Afrika Statistika

Afrika Statistika
Vol. 16 (2), 2021, pages 2719 - 2731.
DOI: http://dx.doi.org/10.16929/jas/2021.2719.180

ISSN 2316-090X

Identification of Minimum Effective Dose based on Ratio of Normally Distributed Data under Heteroscedasticity

Emmanuel Dodzi Kpeglo [1*], **Michael Jackson Adjabui**[2] **and Jakperik Dioggban**[1]

[1]Department of Statistics, CK Tedam University of Technology and Applied Sciences, Navrongo-Ghana
[2]Department of Mathematics, CK Tedam University of Technology and Applied Sciences, Navrongo-Ghana

Received on June 04, 2021; Accepted on July 1, 2021

Abstract. Efficacy and safety study is of practical importance in modern drug development. It is a key component in evaluating the safety of food additives or pesticides, and assessing the effectiveness and safety of drugs. In most of the various statistical procedures, homogeneity of variances among different dose levels was required. This paper without a need for multiplicity adjustment proposes a stepwise confidence set procedure for estimating Minimum Effective Dose (MED) of drugs based on ratio of population means for normally distributed data under heteroscedasticity. The procedure employed the Fieller (1954) method and obtained individual $(1-\alpha)100\%$ confidence intervals for identification of MED. The procedure is applied to a data of an experiment that was published by Ruberg (1989) where the effect of a new compound is measured by an increase in the weight of a particular organ in mice. Simulation study was carried out and results indicate that the procedure controls the family-wise error rate (FWER) strongly. Power of the procedure increases with increasing ratio of means and sample size.

Key words: Confidence set; Fieller's confidence intervals; Heteroscedasticity; Multiple treatments.
AMS 2010 Mathematics Subject Classification Objects : 62-02; 62H15; 62J15.

[*]Emmanuel Dodzi Kpeglo: kpaglo@yahoo.com
Michael Jackson Adjabui : madjabui@uds.edu.gh
Jakperik Dioggban : jdioggban@cktutas.edu.gh

Résumé. (Abstract in French) L'étude de l'efficacité et de l'innocuité est d'une importance pratique dans le développement moderne de médicaments. Il s'agit aussi d'un élément clé dans l'évaluation de l'innocuité des additifs alimentaires ou des pesticides. Dans la plupart des diverses procédures statistiques, l'homogénéité des variances entre les différents niveaux de dose est souvent requise. Dans cet article, il est proposé une procédure par étape pour établir des ensemble de confiance étapes pour estimer la dose minimale efficace (DFD) des médicaments, en présence d'hétéroscédécité, sans qu'il ne soit nécessaire de procéder à des ajsustements multiples. La méthode est basée sur les rapports de moyennes et de la taille de l'échantillon. Elle est ensuite appliquée à un jeu de données disponible dans Fieller (1954). Dans cette application, l'effet d'un nouveau composé est mesuré par une augmentation du poids d'un organe particulier chez la souris. Une étude de simulation a été réalisée et les résultats indiquent que la procédure contrôle fortement le taux d'erreur familial (FWER). La puissance de la procédure augmente avec l'augmentation du rapport entre les moyennes et la taille de l'échantillon.

The authors.
Emmanuel Dodzi Kpeglo, Msc., is preparing a PhD dissertation under the supervision of the third author, at CK Tedam University of Technology and Applied Sciences, Navrongo-Ghana.

Michael Jackson Adjabui, Ph.D., is a Senior Lecturer of Mathematics at CK Tedam University of Technology and Applied Sciences, Navrongo-Ghana.

Jakperik Dioggban, Ph.D,, is a Senior Lecturer of Statistics at Tedam University of Technology and Applied Sciences, Navrongo-Ghana.

1. Introduction

One main objective of undertaking clinical dose-response studies is to identify the minimum effective dose (MED) of drugs. MED is defined as the lowest dose whose effect is significantly different from control Kong *et al.* (2014). The whole process of drug development depends on the correct identification of the dose. This is because, selecting too high a dose can give rise to an unacceptable toxicity profile, while selecting a dose that is too low raises the possibility that the medicinal drug provides insufficient evidence of effectiveness. In searching for the lowest dose whose effect is significantly different from control, many statistical methods have been developed under the assumption of normality and equality of variances. Williams (1971) and Ruberg (1989) both proposed single-step procedures which identify MED consistently and frequently. Tamhane *et al.* (1996) however noted that, stepwise procedures offer a more powerful alternative especially when the interest is only in testing but not to estimate confidence intervals. Tamhane *et al.* (1996) therefore derived several step-down and step-up testing pro-

cedures that are widely used in identifying MED. Bauer (1997), Bretz *et al.* (2003) just to mention a few, developed methods to identify MED. Most of the articles mentioned above proposed methods that used the standard quotation of p-values. However, Hsu and Berger (1999) pointed out that the use of confidence intervals should be preferred because they yield more information about the parameters under investigation than a traditional p-value approach, and that confidence set approach generates methods with meaningful guarantee against incorrect decision.

