

RESEARCH ARTICLE

Open Access



Comparison of alternative approaches for difference, noninferiority, and equivalence testing of normal percentiles

Gwown Shieh

Abstract

Background: Percentiles are widely used in scientific research for determining the comparative magnitude and reference limit of quantitative measurements. The investigations for point and interval estimation of normal percentiles are well documented in the literature. However, the corresponding statistical tests of hypothesis have received relatively little attention.

Methods: To facilitate data analysis and design planning of percentile study, this paper aims to present hypothesis testing procedures and associated power functions for assessing the difference, noninferiority, and equivalence of normal percentiles.

Results: Numerical illustrations about drug dissolution are provided to demonstrate the usefulness of the suggested exact approaches and the deficiency of approximate methods.

Conclusions: The exact approaches are superior to the approximate methods on the basis of control of Type I errors. Computer algorithms are constructed to implement the recommended test procedures and sample size calculations for percentile analysis.

Keywords: Power, Quantile, Reference limit, Sample size

Background

Percentiles are extremely useful for describing the reference threshold and meaningful magnitude of numerical quantities, such as achievement score, developmental index, medical measurement, and physical dimension. The inferential methods for normal means are well documented in the fundamental texts of statistical analysis. However, the methodological aspects and statistical implications of analyzing normal percentiles have been less discussed. It is essential to note that normal percentiles are a linear function of the mean and standard deviation of the underlying population. Because the sample mean and sample variance are complete and sufficient

statistics for the population mean and variance, the minimum variance unbiased estimator of a normal percentile can be readily obtained. Specifically, Royston and Mathews [1] compared the minimum variance unbiased estimator and other useful formulas under the intrinsic criteria of bias and mean square error. More advanced and theoretical treatments of normal percentile estimation are also available in Keating, Mason, and Balakrishnan [2], Keating and Tripathi [3], Parrish [4], Rukhin [5], and Zidek [6, 7].

Both exact and approximate confidence intervals of normal percentiles have been considered in several analytical developments. The exact interval estimation of normal percentiles was presented in Meeker, Hahn, and Escobar [8], Johnson, Kotz, and Balakrishnan [9], and Owen [10]. Note that the exact confidence intervals

Correspondence: gwshieh@mail.nctu.edu.tw
Department of Management Science, National Chiao Tung University,
Hsinchu, Taiwan 30010, Republic of China



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

involve the quantiles of a noncentral t distribution. Such critical values are not commonly available in tabulated forms and the implementation necessitates appropriate computing algorithms. To circumvent the reliance on a noncentral t distribution, approximate methods were considered by using the standardization technique and the regular t distribution. Accordingly, the approximate confidence intervals of Bland and Altman [11] and Chakraborti and Li [12] are computationally simple and the interval calculations do not require specialized software. However, the numerical study of Shieh [13] demonstrated that the confidence limits of the approximate methods generally do not preserve the nominal equal-tailed error rates. The finding provides cautionary counterpoint on the practical value of approximate intervals, especially when the sample sizes are small.

The existing investigations present important inferential methodology for point and interval estimation of normal percentiles. However, the related hypothesis testing problems have not been properly explicated in the literature. It is well known that there exists a direct connection between confidence interval and hypothesis testing. But the two approaches are philosophically different in the outset of precision and power viewpoints. Accordingly, to conduct a significance tests for percentiles, the conclusion can be alternatively obtained by examining whether the specified percentile value is contained in the proper two- or one-sided confidence intervals. It appears that percentile analysis can be performed without explicitly defining the desirable test statistics and associated rejection regions. However, power evaluation and sample size planning for hypothesis testing methodologically differ from the precision and sample size considerations in the context of interval estimation. Consequently, it is of theoretical importance and practical interest to document the exact test procedures, power calculations, and sample size determinations for percentile studies.

To enhance the usage of percentile analysis, this article describes hypothesis testing procedures and associated power functions for assessing the difference, noninferiority, and equivalence of normal percentiles. The difference and noninferiority procedures closely follow the two- and one-tail test formulations. In the conventional studies of the population means, a null hypothesis of zero may be informative to address certain essential research questions. The situations associated with percentile assessment are more sophisticated because the target percentile is unlikely a zero value. The percentile tests for difference and noninferiority require researchers to provide a sensible magnitude that corresponds to the percentile threshold for identifying substantial research finding. Moreover, the importance for establishing equivalence instead of no difference has been emphasized in Blackwelder [14] and Parkhurst [15], among

others. Further details on the design and analysis of non-inferiority and equivalence studies can be found in Fleming et al. [16] and Wellek [17].

Notably, the binomial test of hypotheses concerning quantiles in Mood, Graybill, and Boes ([18], Section 11.3.2) provides an appealing nonparametric alternative. Although the procedure is applicable for all random samples from a continuous distribution, there are not many feasible alpha values for small sample sizes, unless randomized tests are used. In general, the nonparametric tests may be more powerful than their parametric counterparts when normality assumption fails, whereas the nonparametric alternatives are less powerful than the parametric procedures when the conventional assumptions hold. More importantly, the undesirable properties and related problems associated with binomial tests have been addressed in Vos and Hudson [19] and Thulin [20], among others. Comprehensive discussions and reviews for the prevailing Wald large-sample normal test and other alternative interval procedures can be found in Agresti and Coull [21], Newcombe [22], Brown, Cai, and DasGupta [23, 24], and the references therein. The illustrations and appraisals in this article were confined to the test procedures that assume normality of the sampling distribution.

This paper aims to present the exact test procedures for percentile study under the three structural considerations of difference, noninferiority, and equivalence scenarios. For the purpose of providing profound implications in selecting the most appropriate approach, the approximate techniques of Bland and Altman [11] and Chakraborti and Li [12] are also extended to the percentile testing problem. Specifically, Bland and Altman [11] proposed an approximate t distribution for a convenient transformation of the natural, but biased, estimator of the normal percentile. On the other hand, Chakraborti and Li [12] suggested that a standardized minimum variance unbiased estimator also has an approximate t distribution. Note that the simplified considerations proposed in Bland and Altman [11] and Chakraborti and Li [12] may be appealing for inducing computational shortcuts but they do not necessarily maintain the desired accuracy for all settings, especially when the sample sizes are small. Accordingly, it is essential to discern not only which method is most suitable under what circumstances but also the actual differences between the contending test procedures.

