


RESEARCH ARTICLE  

A novel robust meta-analysis model using the t distribution for outlier accommodation and detection

Yue Wang¹, Jianhua Zhao ¹, Fen Jiang¹, Lei Shi¹ and Jianxin Pan²

¹School of Statistics and Mathematics, Yunnan University of Finance and Economics, Kunming, China

²Guangdong Provincial Key Laboratory of Interdisciplinary Research and Application for Data Science, BNU-HKBU United International College, Zhuhai, China

Corresponding author: Jianhua Zhao; Email: jhzhaoynu@gmail.com

Received: 7 June 2024; **Revised:** 10 November 2024; **Accepted:** 19 December 2024

Keywords: expectation maximization; meta-analysis; outlier accommodation; outlier detection; robustness

Abstract

Random effects meta-analysis model is an important tool for integrating results from multiple independent studies. However, the standard model is based on the assumption of normal distributions for both random effects and within-study errors, making it susceptible to outlying studies. Although robust modeling using the t distribution is an appealing idea, the existing work, that explores the use of the t distribution only for random effects, involves complicated numerical integration and numerical optimization. In this article, a novel robust meta-analysis model using the t distribution is proposed (t Meta). The novelty is that the marginal distribution of the effect size in t Meta follows the t distribution, enabling that t Meta can simultaneously accommodate and detect outlying studies in a simple and adaptive manner. A simple and fast EM-type algorithm is developed for maximum likelihood estimation. Due to the mathematical tractability of the t distribution, t Meta frees from numerical integration and allows for efficient optimization. Experiments on real data demonstrate that t Meta is compared favorably with related competitors in situations involving mild outliers. Moreover, in the presence of gross outliers, while related competitors may fail, t Meta continues to perform consistently and robustly.

Highlights

What is already know



Random effects model is a popular tool for handling heterogeneity between studies in meta-analysis. However, the standard model is based on the Gaussian assumption and thus is susceptible to outlying studies.

What is new

A novel robust meta-analysis model using student's t distribution called t Meta is proposed, which is capable of simultaneously accommodating and detecting outlying studies in a simple and adaptive manner. Empirical results show that t Meta is compared favorably with related competitors.

Potential impact for *Research Synthesis Methods* readers

Compared with related competitors, t Meta frees from numerical integration and allows for efficient optimization, which, to our knowledge, offers the first neat solution to robust meta-analysis modeling using the t distribution. Importantly, t Meta provides a simple but powerful robust meta-analysis tool that can accommodate and detect both mild and gross outliers simultaneously.

  This article was awarded Open Data and Open Materials badges for transparent practices. See the Data availability statement for details.

All the authors have contributed equally to this work.

© The Author(s), 2025. 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.

1. Introduction

In meta-analyses, the collected studies often exhibit heterogeneity, characterized by greater variation among studies than can be explained by the variation within each study,¹ which could result in misleading conclusions about the overall treatment effect.^{2,3} The random effects model is a popular tool for handling heterogeneity.^{4,5} However, the standard model assumes normal distributions for both random effects and within-study errors (nMeta), making it susceptible to outlying studies.

Outlier detection is a central research area in meta-analysis. Many methods have been developed. For example, a likelihood ratio test was constructed to identify outliers by detecting inflated variance⁶; a forward search algorithm was developed specifically for this purpose⁷; several outlier and influence diagnostic procedures in meta-regression models were presented.⁸ Subsequently, case deletion diagnostics and local influence analysis using multiple perturbation schemes, were investigated.⁹ Several Bayesian outlier detection measures were also introduced for handling outlying studies in network meta-analysis.¹⁰ Another important methodology for dealing with outliers is outlier accommodation or robust estimation, which can down-weight the influence of outliers. For instance, robust functions like Huber's rho and Tukey's biweight functions were employed to replace the original non-robust objective function, resulting in robust estimates.¹¹

This article focuses on outlier accommodation and detection simultaneously. Several efforts have been made toward this objective. Non-normal alternatives to normal random effects were investigated, and it was found that the t distribution performs the best (t RE-Meta).¹² The shortcoming is that the marginal distribution of y_i in t RE-Meta is mathematically intractable. Consequently, numerical integration is required to evaluate the log-likelihood and numerical optimization methods have to be employed for maximum likelihood (ML) estimation. Subsequently, new models where y_i has a tractable marginal distribution were presented, including the three parameter symmetric marginal model (SYM-Meta) and the four parameter skew marginal model (SKM-Meta).¹³ Nevertheless, numerical optimization has still to be employed to obtain ML estimates. As a tractable model, a variant of a two-component mixture model (MIX-Meta) was proposed, with one component modeling standard studies and the other addressing outlying studies.¹ In MIX-Meta, the marginal distribution of the observed effect y_i is a mixture of two normal distributions. However, MIX-Meta suffers from initialization issues, necessitating multiple runs of the fitting algorithm with different starting values.