Under homogeneity of variances assumption, Hsu and Berger (1999) considered inferences based on differences of treatment means and developed a stepwise confidence set procedure for successive multiple comparisons of dose groups against a control. The problem of heteroscedasticity can arise in clinical dose-response studies since certain biological factors can cause patients in different groups to respond differently with a change of dose level. By employing Stein's two-stage sampling method, Tao *et al.* (2002) proposed a stepwise confidence interval procedure for identifying MED under heteroscedasticity. Using stochastic ordering, Wang and Peng (2014) also established a step-up test procedure to estimate MED under the assumption of unequal variances. Both Tao *et al.* (2002) and Wang and Peng (2014) considered inferences based on difference of treatment means. Medical interpretation of margins based on ratios can easily be made as compared to inferences based on differences. For example, Hauschke and Kieser (2001) stated that the ratio formulation addresses the question, "Is at least Q% of the effectiveness of the reference preserved by the investigated treatment?" In addition to the medical interpretational convenience, Laster and Johnson (2003) demonstrated that the ratio formulation approach is more powerful in the test for non-inferiority of an experimental therapy as compared to that of the difference. Usually when dealing with log-normally distributed data, logarithmic transformation of the ratio test problem can lead to an acceptance range formulated in terms of the difference of means. However, there are many situations that require ratio test formulation in which the (untransformed) observations of the primary variable follow a normal distribution. Bretz *et al.* (2003) considered the case of relevance shifts defined in terms of a ratio of population means, where the original, untransformed data are normally distributed under homoscedasticity.

In this paper, we provide a procedure for identification of the MED when the ratio of treatment versus control is of interest for normally distributed data under heteroscedasticity. The proposed method extends the procedure proposed by Hsu and Berger (1999) into a stepwise confidence set procedure without multiple adjustments by incorporating the partitioning principle. The article is organized as follows. In Section 2, the preliminaries give the problem formulation and define the notation. It also defines the intersection-union principle and Hsu and Berger (1999) stepwise confidence method which will be essential to the proposed procedure. The new stepwise confidence interval procedure with unequal variances will be developed in Section 3. Also in Section 3, we apply the

proposed procedure to examine a real data set. Simulation studies are carried out to assess the power and familywise error rate (FWER) of the proposed procedure. Finally, conclusions are given in Section 4.

2. Statistical background

For $i = 0, 1, \ldots, k$ and $j = 1, \ldots, n_i$, let X_{ij} denote the j^{th} observation at the i^{th} dose level in a one-way layout. Assume that X_{ij} are observed responses of the efficacy of the j^{th} subject in the i^{th} dose group. Assume also that X_{ij} are independent and follows normal distribution with means $E(X_{ij}) = \mu_i$, and possibly unequal variances $Var(X_{ij}) = \sigma_i^2$. The estimators for the means and variances are denoted by $\bar{X}_i$, and $\hat{\sigma}_i^2$ respectively.

The problem of interest is to provide a procedure for the estimation of ratios of unknown means μ_i, $i = 0, 1, \ldots, k$. Without loss of generality and for the purpose of this research, let large values of treatment means μ_i relative to the mean of placebo μ_0 denote high efficacy. Suppose that k doses are tested against a placebo, let $\gamma_i = \mu_i/\mu_0$, $i = 1, \ldots, k$, be the ratios of interest. Let δ be some pre-specified threshold constant for efficacy of a drug. We formulate the problem of identifying the MED as follows:

$$H_{0i} : \gamma_i \leq \delta \text{ versus } H_{ai} : \gamma_i > \delta \text{ for } i = 1, \ldots, k \tag{1}$$

Suppose the random sample X_{ij} observed from the i^{th} dose level has a distribution determined by a parameter $\theta = (\gamma_1, \ldots, \gamma_k)$ with $\theta \in \Theta$. If Θ is the parameter space with $\Theta_i^c = \{\theta : \gamma_i > \delta\}$, $i = 1, \ldots, k$, we can rewrite (1) as follows:

$$H_{0i} : \theta_i \in \Theta_i \text{ versus } H_{ai} : \theta_i \in \Theta_i^c \text{ for } i = 1, \ldots, k \tag{2}$$

