Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/110101

The Relative Potency of Two Drugs Using the Confidence Interval for Ratio of Means of Two Normal Populations with Unknown Coefficients of Variation

Warisa Thangjai [1] *and Sa-Aat Niwitpong* [2,*]

[1]Department of Statistics, Faculty of Science, Ramkhamhaeng University, Bangkok 10240, Thailand
[2]Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok 10800, Thailand

Received: 16 Jun. 2020, Revised: 12 Aug. 2020, Accepted: 3 Sep. 2020
Published online: 1 Jan. 2022

Abstract: In the paper, the relative potency of two drugs using confidence intervals for the ratio of two means of normal distributions with unknown coefficients of variation is considered. The new confidence intervals were constructed using the large sample approach, the method of variance estimates recovery (MOVER) approach, and then compared with the existing approach: the generalized confidence interval (GCI) approach of Lee and Lin [1]. A simulation study showed that the large sample approach can be used to estimate the confidence interval for the ratio of normal means with unknown coefficients of variation when the value of σ_Y/σ_X is small otherwise the GCI approach is recommended when the value of σ_Y/σ_X is large. Applications to drug testing and carboxyhemoglobin test are included.

Keywords: mean, coefficient of variation, normal distribution, large sample approach, MOVER approach

1 Introduction

The ratio of two means has been used in many areas. For example, the relative potency of two drugs or treatments is considered in biological assay and bioequivalence problems. The range of the ratio of means is the relative potency of the test drug to the reference drug. The mean responses of the test drug and the reference drug are used to ensure that the two drugs are equally effective. In practice, the average bioequivalence criterion requires the ratio of two means μ_T/μ_R to be close to 1, where μ_T is the mean response for the test drug and μ_R is the mean response of the reference drug. For more details about the ratio of two means in biological assay and bioequivalence problems, see the research papers of Fieller [2], Finney [3], Chow and Liu [4], Berger and Hsu [5], and Lee and Lin [1]. In statistics, Fieller's theorem gives a confidence region for the ratio of two means. Several researchers have been studied confidence interval estimation for the ratio of two means. Fieller [2,6] constructed a confidence interval for the ratio of the means of two normal distributions. Malley [7] presented the simultaneous confidence intervals for the ratios of means of normal distributions. Cox [8] provided the interval estimates for the ratio of two means of normal distributions with variances related to means. Lee and Lin [1] proposed the generalized confidence interval (GCI) approach to construct the confidence interval for the ratio of means of two normal distributions. Lidong et al. [9] provided the confidence intervals for the ratio of mean of two log-normal distributions. Hannig et al. [10] proposed the simultaneous fiducial generalized confidence intervals for the ratios of means of log-normal distributions. Sadooghi-Alvandi and Malekzadeh [11] proposed a parametric bootstrap approach to construct simultaneous confidence intervals for ratios of means of several log-normal distributions. Abdel-Karim [12] constructed the simultaneous confidence intervals for ratios of means of log-normal distributions based on two approaches using a two-step method of variance estimates recovery (MOVER) approach.

It is well known that the maximum likelihood estimator of the normal population mean μ is the normal sample mean $\bar{x}$. Moreover, $\bar{x}$ is the uniformly minimum variance unbiased estimator of μ. The problem of estimating the mean μ of normal distribution is of interest. Searls [13] provided the minimum mean squared error estimator for estimating the normal mean

* Corresponding author e-mail: sa-aat.n@sci.kmutnb.ac.th

μ with known the coefficient of variation σ/μ. In practice, the coefficient of variation is unknown. Hence, the minimum mean squared error estimator of Searls [13] is not very useful. Srivastava [14] proposed an estimator for mean of normal distribution using the coefficient of variation. Srivastava [15] showed that the estimator of Srivastava [14] is more efficient than the usual estimator $\bar{x}$. The normal mean with unknown coefficient of variation estimator is widely used to estimate the normal population mean μ; see, e.g., Sahai [16], Sahai and Acharya [17], Sodanin et al. [18], and Thangjai et al. [19].

To our knowledge, no paper exists for the confidence interval for the ratio of means of two normal distributions with unknown coefficients of variation. Therefore, this paper extends the paper works of Lee and Lin [1] and Thangjai et al. [19] to construct confidence intervals for the ratio of means of two normal distributions with unknown coefficients of variation. The confidence intervals were constructed based on the large sample approach and the MOVER approach and compared with the existing approach: the GCI approach of Lee and Lin [1]. The generalized confidence interval was introduced by Weerahandi [20] and has successfully used to construct the confidence interval; see, e.g., Wongkhao et al. [21] and Thangjai et al. [19]. The large sample confidence interval is used using the concept of the central limit theorem and the quantile of the normal distribution. The MOVER confidence interval was proposed by Zou and Donner [22], Zou et al. [23], and Donner and Zou [24]. The MOVER approach has been used to construct confidence interval for the parameter by many researches; for example, see Suwan and Niwitpong [25], Wongkhao et al. [21], Niwitpong [26], Niwitpong and Wongkhao [27], Sangnawakij et al. [28], and Sangnawakij and Niwitpong [29].

2 Confidence intervals for the ratio of means of two normal populations with unknown coefficients of variation

Suppose that random variable $X = (X_1, X_2, \ldots, X_n)$ follows a normal distribution with mean μ_X and variance σ_X^2. And suppose that random variable $Y = (Y_1, Y_2, \ldots, Y_m)$ follows a normal distribution with mean μ_Y and variance σ_Y^2. Let X and Y be two independent random variables. The ratio of means of normal distributions is μ_Y/μ_X. Since θ in Equation (1) is the parameter of interest

$$\theta = \frac{\mu_Y}{\mu_X}. \tag{1}$$

The estimator of the parameter of interest is defined by (See, Searls [13])

$$\hat{\theta} = \frac{\hat{\mu}_Y}{\hat{\mu}_X}, \tag{2}$$

where $\hat{\mu}_X = \frac{n\bar{X}}{n+(S_X^2/\bar{X}^2)}$, $\hat{\mu}_Y = \frac{m\bar{Y}}{m+(S_Y^2/\bar{Y}^2)}$, and the sample means and sample variances for X and Y are defined as $\bar{X}$, $\bar{Y}$, S_X^2, and S_Y^2, respectively.

Theorem 1. *Let $X = (X_1, X_2, \ldots, X_n)$ and $Y = (Y_1, Y_2, \ldots, Y_m)$ be independent random variables from $N(\mu_X, \sigma_X^2)$ and $N(\mu_Y, \sigma_Y^2)$, respectively. Let $\bar{X}$, $\bar{Y}$, S_X^2, and S_Y^2 be the sample means and the sample variances of X and Y, respectively. Also, let $\hat{\theta} = \hat{\mu}_Y/\hat{\mu}_X$, where $\hat{\mu}_X = \frac{n\bar{X}}{n+(S_X^2/\bar{X}^2)}$ and $\hat{\mu}_Y = \frac{m\bar{Y}}{m+(S_Y^2/\bar{Y}^2)}$. Therefore, the mean and variance of $\hat{\theta}$ are*

$$E(\hat{\theta}) \cong \left(\frac{E(\hat{\mu}_Y)}{E(\hat{\mu}_X)}\right) \cdot \left(1 + \frac{Var(\hat{\mu}_X)}{(E(\hat{\mu}_X))^2}\right) \tag{3}$$

and

$$Var(\hat{\theta}) \cong \left(\frac{E(\hat{\mu}_Y)}{E(\hat{\mu}_X)}\right)^2 \cdot \left(\frac{Var(\hat{\mu}_Y)}{(E(\hat{\mu}_Y))^2} + \frac{Var(\hat{\mu}_X)}{(E(\hat{\mu}_X))^2}\right), \tag{4}$$

where

$$E(\hat{\mu}_X) = \left(\frac{\mu_X}{1 + \left(\frac{\sigma_X^2}{n\mu_X^2 + \sigma_X^2}\right) \cdot \left(1 + \frac{2\sigma_X^4 + 4n\mu_X^2\sigma_X^2}{(n\mu_X^2 + \sigma_X^2)^2}\right)}\right) \cdot \left(1 + \frac{\left(\frac{n\sigma_X^2}{n\mu_X^2 + \sigma_X^2}\right)^2 \cdot \left(\frac{2}{n} + \frac{2\sigma_X^4 + 4n\mu_X^2\sigma_X^2}{(n\mu_X^2 + \sigma_X^2)^2}\right)}{\left(n + \left(\frac{n\sigma_X^2}{n\mu_X^2 + \sigma_X^2}\right) \cdot \left(1 + \frac{2\sigma_X^4 + 4n\mu_X^2\sigma_X^2}{(n\mu_X^2 + \sigma_X^2)^2}\right)\right)^2}\right)$$

