Research Synthesis Methods (2026), **0**: 1–7
doi:10.1017/rsm.2026.10111

Research Synthesis Methods

RESEARCH-IN-BRIEF

A new method for detecting the skewness of data from the five-number summary

Sanying Feng[1], Hongmei Lin[2], Jiandong Shi[3], Sung Nok Chiu[4] and Tiejun Tong[4]

[1]School of Mathematics and Statistics, Zhengzhou University, China
[2]School of Statistics and Data Science, Shanghai University of International Business and Economics, China
[3]Guangdong Provincial Key Laboratory of Interdisciplinary Research and Application for Data Science, and Department of Statistics and Data Science, Beijing Normal–Hong Kong Baptist University, China
[4]Department of Mathematics, Hong Kong Baptist University, Hong Kong

Corresponding authors: Tiejun Tong and Hongmei Lin; Emails: tongt@hkbu.edu.hk, 7734@suibe.edu.cn

Received: 11 October 2025 UTC; **Revised:** 17 June 2026 UTC; **Accepted:** 22 June 2026 UTC

Keywords: five-number summary; meta-analysis; more powerful test; order statistics; skewness test

Abstract

For clinical trials with continuous outcomes, researchers may opt to report the whole or part of the five-number summary rather than the sample mean and standard deviation, especially when the outcome data are skewed. To include such studies in meta-analysis, several popular methods have been proposed in the literature that convert the five-number summary back to the sample mean and standard deviation. Nevertheless, most existing methods are based on the normality assumption, which may not hold for the clinical studies with the five-number summary being reported. Recently, Shi et al. (2023, *Stat Methods Med Res*, 32, 1338–1360) and Balakrishnan et al. (2023, *Math Methods Stat*, 32, 260–273) proposed methods for detecting the skewness of data that utilize the whole or part of the five-number summary, together with the sample size. In this article, we show that the max-type test of Shi et al. is not only structurally complex but also conservative in controlling the type I error rate, whereas the test of Balakrishnan et al. has difficulty controlling the type I error rate with small sample sizes. Inspired by these findings, we develop a novel test statistic that leverages the ratio of two tails, a measure known for its heightened sensitivity to data skewness. Simulation results demonstrate that our ratio-based test is both less conservative and more powerful compared to the existing methods, especially when the alternative distribution exhibits mild skewness. Additionally, simulated meta-analyses and real-world data examples are presented to demonstrate the utility of our new method in the context of meta-analysis.

Highlights

What is already known?

In recent years, several popular methods have emerged in the literature to estimate the sample mean and standard deviation from the five-number summary; some of these methods rely on the assumption of normality, while others do not. To address this issue, Shi et al.[18] and Balakrishnan et al.[19] recently proposed tests for detecting skewness under three common scenarios, respectively.

What is new?

We propose a novel skewness test, T_3, that leverages the five-number summary and sample size, using a ratio-based statistic for improved sensitivity to skewness. Our approximate formula for significance thresholds, supported by extensive simulations, demonstrates that T_3 maintains proper type I error control

© The Author(s), 2026. Published by Cambridge University Press on behalf of The Society for Research Synthesis Methodology. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

while consistently achieving higher power compared to the methods in Shi et al.[18] and Balakrishnan et al.[19]

Potential impact for RSM readers

The proposed new test has the potential to detect data skewness more accurately compared to existing methods based on the five-number summary, thereby supporting more transparent estimator selection in research synthesis. These results suggest that T_3 may be particularly well suited for meta-analytic applications where only the five-number summary is available.

1. Introduction

In the past several decades, meta-analysis has been widely used to synthesize data from multiple studies for decision making, especially in evidence-based medicine. With accumulated evidence, more reliable information and convincing conclusions can be drawn for scientific research in many different fields. To combine results from different studies, the sample mean and standard deviation are the two most commonly used statistics in meta-analysis for reporting normal data. In contrast, when the data are skewed, researchers may opt to report the whole or part of the five-number summary that consists of the minimum value a, the maximum value b, the first quartile q_1, the third quartile q_3, and the sample median m. Additionally, there are three common scenarios in the literature for reporting the five-number summary as follows:

$$\begin{aligned}\mathcal{S}_1 &= \{a, m, b; n\},\\ \mathcal{S}_2 &= \{q_1, m, q_3; n\},\\ \mathcal{S}_3 &= \{a, q_1, m, q_3, b; n\},\end{aligned}$$

where n is the sample size of the study. Given that full datasets are not always available, a key question is how to estimate individual means and standard deviations from these summary quantities and, further, the global treatment effect.