The article aim to solve the testing problem (2) and make inference for MED by extending the concept of directed confidence set proposed by Hsu and Berger (1999). For our procedure to control the probability of declaring an ineffective dose to be effective, we will introduce the intersection-union principle introduced by Berger (1982). The Intersection-Union Principle involves testing the union of the individual hypotheses against the intersection of the alternative hypotheses.
Thus, if Θ_i is a level α test of H_{0i} for $i = 1, \ldots, k$, then the intersection-union test with rejection region Θ_i^c is a level α test of $H_0 = \bigcup_{i=1}^k \Theta_i$ against $H_a = \bigcap_{i=1}^k \Theta_i^c$.
The major reason behind the intersection–union test is that, when the global null hypothesis H_0 is rejected then each of the individual null hypotheses H_{0i} are rejected. When the intersection-union test is featured, it cancels a need for multiplicity adjustment. This is because; if each individual test is performed at level α, the global test is also performed at level α.
Assume $\Theta_i = \{\theta : \gamma_i \leq \delta\}$, $i = 1, \ldots, k$, are subsets of Θ. Let $\Theta_0 = \bigcup_{i=1}^k \Theta_i$. Note also that the $\Theta_0^c = \bigcap_{i=1}^k \Theta_i^c$. Problem (2) becomes:

$$H_0 : \theta \in \bigcup_{i=1}^{k} \Theta_i \text{ versus } H_a : \theta \in \bigcap_{i=1}^{k} \Theta_i^c \tag{3}$$

The efficacy problem (3) is interpreted as follows: if H_0 is rejected at $m(1 < m < k)$, then there is evidence that doses i, $(i \geq m+1)$ are effective. And we take the lowest dose i for which H_0 is rejected as the MED. If a procedure of testing (3) controls the FWER at α, then the probability of declaring a dose as the MED when either it or a higher dose is ineffective is no more than α.

Definition 2.1. *Let $\Theta^* \subset \Theta$ the parameter space. A confidence set, $C(X)$, for θ is directed towards Θ^* if, for every sample point X either $\Theta^* \subset C(X)$ or $C(X) \subset \Theta^*$.*

For a one sided significant ratio inference, say $\Theta_i^c = \{\mu_i > \mu_0\delta\}$, confidence intervals for $\gamma_i = \mu_i/\mu_0$ of the form $C_i(X)$ are directed toward Θ_i^c for $i = 1, \ldots, k$.
In this article, we consider cases with inferences $\theta \in \Theta_i^c$ and propose a method that provides confidence intervals $C_i(X)$ for θ in a stepwise fashion. Our procedure stops at the first i whenever $C_i(X) \not\subset \Theta_i^c$. To validate our procedure, we partition the parameter space $\Theta = \bigcup_{i=1}^{k} \Theta_i$ into disjoint sets $\Theta_1^*, \Theta_2^*, \ldots,$ and Θ_k^* such that for some index set K, the set $\Theta_k^* \subseteq \Theta : k \in K$ and $\Theta_k^* \cap \Theta_{k'}^* = \phi$ for any $k, k' \in K$ with $k \neq k'$. Partitioning the parameter space into disjoint sets makes it possible for exactly one partition to contain the true parameter θ. This controls FWER by controlling FWER within each $\Theta_k^* \subseteq \Theta : k \in K$.

The following section provides a procedure for identification of the MED. Later, we apply the proposed procedure to examine a real data set and then perform simulation studies to assess the power and *FWER* of the procedure.

3. Results

3.1. The proposed stepwise confidence interval procedure

Consider the problem of identifying MED in problem (3), we first define the MED as $\min\{i : \gamma_i > \delta\}$. We employed Fieller's (1954) method and obtained individual $(1-\alpha)100\%$ confidence intervals $C_i(X)$ for γ_i. Since $\bar{X}_i$ and $\hat{\sigma}_i^2$ are the sample mean and sample variance of the i^{th} group, $i = 0, 1, \ldots, k$, we test $H_0 : \theta \in \bigcup_{i=1}^{k} \Theta_i$ verses $H_1 : \theta \in \bigcap_{i=1}^{k} \Theta_i^c$ using the test statistics:

$$T_i^X = \frac{\bar{X}_i - \delta\bar{X}_0}{\sqrt{\frac{\hat{\sigma}_i^2}{n_i} + \frac{\delta^2\hat{\sigma}_0^2}{n_0}}},$$

where

$$\bar{X}_i = \sum_{j=1}^{n_i} X_{ij}/n_i \quad and \quad \hat{\sigma}_i^2 = \sum_{i=0}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2 \Big/ \sum_{i=0}^{k}(n_i - 1).$$

