
Stepwise procedure for toxicological assessment based on ratio of mean differences for a normally distributed data

**Original Research
Article**

Abstract

We propose a new stepwise confidence set procedure for toxicity study based on ratio of mean difference. Statistical approaches for evaluating toxicity studies that properly control familywise error rate (FWER) for difference of means between treatments and a control already exist. However, in some therapeutic areas, ratio of mean differences is desirable. Therefore, we construct stepwise confidence procedure based on Fieller's confidence intervals for multiple ratio of mean difference without multiplicity adjustment for toxicological evaluation. Simulation study revealed that the FWER is well controlled at prespecified nominal level α . Also, the power of our approach increases with increasing sample size and ratio of mean differences.

Keywords: Balanced design; Fieller; FWER; MED; Stepwise procedure ; Power;

2010 Mathematics Subject Classification: 53C25; 83C05; 57N16

1 Introduction

Toxicity study designed to assess safety of novel drug at various dose levels is a vital concern in recent drug industries. Various statistical procedures for evaluating these toxicological substances for properly controlling the FWER have been proposed. Among them are [14], [8], [13], [12], [2]. In toxicological evaluations, [14] proposed three different testing approaches for comparisons of treatments means and a control mean. [8], proposed a confidence set method for toxicity study under homogeneity of variance across dose groups. However, unknown and equal variances at different dose levels are hardly sustainable in practice. [13] extended [8] procedure under heteroscedastic assumption across dose groups by incorporating two-stage sampling method proposed by [15] for

toxicological studies. [12] extended the confidence procedure for toxicity evaluation based on asymmetric loss function. However, their procedure demands assumption of known variances which is often not reliable for some cases. [2] improved the work of [12] to a case of unknown variances by employing confidence procedure for asymmetric loss function. All these investigations on toxicological equivalence are based on mean differences. However ratio of mean difference is preferred because, apart from being easily interpreted medically for certain therapeutic areas, it is also free of unit of measurement at the endpoint as compared with difference in location parameters. The paper is outlined as follows. Notations, assumptions and formulation of the problem as a testing procedure and we generalized Fierller's confidence intervals for construction of stepwise confidence intervals for ratio of mean differences in Section 2. In Section 3, we propose stepwise procedure for toxicity study. Simulation studies were carried out to investigate the performance of FWER and the power of our procedure in Section 4. Analysis of data set for Bovine Growth Hormone in Toxicity study is presented in Section 5. Section 6 is the conclusion of the article

2 Testing Procedure

To motivate the stepwise confidence procedure under homogeneity of variance across different dose levels for toxicity studies, we will review [7] testing and confidence interval procedure for ratio of means difference in a one-way model and construct a stepwise confidence intervals procedure.

2.1 Testing procedure (Hasler 2012)

Consider a one-way model for practical equivalence problem with k experimental treatments ($E_i, i = 1, 2, \dots, k$), a positive control (E_{k+1}) and a placebo (P). Suppose that $X_{E_i}, X_{E_{k+1}}$ and X_P are the observations for safety endpoints of new treatments, positive control and a placebo respectively. These random variables are mutually independent and follow a normal distribution with treatment means μ_{E_i} , for $i = 1, 2, \dots, k$ with $\mu_{E_{k+1}}$ and μ_P being the means of positive control and placebo respectively. Their respective sample sizes $n_{E_i}, n_{E_{k+1}}$ and n_P are not necessary equal but have a common unknown variance σ^2 . We state the one-way model as:

$$X_{ij} = \mu_i + \epsilon_{ij} \quad i = E_1, E_2, \dots, E_k, E_{k+1}, P, \quad j = 1, \dots, n_i \quad (1)$$

where X_{ij} is the safety response values for the i th dose level. This is formulated in terms of hypothesis as:

$$H_{0i} : \frac{\mu_{E_i} - \mu_P}{\mu_{E_{k+1}} - \mu_P} \leq \delta^{(L)} \text{ or } \frac{\mu_{E_i} - \mu_P}{\mu_{E_{k+1}} - \mu_P} \geq \delta^{(U)} \text{ versus } H_{1i} : \delta^{(L)} < \frac{\mu_{E_i} - \mu_P}{\mu_{E_{k+1}} - \mu_P} < \delta^{(U)} \quad (2)$$

for $i = 1, 2, \dots, k$ where $\delta^{(L)} < 0$ and $\delta^{(U)} > 0$ are some pre-specified quantities. In practice, $\delta^{(L)}$ could be chosen to be $-\delta^{(U)}$ as a relevant safety threshold quantities. Equation (2) can be reformulated as:

$$H_{0i} : \gamma_i \leq \delta^{(L)} \text{ or } \gamma_i \geq \delta^{(U)} \text{ versus } H_{1i} : \delta^{(L)} < \gamma_i < \delta^{(U)}$$

where γ_i is the ratio of difference in means as:

$$\gamma_i = \frac{\mu_{E_i} - \mu_P}{\mu_{E_{k+1}} - \mu_P} \text{ for } i = 1, 2, \dots, k \quad (3)$$

Let the sample mean estimates be:

$$\bar{X}_{E_i} = \frac{1}{n_{E_i}} \sum_{i=1}^{n_{E_i}} X_{E_i}, i = 1, 2, \dots, k \quad \bar{X}_{k+1} = \frac{1}{n_{E_{k+1}}} \sum_{k=1}^{n_{k+1}} X_{E_{k+1},k} \quad \text{and} \quad \bar{X}_P = \frac{1}{n_P} \sum_{j=1}^{n_P} X_{P,j}$$

The unknown and common variance σ^2 can be estimated as:

$$\hat{\sigma}^2 = \frac{(n_{E_i} - 1)S_{E_i} + (n_{E_{k+1}} - 1)S_{E_{k+1}} + (n_P - 1)S_P}{N - (k + 2)} \quad \text{for } i = 1, 2, \dots, k$$

where $\hat{\sigma}^2$ is the pooled estimator of the variance σ^2 and N denotes the total sample size. Then the following modified random variables are obtained from [11] and [7]

$$T_{E_i} = \frac{\bar{X}_{E_i} - (r)\bar{X}_{E_{k+1}} - (1 - (r))\bar{X}_P}{\hat{\sigma} \sqrt{\frac{1}{n_{E_i}} + \frac{(r^2)}{n_{E_{k+1}}} + \frac{(1-r)^2}{n_P}}} \tag{4}$$

where $r = \delta^{(L)}$ or $\delta^{(U)}$ for $i = 1, 2, \dots, k$ are the test statistics for the testing problem in Equation (1), which has t -distribution with $\nu_i = n_i + n_{E_{k+1}} + n_P - 3$ degrees of freedom. Suppose that increasing values of the endpoints represent better treatment effect, then equivalence/safety can be concluded if H_{0i} is rejected. That is:

$$T_{E_i} > t_{k,1-\alpha}(\nu_i) \quad \text{for } i = 1, 2, \dots, k \tag{5}$$

with its corresponding $(1 - \alpha)$ -quantiles $t_{k,1-\alpha}(\nu_i)$ of central k -variate t -distribution with degrees of freedom $\nu_i (i = 1, \dots, k)$.

Therefore, we construct stepwise confidence intervals set procedure based on partitioning principle in Section 2.2.

