Journal of Mathematics Research; Vol. 14, No. 3; June 2022
ISSN 1916-9795 E-ISSN 1916-9809
Published by Canadian Center of Science and Education

Identification of Maximum Safe Dose based on Ratio of Normally Distributed Data under Heteroscedasticity

Emmanuel D. Kpeglo[1], Michael J. Adjabui[2], 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

Correspondence: Emmanuel D. Kpeglo, Department of Statistics, CK Tedam University of Technology and Applied Sciences, Navrongo, Ghana

Received: February 21, 2022 Accepted: March 22, 2022 Online Published: April 24, 2022

doi:10.5539/jmr.v14n3p1 URL: https://doi.org/10.5539/jmr.v14n3p1

Abstract

In most of the various stepwise confidence interval procedures formulated for identifying maximum safe dose (MSD), homogeneity of variances among different dose levels were required. But in practice, homogeneity of variance is often in doubt. This paper proposes a stepwise confidence set procedure for identifying MSD of drugs based on ratio of population means for normally distributed data under heteroscedasticity without the need for multiplicity adjustment. The procedure employed Fieller's method and obtained individual $(1-\alpha)100\%$ confidence intervals for identification of the MSD. We illustrate the procedure with a real life example. In addition, we show that power of the procedure increases with increasing ratio of means, and sample size. Power however decreases with increase in clinical relevance margins. We also illustrate that the new procedure can properly control familywise error rate (FWER).

Keywords: Confidence set, Fiellers confidence intervals, Heteroscedasticity, Multiple treatments

1. Introduction

Dose-finding trials are defined as clinical investigations in which different doses of a pharmaceutical compound bring out therapeutic and toxic responses. These trials with a new drug are differentiated from the placebo of a single dose that produces either therapeutic or toxic response and for which appropriate designs have been suggested. The minimum effective dose (MED) is that which produces a prescribed lowest therapeutic response whereas the maximum safe dose (MSD) is that which elicits a prescribed highest frequency of adverse reactions (Turri & Stein, 1986). These two doses define the interval of therapeutically useful doses. Identifying the MSD level of a pharmaceutical compound is one of the major objectives in toxicological studies and the conventional approach to finding the MSD is to identify the highest dose with toxicity below a certain pre-specified level (Korn *et al.*, 1994).

Under the assumptions of normality, a variety of papers have been published on the identification of the MSD.
Tamhane, Dunnett, Green and Wetherington (2001) proposed a stepwise procedure that used contrasts or order restricted test in estimating the MSD. However, it is a common knowledge that making inferences using confidence interval based approaches provide additional information such as confidence bounds. In this regard, a number of simultaneous confidence intervals methods to estimate MSD were proposed. For example: Hauschke, Kieser and Hothorn (1999), and Hothorn and Hauschke (2000). The stepwise confidence intervals procedures according to Bretz, Hothorn and Hsu (2003) yield more information about the parameters of interest than a traditional p-value method does. Furthermore, once stepwise confidence intervals are derived, it becomes easy to show that directional errors (that is, correct side decisions at each step) are simultaneously controlled with the familywise error rate (FWER). Adjabui, Howard and Luguterah (2017) proposed a simultaneous nonparametric upper bound procedure based on the Wilcoxon Mann-Whitney test in situations when normal assumptions are violated. Their paper incorporated the partitioning principle for confidence sets procedure proposed by Hsu and Berger (1999) for estimating MSD.

Most of the articles mentioned above were proposed under the assumption of equal variance, but the homoscedasticity assumption is rarely satisfied in practice. This is because, certain biological factors can cause patients in different groups to respond differently with a change of dose levels. The following papers: Tao, Guo and Shi (2002), Tamhane and Logan (2004), just to mention a few, developed various multiple comparison procedures under the heteroscedastic setting. Tao, Guo and Shi (2002) proposed a simultaneous confidence intervals approach that makes inferences based on a difference of treatment and control for identification of MSD. However, according to Laster and Johnson (2003), the ratio parameter has the advantage of being available as a dimensionless percentage, easily estimated, and enhances comparison among

studies. Laster and Johnson (2003) further stated that the ratio formulation approach is more powerful in a test for non-inferiority of an experimental therapy as compared to that of the difference. In considering 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 practical situations that require ratio test formulation in which the (untransformed) observations of the primary variable follow a normal distribution. Under homoscedastic setting, Bretz, Hothorn and Hsu (2003) considered the case of relevance shifts defined in terms of ratio of population means, where the original, untransformed data are normally distributed. Tamhane and Logan (2004) developed a contrast testing/decision making approach to estimate MSD. But confidence intervals procedures provide additional information such as a confidence bound for the first unsafe dose. Or, provides an upper confidence bound on how safe the doses are if all the doses are regarded safe.

