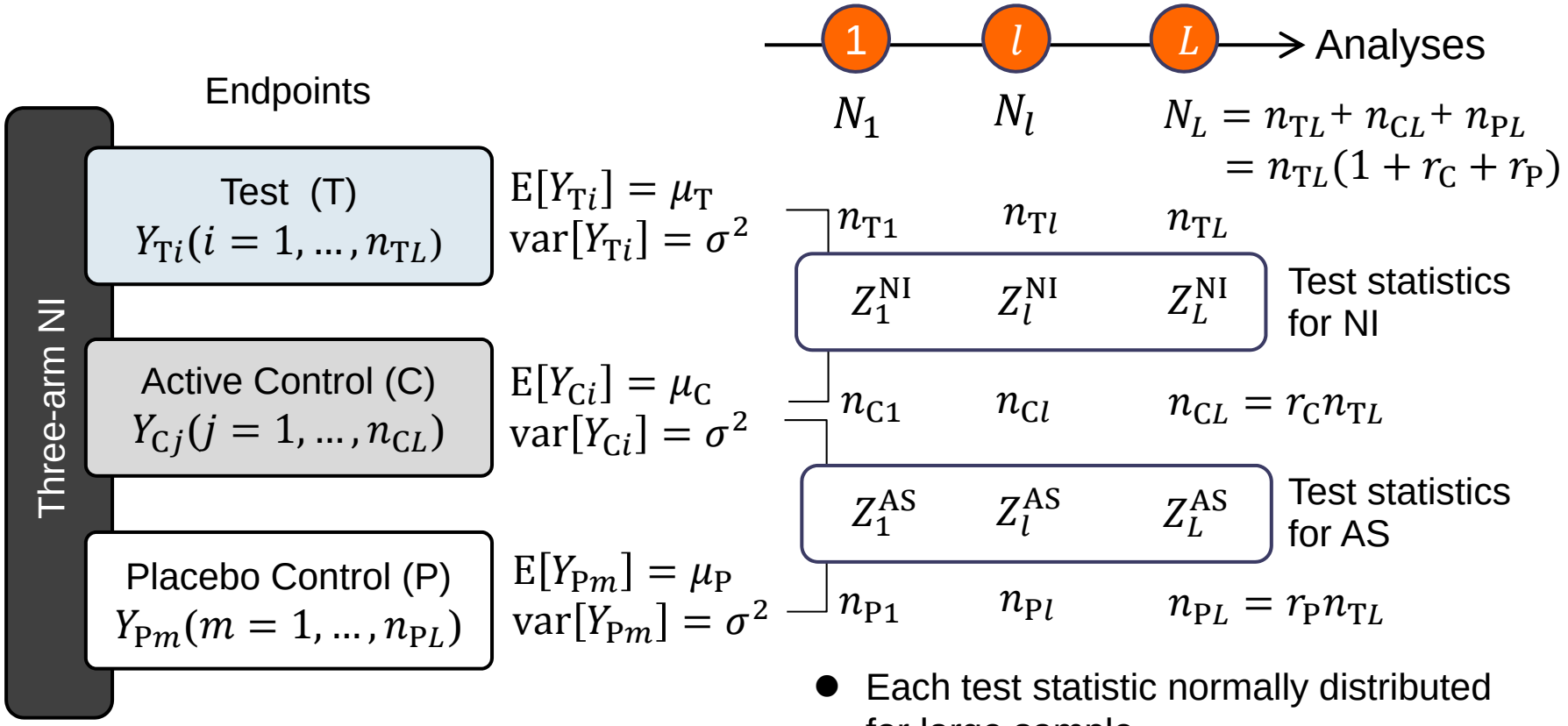


If the AS assumption does not hold, then there will be uncertainty regarding whether a NI result means that they are similarly effective or similarly ineffective.