$$E(\hat{\mu}_Y) = \left(\frac{\mu_Y}{1 + \left(\frac{\sigma_Y^2}{m\mu_Y^2 + \sigma_Y^2} \right) \cdot \left(1 + \frac{2\sigma_Y^4 + 4m\mu_Y^2\sigma_Y^2}{(m\mu_Y^2 + \sigma_Y^2)^2} \right)} \right) \cdot \left(1 + \frac{\left(\frac{m\sigma_Y^2}{m\mu_Y^2 + \sigma_Y^2} \right)^2 \cdot \left(\frac{2}{m} + \frac{2\sigma_Y^4 + 4m\mu_Y^2\sigma_Y^2}{(m\mu_Y^2 + \sigma_Y^2)^2} \right)}{\left(m + \left(\frac{m\sigma_Y^2}{m\mu_Y^2 + \sigma_Y^2} \right) \cdot \left(1 + \frac{2\sigma_Y^4 + 4m\mu_Y^2\sigma_Y^2}{(m\mu_Y^2 + \sigma_Y^2)^2} \right) \right)^2} \right)$$

$$Var(\hat{\mu}_X) = \left(\frac{\mu_X}{1 + \left(\frac{\sigma_X^2}{n\mu_X^2 + \sigma_X^2} \right) \cdot \left(1 + \frac{2\sigma_X^4 + 4n\mu_X^2\sigma_X^2}{(n\mu_X^2 + \sigma_X^2)^2} \right)} \right)^2 \cdot \left(\frac{\sigma_X^2}{n\mu_X^2} + \frac{\left(\frac{n\sigma_X^2}{n\mu_X^2 + \sigma_X^2} \right)^2 \cdot \left(\frac{2}{n} + \frac{2\sigma_X^4 + 4n\mu_X^2\sigma_X^2}{(n\mu_X^2 + \sigma_X^2)^2} \right)}{\left(n + \left(\frac{n\sigma_X^2}{n\mu_X^2 + \sigma_X^2} \right) \cdot \left(1 + \frac{2\sigma_X^4 + 4n\mu_X^2\sigma_X^2}{(n\mu_X^2 + \sigma_X^2)^2} \right) \right)^2} \right)$$

and

$$Var(\hat{\mu}_Y) = \left(\frac{\mu_Y}{1 + \left(\frac{\sigma_Y^2}{m\mu_Y^2 + \sigma_Y^2} \right) \cdot \left(1 + \frac{2\sigma_Y^4 + 4m\mu_Y^2\sigma_Y^2}{(m\mu_Y^2 + \sigma_Y^2)^2} \right)} \right)^2 \cdot \left(\frac{\sigma_Y^2}{m\mu_Y^2} + \frac{\left(\frac{m\sigma_Y^2}{m\mu_Y^2 + \sigma_Y^2} \right)^2 \cdot \left(\frac{2}{m} + \frac{2\sigma_Y^4 + 4m\mu_Y^2\sigma_Y^2}{(m\mu_Y^2 + \sigma_Y^2)^2} \right)}{\left(m + \left(\frac{m\sigma_Y^2}{m\mu_Y^2 + \sigma_Y^2} \right) \cdot \left(1 + \frac{2\sigma_Y^4 + 4m\mu_Y^2\sigma_Y^2}{(m\mu_Y^2 + \sigma_Y^2)^2} \right) \right)^2} \right).$$

*Proof.*Let $\hat{\mu}_X = n\bar{X}/(n + (S_X^2/\bar{X}^2))$ and $\hat{\mu}_Y = m\bar{Y}/(m + (S_Y^2/\bar{Y}^2))$. Following Thangjai et al. [19], the mean and variance of $\hat{\mu}_X$ are

$$E(\hat{\mu}_X) = \left(\frac{\mu_X}{1 + \left(\frac{\sigma_X^2}{n\mu_X^2 + \sigma_X^2} \right) \cdot \left(1 + \frac{2\sigma_X^4 + 4n\mu_X^2\sigma_X^2}{(n\mu_X^2 + \sigma_X^2)^2} \right)} \right) \cdot \left(1 + \frac{\left(\frac{n\sigma_X^2}{n\mu_X^2 + \sigma_X^2} \right)^2 \cdot \left(\frac{2}{n} + \frac{2\sigma_X^4 + 4n\mu_X^2\sigma_X^2}{(n\mu_X^2 + \sigma_X^2)^2} \right)}{\left(n + \left(\frac{n\sigma_X^2}{n\mu_X^2 + \sigma_X^2} \right) \cdot \left(1 + \frac{2\sigma_X^4 + 4n\mu_X^2\sigma_X^2}{(n\mu_X^2 + \sigma_X^2)^2} \right) \right)^2} \right)$$

and

$$Var(\hat{\mu}_X) = \left(\frac{\mu_X}{1 + \left(\frac{\sigma_X^2}{n\mu_X^2 + \sigma_X^2} \right) \cdot \left(1 + \frac{2\sigma_X^4 + 4n\mu_X^2\sigma_X^2}{(n\mu_X^2 + \sigma_X^2)^2} \right)} \right)^2$$
$$\cdot \left(\frac{\sigma_X^2}{n\mu_X^2} + \frac{\left(\frac{n\sigma_X^2}{n\mu_X^2 + \sigma_X^2} \right)^2 \cdot \left(\frac{2}{n} + \frac{2\sigma_X^4 + 4n\mu_X^2\sigma_X^2}{(n\mu_X^2 + \sigma_X^2)^2} \right)}{\left(n + \left(\frac{n\sigma_X^2}{n\mu_X^2 + \sigma_X^2} \right) \cdot \left(1 + \frac{2\sigma_X^4 + 4n\mu_X^2\sigma_X^2}{(n\mu_X^2 + \sigma_X^2)^2} \right) \right)^2} \right).$$

Similarly, the mean and variance of $\hat{\theta}_Y$ are

$$E(\hat{\mu}_Y) = \left(\frac{\mu_Y}{1 + \left(\frac{\sigma_Y^2}{m\mu_Y^2 + \sigma_Y^2} \right) \cdot \left(1 + \frac{2\sigma_Y^4 + 4m\mu_Y^2\sigma_Y^2}{(m\mu_Y^2 + \sigma_Y^2)^2} \right)} \right)$$
$$\cdot \left(1 + \frac{\left(\frac{m\sigma_Y^2}{m\mu_Y^2 + \sigma_Y^2} \right)^2 \cdot \left(\frac{2}{m} + \frac{2\sigma_Y^4 + 4m\mu_Y^2\sigma_Y^2}{(m\mu_Y^2 + \sigma_Y^2)^2} \right)}{\left(m + \left(\frac{m\sigma_Y^2}{m\mu_Y^2 + \sigma_Y^2} \right) \cdot \left(1 + \frac{2\sigma_Y^4 + 4m\mu_Y^2\sigma_Y^2}{(m\mu_Y^2 + \sigma_Y^2)^2} \right) \right)^2} \right)$$

and

$$Var(\hat{\mu}_Y) = \left(\frac{\mu_Y}{1 + \left(\frac{\sigma_Y^2}{m\mu_Y^2 + \sigma_Y^2} \right) \cdot \left(1 + \frac{2\sigma_Y^4 + 4m\mu_Y^2\sigma_Y^2}{(m\mu_Y^2 + \sigma_Y^2)^2} \right)} \right)^2$$
$$\cdot \left(\frac{\sigma_Y^2}{m\mu_Y^2} + \frac{\left(\frac{m\sigma_Y^2}{m\mu_Y^2 + \sigma_Y^2} \right)^2 \cdot \left(\frac{2}{m} + \frac{2\sigma_Y^4 + 4m\mu_Y^2\sigma_Y^2}{(m\mu_Y^2 + \sigma_Y^2)^2} \right)}{\left(m + \left(\frac{m\sigma_Y^2}{m\mu_Y^2 + \sigma_Y^2} \right) \cdot \left(1 + \frac{2\sigma_Y^4 + 4m\mu_Y^2\sigma_Y^2}{(m\mu_Y^2 + \sigma_Y^2)^2} \right) \right)^2} \right).$$

Bain and Engelhardt [30] showed that the sample mean and sample variance are independent when the data come from normal distribution. Therefore, $\bar{X}$ and S_X^2 are independent and $\bar{Y}$ and S_Y^2 are also independent. According to Blumenfeld [31]and Saelee et al. [32], the mean and variance of $\hat{\theta} = \hat{\mu}_Y / \hat{\mu}_X$ are

$$E(\hat{\theta}) \cong \left(\frac{E(\hat{\mu}_Y)}{E(\hat{\mu}_X)} \right) \cdot \left(1 + \frac{Var(\hat{\mu}_X)}{(E(\hat{\mu}_X))^2} \right) \text{ and } Var(\hat{\theta}) \cong \left(\frac{E(\hat{\mu}_Y)}{E(\hat{\mu}_X)} \right)^2 \cdot \left(\frac{Var(\hat{\mu}_Y)}{(E(\hat{\mu}_Y))^2} + \frac{Var(\hat{\mu}_X)}{(E(\hat{\mu}_X))^2} \right),$$

where $E(\hat{\mu}_X)$, $E(\hat{\mu}_Y)$, $Var(\hat{\mu}_X)$, and $Var(\hat{\mu}_Y)$ are defined above.
Hence, Theorem 1 is proved.