The purpose of our article is to provide stepwise confidence intervals for identification of MSD when the ratio of treatment versus control is to be investigated for normally distributed data under unequal variance setting. 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. Simulation studies are carried out in Section 4 to assess the performance of power and FWER of the proposed procedure. In Section 5, we apply the proposed procedure to examine a real data set. Finally, conclusions and recommendations are given in Section 6.

2. Statistical Background

Let $i = 0, 1, \ldots, k$ denote a set of increasing dose levels used in a dose-finding study where 0 corresponds to the zero dose level called the placebo control. Consider a one-way layout setting in which a random sample $Z_{i1}, Z_{i2}, Z_{i3}, \ldots, Z_{in_i}$ is an observed response of toxicity from the i^{th} dose level ($i = 0, 1, \ldots, k$). Assume Z_{ij}, the j^{th} ($1 \leq j \leq n_i$) observation at i^{th} dose level, to be independently normal with means $E(X_{ij}) = \eta_i$ and possibly unequal variances $Var(Z_{ij}) = \tau_i^2$. Let $\bar{Z}_i$, $i = 0, 1, \ldots, k$, be the sample means and let $\hat{\tau}_i^2$ be an estimate of the unequal variances τ_i^2.

In this article, we seek to provide a procedure that will estimate ratios of unknown means η_i, $i = 0, 1, \ldots, k$. Without loss of generality, let large values of treatment means η_i denote high toxicity relative to the mean of the placebo η_0. Suppose that k doses are tested against a placebo, let $\lambda_i = \eta_i/\eta_0$, $i = 1, \ldots, k$, be the ratios of interest. Let ζ be some pre-specified threshold constant for toxicity of a drug. The problem of identifying the MSD is formulated as follows:

$$H_{0i} : \lambda_i \geq \zeta \text{ versus } H_{ai} : \lambda_i < \zeta \text{ for } i = 1, \ldots, k. \tag{1}$$

Assume the random sample Z_{ij} has a distribution determined by the parameter $\theta = (\lambda_1, \ldots, \lambda_k)$ with $\theta \in \Theta$, where Θ is the parameter space, let $\Theta_i = \{\theta : \lambda_i \geq \zeta\}$, $i = 1, \ldots, k$. The problem in (1) can further be written as:

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

In this article, we extend the concept of directed confidence set proposed by Hsu and Berger (1999) to solve the testing problem (2) and make inference for MSD. In order for our procedure to control the probability of declaring unsafe dose as safe, we will introduce the Intersection-Union Principle. The Intersection-Union Principle formulated by Berger (1982) 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 main reason behind the application of the intersectionCunion test is when the global null hypothesis H_0 is rejected each of the individual null hypotheses H_{0i} is rejected. In addition, whenever intersection-union test is introduced, 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 $\Theta_i = \{\theta : \lambda_i \geq \zeta\}$, $i = 1, \ldots, k$, are subsets of Θ. Let $\Theta_0 = \bigcup_{i=1}^k \Theta_i$ and $\Theta_0^c = \bigcap_{i=1}^k \Theta_i^c$. Based on the Intersection-Union Principle, the problem of identifying the MSD can be formulated as follows:

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

Definition 1. *Let Θ be a parameter space. A confidence set, $C(Z)$, for θ is directed toward a subset of the parameter space $\Theta^* \subset \Theta$ if, for every sample point Z either $\Theta^* \subset C(Z)$ or $C(Z) \subset \Theta^*$. For example: in a one sided significant ratio*

inference, say $\Theta_i^c = \{\eta_i < \eta_0\zeta\}$*, confidence intervals for* η_i/η_0 *of the form* $C_i(Z) = (-\infty, B_i)$ *are directed toward* Θ_i^c *for* $i = 1, \ldots, k$.

We consider cases with inferences $\theta \in \Theta_i^c$ and formulate an approach that provides confidence intervals $C_i(Z)$ for θ in a stepwise scenario. The procedure stops at the largest integer i when $C_i(Z) \notin \Theta_i^c$. 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 $\Theta_k^* \cap \Theta_{k'}^* = \varnothing$ for any $k, k' \in K$ with $k \neq k'$. We partition the parameter space into disjoint sets so that exactly one partition contains the true parameter θ. In so doing, the procedure controls FWER by controlling FWER within each $\Theta_k^* \subseteq \Theta : k \in K$.