In order to meta-analyze a collection of studies in which both the sample mean and the sample median are reported, a few new methods have recently emerged that convert the five-number summary back to the sample mean and standard deviation. These methods can be briefly categorized into normal-based methods[1–4] and non-normal-based methods.[5–13] When the data come from a normal distribution, the normal-based estimators consistently perform best.[11] However, they may not work well when the underlying data are non-normal, in which case non-normal-based methods are preferred. Therefore, it is crucial to evaluate the normality of data before conducting a meta-analysis. For complete data, various methods for testing normality have been proposed in the literature. To name a few, the most popular methods include the Kolmogorov–Smirnov test (Lilliefors[14]), the Shapiro–Wilk test (Shapiro and Wilk[15]), the Jarque–Bera test (Jarque and Bera[16]), the histogram test, and the Q–Q plot test. However, these testing methods are not suitable for studies that report only the five-number summary. To further illustrate this issue, one may refer to Altman and Bland,[17] in which the authors stated in their brief note as follows: "*When authors present data in the form of a histogram or scatter diagram then readers can see at a glance whether the distributional assumption is met. If, however, only summary statistics are presented—as is often the case—this is much more difficult.*"

To address this need, Shi et al.[18] and Balakrishnan et al.[19] recently proposed methods for detecting the skewness of data that utilize the whole or part of the five-number summary, together with the sample size. For scenario $\mathcal{S}_1$, both Shi et al.[18] and Balakrishnan et al.[19] adopted $T_1^S = (a + b - 2m)/(b - a)$ as the test statistic. In the case of scenario $\mathcal{S}_2$, both Shi et al.[18] and Balakrishnan et al.[19] adopted $T_2^S = (q_1 + q_3 - 2m)/(q_3 - q_1)$ as the test statistic. Regarding scenario $\mathcal{S}_3$, Shi et al.[18] constructed a

https://doi.org/10.1017/rsm.2026.10111 Published online by Cambridge University Press

max-type test statistic as

$$T_3^S = \max\left\{\frac{2.65\log(0.6n)}{\sqrt{n}}\left|T_1^S\right|,\ \left|T_2^S\right|\right\}. \tag{1.1}$$

Balakrishnan et al.[19] considered the test statistic

$$T_3^B = k_n\frac{(a-2m+b)\omega_a + (q_1-2m+q_3)(1-\omega_a)}{(b-a)\omega_b + (q_3-q_1)(1-\omega_b)}, \tag{1.2}$$

where

$$k_n \approx \begin{cases}0.55, & n \le 200\\ 1.65, & n > 200,\end{cases} \quad \omega_a \approx \begin{cases}1, & n \le 100\\ 1-0.7(\log_{10}(n)-2), & n > 100,\end{cases} \quad \omega_b \approx \begin{cases}0.1, & n \le 200\\ 1, & n > 200.\end{cases}$$

Unlike T_3^S, which is a max-type statistic, T_3^B can be viewed as a weighted combination of numerators and denominators of T_1^S and T_2^S, respectively. Moreover, Shi et al.[18] and Balakrishnan et al.[19] also provided approximate formulas for calculating the critical values. Nevertheless, as shown in Figures 4 and 5 of Shi et al.,[18] the simulation results in their paper indicate that T_3^S with approximated critical values is conservative for all sample sizes and, in some settings, may even be less powerful than T_1^S. Taken together, it is evident that their skewness test for scenario $\mathcal{S}_3$ is not yet optimal and requires further refinement, which serves as the primary motivation for this study. For T_3^B, the simulation studies presented in Section S2 of the Supplementary Material show that this test has difficulty controlling the type I error rate with small sample sizes.