Furthermore, the corresponding power and sample size calculations for advance planning of percentile studies are explicated. Monte Carlo simulation study was also conducted to compare the accuracy of the exact and approximate procedures with respect to the control of Type I error rate. Although an exact technique is theoretically better than the approximate methods, the

actual performance may not guarantee a substantial difference to justify the need for adopting the exact approach that is methodologically sophisticated and computationally demanding. The current study provides detailed analytic explications and numerical evidences to reveal the discrepancy between the exact and approximate procedures for percentile analysis. A drug dissolution problem and accompanying software programs are employed to illustrate the usefulness of suggested procedures for data analysis and design planning.

Methods

Exact test procedures

Assume X_1, \dots, X_N are a sample from a $N(\mu, \sigma^2)$ population with unknown mean μ and variance σ^2 for $N > 1$. The 100pth percentile of the normal distribution $N(\mu, \sigma^2)$ is denoted by θ , where

$$\theta = \mu + z_p\sigma \tag{1}$$

and z_p is the (100·p)th percentile of the standard normal distribution $N(0, 1)$. An intuitive, but biased, estimator of the percentile θ is

$$\hat{\theta}_B = \bar{X} + z_pS, \tag{2}$$

where $\bar{X} = \sum_{i=1}^N X_i/N$ and $S^2 = \sum_{i=1}^N (X_i - \bar{X})^2 / (N-1)$ are the sample mean and sample variance, respectively. Accordingly, the minimum variance unbiased estimator is

$$\hat{\theta}_M = \bar{X} + z_p cS. \tag{3}$$

where $c = (v/2)^{1/2} \Gamma(v/2) / \Gamma\{(v+1)/2\}$ and $v = N - 1$. Further details about the point estimation properties of $\hat{\theta}_B$ and $\hat{\theta}_{MU}$ are available in Royston and Mathews [1]. Also, the recent study of Shieh [13] compared several confidence interval procedures of θ . In contrast, the focus here is on the hypothesis testing of normal percentiles.

Under the prescribed normal setting for the sample $\{X_1, \dots, X_N\}$, standard derivations show that

$$T_E = \frac{\bar{X} - \theta}{(S^2/N)^{1/2}} \sim t(v, -z_p N^{1/2}), \tag{4}$$

where $t(v, -z_p N^{1/2})$ is a noncentral t distribution with degrees of freedom v and noncentrality parameter $-z_p N^{1/2}$. The fundamental properties and related extensions of noncentral t distribution can be found in Johnson, Kotz, and Balakrishnan [9].

Tests for difference

To detect the magnitude of a percentile in terms of the hypotheses

$$H_0 : \theta = \theta_0 \text{ versus } H_1 : \theta \neq \theta_0, \tag{5}$$

the test statistic is of the form

$$T_{E0} = \frac{\bar{X} - \theta_0}{(S^2/N)^{1/2}}, \tag{6}$$

where θ_0 is a constant. The test rejects H_0 at the significance level α if $T_{E0} < \tau_{\alpha/2}$ or $T_{E0} > \tau_{1-\alpha/2}$ where $\tau_{\alpha/2}$ and $\tau_{1-\alpha/2}$ are the lower and upper (100· $\alpha/2$)th quantiles of the distribution $t(v, -z_p N^{1/2})$, respectively, for $0 < \alpha < 0.5$. Accordingly, it can be shown that the power function is of the form

$$\Psi_{DI}(\Delta) = P\{t(v, \Delta) < \tau_{\alpha/2}\} + P\{t(v, \Delta) > \tau_{1-\alpha/2}\}, \tag{7}$$

where $\Delta = (\mu - \theta_0) / (\sigma^2/N)^{1/2}$.

Tests for noninferiority

In addition to the regular test of difference, it is of practical importance to test the hypotheses for noninferiority. The problem of testing noninferiority of percentiles can be presented by the following hypotheses:

$$H_0 : \theta \leq \theta_0 \text{ versus } H_1 : \theta > \theta_0 \tag{8}$$

when larger values of θ are desired and θ_0 is the designated noninferiority threshold. The test procedure rejects the null hypothesis at the significance level α if $T_{E0} > \tau_{1-\alpha}$ and the associated power function is readily obtained as

$$\Psi_{NI}(\Delta) = P\{t(v, \Delta) > \tau_{1-\alpha}\}. \tag{9}$$

On the other hand, if smaller values of θ are preferred, then the following hypotheses should be adopted for the test of noninferiority:

$$H_0 : \theta \geq \theta_0 \text{ versus } H_1 : \theta < \theta_0, \tag{10}$$

where the chosen value θ_0 represents the noninferiority bound. At the significance level α , the rejection region for the lower one-sided test is $T_{E0} < \tau_\alpha$ and the power function is expressed as

$$\Psi_{NI}(\Delta) = P\{t(v, \Delta) < \tau_\alpha\}. \tag{11}$$

Tests for equivalence

Unlike the traditional differences-based procedures, equivalence testing provides a proper method for demonstrating the comparability of target percentile. In general, the null and alternative hypotheses of a test of percentile equivalence can be formulated as

$$\begin{aligned} H_0 : \theta - \theta_T \leq -\delta \text{ or } \theta - \theta_T \geq \delta \text{ versus} \\ H_1 : -\delta < \theta - \theta_T < \delta, \end{aligned} \tag{12}$$

where θ_T and $\delta (> 0)$ are constants. Accordingly, θ_T is the target value and δ represents the minimum threshold for declaring equivalence between the population percentile θ and θ_T . Following the two one-sided tests procedure proposed by Schuirmann [25] and Westlake

[26] for assessing equivalence of mean effects, the null hypothesis is rejected at the significance level α if

$$T_{EL} = \frac{\bar{X} - \theta_T + \delta}{(S^2/N)^{1/2}} > \tau_{1-\alpha} \text{ and } T_{EU} = \frac{\bar{X} - \theta_T - \delta}{(S^2/N)^{1/2}} < \tau_\alpha. \quad (13)$$