The T_i^X rejects H_0, if $T_i^X > t_{1-\alpha,v_i}$, is the $(1-\alpha)$ percentile of the k-variate central t-distribution with v_i degrees of freedom. The degrees of freedom was stated in

Welch (1938) and improved in Satterthwaite (1946). These different degrees of freedom are given by:

$$v_i = \frac{\left(\frac{\hat{\sigma}_i^2}{n_i} + \frac{\delta^2\hat{\sigma}_0^2}{n_0}\right)^2}{\left(\frac{\hat{\sigma}_i^4}{n_i^2(n_i-1)} + \frac{\delta^4\hat{\sigma}_0^4}{n_0^2(n_0-1)}\right)}$$

Note that the degrees of freedom are not exact but estimated because they depend on unknown group variances. Suppose increasing values of the endpoint represent a higher effect of treatment, the decision rule is to reject H_0, if $T_i^X > t_{1-\alpha,v_i}$. The lower limits for the confidence intervals of interest $C_i(X)$ are derived to be the smaller root of the quadratic equation $(T_i^X)^2 = t_{1-\alpha,v_i}^2$. These lower confidence limits are given by:

$$B_i(X) = \frac{\bar{X}_i\bar{X}_0 - \sqrt{a_0\bar{X}_i^2 + a_i\bar{X}_0^2 - a_ia_0}}{\bar{X}_0^2 - a_0}$$

And the $(1-\alpha)100\%$ confidence intervals $C_i(X)$ for γ_i are obtained to be as follows: $C_i(X) = (B_i(X), \infty)$, where $a_i = \hat{\sigma}_i^2 t_{1-\alpha,v_i}/n_i$, $i = 1, \ldots, k$ and $a_0 = \hat{\sigma}_0^2 t_{1-\alpha,v_i}/n_0$.
The proposed stepwise confidence interval procedure takes the following form:

Step 1
If $C_k(X) \subset (\delta, \infty)$ then assert $\gamma_k \in (\delta, \infty)$ and go to step 2. Otherwise conclude $\gamma_k \in C_k(X)$ and stop.
Step 2
If $C_{k-1}(X) \subset (\delta, \infty)$ then assert $\gamma_{k-1} \in (\delta, \infty)$ and go to step 3. Otherwise conclude $\gamma_{k-1} \in C_{k-1}(X)$ and stop.
$\vdots$
Step k-1
If $C_2(X) \subset (\delta, \infty)$ then assert $\gamma_2 \in (\delta, \infty)$ and go to step k. Otherwise conclude $\gamma_2 \in C_2(X)$ and stop.
Step k
If $C_1(X) \subset (\delta, \infty)$ then assert $\gamma_1 \in (\delta, \infty)$ and go to step k+1. Otherwise conclude $\gamma_1 \in C_1(X)$ and stop.
Step k+1
Conclude $\min\{i : \gamma_i\} > \min\left\{i : \frac{\bar{X}_i\bar{X}_0 - \sqrt{a_0\bar{X}_i^2 + a_i\bar{X}_0^2 - a_ia_0}}{\bar{X}_0^2 - a_0}\right\}$ with $i = 1, \ldots, k$.

The above algorithm starts with the highest dose k and in a descending order, sequentially examines the efficacy of each step. Once a statistically insignificant treatment effect is found, we take the lowest dose i for which global hypothesis is rejected as the MED. The flow-chart of the procedure presented above is summarized in Theorem 1.

Theorem 1. *Let $X_{i1}, X_{i2}, \ldots, X_{in_i}$ be the observed data from the i^{th} dose level having a distribution determined by the parameter vector $\theta = (\gamma_1, \ldots, \gamma_k)$ with $\theta \in \Theta$, the parameter space. Let $\theta \in \Theta_i^c = \{\theta : \gamma_i > \delta\}$, $i = 1, \ldots, k$ be a multiple comparison of interest. Let confidence set $C_i(X)$ for θ be based on a set of data X such that $C_i(X)$ is directed toward subsets $\Theta_i^c \subseteq \Theta$. Assume that the stepwise procedure stops at step M, that is, M is the smallest integer i such that $C_i(X) \not\subset \Theta_i^c$ if such i $(1 \leq i \leq k)$ exists; otherwise, let $M = k + 1$. Furthermore, let $\Theta_0 = \phi$, $\Theta_{k+1} = \Theta$, and*

$$C(X) = \Theta_0^c \cap \Theta_1^c \cap \ldots \cap \Theta_{M-1}^c \cap C_M(X).$$

Then for all $\theta \in \Theta$,

$$P(\theta \in C(X)) \geq 1 - \alpha.$$

The proof of Theorem 1 is given in appendix A.

