Doctopic: Statistical Methodology for Data Analysis ARTICLE IN PRESS

Computational Statistics and Data Analysis ••• (••••) ••••••

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

www.elsevier.com/locate/csda

A hierarchical testing procedure for three arm non-inferiority trials

Santu Ghosh [a], Wenge Guo [b], Samiran Ghosh [c,*]

[a] Division of Biostatistics and Data Science, DPHS, Augusta University, United States of America
[b] Department of Mathematical Sciences, New Jersey Institute of Technology, United States of America
[c] Department of Family Medicine and Public Health Sciences, Wayne State University, United States of America

ARTICLE INFO

Article history:
Received 23 June 2021
Received in revised form 27 April 2022
Accepted 27 April 2022
Available online xxxx

Keywords:
Non-inferiority trials
FWER
Extended skew-normal distribution

ABSTRACT

Non-inferiority trials are becoming very popular for comparative effectiveness research. Non-inferiority trials establish that the effect of an experimental treatment is not worse than that of a reference treatment by more than a specified margin. A three-arm non-inferiority trial that includes the placebo, experimental treatment, and a reference treatment is considered. It has been criticized that the conventional approach for three-arm non-inferiority trials loses power for the non-inferiority hypothesis test unless the power of the assay sensitivity test is close to one. In order to overcome this situation, a novel hierarchical testing procedure with two stages for three-arm non-inferiority trials is developed. The family-wise error rate (FWER) is investigated analytically and numerically of the proposed test procedure. Numerical studies indicate that the suggested method controls FWER and has more power than the traditional approach particularly when the power of that assay sensitivity test is not close to one. Through these empirical studies, it is shown that the proposed method can be successfully applied in practice.

1. Introduction

One can show the superiority of an experimental treatment over a reference treatment, or when the superiority of an experimental treatment over a reference treatment is not evident, one can show that the experimental treatment is non-inferior to that reference treatment. An experimental treatment's effect is declared to be non-inferior to the effect of reference treatment when the effect of the experimental treatment on an endpoint is not worse than the effect of a reference treatment on that same endpoint by more than a specified margin. A clinical trial used to evaluate whether an experimental treatment's effect is non-inferior to the effect of a reference treatment is called a non-inferiority trial. A randomized, double-blind placebo-controlled trial is the gold standard in the determination of efficacy and the risk-benefit profile of an investigational drug (EMA (2005); FDA (2016)). A placebo-controlled trial is considered appropriate in the absence of a reference treatment. However, a placebo-controlled trial would be unethical when there is an available reference treatment (WMA (1997)). In the presence of a reference treatment, experimental treatment is compared with a reference treatment. However, the absence of the placebo in an active-controlled trial arises assay sensitivity (D'Agostino et al. (2003)). For a detailed description of the problem, see D'Agostino et al. (2003) and Hung et al. (2003). Assay sensitivity refers to the ability of a trial to distinguish between effective and ineffective treatments. To establish the effectiveness

* Corresponding author at: 421 E. Canfield Avenue Detroit, MI 48201.
E-mail address: sghos@med.wayne.edu (S. Ghosh).

https://doi.org/10.1016/j.csda.2022.107521
0167-9473/

of experimental treatment based on its similarity to the reference treatment, an active-controlled trial assumes constancy assumption, i.e., the patient population in the current active-controlled trial and the past placebo-control trial remains unchanged (D'Agostino et al. (2003)). However, it is not always straightforward to validate this constancy assumption. As a consequence, it is often recommended to include the placebo in an active-controlled trial whenever it is feasible and ethically justifiable, as mentioned in several regulatory guidelines (EMA (2005)). These three-arm non-inferiority trials are often considered to avoid the complications described above. A three-arm non-inferiority trial is optimal in the sense that: (i) it is free from the constancy assumption, (ii) the effect of the experimental treatment can be directly assessed, and (iii) the effect of an experimental treatment can be compared to the effect of active control. Following the footstep of Koch and Rohmel (2004), Pigeot et al. (2003) introduced a systematic approach for constructing a test for three-arm non-inferiority trials, known as fraction margin-based non-inferiority testing, where the non-inferiority margin is the pre-specified negative fraction of the unknown effect size of the reference treatment over placebo in the current three-arm trial. This fraction margin approach is a two-step hierarchical procedure. The first step tests the superiority of the reference treatment over a placebo. The first step gives the internal validation of assay sensitivity (AS). The rejection of the null hypothesis or the confirmation of AS in step one leads to testing the non-inferiority (NI) experimental treatment to the reference treatment. Kieser and Friede (2007) demonstrates that due to this hierarchical structure of the multiple testing problem (AS and NI) in the fraction margin-based approach, the pretest of assay sensitivity may lead to a reduction in power when testing for NI. This finding of Kieser and Friede (2007) motivates us to develop a novel hierarchical testing procedure method for conducting fraction margin-based NI testing. The proposed method tests the NI hypothesis at a significance level α/β whenever the AS hypothesis is significant at level α, where β denotes the power associated with AS test. We demonstrate that the suggested hierarchical testing method controls family-wise error rate (FWER) at level α asymptotically. For the practical implementation, we estimate the power of AS test using the bootstrap method whenever AS hypothesis is significant. Through extensive simulation studies, we provide numerical evidence of the acceptable performance of the proposed procedures in terms of the FWER control and power.

The rest of the article is organized as follows. Section 2 reviews the fraction margin-based non-inferiority testing and develops the proposed hierarchical testing method for three-arm NI trials. Section 3 reports the empirical findings under different scenarios in detail. Section 4 considers a data example. A discussion follows in Section 5. For brevity, derivations of the theoretical results are provided in Appendix.

2. A hierarchical testing method for the assessment of three-arm NI trials

2.1. Fraction margin based non-inferiority testing

To facilitate the discussion of a three-arm non-inferiority trial, we assume that $X_{E,i}$, $X_{R,j}$, and $X_{P,k}$ $(i=1,\ldots,n_E, j=1,\ldots,n_R, k=1,\ldots,n_P)$ denote the observations corresponding to the treatment response in the experimental (E), reference (R), and placebo (P) groups, respectively. For simplicity, we assume that the observations of all three arms are continuous, however, the suggested methods can be extended effortlessly to other types of responses such as binary, count etc. We assume that

$$X_{E,i} \overset{\text{iid}}{\sim} F_E(\cdot),\ \ X_{R,j} \overset{\text{iid}}{\sim} F_R(\cdot),\ \text{ and }\ X_{P,k} \overset{\text{iid}}{\sim} F_P(\cdot),$$

where $F_l(\cdot)$ are continuous distribution functions of X_l, $l \in \{E,R,P\}$. Let $\mu_l = E(X_l)$ denote the mean response of the arm l. We assume that the variances σ_l^2 are finite. Without loss of generality, we assume that the large values of the mean responses represent desirable outcomes. Commonly, non-inferiority is assessed by considering the test problem

$$H_{\text{NI},0}: \mu_E - \mu_R \le \delta \text{ vs } H_{\text{NI},1}: \mu_E - \mu_R > \delta, \tag{1}$$

where $\delta < 0$ denotes a prespecified clinically relevant margin. With the inclusion of the placebo arm to an active-controlled trial, the construction of δ via fraction margin approach (see Pigeot et al. (2003)) can be mathematically expressed as $\delta = f(\mu_R - \mu_P)$ by assuming the assay sensitivity, $\mu_R > \mu_P$; where $f \in (-1,0)$. Reasonable choices of f include $-\frac{1}{2}$, $-\frac{1}{3}$, $-\frac{1}{5}$ but must be guided by practical and clinical considerations. The hypotheses (1) can be rewritten using the expression for δ as follows

$$H_{\text{NI},0}: \frac{\mu_E - \mu_P}{\mu_R - \mu_P} \le \theta \text{ vs } H_{\text{NI},1}: \frac{\mu_E - \mu_P}{\mu_R - \mu_P} > \theta, \tag{2}$$

by assuming $\mu_R - \mu_P > 0$; where $\theta = 1 + f$, known as the retention fraction. The ratio, $(\mu_E - \mu_P)/(\mu_R - \mu_P)$, of the differences in means measures the proportion of efficacy retained by the experimental treatment. A positive but small value of this ratio might lead to the conclusion that the reference treatment should be the standard and the experimental treatment should not be widely adopted although it provides some benefit over placebo. However, when $(\mu_E - \mu_P)/(\mu_R - \mu_P) > 1$ the experimental treatment provides more benefit than the active control. For the derivation of the statistical test procedures for the test problem (2), it is helpful to express (2) as:

$$H_{\mathrm{NI},0}: \mu_E - \theta\mu_R - (1-\theta)\mu_P \le 0 \text{ vs } H_{\mathrm{NI},1}: \mu_E - \theta\mu_R - (1-\theta)\mu_P > 0. \tag{3}$$

The existing frequentist methods (Pigeot et al. (2003); Hasler et al. (2008); Hida and Tango (2011); Ghosh et al. (2017)) for the test problem (3) do not incorporate the pretested AS, $\mu_R - \mu_P > 0$, though NI and AS test statistics are correlated. This disintegration between test problem (3) and the pretested AS condition can lead to a less powerful test method for the NI test (3) unless the power of the pretest is close to unity (Kieser and Friede (2007)). In the next section we introduce a novel procedure for test problem (3) which incorporates the pretested AS condition.