2.1 Large sample confidence interval

The pivotal statistic based on the normal approximation is defined by

$$Z = \frac{\hat{\theta} - E(\hat{\theta})}{\sqrt{Var(\hat{\theta})}} = \frac{\hat{\theta} - \theta}{\sqrt{Var(\hat{\theta})}}. \qquad (5)$$

Therefore, the $100(1-\alpha)\%$ two-sided confidence interval for the ratio of normal means with unknown coefficients of variation based on the large sample approach is

$$CI_{LS} = [L_{LS}, U_{LS}] = \left[\hat{\theta} + z_{\alpha/2}\sqrt{\hat{Var}(\hat{\theta})}, \hat{\theta} + z_{1-\alpha/2}\sqrt{\hat{Var}(\hat{\theta})}\right], \tag{6}$$

where $\hat{\theta}$ is defined in Equation (2), $\hat{Var}(\hat{\theta})$ denotes the estimator of $Var(\hat{\theta})$ which is defined in Equation (4) with μ_X, μ_Y, σ_X^2, and σ_Y^2 replaced by $\bar{X}$, $\bar{Y}$, S_X^2, and S_Y^2, respectively, and $z_{\alpha/2}$ and $z_{1-\alpha/2}$ denote the $(\alpha/2)$-th and $(1-\alpha/2)$-th quantiles of the standard normal distribution, respectively; see Donner and Zou [24] and Thangjai et al. [33].

2.2 Method of variance estimates recovery confidence interval

Again, the estimators of the normal mean with unknown coefficient of variation of X and Y are

$$\hat{\mu}_X = \frac{n\bar{X}}{n + (S_X^2/\bar{X}^2)} \quad \text{and} \quad \hat{\mu}_Y = \frac{m\bar{Y}}{m + (S_Y^2/\bar{Y}^2)}. \tag{7}$$

The variances of $\hat{\mu}_X$ and $\hat{\mu}_Y$ are

$$Var(\hat{\mu}_X) = \left(\frac{\mu_X}{1 + \left(\frac{\sigma_X^2}{n\mu_X^2 + \sigma_X^2}\right) \cdot \left(1 + \frac{2\sigma_X^4 + 4n\mu_X^2\sigma_X^2}{(n\mu_X^2 + \sigma_X^2)^2}\right)}\right)^2 \cdot \left(\frac{\sigma_X^2}{n\mu_X^2} + \frac{\left(\frac{n\sigma_X^2}{n\mu_X^2 + \sigma_X^2}\right)^2 \cdot \left(\frac{2}{n} + \frac{2\sigma_X^4 + 4n\mu_X^2\sigma_X^2}{(n\mu_X^2 + \sigma_X^2)^2}\right)}{\left(n + \left(\frac{n\sigma_X^2}{n\mu_X^2 + \sigma_X^2}\right) \cdot \left(1 + \frac{2\sigma_X^4 + 4n\mu_X^2\sigma_X^2}{(n\mu_X^2 + \sigma_X^2)^2}\right)\right)^2}\right) \tag{8}$$

and

$$Var(\hat{\mu}_Y) = \left(\frac{\mu_Y}{1 + \left(\frac{\sigma_Y^2}{m\mu_Y^2 + \sigma_Y^2}\right) \cdot \left(1 + \frac{2\sigma_Y^4 + 4m\mu_Y^2\sigma_Y^2}{(m\mu_Y^2 + \sigma_Y^2)^2}\right)}\right)^2 \cdot \left(\frac{\sigma_Y^2}{m\mu_Y^2} + \frac{\left(\frac{m\sigma_Y^2}{m\mu_Y^2 + \sigma_Y^2}\right)^2 \cdot \left(\frac{2}{m} + \frac{2\sigma_Y^4 + 4m\mu_Y^2\sigma_Y^2}{(m\mu_Y^2 + \sigma_Y^2)^2}\right)}{\left(m + \left(\frac{m\sigma_Y^2}{m\mu_Y^2 + \sigma_Y^2}\right) \cdot \left(1 + \frac{2\sigma_Y^4 + 4m\mu_Y^2\sigma_Y^2}{(m\mu_Y^2 + \sigma_Y^2)^2}\right)\right)^2}\right). \tag{9}$$

Following Thangjai et al. [19], the confidence intervals for the normal mean with unknown coefficient of variation of X and Y can be written as

$$[l_X, u_X] = [\hat{\mu}_X + z_{\alpha/2}\sqrt{\hat{Var}(\hat{\mu}_X)}, \hat{\mu}_X + z_{1-\alpha/2}\sqrt{\hat{Var}(\hat{\mu}_X)}] \tag{10}$$

and

$$[l_Y, u_Y] = [\hat{\mu}_Y + z_{\alpha/2}\sqrt{\hat{Var}(\hat{\mu}_Y)}, \hat{\mu}_Y + z_{1-\alpha/2}\sqrt{\hat{Var}(\hat{\mu}_Y)}], \tag{11}$$

where $\hat{\mu}_X$ and $\hat{\mu}_Y$ are defined in Equation (7), $\hat{Var}(\hat{\mu}_X)$ and $\hat{Var}(\hat{\mu}_Y)$ denote the estimator of $Var(\hat{\mu}_X)$ and $Var(\hat{\mu}_Y)$ which are defined in Equation (8) and Equation (9) with μ_X, μ_Y, σ_X^2, and σ_Y^2 replaced by $\bar{X}$, $\bar{Y}$, S_X^2, and S_Y^2, respectively, and $z_{\alpha/2}$ and $z_{1-\alpha/2}$ denote the $(\alpha/2)$-th and $(1-\alpha/2)$-th quantiles of the standard normal distribution.

According to Donner and Zou [24], in the case of the ratio of two means, a logarithm transformation prior is applied the use of the difference of two means. The ratio of means is $\theta = \mu_Y/\mu_X$ which is equivalent to $\mu_Y - \theta\mu_X = 0$. We assume positive value parameter $\theta > 0$ as similar steps can be taken for negative value parameter $\theta < 0$. The estimators $\hat{\mu}_X$ and $\hat{\mu}_Y$ are independent. The lower limit and the upper limit of confidence interval for $\mu_Y - \theta\mu_X$ are obtained by

$$l = \hat{\mu}_Y - \theta\hat{\mu}_X - \sqrt{(\hat{\mu}_Y - l_Y)^2 + \theta^2(u_X - \hat{\mu}_X)^2} \tag{12}$$

and

$$u = \hat{\mu}_Y - \theta\hat{\mu}_X + \sqrt{(u_Y - \hat{\mu}_Y)^2 + \theta^2(\hat{\mu}_X - l_X)^2}. \tag{13}$$

By $\mu_Y - \theta\mu_X = 0$, we must have $l = 0$ and $u = 0$, which lead to quadratic equations in θ. Solving Equation (12) and Equation (13) for θ, the lower limit and the upper limit for the ratio are

$$L_{MOVER} = \frac{\hat{\mu}_Y\hat{\mu}_X - \sqrt{(\hat{\mu}_Y\hat{\mu}_X)^2 - l_Y u_X(2\hat{\mu}_Y - l_Y)(2\hat{\mu}_X - u_X)}}{u_X(2\hat{\mu}_X - u_X)} \tag{14}$$

and

$$U_{MOVER} = \frac{\hat{\mu}_Y\hat{\mu}_X + \sqrt{(\hat{\mu}_Y\hat{\mu}_X)^2 - u_Y l_X(2\hat{\mu}_Y - u_Y)(2\hat{\mu}_X - l_X)}}{l_X(2\hat{\mu}_X - l_X)}, \tag{15}$$

where l_X and u_X are defined in Equation (10) and l_Y and u_Y are defined in Equation (11).