It is important to note that the rejection is an intersection of two one-sided segments in terms of the lower and upper (100α) th quantiles τ_α and $\tau_{1-\alpha}$ of the non-central t distribution $t(v, -z_p N^{1/2})$. The rejection region of \bar{X} and S^2/N has an isosceles triangular shape similar to those in Meyners [27] and Schuirmann [28] for the equivalence procedure of two treatment means. Consequently, the power function of the percentile equivalence test can be written as

$$\Psi_{EQ} = P\{\theta_T - \delta + \tau_{1-\alpha}(S^2/N)^{1/2} < \bar{X} < \theta_T + \delta + \tau_\alpha(S^2/N)^{1/2}\}. \quad (14)$$

Moreover, it is clear from the fundamental assumption in Eq. 1 that $Z = (\bar{X} - \mu)/(S^2/N)^{1/2} \sim N(0, 1)$ and $K = vS^2/\sigma^2 \sim \chi^2(v)$, where $\chi^2(v)$ denotes the chi-square distribution with $v = N - 1$ degrees of freedom, and Z and K are independent. Let $H_E = 1$ if $K < \kappa_E$, and $H_E = 0$ if $K \geq \kappa_E$ where $\kappa_E = (4vN\delta^2)/\{\sigma^2(\tau_{1-\alpha} - \tau_\alpha)^2\}$. Then, the exact power function can be expressed by

$$\Psi_{EQ} = E_K[H_E\{\Phi(U_E) - \Phi(L_E)\}], \quad (15)$$

where $U_E = (\theta_T + \delta - \mu)/(\sigma^2/N)^{1/2} + \tau_\alpha(K/v)^{1/2}$, $L_E = (\theta_T - \delta - \mu)/(\sigma^2/N)^{1/2} + \tau_{1-\alpha}(K/v)^{1/2}$, $\Phi(\cdot)$ is the cumulative density function of the standard normal distribution, and the expectation E_K is taken with respect to the distribution K . It is essential to note that the probability $P\{K \geq \kappa_E\} \doteq 0$ in the subsequent numerical assessments under a wide range of model configurations. This phenomenon is similar to the power computations for the equivalence procedure of two treatment means as noted in Siqueira, et al. [29] and Shieh [30]. Therefore, the exact power appraisal can be numerically approximated by

$$\Psi_{AEQ} = P\{t(v, \Delta_U) < \tau_\alpha\} - P\{t(v, \Delta_L) < \tau_{1-\alpha}\}, \quad (16)$$

where $\Delta_U = (\mu - \theta_T - \delta)/(\sigma^2/N)^{1/2}$ and $\Delta_L = (\mu - \theta_T + \delta)/(\sigma^2/N)^{1/2}$.

Approximate methods

For the purpose of method comparisons, two different approaches for testing normal percentiles are also presented next. To construct confidence intervals of normal percentiles, Bland and Altman [11] and Chakraborti and Li [12] considered simple t approximations for the standardized forms of $\hat{\theta}_B$ and $\hat{\theta}_M$, respectively. Their

methods are extended and examined here for the three types of difference, noninferiority, and equivalence testing.

The Chakrabort-Li method

In view of the desirable properties of the minimum variance unbiased estimator $\hat{\theta}_M$, Chakraborti and Li [12] suggested an approximate t distribution for the standardized quantity of $\hat{\theta}_M$:

$$T_M = \frac{\hat{\theta}_M - \theta}{(mS^2/N)^{1/2}} \sim t(v). \quad (17)$$

where $m = 1 + Nz_p^2(c^2 - 1)$ and $t(v)$ is a t distribution with degrees of freedom v . Note that $Var[\hat{\theta}_M] = (m\sigma^2)/N$ and the denominator of T_M is obtained by a direct substitution of σ^2 with S^2 in the standard deviation of $\hat{\theta}_M$.

The simple formulation of T_M provides an alternative test statistic for judging the magnitude of normal percentiles. For the hypothesis test of difference in terms of $H_0: \theta = \theta_0$ versus $H_1: \theta \neq \theta_0$, the null hypothesis can be rejected at the significance level α if $T_{M0} < t_{\alpha/2}$ or $T_{M0} > t_{1-\alpha/2}$, or equivalently $|T_{M0}| > t_{1-\alpha/2}$, where

$$T_{M0} = \frac{\hat{\theta}_M - \theta_0}{(mS^2/N)^{1/2}}, \quad (18)$$

and $t_{\alpha/2}$ and $t_{1-\alpha/2}$ are the lower and upper $100(\alpha/2)$ th quantiles of a t distribution $t(v)$ with degrees of freedom v , respectively. Under the approximate t assumption, the corresponding power function can be derived as

$$\Omega_{DI}(\Delta) = P\{t(v, \Delta) < t_{\alpha/2}m^{1/2} - z_p cN^{1/2}\} + P\{t(v, \Delta) > t_{1-\alpha/2}m^{1/2} - z_p cN^{1/2}\}. \quad (19)$$

Similarly, the test statistic T_{M0} can be applied for hypothesis testing of noninferiority of percentiles in terms of $H_0: \theta \leq \theta_0$ versus $H_1: \theta > \theta_0$. The test procedure rejects the null hypothesis at the significance level α if $T_{M0} > t_{1-\alpha}$ and the associated power function is

$$\Omega_{NI}(\Delta) = P\{t(v, \Delta) > t_{1-\alpha}m^{1/2} - z_p cN^{1/2}\}. \quad (20)$$

Moreover, under the hypotheses: $H_0: \theta \geq \theta_0$ versus $H_1: \theta < \theta_0$, the test of noninferiority is rejected if $T_{M0} < t_\alpha$ and the corresponding power is given by

$$\Omega_{NI}(\Delta) = P\{t(v, \Delta) < t_\alpha m^{1/2} - z_p cN^{1/2}\}. \quad (21)$$

For the case of evaluating percentile equivalence with respect to $H_0: \theta - \theta_T \leq -\delta$ or $\theta - \theta_T \geq \delta$ versus $H_1: -\delta < \theta - \theta_T < \delta$, the null hypothesis is rejected at the significance level α if

$$\begin{aligned}
 T_{ML} &= \frac{\hat{\theta}_M - \theta_T + \delta}{(mS^2/N)^{1/2}} > t_{1-\alpha} \text{ and} \\
 T_{MU} &= \frac{\hat{\theta}_M - \theta_T - \delta}{(mS^2/N)^{1/2}} < t_\alpha.
 \end{aligned}
 \tag{22}$$