2.2. Methodology

Let $\bar{X}_R$, $\bar{X}_P$, and $\bar{X}_E$ denote the sample means corresponding to the samples $\{X_{R,j}\}$, $\{X_{P,k}\}$, and $\{X_{E,i}\}$, respectively. Let $\hat{\sigma}_R^2$, $\hat{\sigma}_P^2$, and $\hat{\sigma}_E^2$ be the sample variances based on the samples $\{X_{R,j}\}$, $\{X_{P,k}\}$, and $\{X_{E,i}\}$, respectively, where

$$\hat{\sigma}_l^2 = (n_l - 1)^{-1} \sum_{i=1}^{n_l} (X_{l,i} - \bar{X}_l)^2,$$

and $l \in \{P, R, E\}$. Let $T_{\mathrm{AS}} = \frac{\bar{X}_R - \bar{X}_P}{\sqrt{\hat{\sigma}_R^2/n_R + \hat{\sigma}_P^2/n_P}}$ be the test statistic for the following assay sensitivity testing

$$H_{\mathrm{AS},0}: \mu_R - \mu_P \le 0 \text{ vs } H_{\mathrm{AS},1}: \mu_R - \mu_P > 0. \tag{4}$$

Let

$$T_{\mathrm{NI}} = \frac{\bar{X}_E - \theta\bar{X}_R - (1-\theta)\bar{X}_P}{\sqrt{\frac{\hat{\sigma}_E^2}{n_E} + \frac{\theta^2\hat{\sigma}_R^2}{n_R} + \frac{(1-\theta)^2\hat{\sigma}_P^2}{n_P}}}$$

denote the test statistic for testing problem (3). Pigeot et al. (2003)'s fractional margin approach based hypothesis testing (3) is performed provided we reject $H_{\mathrm{AS},0}$. Thus, the pretested AS condition acts as a gatekeeper for the non-inferiority testing problem (3). To test the AS and NI hypotheses, the suggested test procedure uses the following sequential steps:

Step 1: To test $H_{\mathrm{AS},0}$, if $T_{\mathrm{AS}} > c_1$, reject $H_{\mathrm{AS},0}$; otherwise, stop.
Step 2: If $H_{\mathrm{AS},0}$ is rejected, test $H_{\mathrm{NI},0}$. If $T_{\mathrm{NI}} > c_2$, reject $H_{\mathrm{NI},0}$; otherwise, accept $H_{\mathrm{NI},0}$.

The pertinent choices for c_1 and c_2 are determined through controlling the familywise error rate (FWER), the probability of making at least one type I errors, when performing hierarchical testing $H_{\mathrm{AS},0}$ and $H_{\mathrm{NI},0}$. The FWER requirement is

$$\mathrm{FWER} = \mathrm{P}\{\text{Reject at least one true } H_{\mathrm{AS},0}, H_{\mathrm{NI},0}\} \le \alpha, \tag{5}$$

where $\alpha \in (0, 1)$.

Lemma 2.1 provides choices for c_1 and c_2. For mathematical easiness, Lemma 2.1 considers controlling the FWER on the boundary conditions, $H^B_{\mathrm{AS},0}: \mu_R - \mu_P = 0$, and $H^B_{\mathrm{NI},0}: \mu_E - \theta\mu_R - (1-\theta)\mu_P = 0$, of the AS and NI AS null hypotheses. We provide a discussion on FWER control in the interior of the AS and NI null hypotheses in Theorem 2.1.

Lemma 2.1. *Let denote $\beta = \beta(\mu_R - \mu_P) = P_{H_{AS,1}}(T_{AS} > c_1)$, the power of the assay sensitivity test at fixed $\mu_R - \mu_P > 0$. Suppose that thresholds c_1 and c_2 are chosen by satisfying the equations*

$$P_{H^B_{AS,0}}(T_{AS} > c_1) = \alpha \ \text{ and } \ P_{H_{AS,1} \cap H^B_{NI,0}}(T_{NI} > c_2 | T_{AS} > c_1) = \frac{\alpha}{\beta}.$$

Let $E = \{$Reject at least one true $H^B_{AS,0}, H^B_{NI,0}\}$. Then

$$P_{H^B_{AS,0} \cap H^B_{NI,0}}(E) = P_{H^B_{AS,0} \cap H^B_{NI,1}}(E) = P_{H_{AS,1} \cap H^B_{NI,0}}(E) = \alpha, \ \text{ and } P_{H_{AS,1} \cap H_{NI,1}}(E) = 0.$$

Lemma 2.1 guarantees that the existence of c_1 and c_2 that give exact control of the FWER at level α under the boundary conditions of $H_{\mathrm{AS},0}$ and $H_{\mathrm{NI},0}$ hypotheses. The thresholds c_1 and c_2 are in Lemma 2.1 are optimal in the sense that c_1 and c_2 cannot be improved without losing control on α, the level of the FWER. However, it is not possible to obtain expressions for c_1 and c_2 when the distributions of the samples from the three-arms are unknown. It is worth mentioning that it would not be easy to compute the finite sample distribution of $T_{NI}|T_{AS} \ge c_1$ for normally distributed samples. Hence, it is required to obtain consistent estimators for c_1 and c_2 for the practical implementation of the suggested method.

c_1 can be easily estimated by z_α, the upper α quantile of the standard normal distribution, since the distribution of T_{AS} is asymptotically N(0,1) as $n_R + n_P \to \infty$. The following Lemma 2.2 gives the asymptotic distribution of $T_{\mathrm{NI}}|T_{\mathrm{AS}} > z_\alpha$ as $\min\{n_E, n_R, n_P\} \to \infty$, which will be used for obtaining a consistent estimator for c_2.

Let us denote $n = n_R + n_P + n_E$ and $\lambda_{n,l} = n_l/n$. The asymptotic distribution of $T_{\mathrm{NI}}|T_{\mathrm{AS}} > z_\alpha$ is based on the following assumptions as $\min\{n_R, n_P, n_E\} \to \infty$:

A1 $\lambda_{n,l} \to \lambda_l \in (0,1)$, for $l \in \{E, R, P\}$.
A2 $\frac{n_R}{n_R+n_P} \to \eta_R$.
A3 $\sqrt{n_R+n_P}(\mu_R - \mu_P) \to \delta \geq 0$.

Lemma 2.2. *Let $F_{T_{NI}|T_{AS}>z_\alpha}(\cdot)$ denote the condition distribution of T_{NI} given $T_{AS} > z_\alpha$. Under the configuration that $H_{AS,1}$ and $H^B_{NI,0}$ are true and under the assumptions A1-A3,*

$$F_{T_{NI}|T_{AS}>z_\alpha}(x) \to F_{ESN}(x) \ \ as \ \ \min\{n_E, n_R, n_P\} \to \infty,$$

for any real x; where $F_{ESN}(x)$ is distribution function of the extended skew-normal distribution.

An excellent review of extended skew-normal distribution can be found in (Azzalini (2005, 2014). In our setup, the extended skew-normal distribution, $F_{\text{ESN}}(x)$, has the following density function

$$f_{\text{ESN}}(x) = \phi(x)\frac{\Phi(\tau\sqrt{1+\gamma^2}+\gamma x)}{\Phi(\tau)}, \ \ x \in \mathcal{R},$$

where $\tau = \frac{\delta}{\sqrt{b}} - z_\alpha$, $\gamma = \rho/(1-\rho^2)$, $\rho = \frac{1}{\sqrt{a_1 a_2}}[(1-\theta)\frac{\sigma_P^2}{\lambda_p} - \theta\frac{\sigma_R^2}{\lambda_r}]$ is the asymptotic correlation between T_{AS} and T_{NI}, and

$$a_1 = \frac{\sigma_E^2}{\lambda_E} + \theta^2\frac{\sigma_R^2}{\lambda_R} + (1-\theta)^2\frac{\sigma_P^2}{\lambda_P}, \ a_2 = \frac{\sigma_R^2}{\lambda_R} + \frac{\sigma_P^2}{\lambda_P}, \ \text{and} \ b = \frac{\sigma_R^2}{\eta_R} + \frac{\sigma_P^2}{\eta_P}.$$

Lemma 2.2 gives a large sample approximation to $(1-\alpha/\beta)$th quantile of the distribution $T_{\text{NI}}|T_{\text{AS}} > z_\alpha$. Let $q_{(1-\alpha/\beta)}$ be the $(1-\alpha/\beta)$ quantile of $F_{\text{ESN}}(x)$, then $q_{(1-\alpha/\beta)}$ is a large sample approximation to the c_2. Based on these large sample estimates to c_1 and c_2, the suggested hierarchical testing method for $\text{H}_{\text{AS},0}$ and $\text{H}_{\text{NI},0}$ is running as follows,

Step 1: To test $H_{\text{AS},0}$, if $T_{\text{AS}} > z_\alpha$, reject $H_{\text{AS},0}$; otherwise, stop.
Step 2: If $H_{\text{AS},0}$ is rejected, test $H_{\text{NI},0}$. If $T_{\text{NI}} > q_{(1-\alpha/\beta)}$, reject $H_{\text{NI},0}$ or declare non-inferiority; otherwise, accept $H_{\text{NI},0}$.

We call this method *asymptotic hierarchical testing* (*AHT*). The next result shows that *AHT* asymptotically controls FWER at any level α.

Theorem 2.1. *Set $c_1 = z_\alpha$ and $c_2 = q_{(1-\alpha/\beta)}$. Let $E = \{$Reject at least one true $H^B_{AS,0}, H^B_{NI,0}\}$. Under the assumptions of Lemma 2.2,*

$$\lim_{\min\{n_E,n_R,n_P\}\to\infty} P_{H^B_{AS,0}\cap H^B_{NI,0}}(E) = \lim_{\min\{n_E,n_R,n_P\}\to\infty} P_{H^B_{AS,0}\cap H_{NI,1}}(E) = \lim_{\min\{n_E,n_R,n_P\}\to\infty} P_{H_{AS,1}\cap H^B_{NI,0}}(E) = \alpha,$$

and

$$\lim_{\min\{n_E,n_R,n_P\}\to\infty} P_{H_{AS,1}\cap H_{NI,1}}(E) = 0.$$

Theorem 2.1 establishes that the suggested *AHT* method asymptotically controls FWER for the boundary cases, $\mu_R - \mu_P = 0$ and $\mu_E - \theta\mu_R - (1-\theta)\mu_P = 0$ of $H_{\text{AS},0}$ and $H_{\text{NI},0}$, respectively. However, when $\mu_R - \mu_P$ is an interior point of $(-\infty, 0]$,

$$\lim_{\min\{n_R,n_P\}\to\infty} P(T_{\text{AS}} > z_\alpha) = 0.$$

This can be observed from fact that

$$T_{\text{AS}} = \frac{\bar{X}_R - \bar{X}_P}{\sqrt{\hat{\sigma}_R^2/n_R + \hat{\sigma}_P^2/n_P}} = \frac{(\bar{X}_R - \mu_R) - (\bar{X}_P - \mu_P)}{\sqrt{\hat{\sigma}_R^2/n_R + \hat{\sigma}_P^2/n_P}} + \frac{\mu_R - \mu_P}{\sqrt{\hat{\sigma}_R^2/n_R + \hat{\sigma}_P^2/n_P}} \to -\infty \text{ in probability},$$

since by CLT, $\frac{(\bar{X}_R-\mu_R)-(\bar{X}_P-\mu_P)}{\sqrt{\hat{\sigma}_R^2/n_R+\hat{\sigma}_P^2/n_P}}$ converges in distribution to N(0,1) distribution, and $\frac{\mu_R-\mu_P}{\sqrt{\hat{\sigma}_R^2/n_R+\hat{\sigma}_P^2/n_P}} \to -\infty$ in probability as $\min\{n_R, n_P\} \to \infty$.