2.2 Construction of stepwise confidence procedure.

For $i = 1, 2, \dots, k$, [3] extended generalized Fieller's theorem [4] to construct simultaneous confidence interval for γ_i . The fact is, we need to solve k quadratic equations in order to devise a new confidence procedure. This results into the following simultaneous $(1 - \alpha)100\%$ confidence limits $Z_{i,1-\alpha}$:

$$Z_{i,1-\alpha} = \left(\frac{-B_i \pm \sqrt{(B_i)^2 - 4A_i C_i}}{2A_i} \right) \quad i = 1, 2, \dots, k$$

$$A_i = (\bar{X}_{E_{k+1}} - \bar{X}_P)^2 - t_{k,(1-\alpha)\nu_i}^2 \hat{\sigma}^2 \left(\frac{1}{n_{E_{k+1}}} + \frac{1}{n_P} \right)$$

$$B_i = -2(\bar{X}_{E_i} - \bar{X}_P)(\bar{X}_{E_{k+1}} - \bar{X}_P) - t_{k,(1-\alpha)\nu_i}^2 \frac{\hat{\sigma}^2}{n_P}$$

$$C_i = (\bar{X}_{E_i} - \bar{X}_P)^2 - t_{k,(1-\alpha)\nu_i}^2 \hat{\sigma}^2 \left(\frac{1}{n_{E_i}} + \frac{1}{n_P} \right)$$

These confidence limits are valid as long as $A_i > 0$. This restriction is fulfilled if and only if $\mu_{E_{k+1}} - \mu_P$ is significantly greater than zero and must be established in the first step in our stepwise procedure for assay sensitivity of the model. In this sequel, we intend to generalize [8] stepwise confidence set-based procedure for mean difference by extending it to ratio of mean difference. To do this, we elicit some results from Hsu and Berger's stepwise confidence procedure for toxicity studies.

Definition 2.1. Suppose that the data X have a distribution determined by a parameter $\Gamma = \{\gamma_1, \gamma_2, \dots, \gamma_k\} \in \Theta$. A confidence set $C(X)$ for Θ is said to be directed towards a subset of the parameter space $\Theta^* \subset \Theta$ if for every sample point X , either $\Theta^* \subset C(X)$ or $C(X) \subset \Theta^*$.

For $(i = 1, \dots, k)$, let

$$D_i^-(X) = \min \left\{ \frac{-B_i - \sqrt{(B_i)^2 - 4A_i C_i}}{2A_i}, 0 \right\}$$

and

$$D_i^+(X) = \max \left\{ \frac{-B_i + \sqrt{(B_i)^2 - 4A_iC_i}}{2A_i}, 0 \right\}$$

Then

$$D_i(X) = \begin{cases} (D_i^-(X), D_i^+(X)), & \text{if } D_i^-(X) < 0 < D_i^+(X), \\ [0, D_i^+(X)], & \text{if } D_i^-(X) = 0, \\ (D_i^-(X), 0], & \text{if } D_i^+(X) = 0 \end{cases}$$

is a $100(1 - \alpha)$ confidence intervals for γ_i .

Let

$$C_i(X) = \begin{cases} D_i(X) & \text{if } D_i(X) \subset (\delta^{(L)}, \delta^{(U)}) \\ D_i(X) \cup (\delta^{(L)}, \delta^{(U)}) & \text{otherwise.} \end{cases}$$

Then $C_i(X)$ is a $100(1 - \alpha)$ confidence intervals for γ_i directed towards $\Theta^* = (\delta^{(L)} < \gamma_i < \delta^{(U)})$ for $i = 1, 2, \dots, k$. In this setting, the confidence set $C_i(X)$ contains the alternative space $\Theta^* = (\delta^{(L)} < \gamma_i < \delta^{(U)})$ or the confidence set is contained in the alternative space $\Theta^* = (\delta^{(L)} \leq \gamma_i$ or $\gamma_i \geq \delta^{(U)})$.

Now the stepwise confidence set procedure for ratio of mean difference is elucidated in Section 3.

3 The proposed procedure

3.1 Stepwise confidence set for toxicological assessment based on ratio of mean differences

To start the stepwise procedure, we make the following two assumptions; Firstly, we assume that A_i is significantly greater than zero. Secondly, dosages of this particular novel drug decreases with increasing dose levels.