3.2. Example

To illustrate the application of the stepwise confidence procedure, we consider the data of an experiment that was published by Ruberg (1989) and also used by Tao *et al.* (2002). The goal of the study is to measure the effect of a new compound by the gain in weight of a particular organ in mice. The summary statistics for a control group and four equally spaced dose groups is given in Table 1. The sample size in the five groups is equal to 12. The variable of interest is the ratio of mean response to the mean of the zero dose. A large value of this ratio is regarded as effective with a threshold of a 10% average increase over its value for the zero dose. Thus, we take $\delta = 1.1$. The assumption of normality of the data and heterogeneity of variances across the dose groups are satisfied (see Tao *et al.* (2002)). If $\delta = 1.1$ and $\alpha = 0.05$, then Table 2 shows the 95% lower confidence limits on μ_i/μ_0, $i = 1, 2, 3, 4$.

Table 1. (see, Ruberg (1989)). Summary Statistics for the Pharmacologic Effect of an Experimental Compound on Relative Organ Weights in Mice

Dose (mg/kg/day)	Mean ± standard deviation
Control	6.20 ± 3.08
10	6.14 ± 2.32
20	6.54 ± 2.77
30	7.67 ± 2.32
40	9.37 ± 1.87

Table 2 shows that the lower confidence limits for the dose groups 10 mg/kg/day, 20 mg/kg/day, and 30 mg/kg/day are lower than the relevance threshold. Thus, all these three doses except 40 mg/kg/day are declared to be statistically insignificant at level α. The MED is correctly specified, if and only if $C_j(X) \subset (\delta, \infty)$ and

Table 2. 95% Lower Confidence Limits for μ_i/μ_0, $i = 1, 2, 3, 4$

Dose (mg/kg/day)	Lower confidence limit
10	0.6848
20	0.7126
30	0.8877
40	1.1246

$C_{i-1}(X) \not\subset (\delta, \infty)$ for $j = i, \ldots, k$.

$C_4(X) = (1.1246, \infty) \subset (1.1, \infty)$ we reject H_{04} (40 mg/kg/day is effective)

$C_3(X) = (0.8877, \infty) \not\subset (1.1, \infty)$ we do not reject H_{03} (30 mg/kg/day is ineffective)

$C_2(X) = (0.7126, \infty) \not\subset (1.1, \infty)$ we do not reject H_{02} (20 mg/kg/day is ineffective)

$C_1(X) = (0.6848, \infty) \not\subset (1.1, \infty)$ we do not reject H_{01} (10 mg/kg/day is ineffective)

In this analysis, our stepwise procedure concluded that the doses 10 mg/kg/day, 20 mg/kg/day and 30 mg/kg/day are ineffective at level α. Furthermore, the 40 mg/kg/day dose and any available higher dose are regarded as effective. Therefore 40 mg/kg/day is recommended as the minimum effective dose (MED).

3.3. Simulation studies

3.3.1. FWER study

FWER is strongly controlled when $\alpha^* = \sup_{H_{0i}}[Pr\{any\ H_{0i}\ is\ rejected\ (1 \leq i \leq k)\}]$ and $max(\alpha^*) \leq \alpha$. Here the test statistics T_i^X rejects H_{0i} for $H_{0i} : \theta_i \in \Theta_i$ versus $H_{ai} : \theta_i \in \Theta_i^c$ when $C_i(X) \subset \Theta_i^c$. Control of FWER is critical in dose finding because the FWER for H_{0i} is assumed to conclude that a clinically relevant treatment effect is present when in fact it is not Hochberg and Tamhane (1987). In this section, we assess the performance of our procedure based on FWER using simulation. The simulation study was conducted to determine the robustness of our procedure at nominal level of 0.025. We used R software for the computation and set the number of iterations at 1,000,000. We also compared FWERs generated from a normal distribution based on the assumption of equal variance among all groups(HOM) alongside unequal variances across dose groups(HET). The result (given in Table 3) indicated that our procedure successfully controls the FWER for HET. However, in the case of HOM, all the simulated values for the FWER exceeds the nominal level of 0.025 and therefore controlling FWER poorly. Our heteroscedastic procedure is therefore robust.