Therefore, the $100(1-\alpha)\%$ two-sided confidence interval for the ratio of normal means with unknown coefficients of variation based on the MOVER approach is

$$CI_{MOVER} = [L_{MOVER}, U_{MOVER}], \tag{16}$$

where L_{MOVER} and U_{MOVER} are defined in equation (14) and equation (15),respectively.

Next, the GCI approach of Lee and Lin [1] is briefly reviewed for constructing the confidence interval for the ratio of means of two normal populations. The generalized pivotal quantity for $\theta = \mu_Y/\mu_X$ is

$$R_{\theta.LL} = \frac{\bar{Y} - Z_Y S_Y/\sqrt{U_Y}}{\bar{X} - Z_X S_X/\sqrt{U_X}}, \tag{17}$$

where Z_X and Z_Y denote the standard normal distributions, U_X denotes the chi-squared distribution with $n-1$ degrees of freedom, and U_Y denotes the chi-squared distribution with $m-1$ degrees of freedom.

Therefore, the $100(1-\alpha)\%$ two-sided confidence interval for the ratio of normal means based on the GCI approach of Lee and Lin [1] is

$$CI_{LL} = [L_{LL}, U_{LL}] = \left[Q^*(\alpha/2), Q^*(1-\alpha/2)\right], \tag{18}$$

where $Q^*(\alpha/2)$ and $Q^*(1-\alpha/2)$ denote the $100(\alpha/2)$-th and $100(1-\alpha/2)$-th percentiles of $R_{\theta.LL}$, respectively.

3 Simulation studies

A Monte Carlo simulation was carried out to evaluate the coverage probabilities and average lengths of the large sample confidence interval (CI_{LS}), the MOVER confidence interval (CI_{MOVER}), and the generalized confidence interval of Lee and Lin [1] (CI_{LL}). The confidence interval was chosen when the coverage probability was greater than or equal to the nominal confidence level $(1-\alpha)$ and having the shortest average length, see Li et al. [34], Xiao et al. [35], and Malekzadeh and Kharrati-Kopaei [36].

The nominal confidence level was chosen to be 0.95. The data were generated from two independent normal distributions, $N(\mu_X, \sigma_X^2)$ and $N(\mu_Y, \sigma_Y^2)$, where the population means were $\mu_X = \mu_Y = 1.0$ and the population standard deviations were $\sigma_X = 1.0$ and $\sigma_Y = 0.3, 0.5, 0.7, 0.9, 1.0, 1.1, 1.3, 1.5, 1.7, 2.0$. Thus, a ratio of coefficients of variation, $(\sigma_Y/\mu_Y)/(\sigma_X/\mu_X)$, was reduced to σ_Y/σ_X. The sample sizes were (n,m) = (30,30), (50,50), (30,50), (100,100), and (50,100).

Coverage probability for $\theta = \mu_Y/\mu_X$ can be computed by the following steps:

Algorithm 1

– Generate $x_1, x_2, \ldots, x_n$ from $N(\mu_X, \sigma_X^2)$, and then compute $\bar{x}$ and s_X^2. And generate $y_1, y_2, \ldots, y_m$ from $N(\mu_Y, \sigma_Y^2)$, and then compute $\bar{y}$ and s_Y^2.
– Use Equation (6) to construct CI_{LS} and record whether or not the value of θ fall in corresponding confidence interval.

–Use Equation (16) to construct CI_{MOVER} and record whether or not the value of θ fall in corresponding confidence interval.

–Repeat Step 1-Step 3 a large number of times, $M = 5000$. The fraction of times that all θ are in their corresponding confidence intervals is the coverage probability.

Table 1 and Figures 1-10 presented the coverage probabilities and the average lengths of 95% two-sided confidence intervals for the ratio of means with unknown coefficients of variation. We can conclude that the coverage probabilities of all the proposed approaches are similarity and satisfactory for all sample sizes. The large sample approach is better than the GCI approach of Lee and Lin [1] in term of average length when the value of σ_Y/σ_X is small ($\sigma_Y/\sigma_X \leq 1.0$), whereas the GCI approach of Lee and Lin [1] is preferable when the value of σ_Y/σ_X is large. However, the large sample approach is easier to use than the GCI approach because the large sample approach uses the simple formula and the GCI approach is a computational approach.

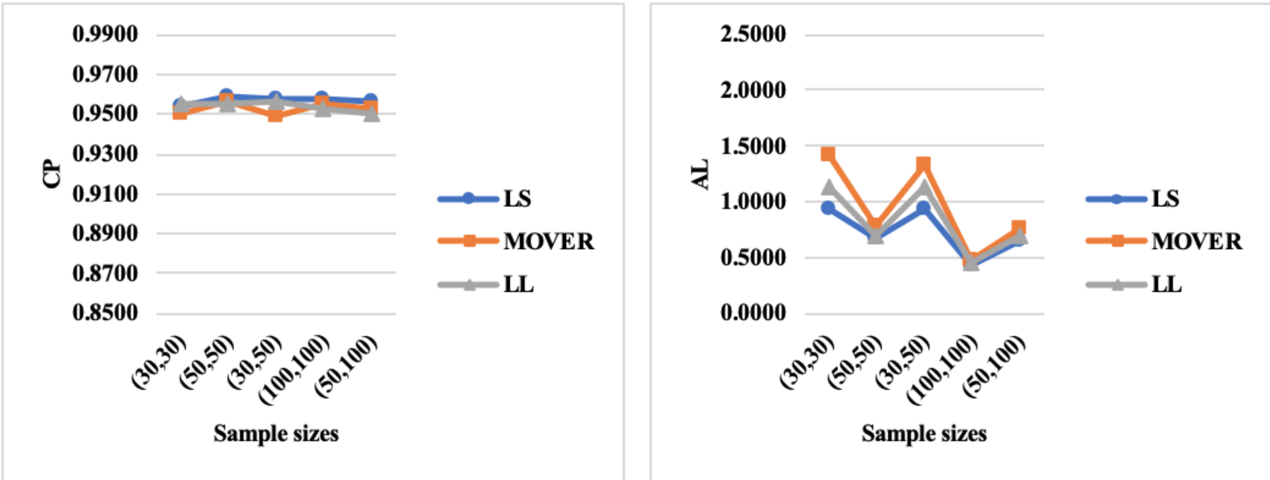

Fig. 1: The CP and AL of 95% two-sided confidence intervals for the ratio of means of normal distributions with unknown coefficients of variation for $\sigma_Y/\sigma_X = 0.3$

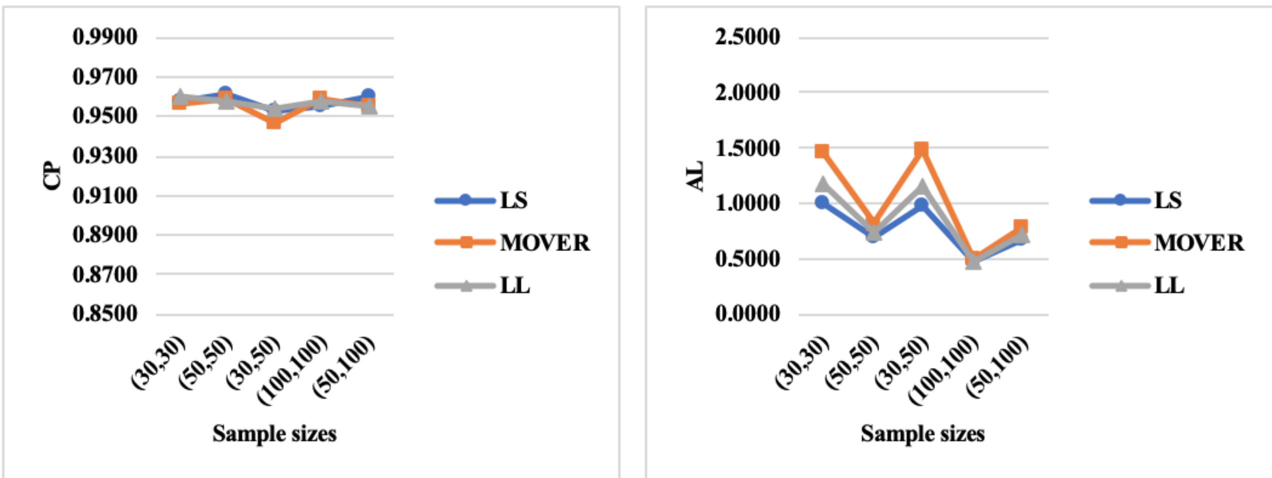

Fig. 2: The CP and AL of 95% two-sided confidence intervals for the ratio of means of normal distributions with unknown coefficients of variation for $\sigma_Y/\sigma_X = 0.5$