For $\mu_E - \theta\mu_R - (1-\theta)\mu_P < 0$ and $\mu_R - \mu_P > 0$,

$$\lim_{\min\{n_R,n_P,n_E\}\to\infty} P(T_{\text{AS}} > q_{1-\alpha/\beta}|T_{\text{AS}} > z_\alpha) = 0.$$

Because

$$T_{\mathrm{NI}} = \frac{(\bar{X}_E - \mu_E) - \theta(\bar{X}_R - \mu_R) - (1-\theta)(\bar{X}_P - \mu_P)}{\sqrt{\frac{\hat{\sigma}_E^2}{n_E} + \frac{\theta^2\hat{\sigma}_R^2}{n_R} + \frac{(1-\theta)^2\hat{\sigma}_P^2}{n_P}}} + \frac{\mu_E - \theta\mu_R - (1-\theta)\mu_P}{\sqrt{\frac{\hat{\sigma}_E^2}{n_E} + \frac{\theta^2\hat{\sigma}_R^2}{n_R} + \frac{(1-\theta)^2\hat{\sigma}_P^2}{n_P}}} \to -\infty$$

in probability as $\min\{n_R, n_P, n_E\} \to \infty$. Thus, we can conclude that the asymptotic FWER of *AHT* is 0 when either $\mu_R - \mu_P$ or $\mu_E - \theta\mu_R - (1-\theta)\mu_P$ is an interior point of $(-\infty, 0]$, and consequently we have

$$\lim_{\min\{n_E, n_R, n_P\}\to\infty} P(\text{Reject at least one true } H_{\mathrm{AS},0}, H_{\mathrm{NI},0}\} \le \alpha.$$

Numerically in Table 3, we show that *AHT* controls FWER for large sample sizes when either $\mu_R - \mu_P$ or $\mu_E - \theta\mu_R - (1-\theta)\mu_P$ is an interior point of $(-\infty, 0]$.

Another crucial observation is that β appears in *AHT* method and Theorem 2.1 is unknown, and hence β is required to estimate to apply the method. To facilitate the discussion of estimation of β, let us take

$$\mu_R - \mu_P = c\sqrt{\frac{\sigma_R^2}{n_R} + \frac{\sigma_P^2}{n_P}},$$

where $c > 0$ is a constant. Under this setup β can be expressed as

$$\beta = P\left(\tilde{T}_{\mathrm{AS}} + \frac{\mu_R - \mu_P}{\sqrt{\frac{\hat{\sigma}_R^2}{n_R} + \frac{\hat{\sigma}_P^2}{n_P}}} > z_\alpha\right) = P\left(\tilde{T}_{\mathrm{AS}} + c\sqrt{\frac{\sigma_R^2/n_R + \sigma_P^2/n_P}{\hat{\sigma}_R^2/n_R + \hat{\sigma}_P^2/n_P}} > z_\alpha\right),$$

where $\tilde{T}_{\mathrm{AS}} = \frac{(\bar{X}_R - \bar{X}_R) - (\bar{X}_P - \bar{X}_P)}{\sqrt{\hat{\sigma}_R^2/n_R + \hat{\sigma}_P^2/n_P}}$. Under the assumption of finiteness of the fourth order moment of the distributions $F_l(\cdot)$,

$$\hat{\sigma}_l^2 = \sigma_l^2 + O_p(n_l^{-1/2}), \tag{6}$$

$l \in \{R, P\}$. From assumption A2, (6) and a simple application of a Taylor series expansion, we may deduce that

$$\sqrt{\frac{\sigma_R^2/n_R + \sigma_P^2/n_P}{\hat{\sigma}_R^2/n_R + \hat{\sigma}_P^2/n_P}} = 1 + O_p((n_R + n_P)^{-1/2}). \tag{7}$$

Using (7), β can be expressed as

$$\begin{aligned}\beta &= P(\tilde{T}_{\mathrm{AS}} + c(1 + O_p((n_R + n_P)^{-1/2})) > z_\alpha) \\ &= P(\tilde{T}_{\mathrm{AS}} + c + O_p((n_R + n_P)^{-1/2} > z_\alpha) \\ &= P(\tilde{T}_{\mathrm{AS}} > z_\alpha - c) + O((n_R + n_P)^{-1/2})),\end{aligned}$$

the last line follows from an application of delta method (Hall (1992)).

We apply the bootstrap to estimate β. To simplify the discussion on the bootstrap estimate of β, let $\bar{X}_R^*$, and $\hat{\sigma}_R^{*2}$ denote the bootstrap version of $\bar{X}_R$, and $\hat{\sigma}_R^2$, computed from a resample $\{X_{R,1}^*, \ldots, X_{R,n_R}^*\}$ from $\{X_{R,1}, \ldots, X_{R,n_R}\}$. Similarly, $\bar{X}_P^*$, and $\hat{\sigma}_P^{*2}$ are the versions of $\bar{X}_P$, and $\hat{\sigma}_P^2$, computed from a $\{X_{P,1}^*, \ldots, X_{P,n_P}^*\}$ from $\{X_{P,1}, \ldots, X_{P,n_P}\}$. The bootstrap version of $\tilde{T}_{\mathrm{AS}} + c\sqrt{\frac{\sigma_R^2/n_R + \sigma_P^2/n_P}{\hat{\sigma}_R^2/n_R + \hat{\sigma}_P^2/n_P}}$ is

$$\tilde{T}_{\mathrm{AS}}^* + \hat{c}\sqrt{\frac{\hat{\sigma}_R^2/n_R + \hat{\sigma}_P^2/n_P}{\hat{\sigma}_R^{*2}/n_R + \hat{\sigma}_P^{*2}/n_P}},$$

where $\tilde{T}_{\mathrm{AS}}^* = \frac{(\bar{X}_R^* - \mu_R) - (\bar{X}_P^* - \mu_P)}{\sqrt{\hat{\sigma}_R^{*2}/n_R + \hat{\sigma}_P^{*2}/n_P}}$ and $\hat{c} = \frac{\bar{X}_R - \bar{X}_P}{\sqrt{\hat{\sigma}_R^2/n_R + \hat{\sigma}_P^2/n_P}}$. The bootstrap estimate of β is

$$\hat{\beta}_B = P_*\left(\tilde{T}_{\mathrm{AS}}^* + \hat{c}\sqrt{\frac{\hat{\sigma}_R^2/n_R + \hat{\sigma}_P^2/n_P}{\hat{\sigma}_R^{*2}/n_R + \hat{\sigma}_P^{*2}/n_P}} > z_\alpha\right),$$

where $P_*(\mathcal{A})$ denotes the conditional probability of the event $\mathcal{A}$ given $\{X_{R,1}, \ldots, X_{R,n_R}\}$, and $\{X_{P,1}, \ldots, X_{P,n_P}\}$. Under the assumption of finiteness of the fourth order moment, we can have

$$\hat{\sigma}_l^{*2} = \hat{\sigma}_l^2 + O_p(n_l^{-1/2}), \tag{8}$$

$l \in \{R, P\}$. Under the assumption A2 along with (8) and a simple application of a Taylor series expansion, we can have the expansion of $\sqrt{\frac{\hat{\sigma}_R^2/n_R+\hat{\sigma}_P^2/n_P}{\hat{\sigma}_R^{*2}/n_R+\hat{\sigma}_P^{*2}/n_P}}$ as below

$$\sqrt{\frac{\hat{\sigma}_R^2/n_R+\hat{\sigma}_P^2/n_P}{\hat{\sigma}_R^{*2}/n_R+\hat{\sigma}_P^{*2}/n_P}} = 1 + O_p((n_R+n_P)^{-1/2}). \tag{9}$$

Using (9), $\hat{\beta}_B$ can be

$$\begin{aligned} P_*\Bigg(\tilde{T}^*_{\text{AS}} + c\sqrt{\frac{\hat{\sigma}_R^2/n_R+\hat{\sigma}_P^2/n_P}{\hat{\sigma}_R^{*2}/n_R+\hat{\sigma}_P^{*2}/n_P}} > z_\alpha\Bigg) &= P_*\Big(\tilde{T}^*_{\text{AS}} + c(1+O_p((n_R+n_P)^{-3/2})) > z_\alpha\Big) \\ &= P_*\Big(\tilde{T}^*_{\text{AS}} + c > z_\alpha\Big) \\ &= P_*\Big(\tilde{T}^*_{\text{AS}} > z_\alpha - c\Big) + O((n_R+n_P)^{-1/2})), \end{aligned} \tag{10}$$

the last line follows from an application of delta method (Hall (1992)). The bootstrap estimate $P_*\Big(\tilde{T}^*_{\text{AS}} > z_\alpha - c\Big)$ of $P\Big(\tilde{T}_{\text{AS}} > z_\alpha - c\Big)$ is second order correct (Hall and Martin (1988)), i.e. $P_*\Big(\tilde{T}^*_{\text{AS}} > z_\alpha - c\Big) = P\Big(\tilde{T}_{\text{AS}} > z_\alpha - c\Big) + O_p((n_R+n_P)^{-1}))$, so (10) implies that

$$P_*\Bigg(\tilde{T}^*_{\text{AS}} + c\sqrt{\frac{\hat{\sigma}_R^2/n_R+\hat{\sigma}_P^2/n_P}{\hat{\sigma}_R^{*2}/n_R+\hat{\sigma}_P^{*2}/n_P}} > z_\alpha\Bigg) = P\Big(\tilde{T}_{\text{AS}} > z_\alpha - c\Big) + O_p((n_R+n_P)^{-1/2})) = \beta + O_p((n_R+n_P)^{-1/2})).$$

Hence $\hat{\beta}_B$ is a consistent estimator of β. We may easily deduce that *AHT* asymptotically controls FWER at level α if is replaced by $\hat{\beta}_B$.