The first step in solving problem (3) is to define the MSD as the $\max\{i : \lambda_i < \zeta\}$. We then employ Fiellers (1954) method to obtain individual $(1 - \alpha)100\%$ confidence intervals $C_i(Z)$ for λ_i. Taking $\bar{Z}_i$ and $\hat{\tau}_i^2$ as the sample mean and sample variance of the i^{th} group, $i = 0, 1, \ldots, k$, we test $H_0 : \theta \in \bigcup_{i=1}^k \Theta_i$ versus $H_a : \theta \in \bigcap_{i=1}^k \Theta_i^c$ using the test statistics:

$$T_i^Z = \frac{\bar{Z}_i - \zeta\bar{Z}_0}{\sqrt{\frac{\hat{\tau}_i^2}{n_i} + \frac{\zeta^2\hat{\tau}_0^2}{n_0}}}.$$

A significant dose-response signal is established, if $T_i^Z < -t_{1-\alpha,v_i}$, where $t_{1-\alpha,v_i}$ is the $(1 - \alpha)$ percentile of the k-variate central t-distribution with v_i degrees of freedom. The degrees of freedom in problems like this was stated in Welch (1938) and improved in Satterthwaite (1946). The different degrees of freedom are given by:

$$\hat{v}_i = \frac{\left(\frac{\hat{\tau}_i^2}{n_i} + \frac{\zeta^2\hat{\tau}_0^2}{n_0}\right)^2}{\left(\frac{\hat{\tau}_i^4}{n_i^2(n_i-1)} + \frac{\zeta^4\hat{\tau}_0^4}{n_0^2(n_0-1)}\right)}.$$

These degrees of freedom, $\hat{v}_i$, are not exact but estimates because they depend on unknown group variances. The upper limits for the confidence intervals of interest $C_i(Z)$ are derived to be the larger root of the quadratic equations $(T_i^Z)^2 = t_{1-\alpha,\hat{v}_i}^2$. These upper confidence limits are given by:

$$B_i = \frac{\bar{z}_i\bar{z}_0 + \sqrt{b_0\bar{z}_i^2 + b_i\bar{z}_0^2 - b_ib_0}}{\bar{z}_0^2 - b_0}.$$

And the $(1 - \alpha)100\%$ confidence intervals $C_i(Z)$ for λ_i are derived as follows: $C_i(Z) = (-\infty, B_i)$, where $b_i = \hat{\tau}_i^2 t_{1-\alpha,\hat{v}_i}/n_i$, $i = 1, \ldots, k$ and $b_0 = \hat{\tau}_0^2 t_{1-\alpha,\hat{v}_i}/n_0$.

3. The Proposed Stepwise Confidence Interval Procedure With Different Variances

To identify the MSD, first compute the upper limits B_i for various dose levels $i = 1, \ldots, k$. Secondly, scan, starting from the lowest dose level ($i = 1$). Sequentially scan doses $i = 2, \ldots, k$ in a step by step manner without adjusting the nominal level α in ascending order searching for the step $i = W$ if it exists, such that $B_W < \zeta$ and $B_{W+1} \geq \zeta$. We take the dose at W for which $B_W < \zeta$ and $B_{W+1} \geq \zeta$ to be the MSD. No MSD is identified if $B_1 \geq \zeta$, indicating that all doses are unsafe. Likewise, a state of no toxicity occurs and all doses are declared as safe if $B_k < \zeta$. Once the dose level at W is identified as MSD, computing the upper confidence limits corresponding to the dose levels at $W + 2, \ldots, k$ becomes unnecessary. In other words, if the test statistic associated with the hypothesis $H_{0\{i\}}$ is statistically significant and the test statistic associated with $H_{0\{i+1\}}$ is statistically insignificant, we can conclude that the doses $1, 2, \ldots, i$ are the only safe doses. As a result, patients may not be subjected to doses $i + 1, \ldots, k$. This reduces the risk of unnecessary exposure of patients to possible toxic effect of a new drug or compound during clinical trials.

Theorem 1. *Suppose that the random sample* $Z_{i1}, Z_{i2}, \ldots, Z_{in_i}$ *is observed from the* i^{th} *dose level having a distribution determined by the parameter* $\theta = (\lambda_1, \ldots, \lambda_k)$ *with* $\theta \in \Theta$*, the parameter space, and* $\theta \in \Theta_i^c = \{\theta : \lambda_i < \zeta\}$*,* $i = 1, \ldots, k$*, a multiple comparison of interest. Consider a confidence set* $C_i(Z)$ *for* θ *based on a set of data* Z *such that* $C_i(Z)$ *is directed toward subsets* $\Theta_i^c \subseteq \Theta$*. Let* W *be the largest integer* i *such that* $C_i(Z) \notin \Theta_i^c$ *if such an* i $(1 \leq i \leq k)$ *exists; otherwise, let* $W = 0$*. Furthermore, let* $\Theta_{k+1} = \varnothing$*,* $\Theta_0 = \Theta$*, and*

$$C(Z) = C_W(Z) \cap \Theta_k^c \cap \Theta_{k-1}^c \cap \ldots \cap \Theta_{W+2}^c \cap \Theta_{W+1}^c.$$

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

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

See Appendix A for the proof of Theorem 1.