Table 1: The coverage probabilities (CP) and average lengths (AL) of 95% two-sided confidence intervals for the ratio of means of normal distributions with unknown coefficients of variation.

n	m	$\frac{\sigma_Y}{\sigma_X}$	CI_{LS} CP	CI_{LS} AL	CI_{MOVER} CP	CI_{MOVER} AL	CI_{LL} CP	CI_{LL} AL
30	30	0.3	0.9542	**0.9461**	0.9500	1.4214	0.9548	1.1309
		0.5	0.9574	**1.0007**	0.9564	1.4702	0.9602	1.1872
		0.7	0.9504	**1.0565**	0.9540	1.5256	0.9592	1.2446
		0.9	0.9490	**1.1558**	0.9596	1.6452	0.9600	1.3392
		1.0	0.9522	**1.2489**	0.9618	1.7913	0.9600	1.4270
		1.1	0.9456	**1.3285**	0.9556	1.8303	0.9548	1.4893
		1.3	0.9410	**1.5218**	0.9578	1.6456	0.9544	1.5932
		1.5	0.9318	1.7012	0.9550	2.2165	0.9604	**1.7236**
		1.7	0.8976	1.8210	0.9274	2.3004	0.9564	**1.8441**
		2.0	0.8680	12.7167	0.9008	2.6116	0.9552	**2.0858**
50	50	0.3	0.9594	**0.6708**	0.9566	0.7824	0.9558	0.7048
		0.5	0.9616	**0.7053**	0.9586	0.8155	0.9578	0.7434
		0.7	0.9588	**0.7625**	0.9590	0.8778	0.9576	0.8084
		0.9	0.9560	**0.8408**	0.9564	0.9579	0.9544	0.8836
		1.0	0.9560	**0.8820**	0.9592	0.9970	0.9554	0.9142
		1.1	0.9592	**0.9467**	0.9640	1.0659	0.9574	0.9611
		1.3	0.9584	1.1011	0.9666	1.2244	0.9530	**1.0524**
		1.5	0.9542	1.2752	0.9688	1.4077	0.9584	**1.1513**
		1.7	0.9414	1.4181	0.9576	1.5522	0.9536	**1.2456**
		2.0	0.9104	1.5634	0.9338	1.7103	0.9540	**1.4166**
30	50	0.3	0.9572	**0.9438**	0.9488	1.3298	0.9564	1.1302
		0.5	0.9522	**0.9739**	0.9466	1.4761	0.9536	1.1483
		0.7	0.9570	**1.0125**	0.9568	1.4388	0.9606	1.1842
		0.9	0.9582	**1.0639**	0.9560	1.4686	0.9600	1.2362
		1.0	0.9536	**1.1092**	0.9566	1.6350	0.9564	1.2796
		1.1	0.9514	**1.1704**	0.9546	1.3030	0.9516	1.3306
		1.3	0.9508	**1.2873**	0.9650	1.7437	0.9550	1.3708
		1.5	0.9580	**1.4677**	0.9734	1.8660	0.9612	1.4905
		1.7	0.9452	1.6087	0.9692	2.0292	0.9624	**1.5751**
		2.0	0.9186	1.7398	0.9500	2.0265	0.9562	**1.7207**
100	100	0.3	0.9572	**0.4384**	0.9558	0.4654	0.9526	0.4446
		0.5	0.9548	**0.4680**	0.9592	0.4954	0.9574	0.4762
		0.7	0.9524	**0.5071**	0.9498	0.5353	0.9530	0.5174
		0.9	0.9502	**0.5623**	0.9514	0.5920	0.9502	0.5718
		1.0	0.9550	**0.5904**	0.9536	0.6201	0.9508	0.5961
		1.1	0.9530	0.6249	0.9556	0.6551	0.9478	**0.6225**
		1.3	0.9652	0.7244	0.9676	0.7566	0.9538	**0.6872**
		1.5	0.9746	0.8526	0.9794	0.8872	0.9574	**0.7541**
		1.7	0.9730	0.9879	0.9798	1.0251	0.9546	**0.8232**
		2.0	0.9618	1.1510	0.9702	1.1920	0.9524	**0.9303**
50	100	0.3	0.9562	**0.6585**	0.9524	0.7682	0.9506	0.6899
		0.5	0.9606	**0.6799**	0.9554	0.7907	0.9548	0.7138
		0.7	0.9578	**0.7077**	0.9532	0.8199	0.9498	0.7440
		0.9	0.9574	**0.7474**	0.9534	0.8583	0.9534	0.7843
		1.0	0.9600	**0.7723**	0.9602	0.8847	0.9578	0.8081
		1.1	0.9570	**0.7928**	0.9602	0.9004	0.9522	0.8193
		1.3	0.9600	**0.8896**	0.9644	1.0061	0.9514	0.8910
		1.5	0.9684	0.9996	0.9736	1.1126	0.9572	**0.9444**
		1.7	0.9708	1.1202	0.9782	1.2323	0.9568	**1.0002**
		2.0	0.9588	1.2803	0.9712	1.3989	0.9494	**1.1055**

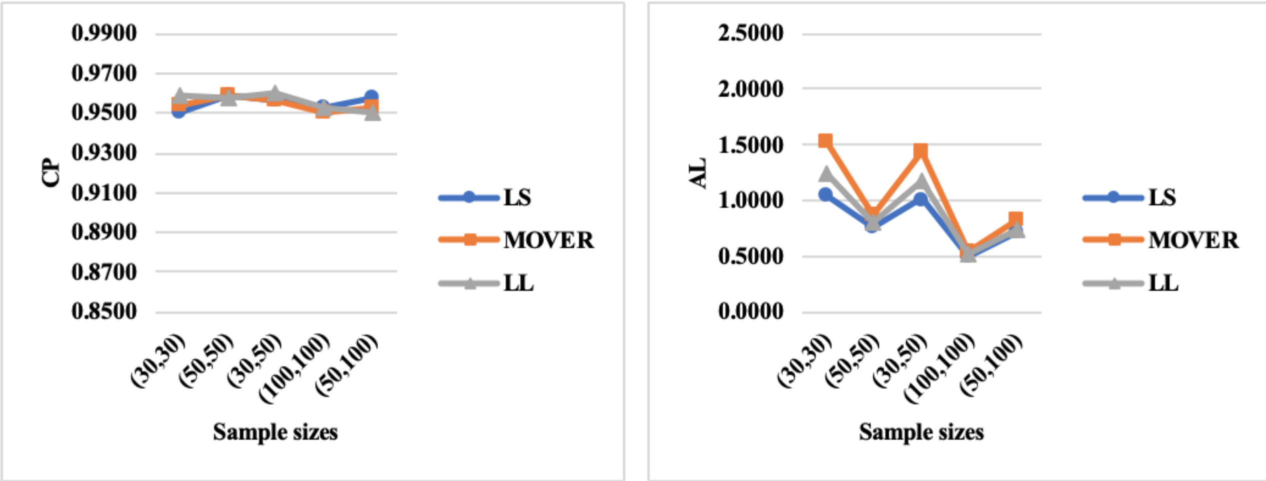

Fig. 3: The CP and AL of 95% two-sided confidence intervals for the ratio of means of normal distributions with unknown coefficients of variation for $\sigma_Y/\sigma_X = 0.7$