Accordingly, the power function can be shown as

$$\Omega_{EQ} = E_K[H_M\{\Phi(U_M) - \Phi(L_M)\}], \tag{23}$$

where $U_M = (\theta_T + \delta - \mu)/(\sigma^2/N)^{1/2} + (t_\alpha m^{1/2} - z_p c N^{1/2})(K/\nu)^{1/2}$, $L_M = (\theta_T - \delta - \mu)/(\sigma^2/N)^{1/2} + (t_{1-\alpha} m^{1/2} - z_p c N^{1/2})(K/\nu)^{1/2}$, and $H_M = 1$ if $K < \kappa_M$, and $H_M = 0$ if $K \geq \kappa_M$ where $\kappa_M = (\nu N \delta^2)/\{m \sigma t_{1-\alpha}^2\}$. Numerically, the power calculation can be simplified just as Ψ_{AEQ} given above:

$$\begin{aligned}
 \Omega_{AEQ} &= P\left\{t(v, \Delta_U) < t_\alpha m^{1/2} - z_p c N^{1/2}\right\} \\
 &\quad - P\left\{t(v, \Delta_L) < t_{1-\alpha} m^{1/2} - z_p c N^{1/2}\right\}.
 \end{aligned}
 \tag{24}$$

The Bland-Altman method

Similar to the test procedures based on the minimum variance unbiased estimator, hypothesis testing of normal percentiles can be conducted with the following transformation of $\hat{\theta}_B$ in Bland and Altman [11]:

$$T_B = \frac{\hat{\theta}_B - \theta}{(bS^2/N)^{1/2}} \sim t(v), \tag{25}$$

where $b = 1 + z_p^2/2$. Specifically, the hypothesis testing of percentile difference in terms of $H_0: \theta = \theta_0$ versus $H_1: \theta \neq \theta_0$ can be rejected at the significance level α if $|T_{B0}| > t_{1-\alpha/2}$ where

$$T_{B0} = \frac{\hat{\theta}_B - \theta_0}{(bS^2/N)^{1/2}}. \tag{26}$$

The associated power function is of the form

$$\begin{aligned}
 \Xi_{DI}(\Delta) &= P\left\{t(v, \Delta) < t_{\alpha/2} b^{1/2} - z_p N^{1/2}\right\} \\
 &\quad + P\left\{t(v, \Delta) < t_{1-\alpha/2} b^{1/2} - z_p N^{1/2}\right\}.
 \end{aligned}
 \tag{27}$$

To perform the hypothesis testing of noninferiority with $H_0: \theta \leq \theta_0$ versus $H_1: \theta > \theta_0$, the test rejects the null hypothesis at the significance level α if $T_{B0} > t_{1-\alpha}$ and the power function is readily obtained as

$$\Xi_{NI}(\Delta) = P\left\{t(v, \Delta) > t_{1-\alpha} b^{1/2} - z_p N^{1/2}\right\}. \tag{28}$$

Likewise, under the hypotheses: $H_0: \theta \geq \theta_0$ versus $H_1: \theta < \theta_0$, the test of noninferiority is rejected if $T_{B0} < t_\alpha$ and the corresponding power is expressed as

$$\Xi_{NI}(\Delta) = P\left\{t(v, \Delta) < t_\alpha b^{1/2} - z_p N^{1/2}\right\}. \tag{29}$$

Moreover, for the equivalence test of normal percentiles under the hypotheses of $H_0: \theta - \theta_T \leq -\delta$ or $\theta - \theta_T \geq \delta$ versus $H_1: -\delta < \theta - \theta_T < \delta$, the null hypothesis is rejected at the significance level α if

$$\begin{aligned}
 T_{BL} &= \frac{\hat{\theta}_B - \theta_T + \delta}{(bS^2/N)^{1/2}} > t_{1-\alpha} \text{ and} \\
 T_{BU} &= \frac{\hat{\theta}_B - \theta_T - \delta}{(bS^2/N)^{1/2}} < t_\alpha.
 \end{aligned}
 \tag{30}$$

In this case, the power function has the following formulation:

$$\Xi_{EQ} = E_K[H_B\{\Phi(U_B) - \Phi(L_B)\}], \tag{31}$$

where $U_B = (\theta_T + \delta - \mu)/(\sigma^2/N)^{1/2} + (t_\alpha b^{1/2} - z_p N^{1/2})(K/\nu)^{1/2}$, $L_B = (\theta_T - \delta - \mu)/(\sigma^2/N)^{1/2} + (t_{1-\alpha} b^{1/2} - z_p N^{1/2})(K/\nu)^{1/2}$, and $H_B = 1$ if $K < \kappa_B$, and $H_B = 0$ if $K \geq \kappa_B$ where $\kappa_B = (\nu N \delta^2)/\{b \sigma^2 t_{1-\alpha}^2\}$. Similar to the other two cases, the power computation can be well approximated by

$$\begin{aligned}
 \Xi_{AEQ} &= P\left\{t(v, \Delta_U) < t_\alpha b^{1/2} - z_p N^{1/2}\right\} \\
 &\quad - P\left\{t(v, \Delta_L) < t_{1-\alpha} b^{1/2} - z_p N^{1/2}\right\}.
 \end{aligned}
 \tag{32}$$

Results

Numerical investigations are presented next to examine and compare the fundamental features of the exact and approximate test procedures of percentiles with respect to the control of Type I error rate and accuracy of power and sample size computation.