Hence, given the $100(1 - \alpha)$ confidence intervals for γ_i :

$$D_i(X) = \begin{cases} (D_i^-(X), D_i^+(X)), & \text{if } D_i^-(X) < 0 < D_i^+(X), \\ [0, D_i^+(X)], & \text{if } D_i^-(X) = 0, \\ (D_i^-(X), 0], & \text{if } D_i^+(X) = 0 \end{cases},$$

for $i = 1, 2, \dots, k$, where k is the total number of dosages to be scanned.

We scan the first toxicological safety/equivalence dose by scanning the highest dose level at $D_k(X)$ for first equivalence drug if its exists and sequentially scan the subsequent doses for $i = k - 1, k - 2, \dots, 1$ without adjusting the α levels in each of the steps in descending fashion searching for the smallest integer M ($1 \leq M \leq k$), if it exists such that $D_M(X) \subset (\delta^{(L)}, \delta^{(U)})$ and $D_{M-1}(X) \not\subset (\delta^{(L)}, \delta^{(U)})$ (this scans the first non-equivalence or unsafe dose). In this set up, doses at $D_k(X), D_{k-1}(X), \dots, D_M(X)$ are established as equivalence while doses at $D_{M-1}(X), D_{M-2}(X), \dots, D_1(X)$ are non-equivalence. Notice that the confidence intervals at each step are computed without multiplicity adjustments.

To elucidate the above procedure, let M ($1 \leq M \leq k$) be the step at which the procedure is terminated. If $1 < M < k$, then a confidence set for γ_M that contains $(\delta^{(L)}, \delta^{(U)})$ is given, and the confidence intervals $\gamma_i \in (\delta^{(L)}, \delta^{(U)})$ for $i = 1, \dots, M$ are given if $M > 1$. If $M = 1$, then a common confidence interval for all $\gamma_i = 1, \dots, k$ which are entirely within the range $(\delta^{(L)}, \delta^{(U)})$ is given. Hence, we state the following proposition.

Proposition 3.1. Let X represent a sample data point and let Θ be the parameter space for parameter vector Γ . For any $i = 1, \dots, k$. let $D_i(X)$ be any $100(1 - \alpha)\%$ confidence interval for γ_i , also let $C_i(X)$

be confidence set directed towards $\delta^{(L)} < \gamma_i < \delta^{(U)}$. Denote M the smallest integer of i such that $D_M(X) \subset (\delta^{(L)}, \delta^{(U)})$ if such an $i(1 \leq i \leq k)$ exists; otherwise let $M = k + 1$. Then for any $\Gamma \in \Theta$

$$P(D_k(X) \subset (\delta^{(L)}, \delta^{(U)}) \cap \dots \cap D_M(X) \subset (\delta^{(L)}, \delta^{(U)}) \cap D_{M-1}(X) \not\subset (\delta^{(L)}, \delta^{(U)}) \cap C_{M-1}(X)) \geq 1 - \alpha$$

Proof. Let step $M(1 \leq M \leq K)$ be the step at which the stepwise procedure stops. If $A_i \leq 0$ then the sensitivity of the experiment is inadequate and the lower confidence bound for γ_{k+1} is given. If for each $D_i(X) \subset (\delta^{(L)}, \delta^{(U)})$, there is a $100(1 - \alpha)\%$ confidence interval for γ_i for $M > 1$, then $C_i(X)$ is a $100(1 - \alpha)$ confidence intervals for γ_i that is directed towards $\Theta^* = (\delta^{(L)} < \gamma_i < \delta^{(U)})$ for $i = 1, 2, \dots, k$. The rest of the proof follows Theorem 1 of [1]. □

Remark 3.1. Proposition 3.1 guarantees that the overall coverage probability is at least $(1 - \alpha)100\%$. In other words, the FWER is properly controlled at prespecified nominal level α .