Remark. *Theorem 1 indicates that the FWER can be controlled in a strong sense at pre-specified nominal level* α. *In other words, all declarations are guaranteed with probability of at least* $100(1-\alpha)\%$.

Hence, from the result of Theorem 1, we state the following proposition.

Proposition 3.1. *The stepwise confidence procedure for normally distributed data under heteroscedaticity strongly controls the FWER at level* α.

Proof. For any subset I of set $\{1, 2, \cdots, k\}$, suppose that $I = \emptyset$, then no FWER will ever exist. Thus assume that $I \neq \emptyset$ and $I = \{i_1, i_2, \cdots, i_m\}$, where $1 \leq i_1 < i_2 < \cdots < i_m \leq k$. Without loss of generality, let

$$\begin{aligned}
&P(\text{ Reject one of } H_{0i}, i \in I \mid H_{0i}, i \in I \text{ is true }) \\
&= 1 - P(\text{ do not reject all } H_{0i}, i \in I \mid H_{0i}, i \in I \text{ is true}) \\
&\leq 1 - P(\text{ do not reject } H_{0i_m} \mid H_{0i}, i \in I \text{ is true}) \\
&\text{the procedure then stops at step } i_m \\
&= 1 - P(\{B_{i_m}(Z) \not\subset (-\infty, \zeta) \mid H_{0i}, i \in I \text{ is true}) \\
&\leq 1 - P(\theta \in C(Z)) \\
&\leq 1 - (1-\alpha)(\text{ By Theorem 1}) \\
&= \alpha.
\end{aligned}$$

□

Remark. *Therefore, Proposition 1 shows that our resulting stepwise procedure strongly control the FWER at level* α.

Hence, we conduct the following simulations to confirm the above theoretical results.

4. Simulation Studies

4.1 FWER Study

Controlling the right type of error rate in multiple testing procedures is a daunting task. There are three competing error rates in multiple comparison methods. Namely, the FWER, false discovering rate (FDR) and per-family error rate (PFER). As pointed out by Huang and Hsu (2007) in Section 7 of their article, FDR cannot be used for the type of clinical trials discussed in this article. Details and examples of failure of FDR to control type I error can also be found in Finner and Roter (2001). In recent times, Lawrence (2019) claimed that PFER is universally ignored and deserved necessary attention. He argued that PFER is more useful in investigations that deal with social and behavioural sciences. Frane (2015) also confirms the previous authors assertion that PFER is more relevant than FWER in the fields of social and behavioural sciences. The reason is that controlling PFER require stricter standard and large sample sizes compared to a procedure that controls the FWER at the same significant level α. Hence PFER is only better than FWER with multiple endpoint. However in a single endpoint, FWER is preferred, especially in our case when multiplicity adjustment is needless.

FWER is strongly controlled at nominal level α when

$$max(\alpha^*) \leq \alpha \text{ where } \alpha^* = \sup_{H_{0i}}[Pr\{any\ H_{0i}\ is\ rejected\ (1 \leq i \leq k)\}].$$

Without loss of generality, we set $\zeta = 0.8$ and $\alpha = 0.025$. Observation were generated using R statistical software with ten thousand replications from a normal distribution based on the assumption of unequal variance across dose groups. This is indicated in Table 1 as HETRO ($\tau_0 \neq \tau_e$). We explored the effect of violation of this assumption as a way of comparing the two situations and this is also indicated in Table 1 as HOMO ($\tau_0 = \tau_e = 6.0$). We used a fixed sample size of the placebo, $n_0 = 15$, and an increasing number of experimental observations $n_e = 4, \cdots, 29$. The mean and standard deviation configuration of the placebo are set at $\eta_0 = 30.0$ and $\tau_0 = 14.0$ respectively, with the experimental treatment mean and standard deviation set at $\eta_e = 24.0$ and $\tau_e = 6.0$. In the simulation study, we considered only $k = 1$ experimental treatment. Results from Table 1 indicated that the FWER is properly controlled at a nominal value $\alpha = 0.025$ in the case of unequal variances while the FWER is not controlled for equal variances because simulated values are far above 0.025, the nominal level.