Due to the nonexistence of explicit expression of $\hat{\beta}_B$, we can use Monte Carlo approximation to $\hat{\beta}_B$. Let $\{\boldsymbol{X}^*_{R,b} = (X^*_{R,1},\dots,X^*_{R,n_R}), b=1,\dots,M\}$, and $\{\boldsymbol{X}^*_{P,b} = (X^*_{P,1},\dots,X^*_{P,n_P}), b=1,\dots,M\}$ denote M independent and identically distributed resamples from $\{X_{R,1},\dots,X_{R,n_R}\}$, and $\{X_{P,1},\dots,X_{P,n_P}\}$, respectively. The Monte Carl approximation bootstrap estimate to $\hat{\beta}_B$ is then

$$\hat{\beta}_B = \frac{1}{M}\sum_{b=1}^{M}\mathcal{I}(\tilde{T}^*_{\text{AS},b} > z_\alpha - \frac{\bar{X}_R - \bar{X}_P}{\sqrt{\hat{\sigma}_R^2/n_R+\hat{\sigma}_P^2/n_P}}).$$

Let $\tilde{T}^*_{\text{AS},1},\dots,\tilde{T}^*_{\text{AS},M}$ be the M values of T_{AS} based M bootstrap samples of where $\mathcal{I}(\cdot)$ is the indicator function, and $\tilde{T}^*_{\text{AS},b} = \frac{(\bar{X}^*_{R,b}-\bar{X}_R)-(\bar{X}^*_{P,b}-\bar{X}_P)}{\sqrt{\hat{\sigma}^{*2}_{R,b}/n_R+\hat{\sigma}^{*2}_{P,b}/n_P}}$.

Alternative to the bootstrap estimate of β is based on the normality approximation to the distribution of T_{AS} under $\text{H}_{\text{AS},1}$. The normal approximation based estimate of β can obtained as follows

$$\hat{\beta}_N = 1 - \Phi\Bigg(z_\alpha - \frac{\bar{X}_R - \bar{X}_P}{\sqrt{\hat{\sigma}_R^2/n_R+\hat{\sigma}_P^2/n_P}}\Bigg).$$

In our simulation, we observe that $\hat{\beta}_B$ gives better control on FWER of the *AHT* than that of $\hat{\beta}_N$. Thus, empirical FWER and power of the *AHT* are reported using the estimate $\hat{\beta}_B$ of β. We now analyze the power of *AHT* method.

Theorem 2.2. *Let* $\pi_{1,n} = P(T_{AS} > z_\alpha, T_{NI} > q_{(1-\alpha/\beta)} | H_{AS,1}, H_{NI,1})$ *denote the power of AHT. Let* $\lim_{\min\{n_E,n_R,n_P\}\to\infty}\pi_{1,n} = \pi_{1,\infty}$. *Then*

(a) $\pi_{1,\infty} = 1$.
(b) *Let* $\mu_E - \theta\mu_R - (1-\theta)\mu_p = \frac{h}{n}$. *Under the assumptions A1-A3, then the asymptotic power of AHT has expression*

$$\pi_{1,\infty} = \Bigg[1 - \Phi\Bigg(z_\alpha - \frac{\delta}{\sqrt{a_2}}\Bigg)\Bigg]\Bigg[1 - F_{ESN}\Bigg(q_{(1-\alpha/\beta)} - \frac{h}{\sqrt{a_1}}\Bigg)\Bigg],$$

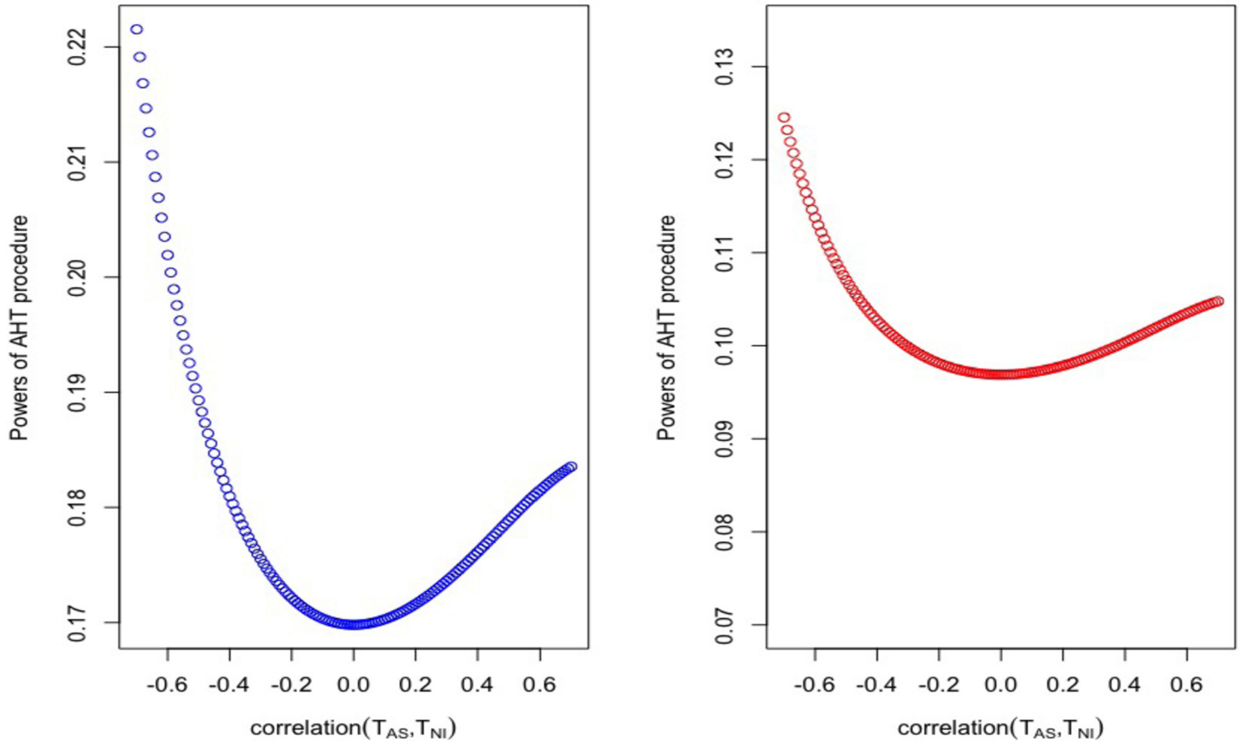

Fig. 1. Power curve of *ATH* procedure as function of ρ; $n_E = 100, n_R = 100, n_P = 100$.

Theorem 2.2(a) confirms that the *AHT* method is consistent against the alternative hypotheses in the sense that when $\min\{n_E, n_R, n_P\} \to \infty$ *AHT* sets forth the non-inferiority of the experiment treatment when both the alternative hypotheses $H_{AS,1}$ and $H_{NI,1}$ are true. Theorem 2.2(b) considers the local asymptotic power of *AHT*. It reflects the characteristic that the *AHT*'s power is between α and one when the values of $\mu_R - \mu_P$, and $\mu_E - \theta\mu_R - (1-\theta)\mu_P$ are close to the boundary between the parameter spaces corresponding to the null and alternative hypotheses. By using the power function in Theorem 2.2(b), we can also compute the approximated power of *AHT* at $\mu_R - \mu_P$ and $\mu_E - \theta\mu_R - (1-\theta)\mu_P$ using equating $\mu_R - \mu_P = \frac{\delta}{n_R+n_p}$, and $\mu_E - \theta\mu_R - (1-\theta)\mu_p = \frac{h}{n}$. The next result discuss the effect of asymptotic correlation ρ, between T_{AS} and T_{NI}, on asymptotic power of *AHT*.

Corollary 2.1. *Under the setup of Theorem 2.2(b),*

$$\pi_{1,\infty} \text{ is an increasing function of } \rho \text{ when } \rho > 0,$$

and

$$\pi_{1,\infty} \text{ is a decreasing function of } \rho \text{ when } \rho < 0.$$

The expression of ρ is discussed in Theorem 2.2. Corollary 2.1 establishes the asymptotic power of *AHT* for fixed values of $\mu_R - \mu_P$ and $\mu_E - \theta\mu_R - (1-\theta)\mu_P$, $\pi_{1,n}$ increases with increased values of ρ when ρ is positive and $\pi_{1,n}$ decreases with increased values of ρ when ρ is negative.

Fig. 1 illustrates Corollary 2.1 for large n_R, n_P, and n_E. The powers $\pi_{1,n}$ of *AHT* are approximated for different values of ρ using the expression in Theorem 2.2(b). Left side of the figure was generated under X_P ∼normal(2,5), X_R ∼normal (2.8, 5), and X_R ∼normal (2.88, 3). Right side of the figure was generated under $X_P \sim$ Exp (1/2), $X_R \sim$ Exp(1/2.8), and X_E ∼Exp (1/2.88). The figure shows that the asymptotic power of *AHT* is a decreasing and also is an increasing function of ρ when $\rho < 0$, and $\rho > 0$, respectively.

3. Simulation study

In this section we present empirical results based on simulated data. We compare the suggested *AHT* method with (Pigeot et al. (2003)) and Hasler et al. (2008)). For our convenience we refer the Pigeot et al. (2003) and Hasler et al. (2008) tests as the traditional (*tradi*) and *Hasler*, respectively. We considered three scenarios to generate the data. In Scenario 1, X_E,

Table 1
The powers of assay sensitivity test at 2.5% level.

Sample Sizes	$(n_E = n_R = n_P = 100)$	$(n_E = 150, n_R = 150, n_P = 75)$	$(n_E = n_R = n_P = 300)$
Scenario 1	0.72	0.72	0.99
Scenario 2	0.71	0.71	0.99
Scenario 3	0.64	0.69	0.98
Sample Sizes	$(n_E = n_R = n_P = 100)$	$(n_E = 150 n_R = 150 n_P = 75)$	$(n_E = n_R = n_P = 300)$
Method = *AHT*	Case 2	Case 2	Case 2
Scenario 1	0.81	0.75	0.99
Scenario 2	0.80	0.80	0.99
Scenario 3	0.72	0.78	0.99

Table 2
Empirical FWER at 2.5% level corresponding to the configuration $H_{\text{AS},1} \cap H^B_{\text{NI},0}$.