4 Simulation studies

4.1 FWER

There are two competitor error rates for toxicological investigation. They are FWER, and the false discovering rate (FDR). [9] claimed that the FDR can not be used for the type of clinical trials discussed in this article but FWER, the details and examples can be found in [5].

We conduct simulation studies to investigate the performance of the FWER of our procedure under the assumption of unknown but equal variances across dose groups. But in practice, the assumption of equal variances is hardly ever sustainable. For this reason, we will compare our procedure with a situation when our assumption of equal variance is violated. Hence the unknown unequal variances testing problem similar to that of (1) can be formulated as two-one-sided test known as TOST. It utilizes the two one-side size- α Welch's approximation t-test to the following hypotheses

$$H_{0i}^1 : \frac{\mu_{E_i} - \mu_P}{\mu_{E_{k+1}} - \mu_P} \leq \delta^{(L)} \text{ vs } H_{1i}^1 : \frac{\mu_{E_i} - \mu_P}{\mu_{E_{k+1}} - \mu_P} > \delta^{(L)}$$

$$H_{0i}^2 : \frac{\mu_{E_i} - \mu_P}{\mu_{E_{k+1}} - \mu_P} \geq \delta^{(U)} \text{ vs } H_{1i}^2 : \frac{\mu_{E_i} - \mu_P}{\mu_{E_{k+1}} - \mu_P} < \delta^{(U)}$$

Therefore in Table 1, we represent this as HET and that of the equal variance assumption as HOM. Without loss of generality we set $\alpha = 0.025$ and $-\delta^{(L)} = \delta^{(U)} = 0.8$ to investigate the performance of our procedure in terms of the FWER while in Table 2 we assessed the power of our procedure. Hence the simulation study is similar to that of [1]. The results of our simulation indicated as in Table 1 that the FWER is well controlled for a situation of equal variances across doses groups (HOM) but liberal when this situation is violated (HET).

4.2 Power Estimation

The power of our procedure is according to the Equation (5) given as:

$$P(T_{E_i} > t_{k, 1-\alpha(\nu_i)}) = \alpha \text{ for } i = 1, 2, \dots, k \tag{6}$$

Table 1: Simulated FWER, given $\alpha = 0.025, n_R = 20, n_P = 20$, and, $-\delta^{(L)} = \delta^{(U)} = 0.8$

$n_{E_1}(n_{E_2})$	HOM	HET
7 (8)	0.0249 (0.02480)	0.0184 (0.0177)
9 (10)	0.0251 (0.0251)	0.0109 (0.0160)
11 (12)	0.0249 (0.0247)	0.0157 (0.0153)
13 (14)	0.0244 (0.0249)	0.0149 (0.0114)
15 (16)	0.0247 (0.0248)	0.0141 (0.0136)
17 (18)	0.0249 (0.0250)	0.0129 (0.0128)
19 (20)	0.0249 (0.0248)	0.0124 (0.0119)
21 (22)	0.0249 (0.0250)	0.0117 (0.0115)
23 (24)	0.0205 (0.0247)	0.0110 (0.0109)
25 (26)	0.0250 (0.0249)	0.0106 (0.0106)
27 (28)	0.0251 (0.0245)	0.0100 (0.0009)
29 (30)	0.0250 (0.0248)	0.0096 (0.0093)

the probability of correctly accepting H_{1i} . Therefore Equation (6) is calculated from non-centrality parameter Θ

$$\Theta_i = \frac{\mu_{E_i} - (r)\mu_{E_{k+1}} - (1 - (r))\mu_P}{\sigma \sqrt{\frac{1}{n_{E_i}} + \frac{(r^2)}{n_{E_{k+1}}} + \frac{(1-r)^2}{n_P}}} \tag{7}$$

We have the expression $\sigma = \epsilon(n_{E_{k+1}} - n_P)$, $\epsilon > 0$ therefore the following representation of non-centrality parameter based on the ratio of mean differences is stated as:

$$\Theta_i = \frac{\gamma_i - r}{\epsilon \sqrt{\left\{ \frac{1}{n_{E_i}} + \frac{r^2}{n_{E-k+1}} + \frac{(1-r)^2}{n_p} \right\}}} \tag{8}$$

From Equation (8), it is clear that the expected values of power is a function of γ_i , the ratio of mean differences and the sample sizes. From Table 2, it can be seen that power increases with increasing γ_i and sample size but decreases with increasing values of ϵ . This is consistent with the results of [11].