Tests for difference

For the purpose of illustration, the null $N(\mu_0, \sigma_0^2)$ distribution is set as $N(0, 1)$ and two different mean values are considered for the alternative distribution $N(\mu, \sigma^2)$: $N(0.4, 1)$ and $N(0.6, 1)$. The corresponding percentiles θ_0 and θ are simplified as $\theta_0 = \mu_0 + z_p \sigma_0 = z_p$ and $\theta = \mu + z_p \sigma = \mu + z_p$, respectively, with $\mu = 0.4$ and 0.6 . For the difference test of percentile in terms of $H_0: \theta = \theta_0$ versus $H_1: \theta \neq \theta_0$, the sample sizes needed to attain the specified power 0.80 for the chosen significance level $\alpha = 0.05$ are determined by the power functions Ψ_{DI} , Ω_{DI} , and Ξ_{DI} for $p = 0.1, \dots, 0.9$. The computed sample sizes for the prescribed three procedures $\{T_{EO}, T_{MO}, T_{B0}\}$ are summarized in Table 1 for all eighteen combined cases of μ and

Table 1 The error between simulated alpha and nominal alpha for the difference tests of percentile $H_0: \theta = \theta_0$ versus $H_1: \theta \neq \theta_0$ with $\mu_0 = 0, \sigma_0 = 1, \sigma = 1$, and $\alpha = 0.05$

μ	p	Exact approach			Chakraborti-Li method			Bland-Altman method					
		N	Lower-tail error	Upper-tail error	Two-tail error	N	Lower-tail error	Upper-tail error	Two-tail error	N	Lower-tail error	Upper-tail error	Two-tail error
0.4	0.1	109	-0.0020	0.0009	-0.0011	97	-0.0101	0.0108	0.0007	95	-0.0108	0.0118	0.0010
	0.2	80	-0.0042	-0.0019	-0.0061	72	-0.0113	0.0060	-0.0053	71	-0.0121	0.0082	-0.0039
	0.3	65	-0.0030	-0.0020	-0.0050	60	-0.0086	0.0045	-0.0041	60	-0.0091	0.0058	-0.0033
	0.4	57	0.0009	0.0012	0.0021	54	-0.0026	0.0040	0.0014	54	-0.0031	0.0046	0.0015
	0.5	52	0.0024	0.0010	0.0034	52	0.0024	0.0010	0.0034	52	0.0024	0.0010	0.0034
	0.6	50	0.0019	0.0005	0.0024	52	0.0056	-0.0029	0.0027	52	0.0056	-0.0032	0.0024
	0.7	51	0.0020	-0.0004	0.0016	56	0.0093	-0.0083	0.0010	57	0.0101	-0.0088	0.0013
	0.8	58	0.0014	-0.0012	0.0002	66	0.0124	-0.0119	0.0005	66	0.0147	-0.0127	0.0020
	0.9	75	-0.0024	-0.0025	-0.0049	87	0.0106	-0.0133	-0.0027	88	0.0133	-0.0138	-0.0005
0.6	0.1	54	0.0004	0.0012	0.0016	46	-0.0137	0.0155	0.0018	44	-0.0148	0.0181	0.0033
	0.2	40	0.0014	-0.0009	0.0005	34	-0.0105	0.0137	0.0032	33	-0.0111	0.0158	0.0047
	0.3	32	-0.0001	-0.0022	-0.0023	29	-0.0094	0.0087	-0.0007	28	-0.0103	0.0104	0.0001
	0.4	27	0.0005	-0.0002	0.0003	26	-0.0048	0.0044	-0.0004	25	-0.0053	0.0052	-0.0001
	0.5	24	0.0016	0.0005	0.0021	24	0.0016	0.0005	0.0021	24	0.0016	0.0005	0.0021
	0.6	23	0.0037	0.0017	0.0054	24	0.0105	-0.0037	0.0068	25	0.0117	-0.0041	0.0076
	0.7	22	0.0028	0.0001	0.0029	26	0.0150	-0.0094	0.0056	26	0.0173	-0.0107	0.0066
	0.8	24	0.0014	-0.0012	0.0002	30	0.0183	-0.0155	0.0028	30	0.0209	-0.0159	0.0050
	0.9	31	0.0009	-0.0015	-0.0006	39	0.0201	-0.0174	0.0027	40	0.0248	-0.0182	0.0066

p . It should be noted that the parameter settings are chosen so that the resulting sample sizes have a reasonable magnitude that is often occurred in practice. Moreover, these situations with small and moderate sample sizes are of great importance in the sense that the contending procedures have the obvious potential of yielding distinct outcomes. Monte Carlo simulation studies of 10,000 iterations were conducted for examining the accuracy of the power functions Ψ_{Dt} , Ω_{Dt} , and Ξ_{Dt} . The results reveal that the simulated powers and the attained powers of all three methods agree to the second decimal place for all cases considered here. To save space, the details are not reported.

Due to the approximate nature of the t distribution associated with the two approximations of Chakraborti and Li [12] and Bland and Altman [11], it is of statistical concern to validate the control of the Type I error rates. Note that the real distribution of the percentile is skewed when sample size is small and p deviates considerably from 0.5. This implies that the symmetric t approximation of the two test statistics T_{M0} , and T_{B0} is presumably unsuitable. In other words, the two critical values $t_{\alpha/2}$ and $t_{1-\alpha/2}$ are theoretically inaccurate when one-sided rejection probability are evaluated. It is constructive to examine three distinctive Type I errors correspond to the lower-tail, upper-tail, and two-sided rejection regions of the difference tests of percentile.

Accordingly, Monte Carlo simulation studies were also performed to compute the simulated Type I error rates of the exact and approximate test procedures for $\theta = \theta_0$ or $\mu = 0$. The simulated Type I error rate was the proportion of the 10,000 replicates whose test statistic fell in the designated rejection region. In the process, the estimates of the lower-tail and upper-tail rejection rates were computed and summed as the overall or two-sided simulated Type I error rate. The accuracy of the control of Type I error rate can be assessed by the differences between the one-sided and two-sided simulation estimates and the nominal values 0.025 and 0.05, respectively. These differences or errors of the three contending test procedures are also reported in Table 1. It can be readily seen from the results in Table 1 that the all three test methods have excellent control of two-sided Type I error rate. The absolute magnitudes of the errors are less than 0.01 for the investigated mean and percentile configurations.

Moreover, the lower-tail and upper-tail rejection rates of the exact approach are also very close to the nominal levels. But the one-sided Type I error rates of the two approximate methods do not maintain the same accuracy especially for low and high percentiles. Despite the desired performance of the approximate tests in overall Type I error rate, the resulting errors of the lower-tail rejection region tend to be negative for small p while

those associated with large p are constantly positive. In contrast, the upper-tail errors have the exactly opposite outcomes. For the particular case with $\mu = 0.6$ and $p = 0.9$, the induced errors for the approximation of Chakraborti and Li [12] are 0.0201 and -0.0174 for lower and upper rejection regions, respectively. The corresponding deviated percentages are $0.0201/0.025 = 80.4\%$ and $0.0174/0.025 = 69.6\%$. To the approximate method of Bland and Altman [11], the lower-tail and upper-tail errors are 0.0248 and -0.0182 with the deviated percentages $0.0248/0.025 = 99.2\%$ and $0.0182/0.025 = 72.8\%$, respectively.