Sample Sizes	$(n_E = n_R = n_P = 100)$	$(n_E = 150 n_R = 150 n_P = 75)$	$(n_E = n_R = n_P = 300)$
Method = *AHT*	Case 1	Case1	Case1
Scenario 1	0.023	0.021	0.025
Scenario 2	0.017	0.021	0.024
Scenario 3	0.023	0.022	0.026
Sample Sizes	$(n_E = n_R = n_P = 100)$	$(n_E = 150 n_R = 150 n_P = 75)$	$(n_E = n_R = n_P = 300)$
Method = *AHT*	Case 2	Case 2	Case 2
Scenario 1	0.023	0.026	0.024
Scenario 2	0.022	0.019	0.027
Scenario 3	0.026	0.022	0.028

Table 3
Empirical FWER at 2.5% level corresponding to the configuration $H_{\text{AS},1} \cap H^I_{\text{NI},0}$ corresponding to Scenario 1.

Sample Sizes	$(n_E = n_R = n_P = 100)$	$(n_E = 150 n_R = 150 n_P = 75)$	$(n_E = n_R = n_P = 300)$
Method=*AHT*	Case 1	Case1	Case1
$\mu_E - \theta\mu_R - (1-\theta)\mu_P = -0.02$	0.021	0.018	0.014
$\mu_E - \theta\mu_R - (1-\theta)\mu_P = -0.04$	0.016	0.012	0.010
$\mu_E - \theta\mu_R - (1-\theta)\mu_P = -0.54$	0	0	0
Sample Sizes	$(n_E = n_R = n_P = 100)$	$(n_E = 150 n_R = 150 n_P = 75)$	$(n_E = n_R = n_P = 300)$
Method = *AHT*	Case 2	Case 2	Case 2
$\mu_E - \theta\mu_R - (1-\theta)\mu_P = -0.02$	0.020	0.016	0.012
$\mu_E - \theta\mu_R - (1-\theta)\mu_P = -0.04$	0.018	0.015	0.012
$\mu_E - \theta\mu_R - (1-\theta)\mu_P = -0.54$	0	0	0

X_R, and X_P were generated from normal$(\mu_E, 3)$, normal$(\mu_R, 5)$, and normal$(\mu_p, 5)$. In Scenario 2, X_E, X_R, and X_P were generated from double-exponential $(\mu_E$, 1.2), double-exponential $(\mu_R$, 1.6), and double-exponential $(\mu_P$, 1.6). In Scenario 3, X_E, X_R, and X_P were generated from $\text{Exp}(1/\mu_E)$, $\text{Exp}(1/\mu_R)$, and $\text{Exp}(1/\mu_p)$. Scenarios 2 and 3 consider heavier tailed distributions and skewed distributions, respectively. Second and third scenarios assess the robustness of our method and its nonparametric extension.

Under each scenario, we considered two cases. In Cases 1, the value of the pair (μ_R, μ_P) was set at (2.8, 2) and in Case 2, we chose (μ_R, μ_P) as (2.9, 2). For each combination of scenario and case, we considered three sets of the triple (n_E, n_R, n_P). They were (100, 100, 100), (150, 150, 75), and (300, 300, 300). The powers of the assay sensitivity test corresponding to the Case 1 and Case 2 are displayed in Table 1.

Under each combination of scenarios, cases, and sample sizes, the empirical FWERs for *AHT* method were computed based on 2000 random samples. And empirical powers of *AHT*, *tradi*, and *Hasler* were also calculated based on 2000 random samples. To evaluate the attained FWER of *AHT*, we computed empirical FWERs under the configuration $H_{\text{AS},1} \cap H_{\text{NI},0}$ with $\theta = 0.8$. Empirical FWERs are reported in Tables 2-3.

For a given value of (μ_R, μ_P), we examined powers of *AHT*, *tradi*, and *Hasler* at $(\mu_E - \mu_P)/(\mu_R - \mu_P) > 0.80$, i.e. the powers were calculated at $(\mu_E - \mu_P)/(\mu_R - \mu_P) = 0.90, 1.00, 1.10, 1.20$, and 1.50. For computing empirical powers of *AHT*, $\beta = \text{P}(T_{\text{AS}} > z_\alpha | \text{H}_{\text{AS},1})$ was replaced by its bootstrap estimate $\hat{\beta}_B$. Along with the empirical powers of *AHT*, we also reported the powers of *AHT* using the asymptotic formula given in Theorem 2.2 (b). They were displayed in parentheses in Tables 4–6.

Results from simulation data presented in Table 2 displays the empirical FWERs of *AHT* method under the configuration $H_{\text{AS},1} \cap H^B_{\text{NI},0}$, where $H^B_{\text{NI},0} : \mu_E - \theta\mu_R - (1-\theta)\mu_R = 0$. The results in Table 2 show that empirical FWER of *AHT* method are close the nominal size 0.025 under the all scenarios and all the cases, and they range from 0.017 to 0.028. Thus, Table 2

Table 4
Empirical Power at 2.5% FWER level corresponding to Scenario 1. The bracketed values are asymptotic powers based on Theorem 2.2(b).

$n_E = n_R = n_P = 100$	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$=0.90	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.00	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.10	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.20	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$=1.50
Method			Case 1/Case 2		
AHT	0.055 (0.049)/0.056 (0.054)	0.093 (0.089)/0.114 (0.104)	0.143 (0.147)/0.171 (0.181)	0.214 (0.222)/0.286 (0.281)	0.466 (0.492)/0.601 (0.617)
tradi	0.028/0.033	0.037/0.073	0.080/0.119	0.143/0.192	0.389/0.550
Hasler	0.030/0.031	0.043/0.069	0.081/0.118	0.140/0.210	0.367/0.521

$n_E = 150, n_R = 150, n_P = 75$	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$=0.90	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.00	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.10	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.20	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.50
Method			Case 1/Case 2		
AHT	0.052 (0.054)/0.064 (0.061)	0.097 (0.106)/0.136(0.128)	0.180 (0.183)/0.235 (0.230)	0.254 (0.282)/0.350 (0.362)	0.555 (0.581)/0.714(0.712)
tradi	0.040/0.045	0.063/0.102	0.139/0.211	0.213/0.306	0.523/0.687
Hasler	0.031/0.040	0.063/0.101	0.141/0.196	0.217/0.312	0.535/0.674

$n_E = n_R = n_P = 300$	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$=0.90	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.00	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.10	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.20	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.50
Method			Case 1/Case 2		
AHT	0.083 (0.079)/0.082(0.089)	0.208 (0.195)/0.249(0.234)	0.389 (0.378)/0.469(0.456)	0.589(0.592)/0.700(0.693)	0.966(0.963)/0.990 (0.989)
tradi	0.072/0.073	0.183/0.240	0.374/0.461	0.578/0.694	0.960/0.990
Hasler	0.080/0.075	0.186/0.230	0.376/0.445	0.585/0.712	0.970/0.987

Table 5
Empirical Power at 2.5% FWER level corresponding to Scenario 2. The bracketed values are asymptotic powers based on Theorem 2.2(b).

$n_E = n_R = n_P = 100$	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$=0.90	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.00	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.10	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.20	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ =1.50
Method			Case 1/Case 2		
AHT	0.048(0.049)/0.051(0.054)	0.076(0.088)/0.111(0.104)	0.141 (0.146)/0.178(0.180)	0.204(0.220)/0.277(0.279)	0.435(0.485)/0.594(0.609)
tradi	0.017/0.027	0.046/0.064	0.071/0.111	0.137/0.215	0.387/0.526
Hasler	0.018/0.036	0.034/0.063	0.076/0.125	0.131/0.212	0.367/0.521
$n_E = 150, n_R = 150, n_P = 75$	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$=0.90	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.00	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.10	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.20	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.50
Method			Case 1/Case 2		
AHT	0.045 (0.054)/0.062 (0.060)	0.092 (0.106)/0.134 (0.127)	0.163 (0.181)/0.219 (0.228)	0.257(0.278)/0.357 (0.357)	0.545 (0.571)/0.690 (0.703)
tradi	0.031/0.044	0.069/0.109	0.127/0.181	0.214/0.305	0.514/0.666
Hasler	0.033/0.047	0.0625/0.091	0.1230/0.191	0.2105/0.304	0.5260 /0.665
$n_E = n_R = 300 = n_P = 300$	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$=0.90	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.00	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.10	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.20	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.50
Method			Case 1/Case 2		
AHT	0.091(0.079)/0.107 (0.089)	0.213 (0.194)/0.229 (0.232)	0.387 (0.375)/0.457 (0.453)	0.588 (0.589)/0.700 (0.690)	0.956 (0.961)/0.986 (0.988)
tradi	0.077/0.098	0.190/0.225	0.377/0.437	0.570/0.690	0.955/0.973
Hasler	0.078/0.097	0.173/0.243	0.359 /0.426	0.570/0.701	0.9640/0.991

Table 6
Empirical Power at 2.5% FWER level corresponding to Scenario 3. The bracketed values are asymptotic powers based on Theorem 2.2(b).