When there is a concern about the AS, to make the evaluation of objective (ii) more interpretable, evaluate a direct comparison of the control intervention (**C**) with the placebo (**P**)

Koch A, Röhmel J (2004) *J Biopharm Stat* **14**:315–325/ Hida E, Tango T (2011) *Stat Med* **30**:224–231/Hida E, Tango T (2013) *J Biopharm Stat* **23**:774–789/ Pigeot I et al (2003) *Stat Med* **22**:883–899

Notations and statistical settings



- Each test statistic normally distributed for large sample
- 2L test statistics 2L-variate normally distributed

Fixed margin approach

	AS (Superiority)	NI
Hypoth.	$\begin{cases} \mathrm{H}_0^{\mathrm{AS}}: \mu_{\mathrm{C}} - \mu_{\mathrm{P}} \leq \omega \\ \mathrm{H}_1^{\mathrm{AS}}: \mu_{\mathrm{C}} - \mu_{\mathrm{P}} > \omega \end{cases}$	$\begin{cases} \mathrm{H}_0^{\mathrm{NI}}: \mu_{\mathrm{T}} - \mu_{\mathrm{C}} \leq -\omega \\ \mathrm{H}_1^{\mathrm{NI}}: \mu_{\mathrm{T}} - \mu_{\mathrm{C}} > -\omega \end{cases}$
Test Stats	$Z_l^{\mathrm{AS}} = \dfrac{\bar{Y}_{\mathrm{C}l} - \bar{Y}_{\mathrm{P}l} - \omega}{\sigma\sqrt{\dfrac{1}{n_{\mathrm{C}l}} + \dfrac{1}{n_{\mathrm{P}l}}}}$	$Z_l^{\mathrm{NI}} = \dfrac{\bar{Y}_{\mathrm{T}l} - \bar{Y}_{\mathrm{C}l} + \omega}{\sigma\sqrt{\dfrac{1}{n_{\mathrm{T}l}} + \dfrac{1}{n_{\mathrm{C}l}}}}$
	$\bar{Y}_{\mathrm{T}l} = n_{\mathrm{T}l}^{-1}\Sigma_{i=1}^{n_{\mathrm{T}l}} Y_{\mathrm{T}i}$, $\bar{Y}_{\mathrm{C}l} = n_{\mathrm{C}l}^{-1}\Sigma_{j=1}^{n_{\mathrm{C}l}} Y_{\mathrm{C}j}$, $\bar{Y}_{\mathrm{P}l} = n_{\mathrm{P}l}^{-1}\Sigma_{m=1}^{n_{\mathrm{P}l}} Y_{\mathrm{P}m}$	
Corr.	$\mathrm{corr}[Z_l^{\mathrm{AS}}, Z_{l'}^{\mathrm{AS}}] = -\sqrt{\dfrac{r_{\mathrm{P}}}{(1+r_{\mathrm{C}})(r_{\mathrm{P}}+r_{\mathrm{C}})}}\ (1 \leq l \leq l' \leq L)$	

- Both hypotheses are tested at α
- Imposes an extra condition on the hypothesis testing for the AS, that is, superiority the C to the P is demonstrated with a NI margin ω
- The inequalities $\mu_{\mathrm{P}} < \mu_{\mathrm{C}} - \omega \leq \mu_{\mathrm{T}}$ hold for any value of ω

Hida E, Tango T (2011) *Stat Med* **30**:224–231/ Hida E, Tango T (2013) *J Biopharm Stat* **23**:774–789