Table 1. Simulated FWER results. Set $\zeta = 0.8$, $\alpha = 0.025$, $n_0 = 15$, $\eta_0 = 30.0$, and $\eta_e = 24.0$

n_e	HETRO	HOMO
4(5)	0.0180(0.0185)	0.0408(0.0380)
6(7)	0.0220(0.0202)	0.0370(0.0318)
8(9)	0.0219(0.0208)	0.0325(0.0351)
10(11)	0.0202(0.0208)	0.0298(0.0296)
12(13)	0.0213(0.0237)	0.0301(0.0300)
14(15)	0.0226(0.0235)	0.0300(0.0312)
16(17)	0.0215(0.0233)	0.0270(0.0278)
18(19)	0.0222(0.0241)	0.0285(0.0270)
20(21)	0.0242(0.0228)	0.0308(0.0267)
22(23)	0.0227(0.0230)	0.0286(0.0251)
24(25)	0.0218(0.0221)	0.0253(0.0297)
26(27)	0.0223(0.0210)	0.0290(0.0300)
28(29)	0.0226(0.0209)	0.0269(0.0288)

Table 2. Simulated power results $\alpha = 0.025$, $\zeta = 1.1, 1.2, 1.3, 1.4, 1.5, 1.6$

Treatment effect(λ)	1.1	1.2	1.3	1.4	1.5	1.6
0.8	0.6662	0.5278	0.3879	0.2632	0.1646	0.0949
0.9	0.7795	0.6583	0.5200	0.3814	0.2586	0.1618
1.0	0.8668	0.7722	0.6499	0.5118	0.3746	0.2538
1.1	0.9268	0.8608	0.7642	0.6410	0.5032	0.3675
1.2	0.9635	0.9225	0.8542	0.7557	0.6316	0.4943
1.3	0.9836	0.9608	0.9176	0.8470	0.7466	0.6217
1.4	0.9933	0.9820	0.9577	0.9123	0.8393	0.7370
1.5	0.9976	0.9926	0.9802	0.9541	0.9065	0.8310

Table 3. Simulated power results $\zeta = 1.15$, $n_0 = n_e = 5, 10, 15, 20, 25, 30$

Treatment effect(λ)	5	10	15	20	25	30
0.8	0.1155	0.2238	0.3289	0.4278	0.5180	0.5983
0.9	0.1396	0.2825	0.4169	0.5364	0.6381	0.7219
1.0	0.1668	0.3477	0.5092	0.6422	0.7459	0.8233
1.1	0.1971	0.4176	0.6009	0.7381	0.8339	0.8975
1.2	0.2304	0.4902	0.6874	0.8187	0.8993	0.9460
1.3	0.2666	0.5631	0.7645	0.8817	0.9436	0.9742
1.4	0.3052	0.6338	0.8298	0.9273	0.9709	0.9889
1.5	0.3460	0.7002	0.8821	0.9581	0.9861	0.9957

4.2 Power Calculation

In recent times, confidence interval approaches for the analysis of a clinical study are frequently becoming insufficient in the design of clinical trials. Power estimation has therefore become a major task in the design phase of clinical trials. This article seeks to correctly estimate MSD S_i for a fixed $i \in 1, \cdots, k$. The MSD is correctly specified, if and only if $C_j(Z) \subset (-\infty, \zeta)$ and $C_{i+1}(Z) \not\subset (-\infty, \zeta)$ for $j = 1, ..., i$. Thus

$$P(MSD = S_i) = P\left(\bigcap_{j=1}^{i} \left\{T_j^Z > t_{1-\alpha,\hat{v}_j}\right\} \cap \left\{T_{i+1}^Z \leq t_{1-\alpha,\hat{v}_j}\right\}\right). \tag{4}$$

The probability of correctly estimating any of the doses, $S_1, \cdots, S_i$, as the true MSD is referred to as the power of the procedure. Here, the power is defined solely in terms of rejecting the incorrect null hypotheses. Equation (4) will therefore be expressed as

$$P(rejecting\ H_{0j}, 1 \le j \le i) = P\left(\bigcap_{j=1}^{i}\left\{T_j^Z > t_{1-\alpha,\hat{\nu}_j}\right\}\right). \tag{5}$$