$n_E = n_R = n_P = 100$	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$=0.90	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.00	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.10	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.20	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ =1.50
Method			Case 1/Case 2		
AHT	0.040 (0.041)/0.037 (0.043)	0.054 (0.063)/0.065 (0.069)	0.073 (0.092)/0.093 (0.104)	0.103 (0.127)/0.137 (0.147)	0.225 (0.255)/0.306 (0.304)
tradi	0.012/0.017	0.021/0.029	0.031/0.049	0.045/0.081	0.154/0.198
Hasler	0.012/0.018	0.021 /0.026	0.038/0.052	0.057/0.068	0.141/0.195
$n_E = 150, n_R = 150, n_P = 75$	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$=0.90	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.00	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.10	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.20	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.50
Method			Case 1/Case 2		
AHT	0.060 (0.066)/0.092 (0.011)	0.098 (0.084)/ 0.112 (0.118)	0.127 (0.113)/0.160 (0.161)	0.159 (0.162)/0.213(0.209)	0.326 (0.342)/ 0.395 (0.410)
tradi	0.043/0.067	0.071/0.085	0.087/0.118	0.127/0.173	0.249/0.331
Hasler	0.035/0.058	0.069 /0.083	0.091 /0.117	0.131/0.176	0.2720/ 0.351
$n_E = n_R = 300 = n_P = 300$	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$=0.90	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.00	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.10	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.20	$\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$ = 1.50
Method			Case 1/Case 2		
AHT	0.062(0.059)/0.074 (0.063)	0.136(0.118)/0.143 (0.132)	0.219 (0.208)/0.233 (0.235)	0.337 (0.322)/0.378 (0.366)	0.702 (0.694)/0.771 (0.778)
tradi	0.045/0.060	0.107/0.126	0.182/0.214	0.317/0.358	0.682/0.764
Hasler	0.055 /0.059	0.100/0.125	0.180/0.232	0.303/0.357	0.6975/0.768

confirms that the suggested method controls FWER. Table 3 reports the FWERs under the configuration $H_{\mathrm{AS},1} \cap H^{I}_{\mathrm{NI},0}$, where $H^{I}_{\mathrm{NI},0} : \mu_E - \theta\mu_R - (1-\theta)\mu_R < 0$, i.e. $H^{I}_{\mathrm{NI},0}$ denotes the interior of $H_{\mathrm{NI},0}$. Table 3 reports empirical FWERs in several interior values in $(-\infty, 0]$ assumed by $\mu_E - \theta\mu_R - (1-\theta)\mu_R$. These empirical FWERs demonstrate that the FWER control of *AHT* in the interior of $H_{\mathrm{NI},0}$. The empirical FWERs for Scenarios 2 and 3 under the same setup of the Table 3, not given for brevity, are broadly similar to Scenario 1.

The empirical powers of *AHT*, *tradi*, and *Hasler* under the different combinations of scenarios and cases are presented in Tables 4–6. These tables show that empirical powers and the asymptotic powers of *AH* are alike. Tables 4–6 indicate that *AHT* method has more power than the *tradi* and *Hasler* when the power of assay sensitivity test is less than or close to 80%. When the power of assay sensitivity close to 1, *AHT* method has slightly more power than the *tradi* and *Hasler* for smaller the values of $\frac{\mu_E-\mu_P}{\mu_R-\mu_P}$. In conclusion, Tables 4–6 demonstrate based on our simulated data that the proposed *AHT* is more powerful than the *tradi* and *Hasler*.

4. Data example

We consider a study on "Mildly Asthmatic Patients to illustrate our proposed method. This study is discussed in Pigeot et al. (2003). The primary outcome variable is FVC (forced vital capacity). The data set consists of experimental ($n_E = 35$), reference ($n_R = 19$) and a placebo ($n_P = 20$) groups. The mean, and standard deviation for the outcome variable are 4.32, and 1.16 for experimental group. The mean, and standard deviation for the outcome variable for the reference group are 4.86, and 1.03. For the placebo group the mean, and standard are 3.14 and 0.97. In order to perform non-inferiority testing, first step is to establish the superiority of the reference group over the placebo. Since $T_{\mathrm{AS}} = 5.36 > z_{0.025} = 1.96$, so the assay sensitivity is established. According to Pigeot et al. (2003), we chose $\theta = 0.5$, for assessment of noninferiority. The box plots of Pigeot et al. (2003) show that the distributions of FVC for the experimental and placebo groups are symmetrical whereas the distribution FVC for the reference group is skewed.

In the absence of study data and for assessing noninferiority of experimental to reference, we independently simulated 5000 samples of sizes 35 and 20 for the experimental and placebo groups from N(4.32,1.16) and N(3.14,0.97), respectively. For the reference group we simulated 5000 samples of size $n_R = 19$ from a gamma (shape parameter=22.09, scale parameter=0.22). Based on these 5000 simulated data sets, the powers of assessing the noninferiority of experimental corresponding to *AHT*, *tradi*, and *Hasler* are 0.309, 0.279, and 0.268, respectively.

5. Discussion

This article presented a novel hierarchical testing procedure for three-arm non-inferiority trials. The developed method does not depend on distributional assumptions on the treatment responses; thus, this procedure is robust. The analytical and numerical studies indicate that the *AHT* procedure maintains FWER. Our simulation study shows that the proposed *AHT* method is more powerful than the traditional testing procedure when the assay sensitivity test's power is not close to one. Both approaches have similar power when the power of the assay sensitivity test is close to one. Thus, we advocate using the proposed method to analyze data from three-arm non-inferiority trials. Though this article concentrates on continuous data, our future research in this direction considers the extension of *AHT* method to other data types, such as binary data and count data.

Acknowledgements

We express our sincere thanks to the Associate Editor and two anonymous reviewers for their useful comments and suggestions on an earlier versions of this manuscript which led to this improved one.

Appendix A

Proof of Lemma 2.1. To prove Theorem 2.1, we consider three configurations, specifically, $H^B_{\mathrm{AS},0} \cap H^B_{\mathrm{NI},0}$, $H^B_{\mathrm{AS},0} \cap H_{\mathrm{NI},1}$, and $H_{\mathrm{AS},1} \cap H^B_{\mathrm{NI},0}$. If $H^B_{\mathrm{AS},0}$ is true then the FWER is controlled for the first two configurations $H^B_{\mathrm{AS},0} \cap H^B_{\mathrm{NI},0}$ and $H^B_{\mathrm{AS},0} \cap H_{\mathrm{NI},1}$, provided the type I error of $H^B_{\mathrm{AS},0}$ is controlled at level α. This observation is true under $H^B_{\mathrm{AS},0} \cap H^B_{\mathrm{NI},0}$ because the event, {rejecting $H^B_{\mathrm{NI},0}$}, occurs only when the event, {rejecting $H^B_{\mathrm{AS},0}$}, occurs. That is {rejecting $H^B_{\mathrm{NI},0}$} is a subset of {rejecting $H^B_{\mathrm{AS},0}$} and the probability of {rejecting $H^B_{\mathrm{AS},0}$} does not depend on the truth or falsity $H^B_{\mathrm{NI},0}$. Under $H^B_{\mathrm{AS},0} \cap H_{\mathrm{NI},1}$, type-I error of $H^B_{\mathrm{AS},0}$ is not associated with the rejection of $H^B_{\mathrm{NI},0}$. Under the third configuration, $H_{\mathrm{AS},1} \cap H^B_{\mathrm{NI},0}$,

$$\begin{aligned}
\mathrm{FWER} &= \mathrm{P}_{H_{\mathrm{AS},1}\cap H^B_{\mathrm{NI},0}}(\{\text{rejecting } H_{\mathrm{AS},1}\} \cap \{\text{rejecting } H^B_{\mathrm{NI},0}\}) \\
&= \mathrm{P}_{H_{\mathrm{AS},1}\cap H^B_{\mathrm{NI},0}}\{T_{\mathrm{AS}} > c_1\}\mathrm{P}_{H_{\mathrm{AS},1}\cap H^B_{\mathrm{NI},0}}\{T_{\mathrm{NI}} > c_2 | T_{\mathrm{AS}} > c_1\} \\
&= \alpha,
\end{aligned}$$

since, by the construction $P_{H_{AS,1}\cap H^B_{NI,0}}\{T_{AS} > c_1\} = \beta$ and $P_{H_{AS,1}\cap H^B_{NI,0}}\{T_{NI} > c_2 | T_{AS} > c_1\} = \frac{\alpha}{\beta}$. Under $H_{AS,1} \cap H_{NI,1}$, FWER=0 since both $H_{AS,0}$ and $H_{NI,0}$ are false. Hence, we have the proof.

Proof of Lemma 2.2. Let us define $Y_{E,i} = \frac{X_{E,i}-\mu_E}{\sigma_E}$, $Y_{R,j} = \frac{X_{R,j}-\mu_R}{\sigma_R}$, and $Y_{P,k} = \frac{X_{P,k}-\mu_P}{\sigma_P}$. Then under $H_{NI,0}$, T_{NI} can be expressed as

$$T_{NI} = \frac{\sqrt{n}[\sigma_E\bar{Y}_E - \theta\sigma_R\bar{Y}_R - (1-\theta)\sigma_P\bar{Y}_P - (\mu_E - \theta\mu_R - (1-\theta)\mu_P)]}{\sqrt{\frac{\sigma_E^2\hat{\tau}_E^2}{\lambda_{E,n}} + \theta^2\frac{\sigma_R^2\hat{\tau}_R^2}{\lambda_{R,n}} + (1-\theta)^2\frac{\sigma_P^2\hat{\tau}_P^2}{\lambda_{P,n}}}}$$
$$= \frac{\sqrt{n}[\sigma_E\bar{Y}_E - \theta\sigma_R\bar{Y}_R - (1-\theta)\sigma_P\bar{Y}_P]}{\sqrt{\frac{\sigma_E^2\hat{\tau}_E^2}{\lambda_{E,n}} + \theta^2\frac{\sigma_R^2\hat{\tau}_R^2}{\lambda_{R,n}} + (1-\theta)^2\frac{\sigma_P^2\hat{\tau}_P^2}{\lambda_{P,n}}}}, \tag{A.1}$$

where $\hat{\tau}_l^2 = (n_l-1)^{-1}\sum_{i=1}^{n_l}(Y_{l,i} - \bar{Y}_l)^2$ and $\lambda_{l,n} = \frac{n_l}{n}$, $l \in \{E, R, P\}$.