Tests for noninferiority

The underlying characteristics of the exact and approximate methods for the noninferiority test of percentile are also assessed. With the same model formulations in the previous scenario of difference test, the required sample sizes are computed for the hypotheses $H_0: \theta \leq \theta_0$ versus $H_1: \theta > \theta_0$ with the power functions Ψ_{NB} , Ω_{NB} , and Ξ_{NI} . As expected, the result reported in Table 2 is relatively smaller than the counterpart in Table 1 with the identical values of μ and p . Moreover, simulation studies were also performed to appraise the actual performance of Type 1 error for $\theta = \theta_0$ or $\mu = 0$. The errors between the simulated rejection rates and nominal value $\alpha = 0.05$ are presented in Table 2. Unlike the exact procedure

with good control of Type error rate, the two approximate tests do not maintain the required performance. Specifically, when $\mu = 0.6$ and $p = 0.9$, the absolute errors (absolute error percentage) can be as large as 0.0258 ($0.0258/0.05 = 51.6\%$) and 0.0260 ($0.0260/0.05 = 52.0\%$) for T_{M0} and T_{B0} of Chakraborti and Li [12] and Bland and Altman [11], respectively. Although the situations improved with increasing sample size as those cases when $\mu = 0.4$, they still suffer some potential deficiency and are outperformed by the exact test.

Tests for equivalence

For the sake of completeness, numerical examination is extended to the equivalence tests of percentile in terms of $H_0: \theta - \theta_T \leq -\delta$ or $\theta - \theta_T \geq \delta$ versus $H_1: -\delta < \theta - \theta_T < \delta$. In this case, the target percentile and threshold are set as $\theta_T = z_p$ and $\delta = 0.6$, respectively. The alternative normal distribution is selected as $N(\mu, 1)$ and the associated percentile is $\theta = \mu + z_p\sigma = \mu + z_p$. Then, the power functions Ψ_{AEQ} , Ω_{AEQ} , and Ξ_{AEQ} are applied to compute the minimum sample sizes required for attaining the nominal power 0.80 at $\alpha = 0.05$. The resulting sample sizes are listed in Table 3 for $\mu = 0$ and 0.3 and $p = 0.1, \dots, 0.9$. It was further justified with simulation studies that the power and sample size calculations of the three procedures are all extremely accurate for all eighteen cases reported here. However, power evaluation is valid

Table 2 The error between simulated alpha and nominal alpha for the non-inferiority tests of percentile $H_0: \theta \leq \theta_0$ versus $H_1: \theta > \theta_0$ with $\mu_0 = 0, \sigma_0 = 1, \sigma = 1$, and $\alpha = 0.05$

μ	p	Exact approach		Chakraborti-Li method		Bland-Altman method	
		N	Error	N	Error	N	Error
0.4	0.1	85	0.0047	77	0.0124	75	0.0168
	0.2	63	0.0017	57	0.0155	56	0.0175
	0.3	51	-0.0011	48	0.0123	47	0.0141
	0.4	44	0.0033	43	0.0046	42	0.0070
	0.5	41	-0.0010	41	0.0004	41	0.0003
	0.6	39	-0.0017	41	-0.0048	41	-0.0077
	0.7	40	-0.0004	44	-0.0112	45	-0.0120
	0.8	46	0.0001	51	-0.0154	52	-0.0183
	0.9	60	-0.0019	68	-0.0170	69	-0.0182
0.6	0.1	42	-0.0027	36	0.0197	35	0.0274
	0.2	31	-0.0042	27	0.0192	26	0.0278
	0.3	25	0.0010	23	0.0119	22	0.0215
	0.4	21	-0.0002	20	0.0052	20	0.0072
	0.5	19	-0.0006	19	-0.0008	19	-0.0014
	0.6	18	-0.0008	19	-0.0086	19	-0.0092
	0.7	18	-0.0013	20	-0.0154	21	-0.0195
	0.8	20	-0.0026	23	-0.0220	24	-0.0252
	0.9	25	-0.0016	31	-0.0258	31	-0.0260

Table 3 The error between simulated alpha and nominal alpha for the equivalence tests of percentile $H_0: \theta - \theta_T \leq -\delta$ or $\theta - \theta_T \geq -\delta$ versus $H_1: -\delta < \theta - \theta_T < \delta$ with $\delta = 0.6$, $\theta_T = z_p$, $\sigma = 1$, and $\alpha = 0.05$

μ	ρ	Exact approach		Chakraborti-Li method		Bland-Altman method	
		<i>N</i>	Error	<i>N</i>	Error	<i>N</i>	Error
0	0.1	47	-0.0006	45	0.0196	45	0.0257
	0.2	35	0.0003	34	0.0197	34	0.0162
	0.3	30	-0.0015	29	0.0121	29	0.0143
	0.4	27	-0.0039	26	0.0070	26	0.0056
	0.5	26	-0.0006	26	-0.0006	26	-0.0033
	0.6	27	-0.0018	26	-0.0071	26	-0.0111
	0.7	30	-0.0033	29	-0.0144	29	-0.0154
	0.8	35	-0.0047	34	-0.0188	34	-0.0191
	0.9	47	-0.0034	45	-0.0205	45	-0.0212
0.3	0.1	110	-0.0024	121	0.0117	123	0.0205
	0.2	84	0.0000	91	0.0097	92	0.0103
	0.3	73	0.0024	78	0.0063	78	0.0069
	0.4	69	0.0001	72	0.0023	72	0.0043
	0.5	71	0.0015	71	-0.0014	71	0.0012
	0.6	76	0.0008	74	-0.0081	74	-0.0069
	0.7	87	-0.0050	82	-0.0068	81	-0.0075
	0.8	106	-0.0041	99	-0.0112	97	-0.0112
	0.9	144	-0.0005	133	-0.0123	131	-0.0152

and informative only when the critical value satisfies the nominal Type I error rate. Additional simulation studies were employed to assess the control of Type I error rates of the equivalence tests $\{T_{EL}, T_{EU}\}$, $\{T_{ML}, T_{MU}\}$, and $\{T_{BL}, T_{BU}\}$ for $\theta = \theta_T - \delta$ or $\mu = -\delta = -0.6$. The errors between the simulated and nominal Type I error rates are presented in Table 3. The assessments show that the two approximate tests $\{T_{ML}, T_{MU}\}$, and $\{T_{BL}, T_{BU}\}$ are not as good as the exact procedure $\{T_{EL}, T_{EU}\}$. The deficiency of the two simple *t* distributions is particularly more prominent when the sample size is small and $|\rho - 0.5|$ is large.