Equation in (5) can be interpreted as the probability of correctly estimating any of the doses, $S_1, \cdots, S_i$, as the true MSD. (4) and (5) gives similar results. Nevertheless, the stepwise procedure presented in Section 3 controls the familywise error rate indicating that the probability of incorrectly estimating any of the $S_{i+1}, \cdots, S_k$ to be the true MSD is still controlled. The probability in (5) can be evaluated using existing numerical methods published in Genz and Bretz (2002) for the computation of non-central multivariate t-probabilities. Expression in (5) involves a $k - i + 1$ variate non-central t-distribution with $\hat{\nu}$ degrees of freedom as stated in section 2 above. The non-centrality parameters are given by

$$\Theta_j = \frac{\eta_j - \zeta\eta_0}{\sqrt{\frac{\tau_j^2}{n_j} + \frac{\zeta^2\tau_0^2}{n_0}}}, \quad j = 1, ..., i.$$

A simulation study concerning the impact of various parameters on the power was done setting $\zeta = 1.1$ and $\alpha = 0.025$. Results for different ratio of means and clinical relevance margins are presented in Table 2. As expected, power increases with increasing ratio of means (incresing effect, λ), and decreases for higher clinical relevance margins (that is, larger values of ζ). In Table 3, power increases with an increase in both the ratio of means, and the sample size.

5. Practical Application

To illustrate our stepwise confidence interval procedure, we consider an example from Tamhane and Logan (2004). In the study, 90-day routine rat study was conducted to identify the toxicity of a crop protection compound. Test substance was added directly to the rodent diet and was thoroughly mixed to ensure homogeneous distribution. Three doses of the compound were compared with a zero dose control. The goal of the study is to measure the kidney weight to the body weight ratio. A large value of this ratio is regarded as unsafe with a threshold of a 15% average increase over its value for the zero dose. Thus, we take $\zeta = 1.15$. The summary statistics for the zero dose control and kidney weight is given in Table 4. The assumption of normality of the data and heterogeneity of variances across the dose groups are satisfied, (see Tamhane and Logan (2004)). Given that $\zeta = 1.15$ and $\alpha = 0.05$, Table 5 shows the 95% lower confidence limits on η_i/η_0, $i = 1, 2, 3$.

Table 4. Kidney Wt./Body Wt.$\times 10^3$

Treatment group	n_i	Mean	Standard deviation
Control	18	6.5606	0.5064
1	20	6.9975	0.5755
2	19	7.6778	0.5949
3	18	9.2606	1.0052

Table 5. 95% Upper Confidence Limits for η_i/η_0, $i = 1, 2, 3$

Treatment group	Upper confidence limit
1	1.1243
2	1.2326
3	1.5047

The results in Table 5 show that the upper confidence limit for dose 1 is lower than the relevant toxicity threshold and therefore is declared to be statistically significant at level 0.05. Whereas upper confidence limits for doses 2 and 3 are however greater than the relevant threshold and therefore are declared to be statistically insignificant at level 0.05. Note that MSD is correctly specified, if and only if $C_j(Z) \subset (-\infty, \zeta)$ and $C_{i+1}(Z) \not\subset (-\infty, \zeta)$ for $j = 1, ..., i$.

$C_1(Z) = (-\infty, 1.1243) \subset (-\infty, 1.15)$ we reject H_{01} (dose 1 is safe)

$C_2(Z) = (-\infty, 1.2326) \not\subset (-\infty, 1.15)$ we do not reject H_{02} (dose 2 is unsafe)

The stepwise confidence interval procedure stops at step 2, and will however need no further testing. In this analysis, our procedure concluded that dose 1 is safe while doses 2 and 3 are unsafe at level 0.05. Therefore, dose 1 is recommended as the MSD.

6. Conclusions

Under the assumption of heteroscedasticity, this article established a general way of obtaining a stepwise confidence interval for the ratio of means. Most of the procedures formulated to identify MSD of drugs were done under the assumption of equal variance, but the homoscedasticity assumption is rarely satisfied in practice. This is as a result of certain biological factors that can cause patients in different dose groups to respond differently with a change in dose levels. In many standard situations, inferences on differences are also not a suitable way to investigate the data. Ratios, therefore, are mostly a better alternative measure of toxicity and often easier to interpret. Ratio formulation approaches are more powerful in a test for non-inferiority of an experimental therapy as compared to that of the difference, and the ratio parameters have the advantage of being available as a dimensionless percentage, easily estimated, and enhance comparison among studies. 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. Furthermore, our procedure provides additional information such as a confidence bound for the first unsafe dose. Or, provides an upper confidence bound on how safe the doses are if all the doses are regarded safe. In order to assess the performance of the proposed procedure, a simulation study is carried out. Results from the simulation study 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. It is also found that all the stepwise confidence intervals control the familywise error rate.