Motivated by these observations, we construct a new test statistic using the ratio of two tails, which is known to be more sensitive to the skewness of data. Our new test under scenario $\mathcal{S}_3$ not only yields more accurate control for the type I error rate across all sample sizes, but also provides a higher power than T_3^S and T_3^B in a wide range of settings, especially when the underlying data are not severely skewed. Therefore, the proposed method is better suited for detecting skewness in meta-analytic studies and guiding the selection of appropriate estimation approaches for subgroup analysis of normally distributed and skewed data. By serving as an auxiliary tool for pre-meta-analysis exploratory analysis of distributional characteristics, it helps researchers select mean and standard deviation estimation suited to the specific features of the summary data, thereby producing more accurate and reliable results.

2. A new skewness test for scenario $\mathcal{S}_3$

For ease of presentation, let $X_1, X_2, \ldots, X_n$ be a random sample of size n from the normal distribution with mean μ and standard deviation σ, and $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$ be the corresponding order statistics. For simplicity, we let $n = 4Q + 1$ with Q being a positive integer. The five-number summary can then be represented as $a = X_{(1)}$, $q_1 = X_{(Q+1)}$, $m = X_{(2Q+1)}$, $q_3 = X_{(3Q+1)}$, and $b = X_{(n)}$. Lastly, by standardization, we let $X_i = \mu + \sigma Z_i$ for $i = 1, 2, \ldots, n$, or equivalently,

$$X_{(i)} = \mu + \sigma Z_{(i)}, \tag{2.1}$$

where $Z_1, Z_2, \ldots, Z_n$ are independent random variables from the standard normal distribution, with $Z_{(1)} \le Z_{(2)} \le \cdots \le Z_{(n)}$ being the order statistics.

To address the limitations of the test statistics T_3^S and T_3^B under scenario $\mathcal{S}_3 = \{a, q_1, m, q_3, b; n\}$, we introduce a new test statistic that integrates the five-number summary in a unified manner. By Lemma S1 in the Supplementary Material, we have $E(b - q_3) = E(q_1 - a)$ for normal data, showing that $\lambda = E(b - q_3)/E(q_1 - a)$ can serve as an effective measure of data skewness. Accordingly, we

https://doi.org/10.1017/rsm.2026.10111 Published online by Cambridge University Press

recommend testing the following hypotheses:

$$H_0 : \lambda = 1 \quad \text{versus} \quad H_1 : \lambda \neq 1.$$

If the null hypothesis is rejected, we then conclude that the data are significantly skewed and, consequently, not normally distributed.

To construct the test statistic, we take the sample ratio $(b - q_3)/(q_1 - a)$ as a proxy of λ. The sample ratio approaches one when the underlying distribution is symmetric. For testing whether $\lambda = 1$ under scenario $\mathcal{S}_3$, our new proposal for the test statistic is as follows:

$$T_3 = \log\left(\frac{b - q_3}{q_1 - a}\right). \tag{2.2}$$

Note that, by invoking (2.1), the above test statistic can also be expressed as

$$T_3 = \log\left(\frac{Z_{(n)} - Z_{(3Q+1)}}{Z_{(Q+1)} - Z_{(1)}}\right). \tag{2.3}$$

Since the right-hand side of (2.3) is a function of the order statistics of the standard normal distribution, it follows that the null distribution of T_3 depends only on the sample size n, but not on the parameters μ and σ. Additionally, it is noteworthy that the test statistic T_3^S in Shi et al.[18] is to compare the difference between the quantile distances, whereas our new test statistic T_3 is to compare the ratio between the quantile distances. This serves as another distinction between the two types of test statistics.