An example

To demonstrate the usefulness of the suggested techniques and accompanying programs, a quality control application in pharmaceutical products is exemplified and analyzed with the hypothesis testing and sample size procedures. Suppose a sample of the selected batch of tablets is obtained and tested according to the acceptance sampling plan. Specifically, the dissolution performance is assessed in terms of the percentage of tablets dissolved less than a specified amount at a certain time period.

For illustration, the summary statistics of the dissolution values are $\bar{X} = 50.10$ and $S = 1.31$ for $N = 15$. Sup-

pose that the experimenter is interested in the 90th percentile of the distribution of the dissolution quantity. Then, it follows that $z_{0.9} = 1.2816$, $c = 1.0180$, and the minimum variance unbiased estimator is $\hat{\theta}_M = 51.8091$. Using the working settings $\mu_0 = 49.3$ and $\sigma_0 = 1.2$, the associated 90th percentile value is computed as $\theta_0 = \mu_0 + z_p\sigma_0 = 50.8379$. The test statistic T_{E0} has a value of -2.1816 . For testing the hypotheses of $H_0: \theta = 50.8379$ versus $H_1: \theta \neq 50.8379$, the two critical values or the lower and upper 2.5th quantiles of the noncentral *t* distribution $t(14, -4.9636)$ are $\tau_{0.025} = -8.7695$ and $\tau_{0.975} = -2.7909$, respectively. The null hypothesis is rejected and it implies that the 90th percentile of dissolution amount is not 50.8379 at $\alpha = 0.05$. Also, it can be shown that the values of the two approximate test statistics for Chakraborti and Li [12] and Bland and Altman [11] are $T_{M0} = 2.0857$ and $T_{B0} = 2.0614$, respectively, with the critical value $t_{0.975} = 2.1448$. Hence, the two approximate tests suggest that the null hypothesis cannot be rejected for $\alpha = 0.05$.

Moreover, a noninferiority test can be formed as $H_0: \theta \leq 50.8379$ versus $H_1: \theta > 50.8379$. With a critical value $\tau_{0.05} = -3.1072$, it indicates that the 90th percentile of dissolution distribution is higher than 50.8379 at the 5% level of significance. In this case, the critical value for the two approximate methods is $t_{0.95} = 1.7613$. Thus, the two approximate tests lead to the same result as the

exact procedure. Assume a equivalence test of the 90th percentile is expressed as $H_0: \theta - \theta_T \leq -\delta$ or $\theta - \theta_T \geq \delta$ versus $H_1: -\delta < \theta - \theta_T < \delta$ with $\theta_T = 50 + (1.2816)(1.3) = 51.6660$ and $\delta = 1.2$. The test statistics are $\{T_{EL}, T_{EU}\} = \{-1.0821, -8.1776\}$ and the corresponding critical values are $\{\tau_{0.95}, \tau_{0.05}\} = \{-3.1072, -8.0108\}$. Thus, there is sufficient evidence to suggest that the 90th percentile is practically 51.6660 for the 5% level of significance. The resulting values for the approximate tests are $\{T_{ML}, T_{MU}\} = \{2.8845, -2.2700\}$ and $\{T_{BL}, T_{BU}\} = \{2.8761, -2.3817\}$. With the associated critical values $t_{0.95} = 1.7613$ and $t_{0.05} = -1.7613$, they also reach the same equivalence outcome. Supplemental computer programs are provided to take advantage of the embedded statistical functions in the interactive matrix language software of Statistical Analysis System (SAS/IML) [31] for performing the prescribed exact test procedures.

For planning future drug dissolution study, sample size calculations should be considered so that the tests have enough power to confirm meaningful magnitude of percentile. It is commonly assumed that typical sources like published findings or expert opinions can offer plausible and reasonable values for the vital characteristics of future study. Hence, the sample statistics of the summary statistics are used as parameter values $\mu = 50.1$ and $\sigma = 1.31$. To achieve the nominal power 0.80 with $\alpha = 0.05$, the constructed SAS/IML programs reveal that the required sample sizes are $N = 21$ and 17 for the test of difference: $H_0: \theta = 50.8379$ versus $H_1: \theta \neq 50.8379$ and the test of noninferiority: $H_0: \theta \leq 50.8379$ versus $H_1: \theta > 50.8379$, respectively. Moreover, for the abovementioned test of equivalence with $\theta_T = 51.6660$ and $\delta = 1.2$, sample size $N = 25$ is needed to attaining the nominal power 0.80 at $\alpha = 0.05$. Note that the exemplifying configurations are included in the user specifications of the SAS/IML programs presented in the supplemental files. Accordingly, users can easily modify the input values in these statements to accommodate their own model specifications.

Discussion

The present investigation generalizes and expands current results in the statistical literature by describing both exact and approximate procedures for the three different percentile tests of difference, noninferiority, and equivalence. The exact approach employs a noncentral t distribution, while the approximate techniques follow the familiar t distribution as considered in Chakraborti and Li [12] and Bland and Altman [11]. Regarding the two approximate procedures, the results of the conventional tests for difference show that the lower critical value $t_{\alpha/2}$ is generally too small for lower normal percentiles and is typically too large for the higher normal percentiles. On the other hand, the upper critical value $t_{1-\alpha/2}$

overestimate and underestimate the correct one for small and large p , respectively. Even the overall Type I error is not an issue, it is statistically improper to recommend a two-sided test procedure on the basis of a combination of some noticeable under- and over-sized critical values and rejection regions. Moreover, despite the relatively involved analytic assessments and computational requirements, the comprehensive numerical appraisals show that the exact approach is superior to the approximate methods on the basis of control of Type I errors.