Table 3. Simulated FWER results for $\delta = 1.1$, $\alpha = 0.025$, $n_0 = 8$, $\mu_0 = 6.2$, $\mu_1 = 6.75$, $\sigma_0 = 3.08$, $\sigma_1 = 2.32$

n_1	HET	HOM
10(11)	0.0224(0.0224)	0.0274(0.0286)
12(13)	0.0229(0.0231)	0.0302(0.0315)
14(15)	0.0229(0.0229)	0.0323(0.0332)
16(17)	0.0232(0.0230)	0.0343(0.0353)
18(19)	0.0235(0.0231)	0.0367(0.0368)
20(21)	0.0232(0.0233)	0.0377(0.0384)
22(23)	0.0233(0.0231)	0.0390(0.0396)
24(25)	0.0236(0.0229)	0.0407(0.0406)
26(27)	0.0233(0.0234)	0.0414(0.0420)
28(29)	0.0231(0.0234)	0.0424(0.0431)
30(31)	0.0233(0.0233)	0.0434(0.0438)
32(33)	0.0234(0.0233)	0.0442(0.0447)
34(35)	0.0235(0.0235)	0.0454(0.0456)

Table 4. Simulated power results for $\alpha = 0.025$, $\delta = 1.2, 1.3, 1.4, 1.5, 1.6$

Ratio(γ)	1.2	1.3	1.4	1.5	1.6
0.8	0.7489	0.5359	0.3368	0.1912	0.1010
0.9	0.8659	0.6907	0.4848	0.3041	0.1749
1.0	0.9385	0.8173	0.6347	0.4395	0.2761
1.1	0.9760	0.9052	0.7661	0.5821	0.3995
1.2	0.9920	0.9571	0.8659	0.7144	0.5338
1.3	0.9978	0.9831	0.9316	0.8225	0.6640
1.4	0.9995	0.9943	0.9691	0.9002	0.7768
1.5	0.9999	0.9983	0.9877	0.9495	0.8641

Table 5. Simulated power results for $\delta = 1.7$, $n_0 = n_i = 10, 15, 20, 25, 30$

Ratio(γ)	10	15	20	25	30
0.8	0.0509	0.0604	0.0693	0.0778	0.0861
0.9	0.0946	0.1266	0.1581	0.1894	0.2205
1.0	0.1610	0.2313	0.3001	0.3667	0.4301
1.1	0.2521	0.3719	0.4817	0.5787	0.6621
1.2	0.3645	0.5320	0.6671	0.7697	0.8443
1.3	0.4899	0.6870	0.8184	0.8991	0.9459
1.4	0.6160	0.8144	0.9172	0.9653	0.9861
1.5	0.7307	0.9033	0.9688	0.9907	0.9974

3.3.2. Power calculation

Confidence interval procedures for analyzing clinical trials are frequently becoming insufficient in the design of clinical studies. Power calculation has therefore become a major task in the design phase of a clinical study. This article considers the correct estimation of minimum effective dose i, $i = 1, \cdots, k$, for normally distributed

data. The MED is correctly specified if and only if $C_j(X) \subset (\delta, \infty)$ and $C_{i-1}(X) \not\subset (\delta, \infty)$ for $j = i, ..., k$. Thus

$$P(\text{MED} = i) = P\left(\bigcap_{j=i}^{k} \left\{T_j^X > t_{1-\alpha,v_i}\right\} \cap \left\{T_{i-1}^X \leq t_{1-\alpha,v_i}\right\}\right). \tag{4}$$

Here, we define power as the probability of correctly estimating any of the doses, i, to be the true MED. This is the same as finding the probability of rejecting at least one of the false hypotheses. Equation (4) will therefore be expressed as

$$P(rejectingH_{0j}, i \leq j \leq k) = P\left(\bigcap_{j=i}^{k} \left\{T_j^X > t_{1-\alpha,v_i}\right\}\right) \tag{5}$$

where

$$v_i = \frac{\left(\frac{\hat{\sigma}_i^2}{n_i} + \frac{\delta^2\hat{\sigma}_0^2}{n_0}\right)^2}{\left(\frac{\hat{\sigma}_i^4}{n_i^2(n_i-1)} + \frac{\delta^4\hat{\sigma}_0^4}{n_0^2(n_0-1)}\right)}.$$

Simulation study is performed in order to assess the impact of several parameters on the power to estimate the true MED. Results for different ratio of means, clinical relevance margins, and the sample sizes are presented in Tables 5 and 5 respectively. In Table 4, power increases with increasing ratio of means and decreases for higher clinical relevance margins while in Table 5, power increases with increase in both the ratio of means and sample size. These simulation results indicated a very good performance for our approach.