Now,

$$a_{1,n} = \frac{\sigma_E^2\hat{\tau}_E^2}{\lambda_{E,n}} + \frac{\sigma_R^2\hat{\tau}_R^2}{\lambda_{R,n}} + \frac{\sigma_P^2\hat{\tau}_P^2}{\lambda_{P,n}} = \frac{\sigma_E^2}{\lambda_{E,n}}(n_E-1)^{-1}(\sum_{i=1}^{n_E}Y_{E,i}^2 - n_E\bar{Y}_E^2) + \frac{\theta^2\sigma_R^2}{\lambda_{R,n}}(n_R-1)^{-1}(\sum_{i=1}^{n_R}Y_{R,i}^2 - n_R\bar{Y}_R^2)$$
$$+ \frac{(1-\theta)^2\sigma_P^2}{\lambda_{P,n}}(n_P-1)^{-1}(\sum_{i=1}^{n_P}Y_{P,i}^2 - n_P\bar{Y}_P^2) = n_E(n_E-1)^{-1}\frac{\sigma_E^2}{\lambda_{E,n}} + n_R(n_R-1)^{-1}\frac{\theta^2\sigma_R^2}{\lambda_{R,n}} + n_P(n_P-1)^{-1}\frac{(1-\theta)^2\sigma_P^2}{\lambda_{P,n}}$$
$$+ \frac{\sigma_E^2}{\lambda_{E,n}}(n_E-1)^{-1}(\sum_{i=1}^{n_E}Y_{E,i}^2 - 1) - \frac{\sigma_E^2}{\lambda_{E,n}}(n_E-1)^{-1}n_E\bar{Y}_E^2 + \frac{\theta^2\sigma_R^2}{\lambda_{R,n}}(n_R-1)^{-1}\sum_{i=1}^{n_R}(Y_{R,i}^2 - 1) - \frac{\theta^2\sigma_R^2}{\lambda_{R,n}}(n_R-1)^{-1}n_R\bar{Y}_R^2$$
$$+ \frac{(1-\theta)^2\sigma_P^2}{\lambda_{P,n}}(n_P-1)^{-1}\sum_{i=1}^{n_P}(Y_{P,i}^2 - 1) - \frac{(1-\theta)^2\sigma_P^2}{\lambda_{P,n}}(n_P-1)^{-1}n_P\bar{Y}_P^2$$
$$= \frac{\sigma_E^2}{\lambda_{E,n}} + \frac{\theta^2\sigma_R^2}{\lambda_{R,n}} + \frac{(1-\theta)^2\sigma_P^2}{\lambda_{P,n}} + O_p(N^{-1/2}), \tag{A.2}$$

the last line follows from the facts that $n_l^{-1}\sum_{i=1}^{n_l}(Y_{R,i}^2 - 1) = O_p(n_l^{-1/2})$, $\bar{Y}_l^2 = O_p(n_l^{-1})$, and $\frac{n_l}{n} = O(1)$.

Under $H_{AS,1}$, T_{AS} can be expressed as

$$T_{AS} = \frac{(\sigma_R\bar{Y}_R - \sigma_P\bar{Y}_P)}{\sqrt{\frac{\sigma_R^2\hat{\tau}_R^2}{n_R} + \frac{\sigma_P^2\hat{\tau}_P^2}{n_P}}} + \frac{(\mu_R - \mu_P)}{\sqrt{\frac{\sigma_R^2\hat{\tau}_R^2}{n_R} + \frac{\sigma_P^2\hat{\tau}_P^2}{n_P}}},$$
$$= \frac{\sqrt{n}[(\sigma_R\bar{Y}_R - \sigma_P\bar{Y}_P)]}{\sqrt{\frac{\sigma_R^2\hat{\tau}_R^2}{\lambda_{R,n}} + \frac{\sigma_P^2\hat{\tau}_P^2}{\lambda_{P,n}}}} + \frac{\sqrt{n'}(\mu_R - \mu_P)}{\sqrt{\frac{\sigma_R^2\hat{\tau}_R^2}{\eta_{R,n'}} + \frac{\sigma_P^2\hat{\tau}_P^2}{\eta_{P,n'}}}}, \tag{A.3}$$

where $n' = n_R + n_P$, $\eta_{R,n'} = \frac{n_R}{n'}$, and $\eta_{P,n'} = \frac{n_P}{n'}$. Similarly, we can have

$$a_{2,n} = \frac{\sigma_R^2\hat{\tau}_R^2}{\lambda_{R,n}} + \frac{\sigma_P^2\hat{\tau}_P^2}{\lambda_{P,n}} = \frac{\sigma_R^2}{\lambda_{R,n}} + \frac{\sigma_P^2}{\lambda_{P,n}} + O_p(n^{-1/2}), \tag{A.4}$$

and

$$b_{2,n'} = \frac{\sigma_R^2\hat{\tau}_R^2}{\eta_{R,n'}} + \frac{\sigma_P^2\hat{\tau}_P^2}{\eta_{P,n'}} = \frac{\sigma_R^2}{\eta_{R,n'}} + \frac{\sigma_P^2}{\eta_{P,n'}} + O_p(n^{-1/2}), \tag{A.5}$$

since $\frac{n}{n'} = O(1)$.

Let us introduce some notations $a_1 = \frac{\sigma_E^2}{\lambda_E} + \frac{\theta^2\sigma_R^2}{\lambda_R} + \frac{(1-\theta)^2\sigma_P^2}{\lambda_P}$, $a_2 = \frac{\sigma_R^2}{\lambda_R} + \frac{\sigma_P^2}{\lambda_P}$, $U_{NI} = \frac{\sqrt{n}}{\sqrt{a_1}}[\sigma_E\bar{Y}_E - \theta\sigma_R\bar{Y}_R - (1-\theta)\sigma_P\bar{Y}_P]$, and $V_{NI} = \frac{\sqrt{n}[(\sigma_R\bar{Y}_R - \sigma_P\bar{Y}_P)]}{\sqrt{a_2}}$. Based on Equations (A.1)-(A.5), we can have following expression for $(T_{NI}, T_{AS})^\top$

$$(T_{NI}, T_{AS})^\top = A_n(U_{NI}, V_{AS})^\top + (0, \frac{\sqrt{n'}[\mu_R - \mu_p]}{\sqrt{b_{2,n'}}})^\top + O_p(n^{-1/2}), \tag{A.6}$$

where

$$A_n = \begin{bmatrix} \sqrt{\frac{a_1}{a_{1,n}}} & 0 \\ 0 & \sqrt{\frac{a_2}{a_{2,n}}}, \end{bmatrix}.$$

To obtain the distribution of $(T_{\text{NI}}, T_{\text{AS}})^\top$, we now consider the joint distribution of $(U_{\text{NI}}, V_{\text{AS}})^\top$. Under the assumption of Theorem 2.1, $\sqrt{n_l}\bar{Y}_l \to \text{N}(0, \lambda_l^{-1})$ in distribution as $\min\{n_E, n_R, n_P\} \to \infty$, where $l \in \{E, R, P\}$. By applying the delta method (Serfling (1980), pp 122) to sequence $\{(U_{\text{NI}}, V_{\text{AS}})^\top\}$, we can show that $(U_{\text{NI}}, V_{\text{AS}})^\top \to \text{N}_2(\mathbf{0}, \Sigma)$ in distribution $\min\{n_E, n_R, n_P\} \to \infty$, where $\mathbf{0}$ is zero vector, and

$$\Sigma = \begin{bmatrix} 1 & (a_1a_2)^{-1/2}[\frac{(1-\theta)\sigma_P^2}{\lambda_P} - \frac{\theta\sigma_R^2}{\lambda_R}] \\ (a_1a_2)^{-1/2}[\frac{(1-\theta)\sigma_P^2}{\lambda_P} - \frac{\theta\sigma_R^2}{\lambda_R}] & 1. \end{bmatrix}$$

By using the above limiting distribution of $(U_{\text{NI}}, V_{\text{AS}})^\top$ along with Theorem 4.19 in Polansky (2011) in (A.6), we can conclude that

$$(T_{\text{NI}}, T_{\text{AS}})^\top \to N_2(\mu, \Sigma) \text{ as } \min\{n_E, n_R, n_P\} \to \infty,$$

where $\mu = (0, \frac{\delta}{\sqrt{b_2}})^\top$, where $b_2 = \frac{\sigma_R^2}{\eta_R} + \frac{\sigma_P^2}{\eta_P}$ is the limit of $b_{2,n'}$ as $\min\{n_R, n_P\} \to \infty$.

Since as $n \to \infty$, $(T_{\text{NI}}, T_{\text{AS}})^\top \to N_2(\mu, \Sigma)$, thus

$$T_{\text{NI}}|T_{\text{AS}} \sim \text{ESN}(0, 1, \gamma, \tau),$$

as $\min\{n_E, n_R, n_P\} \to \infty$ from the definition of extended skew-normal distribution.

Proof of Theorem 2.1. The arguments of Lemma 2.1 indicates that the suggested AHT method controls FWER asymptotically under the configurations $H^B_{\text{AS},0} \cap H^B_{\text{NI},0}$, $H^B_{\text{AS},0} \cap H_{\text{NI},1}$ since the test $T_{\text{AS}} > z_\alpha$ is asymptotically of level α. Let's consider the FWER Under configuration $H_{\text{AS},1} \cap H^B_{\text{NI},0}$,

$$\begin{aligned} \text{FWER} &= \text{P}_{H_{\text{AS},1} \cap H^B_{\text{NI},0}}\{T_{\text{AS}} > z_\alpha, T_{\text{NI}} > q_{1-\alpha/\beta}\} \\ &= \beta \times \text{P}_{H_{\text{AS},1} \cap H^B_{\text{NI},0}}\{T_{\text{NI}} > q_{1-\alpha/\beta}|T_{\text{AS}} > z_\alpha\}. \end{aligned}$$

Under $\text{H}_{\text{AS},1}$, $\beta \to \left[1 - \Phi(z_\alpha - \frac{\delta}{a_2})\right]$ and $\frac{\alpha}{\beta} \to \alpha\left[1 - \Phi(z_\alpha - \frac{\delta}{a_2})\right]^{-1} = \tau$, as $\min\{n_E, n_R, n_P\} \to \infty$. Since F^{-1}_{ESN} is a continuous function on (0,1), thus $q_{\alpha/\beta} = F^{-1}_{\text{ESN}}(\alpha/\beta) \to q_\tau$ as $\min\{n_E, n_R, n_P\} \to \infty$. Based Lemma 2.2, we can conclude that

$$\text{P}_{H_{\text{AS},1} \cap H^B_{\text{NI},0}}\{T_{\text{NI}} > q_{1-\alpha/\beta}|T_{\text{AS}} > z_\alpha\} \to \tau,$$

as $\min\{n_E, n_R, n_P\} \to \infty$. Thus,

$$\text{FWER} \to \left[1 - \Phi(z_\alpha - \frac{\delta}{a_2})\right]\tau = \alpha$$

as $\min\{n_E, n_R, n_P\} \to \infty$.

Under $H_{\text{AS},1} \cap H_{\text{NI},1}$, FWER=0 since both $H^B_{\text{AS},0}$ and $H^B_{\text{NI},0}$ are false. Hence, we have the proof.