Fraction approach

	AS (Superiority)	NI
Hypoth.	$\begin{cases} \mathrm{H}_0^{\mathrm{AS}}: \mu_C - \mu_P \le 0 \\ \mathrm{H}_1^{\mathrm{AS}}: \mu_C - \mu_P > 0 \end{cases}$	$\begin{cases} \mathrm{H}_0^{\mathrm{NI}}: \dfrac{\mu_T - \mu_C}{\mu_C - \mu_C} \le \theta \\ \mathrm{H}_1^{\mathrm{NI}}: \dfrac{\mu_T - \mu_C}{\mu_C - \mu_C} > \theta \end{cases}$
Test Stats	$Z_l^{AS} = \dfrac{\bar{Y}_{Cl} - \bar{Y}_{Pl}}{\sigma\sqrt{\dfrac{1}{n_{Cl}} + \dfrac{1}{n_{Pl}}}}$	$Z_l^{\mathrm{NI}} = \dfrac{\bar{Y}_{\mathrm{T}l} - \theta\bar{Y}_{Cl} + (1-\theta)\omega}{\sigma\sqrt{\dfrac{1}{n_{\mathrm{T}l}} + \dfrac{\theta^2}{n_{\mathrm{C}l}} + \dfrac{(1-\theta)^2}{n_{\mathrm{C}l}}}}$
	$\bar{Y}_{\mathrm{T}l} = n_{\mathrm{T}l}^{-1}\Sigma_{i=1}^{n_{\mathrm{T}l}} Y_{\mathrm{T}i}$, $\bar{Y}_{\mathrm{C}l} = n_{\mathrm{C}l}^{-1}\Sigma_{j=1}^{n_{\mathrm{C}l}} Y_{\mathrm{C}j}$, $\bar{Y}_{\mathrm{P}l} = n_{\mathrm{P}l}^{-1}\Sigma_{m=1}^{n_{\mathrm{P}l}} Y_{\mathrm{P}m}$	
Corr.	$\mathrm{corr}[Z_l^{\mathrm{AS}}, Z_{l'}^{\mathrm{NI}}] = \dfrac{-\dfrac{\theta}{r_{\mathrm{C}}} + \dfrac{(1-\theta)}{r_{\mathrm{P}}}}{\sqrt{1 + \dfrac{\theta^2}{r_{\mathrm{C}}} + \dfrac{(1-\theta)^2}{r_{\mathrm{P}}}}\sqrt{\dfrac{1}{r_{\mathrm{C}}} + \dfrac{1}{r_{\mathrm{P}}}}}$ $(1 \le l \le l' \le L)$	

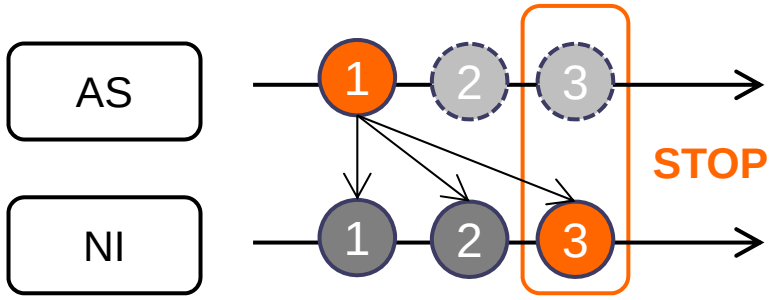
- Both hypotheses are tested at α
- $\theta(0 < \theta < 1)$ is prespecified and determined by $\theta = 1 - \omega/(\mu_{\mathrm{C}} - \mu_{\mathrm{P}})$ as a fraction of difference between μ_{C} and μ_{P}, using the NI margin ω.
- Hypothesis testing is logically ordered: AS → NI

Pigeot I et al (2003) *Stat Med* **22**:883–899

Features in the two approaches

Fixed margin	Fraction
● The correlation is determined by the sampling ratio, but the two test statistics are always negatively correlated	● The two test statistics are positively or negatively correlated depending on sampling ratio and fraction
● Can indirectly demonstrate the superiority of the experimental intervention relative to the placebo if $\mathrm{H}_0^{\mathrm{AS}}$ and $\mathrm{H}_0^{\mathrm{NI}}$ are rejected, without direct comparison of the experimental intervention to the placebo.	● Can demonstrate $\mu_{\mathrm{T}} > \mu_{\mathrm{P}}$ irrespective of θ since $\mu_{\mathrm{T}} - \mu_{\mathrm{P}} > \theta(\mu_{\mathrm{C}} - \mu_{\mathrm{P}}) > 0$ if both null hypotheses $\mathrm{H}_0^{\mathrm{AS}}$ and $\mathrm{H}_0^{\mathrm{NI}}$ are rejected
● Can reject $\mathrm{H}_0^{\mathrm{NI}}$ when $\mu_{\mathrm{C}} - \Delta < \mu_{\mathrm{P}} < \mu_{\mathrm{C}}$ is true	● Cannot reject $\mathrm{H}_0^{\mathrm{NI}}$ when $\mu_{\mathrm{C}} - \Delta < \mu_{\mathrm{P}} < \mu_{\mathrm{C}}$ is true- whether the fraction approach can allow demonstration of noninferiority of the experimental intervention to the control intervention is questionable under $\mu_{\mathrm{C}} - \Delta < \mu_{\mathrm{P}}$.