4. Conclusions

In this paper, we considered a general way of obtaining stepwise confidence intervals for the ratio of means under heteroscedasticity. Many testing procedures have been proposed to identify the MED of drugs under the assumption of equal variance. The main difficulty with most of these procedures is that they fail to address problems where the assumption of equal variance is not practicable due to certain biological factors. In many standard situations, inferences on differences are also not a suitable way to investigate the data. Ratios, therefore, are often a better alternative measure of efficacy and often easier to interpret. The strength of the proposed procedure is to provide stepwise confidence intervals using an extended intersection-union principle which in effect cancels a need for multiplicity adjustment. Simulation study is carried out to assess the performance of the proposed procedure. It is found that all the stepwise confidence intervals control FWER. Simulation study results also indicated that

the power of the procedure increases with increasing ratio of means and sample size. Meanwhile the power decrease with increase in clinical relevance margins.

Appendix A: Proof of Theorem 1

Consider the following sets $\Theta_1^*, \Theta_2^*, \ldots, \Theta_{k+1}^*$, as partitions of the parameter space Θ.

$$\begin{aligned}
\Theta_1^* &= \Theta_1 \\
\Theta_2^* &= \Theta_2 \cap \Theta_1^c \\
&\vdots \\
\Theta_i^* &= \Theta_i \cap \Theta_1^c \cap \ldots \cap \Theta_{i-2}^c \cap \Theta_{i-1}^c \\
&\vdots \\
\Theta_k^* &= \Theta_k \quad \cap \Theta_1^c \cap \ldots \cap \Theta_{k-2}^c \cap \Theta_{k-1}^c \\
\Theta_{k+1}^* &= \Theta_{k+1} \cap \Theta_1^c \cap \ldots \cap \Theta_{k-1}^c \cap \Theta_k^c
\end{aligned}$$

If $\theta \in \Theta_i^*$ then clearly a $100(1-\alpha)\%$ confidence set for θ will be

$$C(X) = \bigcup_{i=1}^{k+1} (C_i(X) \cap \Theta_i^*). \tag{6}$$

For all $i < M$ (if such i exists) $C_i(X) \cap \Theta_i^* = \phi$ since $\Theta_i^* \subset \Theta_i$. Then

$$C(X) = \bigcup_{i=M}^{k+1} (C_i(X) \cap \Theta_i^*). \tag{7}$$

Similarly, for all $i > M$ (if such i exists) we have $\Theta_i^* \subset \Theta_0^c \cap \ldots \cap \Theta_{M-1}^c \cap \Theta_M^c$ and equation (7) becomes:

$$\begin{aligned}
C(X) &= \bigcup_{i=M}^{k+1} (C_i(X) \cap \Theta_i^*) \\
&\subset (C_M(X) \cap \Theta_M^*) \cup (\Theta_0^c \cap \ldots \cap \Theta_{M-1}^c \cap \Theta_M^c) \\
&= (\Theta_0^c \cap \ldots \cap \Theta_{M-1}^c \cap \Theta_M \cap C_M(X)) \cup (\Theta_0^c \cap \ldots \cap \Theta_{M-1}^c \cap \Theta_M^c)
\end{aligned} \tag{8}$$

If $M < k+1$, then $\Theta_M^c \subset C_M(X)$ and $C_M(X) \not\subset \Theta_M^c$ will imply $\Theta_M^c \subset C_M(X)$. Hence

$$\begin{aligned}
C(X) &= \bigcup_{i=M}^{k+1} (C_i(X) \cap \Theta_i^*) \\
&\subset (C_M(X) \cap \Theta_M^*) \cup (\Theta_0^c \cap \ldots \cap \Theta_{M-1}^c \cap \Theta_M^c) \\
&= (\Theta_0^c \cap \ldots \cap \Theta_{M-1}^c \cap \Theta_M \cap C_M(X)) \cup (\Theta_0^c \cap \ldots \cap \Theta_{M-1}^c \cap \Theta_M^c) \\
&= (\Theta_0^c \cap \ldots \cap \Theta_{M-1}^c \cap \Theta_M \cap C_M(X)) \cup (\Theta_0^c \cap \ldots \cap \Theta_{M-1}^c \cap C_M(X)) \\
&= (\Theta_0^c \cap \ldots \cap \Theta_{M-1}^c \cap \Theta_M \cap C_M(X))
\end{aligned} \tag{9}$$

For all $\theta \in \Theta_i^*$,

$$\begin{aligned} P(\theta \in C(X)) &= P(\theta \in (C_i(X) \cap \Theta_i^*)) \\ &= P(\theta \in C_i(X)) \geq 1 - \alpha. \end{aligned}$$

This completes the proof of Theorem 1. ■

Acknowledgments. We thank editor-in-chief and the reviewers for their constructive comments that have improved the article significantly.