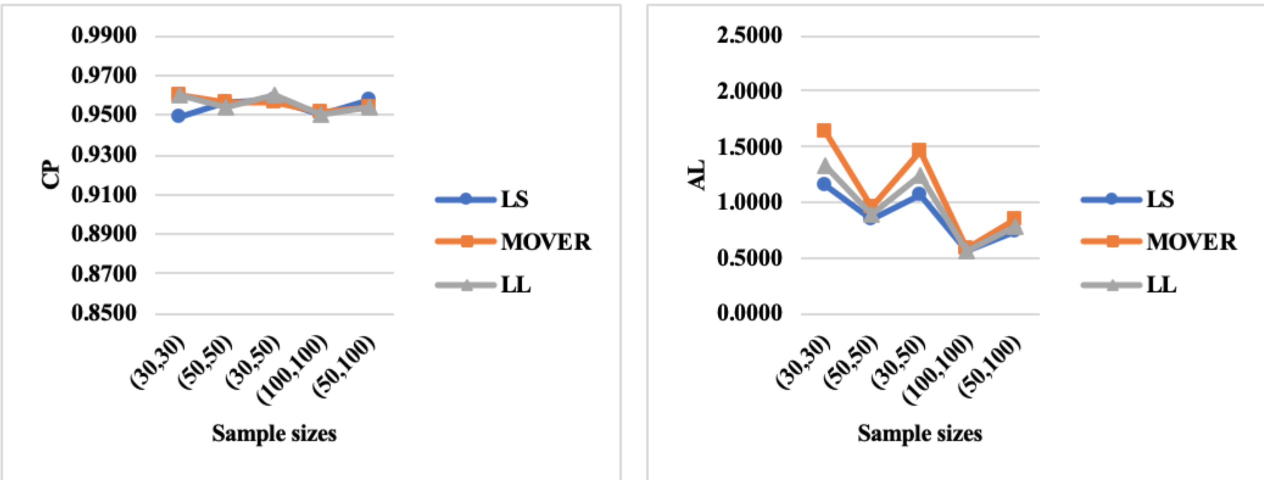

Fig. 4: The CP and AL of 95% two-sided confidence intervals for the ratio of means of normal distributions with unknown coefficients of variation for $\sigma_Y/\sigma_X = 0.9$

4 Empirical applications

In this section, two examples are considered to illustrate the proposed approaches.

*Example 1.*Let $x = (x_1, x_2, \dots, x_n)$ and $y = (y_1, y_2, \dots, y_m)$ be two independent sets of observations for the potency of a reference drug and a test drug, respectively. The data example is taken from Lee and Lin [1] for testing the potency of a reference drug and a test drug. The data was assumed to construct confidence intervals for the ratio of two means. Table 2 shows the statistics of the reference drug and test drug data. The data can be approximated by the normal distribution. The ratio of two means was $\bar{y}/\bar{x} = 1.3684$. The confidence intervals for the ratio of two means with unknown coefficients of variation was presented in Table 3. From Table 3, the large sample approach is better than the other approaches in term of the interval length. The results support the simulation results when the value of σ_Y/σ_X is small.

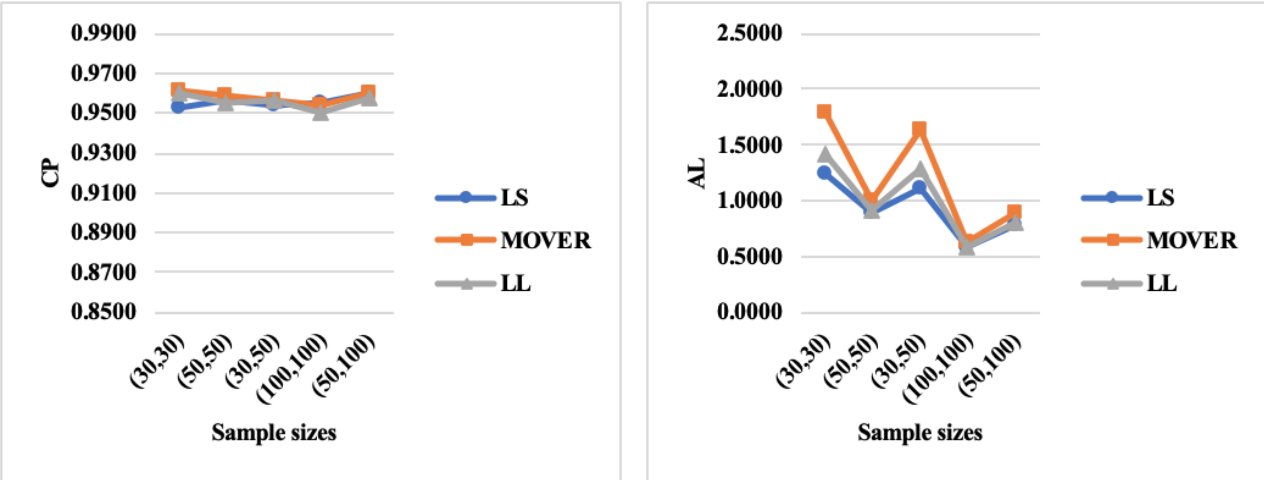

Fig. 5: The CP and AL of 95% two-sided confidence intervals for the ratio of means of normal distributions with unknown coefficients of variation for $\sigma_Y/\sigma_X = 1.0$

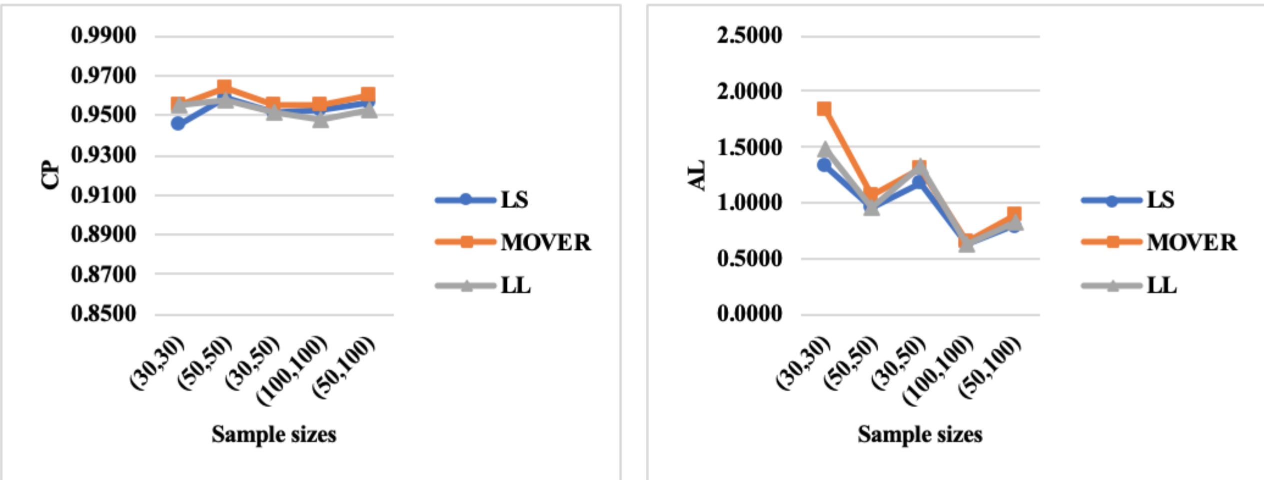

Fig. 6: The CP and AL of 95% two-sided confidence intervals for the ratio of means of normal distributions with unknown coefficients of variation for $\sigma_Y/\sigma_X = 1.1$

Table 2: Sample statistics of the reference drug and test drug data.

Statistics	Reference drug	Test drug
Sample size	50	50
Sample mean	0.3800	0.5200
Sample variance	0.2309	0.2446
Coefficient of variation	1.2645	0.9511

*Example 2.*A real data example provided by Jarvis et al. [37], Pagano and Gauvreau [38], and Lee and Lin [1] is used. The data represents carboxyhemoglobin levels for nonsmokers and cigarette smokers. Table 4 shows the statistics of the nonsmokers and cigarette smokers. The data is fitted by normal distribution. The ratio of two means was $\bar{y}/\bar{x} = 3.1538$. The 95% two-sided confidence intervals for the ratio of two means with unknown coefficients of variation were evaluated. From Table 5, it can be seen that the length of CI_{LL} are shorter than those of CI_{LS} and CI_{MOVER}. Hence, the GCI approach

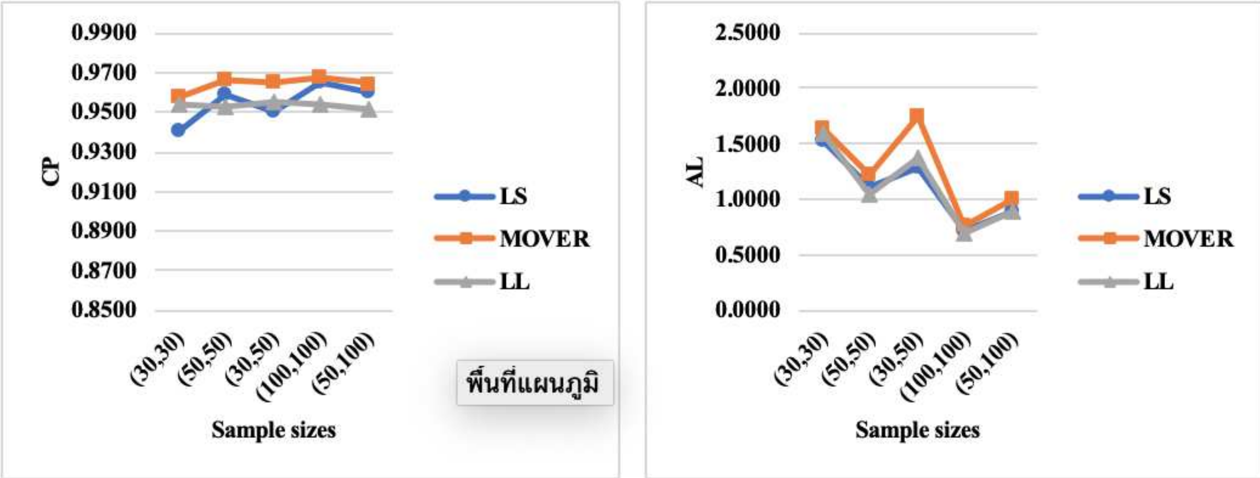

Fig. 7: The CP and AL of 95% two-sided confidence intervals for the ratio of means of normal distributions with unknown coefficients of variation for $\sigma_Y/\sigma_X = 1.3$