Decision-making frameworks for GS designs: DF-A

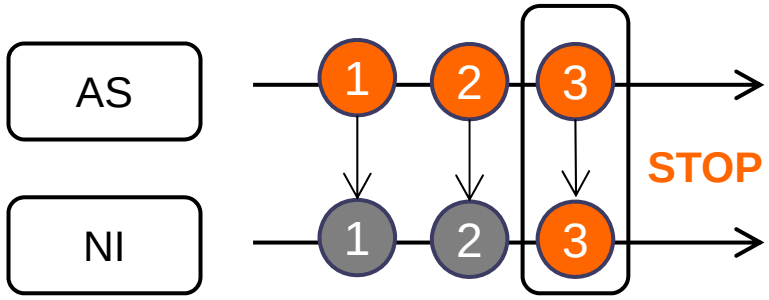


Stopping rule for DF-A
At the lth analysis ($l = l', \ldots, L-1$) if $Z_{l'}^{\mathrm{AS}} > c_{l'}^{\mathrm{AS}}(\alpha)$ for some $l'(1 \le l' \le l)$ ($\mathrm{H}_0^{\mathrm{AS}}$ has been rejected), and $Z_l^{\mathrm{NI}} > c_l^{\mathrm{NI}}(\alpha)$, then reject $\mathrm{H}_0^{\mathrm{NI}}$ and stop the trial otherwise, continue the trial At the Lth analysis if $Z_{l'}^{\mathrm{AS}} > c_{l'}^{\mathrm{AS}}(\alpha)$ for some $l'(1 \le l' \le L)$, and $Z_L^{\mathrm{NI}} > c_L^{\mathrm{NI}}(\alpha)$, then reject $\mathrm{H}_0^{\mathrm{NI}}$ otherwise do not reject $\mathrm{H}_0^{\mathrm{NI}}$

- NI is evaluated only after the AS is demonstrated.
- A trial stops if the AS and the NI are achieved at any interim analysis, i.e., not necessarily simultaneously.
- If AS is demonstrated but NI is not, then the trial continues and subsequent hypothesis testing is repeatedly conducted only for NI until the NI is demonstrated.

$c_l^{\mathrm{AS}}(\alpha)$ and $c_l^{\mathrm{NI}}(\alpha)$ are critical values separately prespecified, using any GS methods

Decision-making frameworks for GS designs: DF-B



Stopping rule for DF-B

At the lth analysis ($l = 1, \ldots, L-1$)
if $Z_l^{\mathrm{AS}} > c_l^{\mathrm{AS}}(\alpha)$ and $Z_l^{\mathrm{NI}} > c_l^{\mathrm{NI}}(\alpha)$, then reject $\mathrm{H}_0^{\mathrm{AS}}$ and $\mathrm{H}_0^{\mathrm{NI}}$ and stop the trial
otherwise, continue the trial
At the Lth analysis
if $Z_L^{\mathrm{AS}} > c_L^{\mathrm{AS}}(\alpha)$ and $Z_L^{\mathrm{NI}} > c_L^{\mathrm{NI}}(\alpha)$, then reject $\mathrm{H}_0^{\mathrm{AS}}$ and $\mathrm{H}_0^{\mathrm{NI}}$ and stop the trial
otherwise do not reject $\mathrm{H}_0^{\mathrm{AS}}$ and $\mathrm{H}_0^{\mathrm{NI}}$

- A special case of DF-A, NI is evaluated only after the AS is demonstrated,
- A trial stops only if AS and NI are demonstrated at the same interim analysis simultaneously.
- Otherwise, the trial will continue and the subsequent hypothesis testing is repeatedly conducted for both AS and NI until simultaneous significance is reached.