Conclusion

In view of the conceptual simplicity and context-free feature, percentiles are widely used for determining the relative magnitude and substantial importance of quantitative measurements in all scientific fields. Accordingly, much of the literature has provided the inferential procedures for point and interval estimation of normal percentiles. To extend the applicability of percentile analysis, this article addresses the hypothesis testing problem for the percentiles of a normal distribution. The recommended test procedures and derived power functions are also empirically justified for percentile score assessments and sample size determinations. In order to facilitate data analysis and study planning, specialized computer programs are presented for conducting hypothesis testing and sample size calculation in percentile research.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12874-020-00933-z>.

Additional file 1. SAS/IML program for conducting percentile test of difference.

Additional file 2. SAS/IML program for computing required sample size for percentile test of difference.

Additional file 3. SAS/IML program for conducting percentile test of noninferiority.

Additional file 4. SAS/IML program for computing required sample size for percentile test of noninferiority.

Additional file 5. SAS/IML program for conducting percentile test of equivalence.

Additional file 6. SAS/IML program for computing required sample size for percentile test of equivalence.

Abbreviation

SAS/IML: The interactive matrix language software of Statistical Analysis System

Acknowledgements

The author would like to thank the Associate editor and two reviewers for their constructive comments that led to an improved article.

Authors' contributions

GS conceived of the study, conducted the theoretical examination, carried out the numerical computations, and drafted the manuscript. The author read and approved the final manuscript.

Funding

This work was supported by a grant from the Ministry of Science and Technology (MOST 107-2410-H-009-024-MY3).

Availability of data and materials

All data generated or analysed during this study are included in this published article.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The author declares that he has no competing interests.

Received: 13 September 2019 Accepted: 19 February 2020

Published online: 13 March 2020

References

1. Royston P, Matthews JNS. Estimation of reference ranges from normal samples. *Stat Med*. 1991;10:691–5.
2. Keating JP, Mason RL, Balakrishnan N. Percentile estimators in location-scale parameter families under absolute loss. *Metrika*. 2010;72:351–67.
3. Keating JP, Tripathi RC. Percentiles, estimation of, *Encyclopedia of statistical sciences*, vol. VI. New York: Wiley; 1985. p. 668–74.
4. Parrish RS. Comparison of quantile estimators in normal sampling. *Biometrics*. 1990;46:247–57.
5. Rukhin AL. A class of minimax estimators of a normal quantile. *Stat Probability Letts*. 1983;1:217–21.
6. Zidek JV. Inadmissibility of the best invariant estimator of extreme quantiles of the normal law under squared error loss. *Ann Math Stat*. 1969;40:1801–8.
7. Zidek JV. Inadmissibility of a class of estimators of a normal quantile. *Ann Math Stat*. 1971;42:1444–7.
8. Meeker WQ, Hahn GJ, Escobar LA. *Statistical intervals: a guide for practitioners and researchers*. New York: Wiley; 2017.
9. Johnson NL, Kotz S, Balakrishnan N. *Continuous univariate distributions*, vol. 2. 2nd ed. New York: Wiley; 1995.
10. Owen DB. A survey of properties and applications of the noncentral *t*-distribution. *Technometrics*. 1968;10:445–78.
11. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res*. 1999;8:135–60.
12. Chakraborti S, Li J. Confidence interval estimation of a normal percentile. *Am Stat*. 2007;61:331–6.
13. Shieh G. The appropriateness of Bland-Altman's approximate confidence intervals for limits of agreement. *BMC Med Res Methodol*. 2018;18:45.
14. Blackwelder WC. "Proving the null hypothesis" in clinical trials. *Control Clin Trials*. 1982;3:345–53.
15. Parkhurst DF. Statistical significance tests: equivalence and reverse tests should reduce misinterpretation. *BioScience*. 2001;51:1051–7.
16. Fleming TR, Odem-Davis K, Rothmann MD, Shen YL. Some essential considerations in the design and conduct of non-inferiority trials. *Clin Trials*. 2011;8:432–9.
17. Wellek S. *Testing statistical hypotheses of equivalence and noninferiority*. 2nd ed. New York: Chapman and Hall/CRC; 2010.
18. Mood AM, Graybill FA, Boes DC. *Introduction to the theory of statistics*. 3rd ed. New York: McGraw-Hill; 1974.
19. Vos PW, Hudson S. Problems with binomial two-sided tests and the associated confidence intervals. *Aust N Z J Stat*. 2008;50:81–9.
20. Thulin M. The cost of using exact confidence intervals for a binomial proportion. *Electron J Stat*. 2014;8:817–40.
21. Agresti A, Coull BA. Approximate is better than "exact" for interval estimation of binomial proportions. *Am Stat*. 1998;52:119–26.
22. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med*. 1998;17:857–72.
23. Brown LD, Cai TT, DasGupta A. Interval estimation for a binomial proportion. *Stat Sci*. 2001;16:101–17.
24. Brown LD, Cai TT, Dasgupta A. Confidence intervals for a binomial proportion and asymptotic expansions. *Ann Stat*. 2002;30:160–201.
25. Schuurmann DL. On hypothesis testing to determine if the mean of a normal distribution is contained in a known interval. *Biometrics*. 1981;37:617.
26. Westlake WJ. Response to T.B.L. Kirkwood: Bioequivalence testing—a need to rethink. *Biometrics*. 1981;37:589–94.
27. Meyners M. Equivalence tests—a review. *Food Qual Prefer*. 2012;26:231–45.
28. Schuurmann DJ. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *J Pharmacokinetics Biopharm*. 1987;15:657–80.
29. Siqueira AL, Whitehead A, Todd S, Lucini MM. Comparison of sample size formula for 2 × 2 cross-over designs applied to bioequivalence studies. *Pharm Stat*. 2005;4:233–43.
30. Shieh G. Exact power and sample size calculations for the two one-sided tests of equivalence. *PLoS One*. 2016;11:e0162093.
31. SAS Institute. *SAS/IML User's Guide, Version 9.3*. Cary: SAS Institute Inc; 2017.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