The null distribution of the test statistic T_3 is presented in Theorem S1 in the Supplementary Material, based on properties of order statistics.[20,21] Unlike T_3^S in (1.1) and T_3^B in (1.2) whose null distributions remain unknown, Theorem S1 in the Supplementary Material shows that the null distribution of the ratio-based test statistic T_3 can be explicitly derived; however, it is noted that the formula in the theorem involves several integrals and may not be readily accessible for practitioners. To obtain the critical values of $c_{3,\alpha/2}(n)$ for practical use, we apply the sampling method to numerically compute $c_{3,\alpha/2}(n)$ for $n = 4Q + 1$ with Q up to 100 and $\alpha = 0.05$, and present them in Table S1 in the Supplementary Material, computed from 1,000,000 simulations. For ease of implementation, an approximate formula for the sample size adjusted critical values is also provided as follows:

$$c_{3,0.025}(n) \approx 3.23/\log(n^{1.25} - 3.1) + 8.45/(n + 1.5). \tag{2.4}$$

Figure S1 in the Supplementary Material shows that the approximation accuracy of $c_{3,0.025}(n)$ is well guaranteed. Lastly, we note that the approximate formula in (2.4) is also applicable for any sample size n with non-integer Q. For example, when $n = 11$ with $Q = (11 - 1)/4 = 2.5$, the critical value will be given as $c_{3,0.025}(11) \approx 3.23/\log(11^{1.25} - 3.1) + 8.45/(11 + 1.5) = 1.8176$.

The simulated meta-analyses in Section S3 of the Supplementary Material were designed to evaluate two related issues: 1) whether using our new test to guide the estimator selection helps improve the estimation in meta-analysis compared with the benchmark methods and 2) how different skewness tests perform in the test-guided procedure for meta-analysis. Our findings show that the skewness test-guided approach produces pooled estimates that closely approximate those of the ideal case, where the true underlying distribution is known. In most scenarios, this strategy reduces bias and mean squared error for pooled effects, increases coverage probabilities, and yields narrower confidence intervals compared to methods that do not account for skewness. We have further applied our new test to real datasets in Section S4–S5 of the Supplementary Material to highlight the utility of the skewness test as a diagnostic tool for both guiding estimator selection and interpreting meta-analytic results.

https://doi.org/10.1017/rsm.2026.10111 Published online by Cambridge University Press

Table 1. *Summary table of the skewness tests under the three common scenarios.*

Scenario	Test statistic	Critical region of size 0.05
$\mathcal{S}_1 = \{a, m, b; n\}$	$T_1^S = \dfrac{a+b-2m}{b-a}$	$\lvert T_1^S\rvert > \dfrac{1}{\log(n+9)} + \dfrac{2.5}{n+1}$
$\mathcal{S}_2 = \{q_1, m, q_3; n\}$	$T_2^S = \dfrac{q_1+q_3-2m}{q_3-q_1}$	$\lvert T_2^S\rvert > \dfrac{2.65}{\sqrt{n}} + \dfrac{6}{n^2}$
$\mathcal{S}_3 = \{a, q_1, m, q_3, b; n\}$	$T_3 = \log\left(\dfrac{b-q_3}{q_1-a}\right)$	$\lvert T_3\rvert > \dfrac{3.23}{\log(n^{1.25}-3.1)} + \dfrac{8.45}{n+1.5}$

3. Conclusion and discussion

For clinical studies with continuous outcomes, meta-analytic approaches typically require the sample mean and standard deviation from each included study. However, some studies report alternative summary statistics, such as the minimum, maximum, median, and the first and third quartiles, especially when the outcome distribution is skewed. To better detect the skewness of data and thereby select more appropriate methods for converting the five-number summary back to the sample mean and standard deviation, we propose a novel skewness detection test, T_3, which utilizes the five-number summary and sample size. The new test statistic is constructed as the ratio of $b - q_3$ to $q_1 - a$, providing enhanced sensitivity to data skewness. Numerical simulations demonstrate that, compared to Shi et al.,[18] our new test appropriately controls the type I error rate while consistently achieving higher statistical power across all unimodal distributions. Given that the test statistic can be easily computed and interpreted without specialized software, it is particularly well suited for clinical researchers and evidence synthesis practitioners who may lack formal training in mathematical statistics. Accordingly, we recommend the use of T_3 under scenario $\mathcal{S}_3$ to detect the skewness of data from the five-number summary within a meta-analytic framework. Lastly, Table 1 summarizes the best test statistics for the three common scenarios, with the first two adopted from Shi et al.[18] and the last one following formula (2.2) of the present work.