$c_l^{\mathrm{AS}}(\alpha)$ and $c_l^{\mathrm{NI}}(\alpha)$ are critical values separately prespecified, using any GS methods

Approaches and decision-making frameworks

	Fixed Margin	Fraction
DF-A	● Flexible, and slightly powerful than DF-B ● Allows for dropping placebo group if AS is demonstrated at the interim	● Flexible, and slightly powerful than DF-B ● Not allow for dropping placebo group as the test statistics for the NI includes the amount of $\overline{Y}_{\mathrm{P}m}$.
DF-B	● Simple, but less powerful than DF-A ● Not allow for dropping of the placebo group even if AS is demonstrated at the interim	● Simple, but less powerful than DF-A

An illustration: Rotigotine trial (Mizuno et al 2014)

Objective To evaluate the **superiority** of transdermal rotigotine to placebo, and to evaluate **NI** to ropinirole, in Japanese Parkinson's disease patients on concomitant levodopa therapy.

Primary endpoints The change in the unified Parkinson's disease rating scale (UPDRS) Part III (ON state) sum score from baseline to week 16 of the treatment period

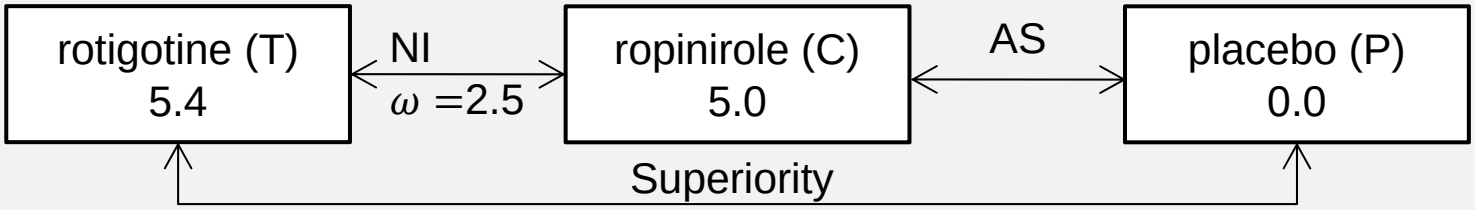


Illustration Calculate the maximum sample size (**MSS**) and average sample number (**ASN**) under th alternative hypothesis for evaluating AS and NI

- 80 % power at the 2.5 % significance level for a one-sided test
- # of planned analyses: L =1 and 2.
- Critical boundaries determined using Lan-DeMets error spending method (Lan and DeMets, 1983) with equally-spaced increment of information.
- Allocation ratio: $1: r_C: r_P = 1:1:1$

Mizuno Y et (2014) Parkinsonism and Related Disorders **20**:466 1388–1393/ Lan KKG, DeMets DL (1983) Biometrika **70**:659-663