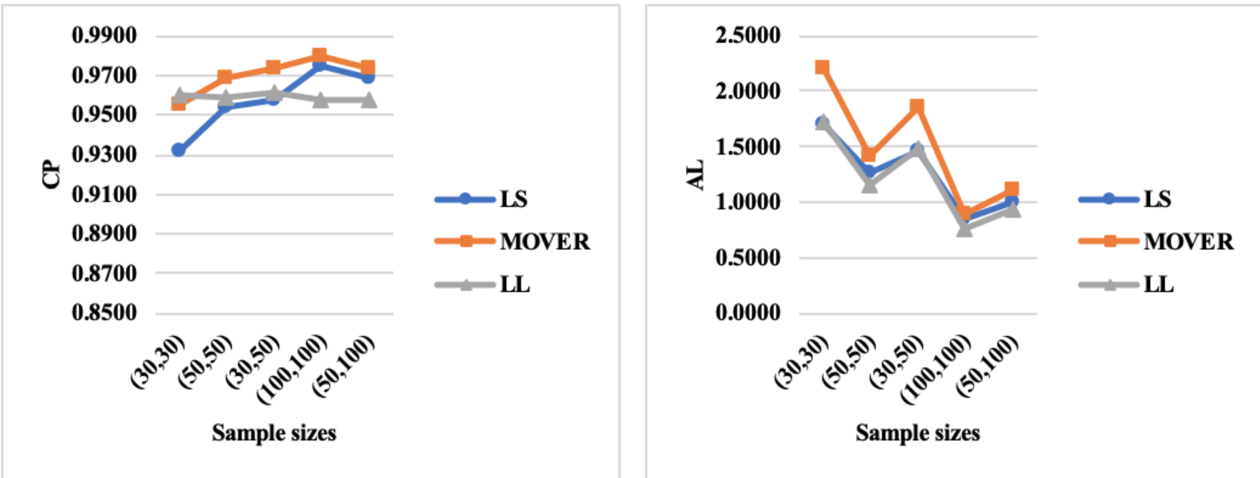

Fig. 8: The CP and AL of 95% two-sided confidence intervals for the ratio of means of normal distributions with unknown coefficients of variation for $\sigma_Y/\sigma_X = 1.5$

of Lee and Lin [1] performs well in term of the interval length. The results confirm the simulation results in the previous section when the value of σ_Y/σ_X is large.

5 Discussion and conclusions

This paper is extension of previous works of Lee and Lin [1] and Thangjai et al. [19]. Lee and Lin [1] proposed the confidence intervals for the ratio of means of normal distributions using the GCI approach. The results of Lee and Lin [1] showed that the GCI approach is valuable, especially when the two variances are quite different.

For all sample sizes, all the proposed confidence intervals provide the similar coverage probabilities. The generalized confidence interval yields the shortest average lengths when the value of σ_Y/σ_X is large. However, the large sample approach is better than the GCI approach of Lee and Lin [1] in term of average length when the value of σ_Y/σ_X is small.

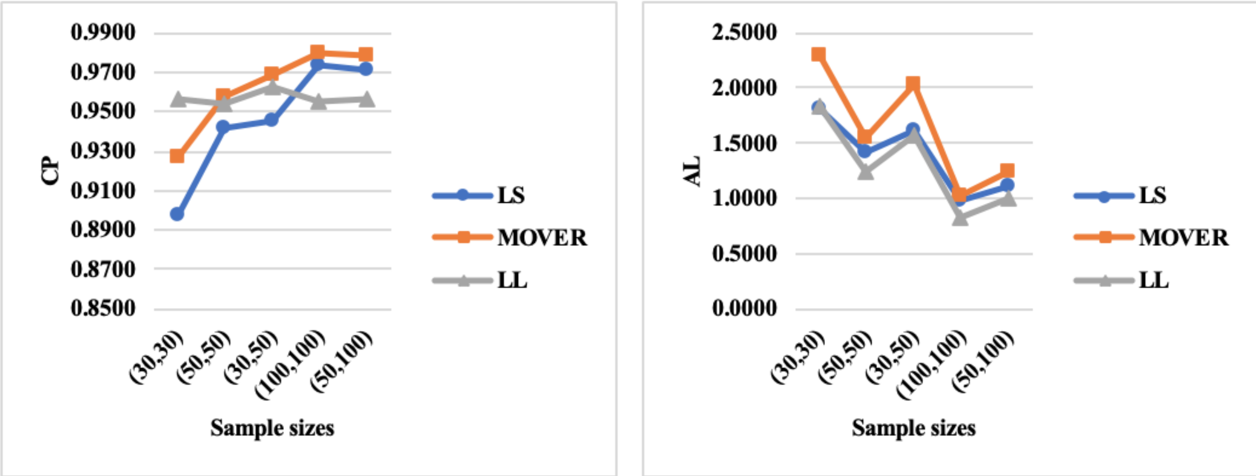

Fig. 9: The CP and AL of 95% two-sided confidence intervals for the ratio of means of normal distributions with unknown coefficients of variation for $\sigma_Y/\sigma_X = 1.7$

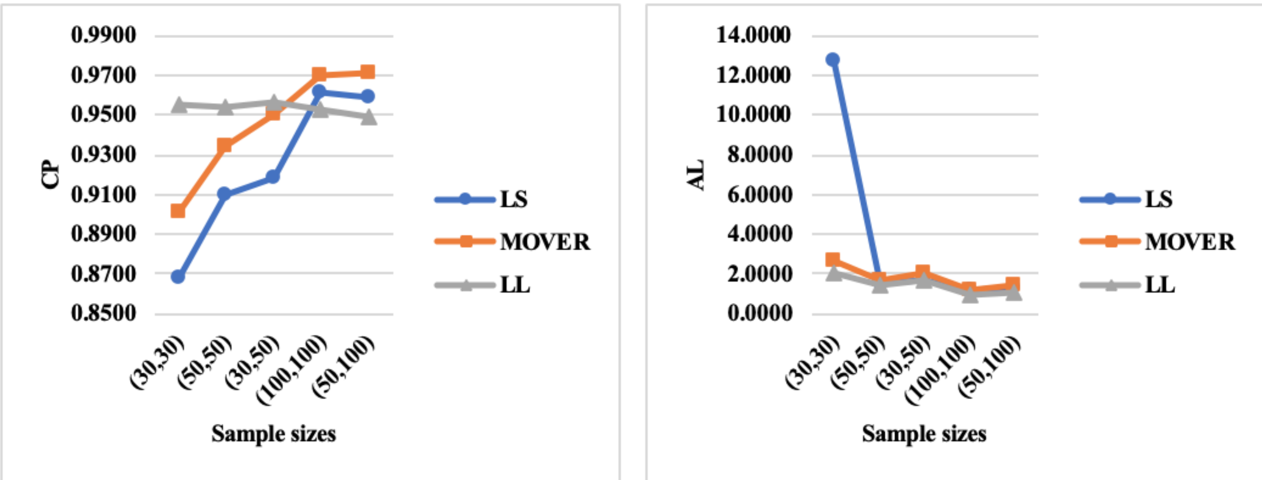

Fig. 10: The CP and AL of 95% two-sided confidence intervals for the ratio of means of normal distributions with unknown coefficients of variation for $\sigma_Y/\sigma_X = 2.0$

Table 3: The 95% two-sided confidence intervals for the ratio of two means with unknown coefficients of variation using the potency of reference drug and test drug.