Table 2: Power Estimation of the confidence intervals for $\sigma_R = 10, \sigma_P = 10, \sigma_{E_i} = 10, i = 1, 2.$

Ratio(γ_i)	$n_{E_i=1,2}$	$\epsilon = 0.25$	$\epsilon = 0.5$	$\epsilon = 1$
0.85	5	0.0623	0.0402	0.0319
0.85	20	0.1161	0.0574	0.00385
0.85	30	0.1409	0.0644	0.0410
0.85	40	0.1606	0.700	0.00429
0.90	5	0.1332	0.0623	0.0402
0.90	20	0.3336	0.1410	0.0573
0.90	30	0.4234	0.1409	0.0645
0.90	40	0.4903	0.1606	0.0700
0.95	5	0.2460	0.0928	0.0503
0.95	20	0.6312	0.2082	0.0828
0.95	30	0.7550	0.2627	0.1085
0.95	40	0.8273	0.3056	0.1086
1.00	5	0.3964	0.1332	0.0623
1.00	20	0.8643	0.3336	0.1161
1.00	30	0.9422	0.4230	0.1409
1.00	40	0.9720	0.4903	0.1606
1.05	5	0.5641	0.11842	0.0764
1.05	20	0.9689	0.4830	0.1578
1.05	30	0.9930	0.5982	0.1961
1.05	40	0.9980	0.6771	0.2266
1.10	5	0.7300	0.2460	0.0928
1.10	20	0.9957	0.6312	0.2082
1.10	30	0.9996	0.7550	0.2630
1.10	40	0.9999	0.8273	0.3057
1.15	5	0.8437	0.3124	0.1113
1.15	20	0.9965	0.7635	0.2672
1.15	30	0.9996	0.9232	0.3335
1.15	40	0.9999	0.9232	0.3951
1.20	5	0.9242	0.3963	0.1332
1.20	20	0.9999	0.8643	0.3335
1.20	30	0.9999	0.9422	0.4235
1.20	40	1.0000	0.9730	0.4903

5 Example: Bovine Growth Hormone Toxicity study

To illustrate the procedure discussed in this article, we used bovine growth hormone for toxicity assessment. Writing for Food and Drug Administration (FDA), [10] reported on a number of experiments that did not indicate bovine growth hormones are harmful if present in milk consumed by humans. A subset of this data was considered by [8]. Data from one of the experiments in that article gave absolute weight of various organs measured from control hypophysectomized rats and hypophysectomized rats treated orally with peptide hormone recombinant insulin-like growth factor- I(rIGF-I). In addition to groups given rIGF-I orally, one group was given negative "saline control" and another group was given rIGF-I via subcutenously (sc) implanted osmotic minipump as a positive control. Spleen weights of rats treated for either 17 days by gavage or 15 days by continuous sc infusion are given in Table 3.

Table 3: Spleen weight (g) of male rats ([10])

Tretament label	Dose (mg/kg)	Sample Size	Mean Weight	Std.dev. Weight
1 =non(Saline)	0	20	147.6	8.8
2 = oral rIGF-I	0.01	20	147.2	5.7
3 =oral rIGF-I	.1	20	149.66	5.8
4 =oral rIGF-I	1.	20	147.1	6.6
5 = Sc infusion rIGF-I	1.0	10	239.5	17.9

$$A_i > 0$$

$$D_1(X) = (-0.0368, 0.0288)$$

$$D_2(X) = (-0.0116, 0.05573)$$

$$D_3(X) = (-0.0431, 0.0778)$$

Since $A_i > 0$, we can claim that the experiment is sufficiently sensitive to distinguish between positive and negative control. It could be observed that all the $D_i(X)$ for $i = 1, 2, 3$ and 4 lies entirely in the range $(\delta^{(L)}, \delta^{(U)})$. Therefore, practical equivalence has been established and the result is consistent with FDA conclusion that "the use of rbGH in diary cattle presents no increased health risk to consumer" [10]