References

Adjabui, M. J., Howard, N., & Luguterah, A. (2017). Nonparametric stepwise procedure for identification of maximum safe dose (MSD). *Asian Research Journal of Mathematics, 6*(3), 1-12. https://doi.org/10.9734/ARJOM/2017/36377

Berger, R. L. (1982). Multiparameter hypothesis testing and acceptance sampling. *Technometrics, 24*, 295-300. https://doi.org/10.2307/1267823

Bretz, F., Hothorn, L. A., & Hsu, J. C. (2003). Identifying effective and/or safe doses by stepwise confidence intervals for ratios. *Statistics in Medicine, 22*, 847-858. https://doi.org/10.1002/sim.1449

Fieller, E. C. (1954). Some problems in interval estimation. *Journal of the Royal Statistical Society, Series B 16*, 175-185. https://doi.org/10.1111/j.2517-6161.1954.tb00159.x

Finner, H., & Roter, M. (2001). On the false discovering rate and expected type I error. *Biometrical J., 43*, 985-1005. https://doi.org/10.1002/1521-4036(200112)43:8¡985::AID-BIMJ985¿3.0.CO;2-4

Frane, A. V. (2015). Are per-family type I error rate relevant in social and behavioural science. *J. Mod Appl Stat method, 14*(1), 5. https://doi.org/10.22237/jmasm/1430453040

Genz, A., & Bretz, F. (2002). Methods for the computation of multivariate t-probabilities. *Journal of Computational and Graphical Statistics, 11*, 950-971. https://doi.org/10.1198/106186002394

Hauschke, D., Kieser, M., & Hothorn, L. A. (1999). Proof of safety in toxicology based on the ratio of two means for normally distributed data. *Biometrical Journal, 41*(3), 295-304. https://doi.org/10.1002/(SICI)1521-4036(199906)41:3¡295::AID-BIMJ295¿3.0.CO;2-2

Hothorn, L. A., & Hauschke, D. (2000). Identifying the maximum safe dose: a multiple testing approach. *Journal of Biopharmaceutical Statistics, 10*, 15-30. https://doi.org/10.1081/BIP-100101010

Hsu, J. C., & Berger, R. (1999). Stepwise confidence intervals without multiplicity adjustment for dose-response and toxicity studies. *Journal of the American Statistical Association, 94*, 468-482. https://doi.org/10.2307/2670167

Huang, Y., & Hsu, J. C. (2007) Hochber's step-up method: Cutting corner's off Holm's stepdown methods. *Biometrika, 94*(4), 965-975. https://doi.org/10.1093/biomet/asm067

Korn, E. L., Midthune, D., Chen, T. T., Rubinstein, L. V., Christian, M. C., & Simon, R. M. (1994). A comparison of two phase I trial designs. *Statistics in Medicine*, 1799-1806.

Laster, L. L., & Johnson, M. F. (2003). Non-inferiority trials: the at least as good as criterion *Statistics in Medicine, 22*, 187-200. https://doi.org/10.1002/sim.1137

Lawrence, J. (2019). Familywise and per-familywise error rate of multiple comparisons procedures. *Statistics in medicine, 38*, 3586-3598. https://doi.org/10.1002/sim.8190

Satterthwaite, F. E. (1946). An approximate distribution of estimates variances component. *Biometrics, 2*, 110-114. https://doi.org/10.2307/3002019

Tamhane, A. C., & Logan, B. R. (2004). Finding the multiple safe dose level for heteroscedastic data. *Journal of Biopharmaceutical Statistics, 14*(4), 843-856. https://doi.org/10.1081/BIP-200035413

Tamhane, A. C., Dunnett, C. W., Green, J. W., & Wetherington, J. F. (2001). Multiple test procedures for identifying the maximum safe dose. *Journal of the American Statistical Association, 96*, 835-843. https://doi.org/10.1198/016214501753208546

Tao, J., Guo, J.-H., & Shi, N.-Z. (2002). Stepwise Procedures under Unknown Variances for Toxicological Evaluation. *Biometrical Journal, 44*, 149-160. https://doi.org/10.1002/1521-4036(200203)44:2¡149::AID-BIMJ149¿3.0.CO;2#

Turri, M., & Stein, G. (1986). The determination of practically useful doses of new drugs: some methodological considerations. *Statistics in Medicine, 5*, 449-457. https://doi.org/10.1002/sim.4780050509

Welch, B. L. (1938). The significance of difference between two means when the population variances unequal. *Biometrika, 29*, 350-362. https://doi.org/10.2307/2332010

Appendix A: Proof of Theorem 1.