Proof of Theorem 2.2(a). $\pi_{1,n}$ can be expressed

$$\begin{aligned} \pi_{1,n} &= \text{P}(T_{\text{AS}} > z_\alpha, T_{\text{NI}} > q_{(1-\alpha/\beta)}|\text{H}_{\text{AS},1}, \text{H}_{\text{NI},1}) = \\ &\text{P}(T^*_{\text{AS}} > z_\alpha - \frac{\mu_R - \mu_P}{\sqrt{\hat{\sigma}_R^2/n_R + \hat{\sigma}_P^2/n_P}}, T^*_{\text{NI}} > q_{(1-\alpha/\beta)} - \frac{\mu_E - \theta\mu_R - (1-\theta)\mu_P}{\sqrt{\hat{\sigma}_E^2/n_E + \theta^2\hat{\sigma}_R^2/n_R + (1-\theta)^2\hat{\sigma}_P^2/n_P}}|\text{H}_{\text{AS},1}, \text{H}_{\text{NI},1}), \end{aligned} \quad \text{(A.7)}$$

where $T^*_{\text{NI}} = \frac{(\bar{X}_E - \mu_E) - \theta(\bar{X}_R - \mu_R) - (1-\theta)(\bar{X}_P - \mu_P)}{\sqrt{\frac{\hat{\sigma}_E^2}{n_E} + \theta^2\frac{\hat{\sigma}_R^2}{n_R} + (1-\theta)^2\frac{\hat{\sigma}_P^2}{n_P}}}$, and $T^*_{\text{AS}} = \frac{(\bar{X}_R - \mu_R) - (\bar{X}_P - \mu_P)}{\sqrt{\frac{\hat{\sigma}_R^2}{n_R} + \frac{\hat{\sigma}_P^2}{n_P}}}$. The arguments in the proof of Lemma 2.2 conclude that $(T^*_{\text{NI}}, T^*_{\text{AS}})^\top \to \text{N}_2(\mathbf{0}, \Sigma)$ as $n \to \infty$. Again, $z_\alpha - \frac{\mu_R - \mu_P}{\sqrt{\hat{\sigma}_R^2/n_R}}$ and $q_{(1-\alpha/\beta)} - \frac{\mu_E - \theta\mu_R - (1-\theta)\mu_P}{\sqrt{\hat{\sigma}_E^2/n_E + \theta^2\hat{\sigma}_R^2/n_R + (1-\theta)^2\hat{\sigma}_P^2/n_P}}$ both diverge in probability to $-\infty$. Thus, it follows that $\pi_{1,n} \to 1$ as $n \to \infty$.

Proof of Theorem 2.2(b). $\pi_{1,n}$ can be written

$$\pi_{1,n} = \mathrm{P}(T_{\mathrm{AS}} > z_\alpha, T_{\mathrm{NI}} > q_{(1-\alpha/\beta)} | \mathrm{H}_{\mathrm{AS},1}, \mathrm{H}_{\mathrm{NI},1}) = \mathrm{P}_{\mathrm{H}_{\mathrm{AS},1}}(T^*_{\mathrm{AS}} > z_\alpha - \frac{\mu_R - \mu_P}{\sqrt{\hat\sigma_R^2/n_R + \hat\sigma_P^2/n_P}}) \times$$

$$\mathrm{P}_{\mathrm{H}_{\mathrm{AS},1},\mathrm{H}_{\mathrm{NI},1}}\left(T^*_{\mathrm{NI}} > q_{(1-\alpha/\beta)} - \frac{\mu_E - \theta\mu_R - (1-\theta)\mu_P}{\sqrt{\hat\sigma_E^2/n_E + \theta^2\hat\sigma_R^2/n_R + (1-\theta)^2\hat\sigma_P^2/n_P}} | T_{AS} > z_\alpha\right). \quad \text{(A.8)}$$

Asymptotically, $\mathrm{P}_{\mathrm{H}_{\mathrm{AS},1}}(T^*_{\mathrm{AS}} > z_\alpha - \frac{\mu_R-\mu_P}{\sqrt{\hat\sigma_R^2/n_R+\hat\sigma_P^2/n_P}}) \to 1 - \Phi(z_\alpha - \frac{\delta}{a_2})$ and using the similar arguments of Lemma 2.2 show that

$$\mathrm{P}_{\mathrm{H}_{\mathrm{AS},1},\mathrm{H}_{\mathrm{NI},1}}\left(T^*_{\mathrm{NI}} > q_{(1-\alpha/\beta)} - \frac{\mu_E - \theta\mu_R - (1-\theta)\mu_P}{\sqrt{\hat\sigma_E^2/n_E + \theta^2\hat\sigma_R^2/n_R + (1-\theta)^2\hat\sigma_P^2/n_P}} | T_{AS} > z_\alpha\right) \to 1 - F_{\mathrm{ESN}}(q_{(1-\alpha/\beta)} - \frac{h}{\sqrt{a_1}})$$

as $n \to \infty$. Thus, from (A.8) we have

$$\pi_{1,n} \to \left[1 - \Phi(z_\alpha - \frac{\delta}{a_2})\right]\left[1 - F_{\mathrm{ESN}}(q_{(1-\alpha/\beta)} - \frac{h}{\sqrt{a_1}})\right]$$

as $n \to \infty$.

Proof of Corollary 2.1. From Theorem 2.2(b), we know that

$$\pi_{1,n} \to \left[1 - \Phi\left(z_\alpha - \frac{\delta}{\sqrt{\frac{\sigma_R^2}{\eta_R} + \frac{\sigma_P^2}{\eta_P}}}\right)\right]\left[1 - F_{\mathrm{ESN}}\left(q_{(1-\alpha/\beta)} - \frac{h}{\sqrt{a_1}}\right)\right] \quad \text{(A.9)}$$

as $n \to \infty$. Since $1 - F_{\mathrm{ESN}}\left(q_{(1-\alpha/\beta)} - \frac{h}{\sqrt{a_1}}\right)$ in (A.9) depends on ρ, so it is sufficient to show that $F_{\mathrm{ESN}}\left(q_{(1-\alpha/\beta)} - \frac{h}{\sqrt{a_1}}\right)$ is an increasing function of ρ. Using (2.48) of Azzalini (2005), we can express $F_{\mathrm{ESN}}\left(q_{(1-\alpha/\beta)} - \frac{h}{\sqrt{a_1}}\right)$ as

$$F_{\mathrm{ESN}}\left(q_{(1-\alpha/\beta)} - \frac{h}{\sqrt{a_1}}\right) = \frac{\Phi_{\mathrm{B}}(q_{(1-\alpha/\beta)} - \frac{h}{\sqrt{a_1}}, \tau; -\rho)}{\Phi(\tau)}, \quad \text{(A.10)}$$

where $\Phi_{\mathrm{B}}(u, v; \rho)$ is the standard bivariate normal distribution function. Since $\Phi_{\mathrm{B}}(u, v; \rho)$ is an increasing function of ρ (see Sibuya (1960)), thus $\Phi_{\mathrm{B}}(q_{(1-\alpha/\beta)} - \frac{h}{\sqrt{a_1}}, \tau; -\rho)$ is an increasing function and a is decreasing function of ρ when $\rho < 0$, and $\rho > 0$, respectively. Hence, consequently, $1 - F_{\mathrm{ESN}}\left(q_{(1-\alpha/\beta)} - \frac{h}{\sqrt{a_1}}\right)$ is a decreasing and is an increasing function of ρ when $\rho < 0$, and $\rho > 0$, respectively.

References

Azzalini, A., 2005. The skew-normal distribution and related multivariate families. Scand. J. Stat. 32, 159–188.
Azzalini, A., 2014. The Skew-Normal and Related Families. Institute of Mathematical Statistics Monographs. Cambridge University Press.
D'Agostino, R., Massaro, J., Sullivan, L., 2003. Noninferiority trials: design concepts and issues-the encounters of academic consultants in statistics. Stat. Med. 22, 169–186.
EMA, 2005. Guideline on the choice of the noninferiority margin (doc. ref. emea/cpmp/ewp/2158/99). European Medicines Agency: Pre-authorisation Evaluation of Medicines for Human Use.
FDA, 2016. Non-inferiority Clinical Trials to Establish Effectiveness: Guidance for Industry. Silver Spring, MD: US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER). Center for Biologics Evaluation and Research (CBER).
Ghosh, S., Chatterjee, A., Ghosh, S., 2017. Non-inferiority test based on transformations for non-normal distributions. Comput. Stat. Data Anal. 113, 73–87.
Hall, P., 1992. The Bootstrap and Edgeworth Expansion. Springer, New York.
Hall, P., Martin, M., 1988. On the bootstrap and two-sample problems. Aust. J. Stat. 30, 179–192.
Hasler, M., Vonk, R., Hothorn, L., 2008. Assessing non-inferiority of a new treatment in a three-arm trial in the presence of heteroscedasticity. Stat. Med. 27, 490–503.
Hida, E., Tango, T., 2011. Assessing non-inferiority of a new treatment in a three-arm trial in the presence of heteroscedasticity. Stat. Med. 30, 224–231.
Hung, H., Wang, S-J., Tsong, Y., Lawrence, J., Robert, T.O., 2003. Some fundamental issues with non-inferiority testing in active controlled trials. Stat. Med. 22, 213–225.
Kieser, M., Friede, T., 2007. Planning and analysis of three-arm non-inferiority trials with binary endpoints. Stat. Med. 26, 253–273.
Koch, A., Rohmel, J., 2004. Hypothesis testing in the 'gold standard' design for proving the efficacy of an experimental treatment relative to placebo and a reference. J. Biopharm. Stat. 14, 315–325.
Pigeot, I., Juliane, S., Joachim, R., Dietre, H., 2003. Assessing non-inferiority of a new treatment in a three-arm clinical trial including a placebo. Stat. Med. 22, 883–889.

Polansky, A., 2011. Introduction to Statistical Limit Theory. Chapman & Hall/CRC, Boca Raton, FL.
Serfling, R.J., 1980. Approximation Theorems of Mathematical Statistics. John Wiley and Series, New York.
Sibuya, M., 1960. Bivariate extreme statistics. Ann. Inst. Stat. Math. 11, 195–210.
WMA, 1997. World Medical Association declaration of Helsinki. Recommendations guiding physicians in biomedical research involving human subjects. J. Am. Med. Assoc. 227, 925–926.