Numerous methods exist for estimating the mean and standard deviation from summary data, some assuming normality and others accommodating non-normality. In contrast, the present work proposes a testing method based on the five-number summary, helping researchers select appropriate estimators and thereby enabling more reasonable meta-analysis. As such, it serves as an exploratory data analysis tool in the pre-meta-analysis stage. Based on our numerical results, we recommend routinely checking summary data for skewness prior to meta-analysis to guide estimator selection. If significant skewness is detected, subgroup analysis should be conducted to improve rigor and transparency. The proposed test can be applied to the overall meta-analysis or to predefined subgroups, and it detects true skewness more reliably than informal visual inspection.

High-quality meta-analysis represents the highest level of evidence in evidence-based medicine. By improving data diagnostic tools, our approach can indirectly reduce the risk of misleading conclusions derived from skewed five-number summary data, which have substantial implications for clinical practice and public health decision-making. However, it is important to acknowledge a limitation: the proposed statistic primarily captures differences in tail heaviness. When skewness mainly arises from a shift of central mass (e.g., mild skewness with nearly symmetric tails), the test's power may decrease. Therefore, results should be interpreted cautiously in such scenarios.

In addition, the conventional significance level of 0.05 is adopted in this article, although other levels can also be taken. The choice of significance level reflects a trade-off between type I error control and statistical power. A smaller level, such as 0.01, offers stricter control of the type I error but may reduce statistical power and increase the risk that some truly skewed studies are incorrectly classified as non-skewed. Applying methods designed for non-skewed data to truly skewed studies may introduce bias in the estimated means and standard deviations, thereby affecting the meta-analytic results. Conversely,

https://doi.org/10.1017/rsm.2026.10111 Published online by Cambridge University Press

using a larger significance level, such as 0.1, increases the ability to detect the skewness of data but also raises the risk of incorrectly classifying non-skewed studies as skewed. In such cases, methods intended for skewed data may be unnecessarily applied, potentially leading to a loss of efficiency in both study-level estimation and subsequent meta-analysis.

In future work, we will incorporate kurtosis features to develop a more comprehensive testing framework that addresses the current limitation regarding central mass shifts. By combining skewness and kurtosis diagnostics, the forthcoming framework aims to provide an even more robust and sensitive tool for assessing distributional assumptions in meta-analysis. We believe that this line of research will further strengthen the evidence base for systematic reviews and healthcare decision-making. Furthermore, although our numerical results show that the proposed test performs better than Shi et al.[18] and Balakrishnan et al.[19] for unimodal distributions, a rigorous theoretical justification is not yet available. Hence, establishing the theoretical superiority of the proposed test remains a promising direction for future work.

Acknowledgements. The authors would like to thank the editor, the associate editor, and two reviewers for their valuable suggestions, which considerably improved the manuscript.

Author contributions. Conceptualization: S.F. and T.T.; Formal analysis: S.F., H.L., and J.S.; Methodology: S.F., H.L., S.N.C., and T.T.; Supervision: S.N.C.; Writing—original draft: S.F.; Writing—review and editing: H.L., J.S., S.N.C., and T.T.

Competing interests. The authors declare that no competing interests exist.

Data availability statement. The R code used for the simulations and real-world data examples can be found at https://zenodo.org/records/21047915.

Funding statement. S.F.'s research was supported by the National Social Science Foundation of China (23BTJ061), the Humanities and Social Science Project of Ministry of Education of China (21YJC910003), the Natural Science Foundation of Henan Province (262300421873), and the Postgraduate Education Reform and Quality Improvement Project of Henan Province (YJS2026AL016). H.L.'s research was supported by the National Natural Science Foundation of China (12171310), the Shanghai "Project Dawn 2022" (22SG52), and the Basic Research Project of Shanghai Science and Technology Commission (22JC1400800). S.N.C.'s research was supported by the Research Matching Grant Scheme (RMGS-2022-11-08 and RMGS-2023-15-07) from the Research Grants Council of Hong Kong. T.T.'s research was supported by the General Research Fund of Hong Kong (HKBU12300123 and HKBU12303421) and the Initiation Grant for Faculty Niche Research Areas of Hong Kong Baptist University (RC-FNRA-IG/23-24/SCI/03).