Approach	Confidence interval		
	Lower limit	Upper limit	Interval length
LS	0.7597	2.0144	1.2547
MOVER	0.8925	2.2954	1.4029
LL	0.8871	2.3043	1.4172

Therefore, the large sample approach is chosen when the value of σ_Y/σ_X is small ($\sigma_Y/\sigma_X \leq 1.0$) otherwise the GCI approach of Lee and Lin [1] is preferable when the value of σ_Y/σ_X is large.

It is noted that the large sample approach is based on the Central Limit Theorem. Therefore, the large sample approach is satisfactory when the sample sizes are moderate sizes and large sizes ($(n,m) \geq (30,30)$). Furthermore, the coverage probability of the GCI approach of Lee and Lin [1] do not depend on the value of σ_Y/σ_X. From the simulation results,

Table 4: Sample statistics of the nonsmokers and cigarette smokers.

Statistics	Nonsmokers	Cigarette smokers
Sample size	121	75
Sample mean	1.3000	4.1000
Sample variance	1.7040	4.0540
Coefficient of variation	1.0041	0.4911

Table 5: The 95% two-sided confidence intervals for the ratio of two means with unknown coefficients of variation using carboxyhemoglobin levels for the nonsmokers and cigarette smokers.

Approach	Confidence interval		
	Lower limit	Upper limit	Interval length
LS	1.1717	5.1681	3.9964
MOVER	1.2468	5.3778	4.1310
LL	2.5630	3.9108	1.3478

which were not shown here, the GCI approach of Lee and Lin [1] is a preferable method for small sample sizes, i.e. $(n,m) < (30,30)$.

Funding

This research was funded by King Mongkut's University of Technology North Bangkok. Grant number: KMUTNB-FF-65-22.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] J.C. Lee and S.H. Lin, Generalized confidence intervals for the ratio of means of two normal populations, *J. Stat. Plan. Infer.*, **123**, 49-60 (2004).
[2] E.C. Fieller, A fundamental formula in the statistics of biological assay, and some applications, *Q. J. Pharm. Pharmacol.*, **17**, 117-123 (1944).
[3] D.J. Finney, *Statistical Method in Biological Assay*. Charles Griffin and Co., Ltd., London UK, (1978).
[4] S.C. Chow and J.P. Liu, *Design and Analysis of Bioavailability and Bioequivalence Studies*. Chapman and Hall/CRC, New York US, (1992).
[5] R.L. Berger and J.C. Hsu, Bioequivalence trials, intersection union tests and equivalence confidence sets, *Stat. Sci.*, **11**, 283-319 (1996).
[6] E.C. Fieller, Some problems in interval estimation, *J. Royal Stat. Soc. Ser. B.*, **16**, 175-185 (1954).
[7] J.D. Malley, Simultaneous confidence intervals for ratios of normal means. *J. Am. Stat. Assoc.*, **77**, 170-176 (1982).
[8] C.P. Cox, Interval estimates for the ratio of the means of two normal populations with variances related to means, Biometrics, **41**, 261-265 (1985).
[9] E. Lidong, J. Hannig and H.K. Iyer, *Fiducial Generalized Confidence Intervals for Functions of Lognormal Means and Variances*. Colorado State University, Fort Collins US, (2005).
[10] J. Hanning, E. Lidong, A. Abdel-Karim and H. Iyer, Simultaneous fiducial generalized confidence intervals for ratios of means of lognormal distributions, *Austrian J. Stat.*, **35**, 261-269 (2006).
[11] S.M. Sadooghi-Alvandi and A. Malekzadeh, A. Simultaneous confidence intervals for ratios of means of several lognormal distributions: A parametric bootstrap approach, *Comp. Stat. Data Anal.*, **69**, 133-140 (2014).
[12] A.H. Abdel-Karim, Construction of simultaneous confidence intervals for ratios of means of lognormal distributions, *Commun. Stat. Simul. Comp.*, **44**, 271-283 (2015).
[13] D.T. Searls, The utilization of a known coefficient of variation in the estimation procedure, *J. Am. Stat. Assoc.*, **59**, 1225-1226 (1964).
[14] V.K. Srivastava, On the use of coefficient of variation in estimating mean, *J. Indian Soc. Agric. Stat.*, **26**, 33-36 (1974).

[15] V.K. Srivastava, A note on the estimation of mean in normal population, *Metrika*, **27**, 99-102 (1980).
[16] A. Sahai. On an estimator of normal population mean and UMVU estimation of its relative efficiency, *Applied Math. Comp.*, **152**, 701-708 (2004).
[17] A. Sahai and R.M. Acharya, Iterative estimation of normal population mean using computational-statistical intelligence, *Comp. Sci. Tech.*, **4**, 500-508 (2016).
[18] S. Sodanin, S. Niwitpong and S. Niwitpong, *Generalized confidence intervals for the normal mean with unknown coefficient of variation*, in AIP Conf. Proc., 030043-1-030043-8, (2016).
[19] W. Thangjai, S. Niwitpong and S. Niwitpong, Confidence intervals for mean and difference of means of normal distributions with unknown coefficients of variation, *Mathematics*, **5**, 1-23 (2017).
[20] S. Weerahandi, Generalized confidence intervals, *J. Am. Stat. Assoc.*, **88**, 899-905 (1993).
[21] A. Wongkhao, S. Niwitpong and S. Niwitpong, Confidence intervals for the ratio of two independent coefficients of variation of normal distribution, *Far East J. Math. Sci.*, **98**, 741-757 (2015).
[22] G.Y. Zou and A. Donner, Construction of confidence limits about effect measures: A general approach, *Stat. Med.*, **27**, 1693-1702 (2008).
[23] G.Y. Zou, J. Taleban and C.Y. Huo, Confidence interval estimation for lognormal data with application to health economics, *Comp. Stat. Data Anal.*, **53**, 3755-3764 (2009).
[24] A. Donner and G.Y. Zou, Closed-form confidence intervals for function of the normal mean and standard deviation, *Stat. Methods Med. Res.*, **21**, 347-359 (2010).
[25] S. Suwan and S. Niwitpong, Estimated variance ratio confidence interval of nonnormal distributions, *Far East J. Math. Sci.*, **4**, 339-350 (2013).
[26] S. Niwitpong, Confidence intervals for the difference between coefficients of variation of normal distribution with bounded parameters, *Far East J. Math. Sci.*, **98**, 649-663 (2015).
[27] S. Niwitpong and A. Wongkhao, Confidence intervals for the difference between inverse of normal means, *Adv. Appl. Stat.*, **48**, 337-347 (2016).
[28] P. Sangnawakij, S. Niwitpong and S. Niwitpong, Confidence intervals for the ratio of coefficients of variation of the gamma distributions, *Lecture Notes in Computer Science*, **9376**, 193-203 (2016).
[29] P. Sangnawakij and S. Niwitpong, Confidence intervals for coefficients of variation in two-parameter exponential distribution, *Commun. Stat. Simul. Comp.*, **46**, 6618-6630 (2016).
[30] L.J. Bain and M. Engelhardt, *Introduction to probability and mathematical statistics*. Duxbury Press, California US, (1991).
[31] D. Blumenfeld, *Operations research calculations*. CRC Press, New York US, (2001).
[32] V. Saelee, S. Kornpetpanee and P. Wongkasem, An examination of the variability of agricultural production using a new approximate confidence interval for the coefficient of variation, *Sci. Asia.*, **39**, 440-443 (2013).
[33] W. Thangjai, S. Niwitpong and S. Niwitpong, Confidence intervals for the common mean of several normal populations, *Studies in Computational Intelligence*, **692**, 321-331 (2017).
[34] H.Q. Li, M.L. Tang and W.K. Wong, Confidence intervals for ratio of two Poisson rates using the method of variance estimates recovery, *Comput. Stat.*, **29**, 869-889 (2014).
[35] M. Xiao, T. Jiang, H. Zhang and G. Shan, Exact one-sided confidence limit for the ratio of two Poisson rates, *Stat. Biopharm. Res.*, **9**, 180-185 (2017).
[36] A. Malekzadeh and M. Kharrati-Kopaei, Inferences on the common mean of several heterogeneous log-normal distributions, *J. Appl. Stat.*, **46**, 1066-1083 (2019).
[37] M.J. Jarvis, H. Tunstall-Pedoe, C. Feyerabend, C. Vesey and Y. Saloojee, Comparison of tests used to distinguish smokers from nonsmokers, *Am. J. Public Health*, **77**, 1435-1438 (1987).
[38] M. Pagano and K. Gauvreau, *Principles of Biostatistics*. Duxbury Press, California US, (1993).