References

Adjabui, M. J., Howard, N., and Luguterah, A. (2016). Nonparametric stepwise procedure for identification of minimum effective dose(MED). *International Journal of Statistics and System*, 11(1), 77-88.

Bauer, P. (1997). A note on multiple testing procedures in dose finding. *Biometrics*, 53, 1125–1128.

Berger, R. L. (1982). Multiparameter hypothesis testing and acceptance sampling. *Technometrics*, 24, 295–300.

Bretz, F., Hothorn, L. A., and Hsu, J. C. (2003). Identifying effective and/or safe doses by stepwise confidence intervals for ratios. *Statistics in Medicine*, 22, 847-858.

Bretz, F., Pinheiro, J. C., and Branson, M. (2005). Combining multiple comparisons and modeling techniques in dose response studies. *Biometrics*, 61, 738–748.

Chen, J. T. (2008). A two-stage stepwise estimation procedure. *Biometrics*, 64, 406–412.

Fieller, E. C. (1954). Some problems in interval estimation. *Journal of the Royal Statistical Society*, Series B 16, 175 –185.

Finner, H., and Strassburger, K. (2002). The partitioning principle: a powerful tool in multiple decision theory. *Annals of Statistics*, 30, 1194 –1213.

Hauschke, D., and Kieser, M. (2001). Multiple testing to establish non-inferiority of k treatments with a reference based on the ratio of two means. *Drug Information Journal*, 35, 1247–1251.

Hochberg, Y. and Tamhane, A. C. (1987). *Multiple comparison procedures*. New York: Wiley.

Hothorn, L. A., and Hauschke, D. (2000). Identifying the maximum safe dose: a multiple testing approach. *Journal of Biopharmaceutical Statistics*, 10, 15 –30.

Hsu, J. C. and Berger, R. (1999). Stepwise confidence intervals without multiplicity adjustment for dose-response and toxicity studies. *Journal of the American Statistical Association*, 94, 468–482.

Kong, M., Rai, S., and Bolli, R. (2014). Statistical methods for selecting maximum effective dose and evaluating treatment effect when dose - response is monotonic. *Statistics in Biopharmaceutical Research*, 6 (1), 16-29.

Laster, L.L., and Johnson, M.F. (2003). Non-inferiority trials: the 'at least as good as' criterion *Statistics in Medicine*, 22, 187–200.

Ruberg, S. J. (1989). Contrasts for identifying the minimum effective dose. *Journal of the American Statistical Association*, 84, 816–822.

Satterthwaite, F. E. (1946). An approximate distribution of estimates variances component. *Biometrics*, 2, 110-114.

Tamhane, A. C., Hochberg, Y., and Dunnett, C. W. (1996). Multiple test procedures for dose finding. *Biometrics*,52, 21–37.

Tamhane, A. C., Dunnett, C. W., Green, J. W., and Wetherington, J. F. (2001). Multiple test procedures for identifying the maximum safe dose. *Journal of the American Statistical Association*,96, 835–843.

Tao, J., Shi, N.-Z., Guo, J.-H. and Gao W.(2002). Stepwise Procedures for the Identification of Minimum Effective Dose with Unknown Variances. *Statistics and Probability Letters*, 57, 121-131.

Tao, J., Guo, J.-H. and Shi, N.-Z.(2002). Stepwise Procedures under Unknown Variances for Toxicological Evaluation. *Biometrical Journal*, 44, 149-160.

Thall, P. F., and Russell, K. E. (1998). A strategy for dose-finding and safety monitoring based on efficacy and adverse outcomes in phase I/II clinical trials. *Biometrics*, 54, No. 1, 251-264.

Turri, M., and Stein, G. (1986). The determination of practically useful doses of new drugs: some methodological considerations. *Statistics in Medicine*, 5, 449–457.

Wang, W. and Peng, J. (2014). A step-up test procedure to find the minimum effective dose. *Journal of Biopharmaceutical Statistics*, 25, 525–538.

Williams, D. A. (1971). A test for difference between treatment means when several doses are compared with a zero dose. *Biometrics*, 27, 103–117.

Welch, B. L. (1938). The significance of difference between two means when the population variances unequal. *Biometrika*, 29, 350-362.