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/rsm.2026.10111.

References

[1] Alba DE, Fernández-Durán JJ, Gregorio-Dominguez MM. Bayesian inference for the mean and standard deviation of a normal population when only the sample size, mean and range are observed. *J Appl Stat*. 2006;33: 89–99.
[2] Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol* 2014;14: 135.
[3] Luo D, Wan X, Liu J, Tong T. Optimally estimating the sample mean from the sample size, median, mid-range and/or mid-quartile range. *Stat Methods Med Res*. 2018;27: 1785–1805.
[4] Shi J, Luo D, Weng H, et al. Optimally estimating the sample mean and standard deviation from the five-number summary. *Res Synth Methods*. 2020;11: 641–654.
[5] Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol*. 2005;5: 13.
[6] Bland M. Estimating mean and standard deviation from the sample size, three quartiles, minimum, and maximum. *Int J Stat Med Res*. 2015;4: 57–64.
[7] McGrath S, Zhao X, Steele R, Thombs BD, Benedetti A. Estimating the sample mean and standard deviation from commonly reported quantiles in meta-analysis. *Stat Methods Med Res*. 2020;29: 2520–2537.
[8] Shi J, Tong T, Wang Y, Genton MG. Estimating the mean and variance from the five-number summary of a log-normal distribution. *Stat Interface*. 2020;13: 519–531.
[9] Cai S, Zhou J, Pan J. Estimating the sample mean and standard deviation from order statistics and sample size in meta-analysis. *Stat Methods Med Res*. 2021;30: 2701–2719.
[10] Kwon D, Reddy RRS, Reis IM. ABCMETAapp: R shiny application for simulation-based estimation of mean and standard deviation for meta-analysis via approximate Bayesian computation. *Res Synth Methods*. 2021;12: 842–848.

[11] Balakrishnan N, Rychtář J, Taylor D, Walter SD. Unified approach to optimal estimation of mean and standard deviation from sample summaries. *Stat Methods Med Res*. 2022;31: 2087–2103.
[12] Yang X, Hutson AD, Wang DL. A generalized BLUE approach for combining location and scale information in a meta-analysis. *J Appl Stat*. 2022;49: 3846–3867.
[13] Tang X, Tong T, Zhang X, Chu H. Minimum distance estimation of mean and standard deviation from reported quantiles. *Res Synth Methods*. 2026; In press. https://doi.org/10.1017/rsm.2026.10090.
[14] Lilliefors HW. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *J Am Stat Assoc*. 1967;62: 399–402.
[15] Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). *Biometrika*. 1965;52: 591–611.
[16] Jarque CM, Bera AK. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Econ Lett*. 1980;6: 255–259.
[17] Altman DG, Bland JM. Statistics notes: detecting skewness from summary information. *Br Med J*. 1996;313: 1200.
[18] Shi J, Luo D, Wan X, et al. Detecting the skewness of data from the five-number summary and its application in meta-analysis. *Stat Methods Med Res*. 2023;32: 1338–1360.
[19] Balakrishnan N, Rychtář J, Taylor D. Estimating sample skewness from sample data summaries and associated evaluation of normality. *Math Methods Stat*. 2023;32: 260–273.
[20] Ahsanullah M, Nevzorov VB, Shakil M. *An Introduction to Order Statistics*. Springer; 2013.
[21] Reiss RD. *Approximate Distributions of Order Statistics*. Springer; 1989.

Cite this article: Feng S, Lin H, Shi J, Chiu SN, Tong T. A new method for detecting the skewness of data from the five-number summary. *Research Synthesis Methods*. 2026;0: 1–7. https://doi.org/10.1017/rsm.2026.10111

https://doi.org/10.1017/rsm.2026.10111 Published online by Cambridge University Press