6 Conclusion

We have constructed a confidence set-based stepwise procedure under homogeneity of variances across dosages based on ratio of means differences in toxicity assessment. Our procedure controls the FWER at or below a pre-assigned nominal level α . This is a central requirement by FDA for statistical methodologies for toxicological evaluations. Simulation studies showed that the FWER is well controlled in the case of homogeniety of variances but quite liberal when this assumption is violated. Also , our simulation study indicated that increases in the samples sizes and the ratio of means differences implies greater power. Our procedure can be employed in non-inferiority clinical trials for toxicity and efficacy investigation.

Data Availability

The data used to support the findings of this study is available in [10]

References

- [1] Adjabui J. M., Howard K. N., Akamba M. Biostatistical assessment of mutagenicity studies; A stepwise confidence procedure. *Journal of probability and statistics*, Hindawi 2019: 1-8
- [2] Cao, L., Tao, J., Shi N. Liu W. . A stepwise confidence interval procedure under unknown variance based on an asymmetric loss function for toxicological evaluation *Aust. N. Z. J. Stat.* 2015 57(1), 73-98.
- [3] Dilba, G., Bretz, E., Guiard, V. Simultaneous confidence sets confidence intervals for multiple ratio. *Journal of statistical planning and inferences* 136(8):2640-2658 2006
- [4] Fieller, E. C. Some problems in interval estimation *Journal of Royal Statistical Society series B* 1954 16:175-185.
- [5] Finner H, Rothers M. On false discovering error rate and expected type error *Biometrical J.* 2001; 43,985-1005
- [6] Finner, H., Strassburger, K. The partitioning principle : a powerful tool in multiple decision theory. *Annals of statistics*, 2002 1194-1213.
- [7] Hasler M. Multiple comparisons to both negative and positive control *Pharmaceutical Statistics* 2012 11:74-81 .
- [8] Hsu, J . C , and Berger , R . L. Stepwise confidence interval without multiplicity adjustment for dose response and toxicity studies. *Journal of America Statistics Association* 1999 94: 468 -482
- [9] Haung, Y., Jason C. H. Hochber's step-up method: Cutting corner's off Holm's stepdown methods *Biometrika* 2007, 94(4);965-975
- [10] Juskevic, J. C. and Guyer, C. G. Bovine growth hormone: Human food safety evaluation. *Science* 249, 875-884 1990
- [11] Pigeot, I. Shaffer, J., Rohmel J., and Hauschke D. Assessing non-inferiority of new treatment in a three-arm clinical trial including a placebo *Statist Med.* 2003 22, pp. 883-889.
- [12] Tao, J., Tang, M., Shi N. Stepwise confidence interval procedure based on an asymmetric loss functio with application to toxicological evaluation *Aust. N. Z stat.* 52(1), 93-107 2010
- [13] Tao, J., Shi, N., Guo, J. Gao W. Stepwise Procedure for identification of minimum effective dose with Unknown variances *Statistics and Probability Letters* 57: 121-131 2002
- [14] Shirley, E. The comparisons of treatments with control groups means in toxicogical studies. *Appl Stat* 28 144-151 1979
- [15] Stein, C. A two sample test for a linear hypothesis whose power is independent of the variances *Ann. Math. Statist.* 16, 243-258. 1945
- [16] The significance of difference between two means when the population variances unequal. *Biometrika*, 1938 29, 350-362.