For any $\theta \in \Theta$, $\theta \in \bigcup_{i=0}^{k} \Theta_i$.
Consider the following sets $\Theta_0^*, \Theta_1^*, \ldots, \Theta_k^*$, as partitions of the parameter space Θ.

$$\begin{aligned}
\Theta_k^* &= \Theta_k \\
\Theta_{k-1}^* &= \Theta_{k-1} \cap \Theta_k^c \\
\Theta_{k-2}^* &= \Theta_{k-2} \cap \Theta_k^c \cap \Theta_{k-1}^c \\
&\vdots \\
\Theta_i^* &= \Theta_i \cap \Theta_k^c \cap \Theta_{k-1}^c \cap \ldots \cap \Theta_{i+1}^c \\
&\vdots \\
\Theta_1^* &= \Theta_1 \cap \Theta_k^c \cap \Theta_{k-1}^c \cap \ldots \cap \Theta_2^c \\
\Theta_0^* &= \Theta_0 \cap \Theta_k^c \cap \Theta_{k-1}^c \cap \ldots \cap \Theta_1^c
\end{aligned}$$

Thus

$$\Theta = \bigcup_{i=0}^{k} \Theta_i = \bigcup_{i=0}^{k} \Theta_i^*. \tag{6}$$

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

$$C(Z) = \bigcup_{i=0}^{k} (C_i(Z) \cap \Theta_i^*). \tag{7}$$

Here

$$\bigcup_{i=0}^{k} (C_i(Z) \cap \Theta_i^*) = \bigcup_{i=0}^{k} (C_i(Z) \cap \Theta_i \cap \Theta_k^c \cap \Theta_{k-1}^c \cap \ldots \cap \Theta_{i+1}^c). \tag{8}$$

For some non negative integer W, (8) can be written as

$$\bigcup_{i=0}^{k} (C_i(Z) \cap \Theta_i^*) = \bigcup_{i=0}^{W-1} \bigcup \bigcup_{i=W} \bigcup \bigcup_{i=W+1}^{k} (C_i(Z) \cap \Theta_i \cap \Theta_k^c \cap \Theta_{k-1}^c \cap \ldots \cap \Theta_{i+1}^c). \tag{9}$$

From the Theorem 1, W is the largest integer i such that $C_i(Z) \not\subset \Theta_i^c$, then for $W+1 \leq i \leq k$, $C_i(Z) \subset \Theta_i^c$. Thus $C_i(Z)$ is a subset of Θ_i^c that has no common element with Θ_i.
Note that

$$\bigcup_{i=W+1}^{k} (C_i(Z) \cap \Theta_i \cap \Theta_k^c \cap \Theta_{k-1}^c \cap \ldots \cap \Theta_{i+1}^c) = \varnothing. \tag{10}$$

We can write

$$\begin{aligned} C(Z) &= \bigcup_{i=0}^{k}(C_i(Z) \cap \Theta_i^*) \\ &= \bigcup_{i=0}^{W-1}\bigcup\bigcup_{i=W}(C_i(Z) \cap \Theta_i \cap \Theta_k^c \cap \Theta_{k-1}^c \cap \ldots \cap \Theta_{i+1}^c). \end{aligned}$$

Meaning

$$C(Z) = \bigcup_{i=0}^{W-1}(C_i(Z) \cap \Theta_i \cap \Theta_k^c \cap \Theta_{k-1}^c \cap \ldots \cap \Theta_W^c \cap \ldots \cap \Theta_{i+1}^c) \cup (C_W(Z) \cap \Theta_W \cap \Theta_k^c \cap \Theta_{k-1}^c \cap \ldots \cap \Theta_{W+1}^c)$$

$$\subseteq \quad (\Theta_k^c \cap \Theta_{k-1}^c \cap \ldots \cap \Theta_W^c) \cup (C_W(Z) \cap \Theta_W \cap \Theta_k^c \cap \Theta_{k-1}^c \cap \ldots \cap \Theta_{W+2}^c \cap \Theta_{W+1}^c)$$

$$= \quad C_W(Z) \cap \Theta_k^c \cap \Theta_{k-1}^c \cap \ldots \cap \Theta_{W+2}^c \cap \Theta_{W+1}^c.$$

Confidence set $C_i(Z)$ is directed toward parameter set Θ_i^c, therefore for all $\theta \in \Theta_i^*$,

$$P(\theta \in C_W(Z) \cap \Theta_k^c \cap \Theta_{k-1}^c \cap \ldots \cap \Theta_{W+2}^c \cap \Theta_{W+1}^c) \geq P(\theta \in C(Z)) \geq 1 - \alpha. \tag{11}$$

This completes the proof of Theorem 1. □