Rotigotine trial: MSS and ASN based on DF-A and DF-B

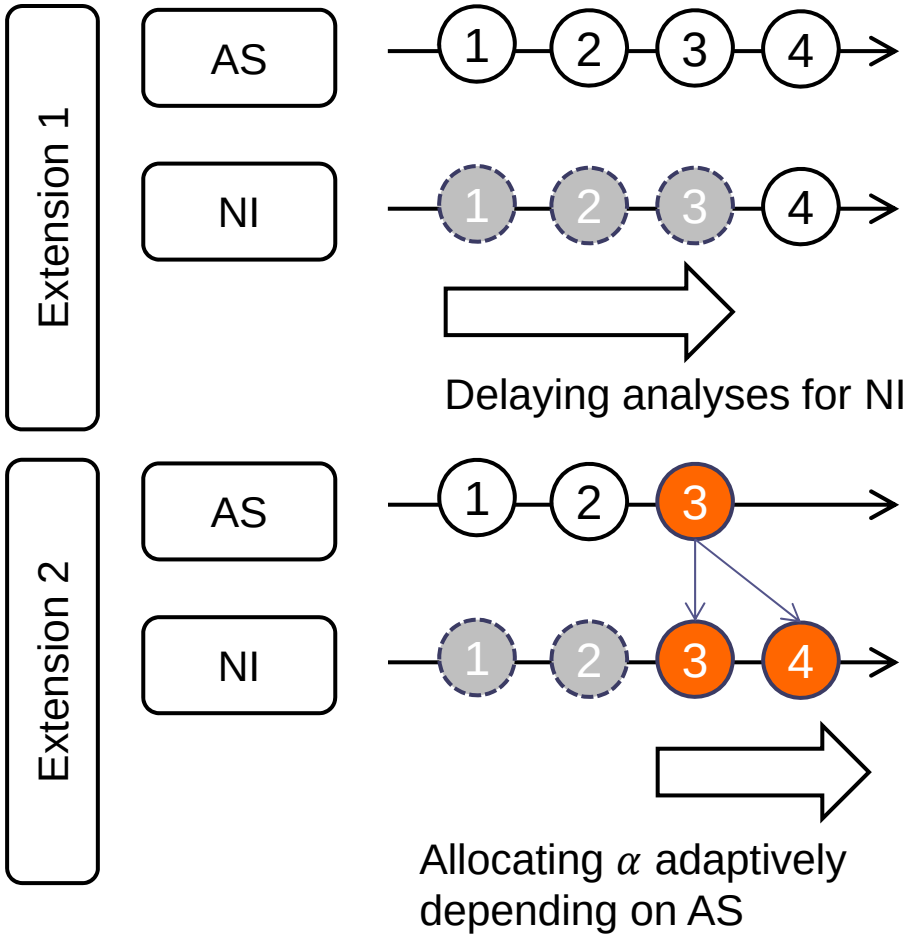
# of analyses and decision-making frameworks		Bound. Func. (AS-NI)	Fixed Margin Approach			Fraction Approach		
			MSS	ASN1	ASN2	MSS	ASN1	ASN2
$L=1$			717			351		
$L=2$	DF-A	OF-OF	720	713	694	351	336	NA
		PC-PC	801	681	655	393	312	NA
		OF-PC	759	726	713	396	339	NA
		PC-OF	771	730	681	354	330	NA
	DF-B	OF-OF	720	713	NA	351	336	NA
		PC-PC	810	686	NA	393	312	NA
		OF-PC	759	726	NA	396	339	NA
		PC-OF	789	743	NA	357	332	NA

MSS: Maximum sample size
ASN1: Average sample number without dropping P after AS is demonstrated.
ASN2: Average sample number with dropping P after AS is demonstrated

Summary and findings

1. DF-A and DF-B for the fixed margin and the fraction approaches provide the possibility of stopping a trial early when evidence is overwhelming, thus offering efficiency (e.g., an ASN potentially 4–15 % fewer than the fixed-sample designs with equally sized groups and four analyses).
2. There are no major differences in both MSS and ASN between DF-A and DF-B for the fixed margin and the fraction approaches, although DF-A is slightly more powerful than DF-B. By using the DF-A for the fixed margin approach, the time that participants are exposed to placebo can be minimized as the DF-A allows dropping of the placebo group if AS has been demonstrated at an interim analysis
3. For the fixed margin approach, selecting the OF-type boundary for both AS and NI could lead to fewer participants for the MSS and the ASN compared with other critical boundary combinations.
4. For the fraction approach, selecting the OF-type boundary for both AS and NI, or the PC-type boundary for AS and the OF-type boundary for NI provides better efficiency with respect to the MSS and the ASN compared with other critical boundary combinations.

Possible extensions for fixed margin approach



- Can improve the power and decrease MSS, but lose the efficiency (increase ASN)
- Can apply the fraction approach

- Can improve the power and decrease MSS.
- Generally adaptive alpha-allocation inflate the Type I error when the two test statistics are positively correlated, but this is not happen when they are negatively correlated.
- Can not apply the fraction approach

Thank you for your kind attention

If you have any questions, please **e-mail to**

toshi.hamasaki@ncvc.go.jp

Ochiai T, Hamasaki T, Evans SR, Asakura K, Ohno Y. Group-sequential three arm noninferiority clinical trials designs. Journal of Biopharmaceutical Statistics 2016 (First Published online on: 18 February 2016 as doi: 10.1080/10543406.2016.1148710)