The common feature of these methods is that the error terms are assumed to follow the normal distribution. In this article, we break this limitation as the marginal distribution of error term in our proposed model follows the t distribution. It is known that the t distribution includes the normal distribution as a special case when the degrees of freedom ν goes to infinity. This means that t Meta offers greater flexibility and applicability than the conventional normal assumption. The main contributions of this article are as follows.

- (i) The marginal distribution of the effect size y_i in t Meta follows the t distribution, enabling it to simultaneously accommodate and detect outliers in a simple and adaptive manner. 1) The t distribution offers an additional robustness tuning parameter which can adaptively down-weight outlying studies. 2) The expected weights follow in proportion to a Beta distribution, providing a useful critical value for outlier detection.
- (ii) t Meta provides a simple but powerful robust meta-analysis tool that can accommodate and detect both mild and gross outliers simultaneously. As can be seen from Section 4, 1) t Meta versus SYM-Meta and SKM-Meta. Both the three-parameter SYM-Meta and four-parameter SKM-Meta fail in most of the outlier detection tasks, though they have good performance in outlier accommodation. 2) t Meta versus t RE-Meta and MIX-Meta. While all the three methods can be used to detect mild outliers, t Meta performs the best in outlier accommodation. More importantly, in the presence of gross outliers, both t RE-Meta and MIX-Meta could fail while t Meta still performs satisfactorily.

(iii) Due to its mathematical tractability, *tMeta* frees from numerical integration and allows for efficient optimization. In contrast, *tRE-Meta* requires both complicated numerical integration and numerical optimization; *SYM-Meta* and *SKM-Meta* involve complex numerical optimization¹³; *MIX-Meta* requires multiple runs of the fitting algorithm due to the sensitivity issue of mixture models to initialization.¹ To our knowledge, *tMeta* offers the first neat solution to robust meta-analysis modeling using the *t* distribution.

The rest of this article is organized as follows. Section 2 reviews some related works. Section 3 proposes our new model *tMeta*. Section 4 conducts case studies to compare *tMeta* with several closely related competitors. Section 5 offers a summary of the entire article.

2. Background

In this section, we briefly review some fundamental results concerning the standard model *nMeta* and Student’s *t* distribution.

2.1. Normal meta-analysis model (*nMeta*)

In *nMeta*, the effect size y_i for the *i*-th study is defined as follows

$$y_i = \mu + b_i + e_i, \quad i = 1, \dots, N, \tag{1}$$

where the random effects b_i captures heterogeneity across studies and follows $\mathcal{N}(0, \sigma^2)$, the within-study error e_i follows $\mathcal{N}(0, s_i^2)$ and they are independent of each other. Here, μ is the overall effect size, σ^2 is the unknown between-study variance and s_i^2 is the known within-study variance.

From (1), we have $y_i \sim \mathcal{N}(\mu, \sigma^2 + s_i^2)$. Estimates for the parameters μ and σ^2 can be obtained through maximum likelihood methods.¹⁴

2.2. Student’s *t* distribution

Suppose that a random variable y follows the univariate *t* distribution $t(\mu, \sigma^2, \nu)$, with center $\mu \in \mathbb{R}$, scale parameter $\sigma^2 \in \mathbb{R}^+$, and degrees of freedom $\nu > 0$, then the probability density function (p.d.f.) of y is given by

$$f(y; \mu, \sigma^2, \nu) = \frac{\sigma^{-1} \Gamma(\frac{\nu+1}{2})}{(\pi\nu)^{\frac{1}{2}} \Gamma(\frac{\nu}{2})} \left\{ 1 + \frac{\delta^2(\mu, \sigma^2)}{\nu} \right\}^{-\frac{(\nu+1)}{2}},$$

where $\Gamma(\cdot)$ is the gamma function and $\delta^2(\mu, \sigma^2) = (y - \mu)^2 / \sigma^2$ is the squared Mahalanobis distance of y from the center μ with respect to σ^2 . If $\nu > 1$, $\mathbb{E}[y] = \mu$; if $\nu > 2$, $\text{Var}(y) = \nu\sigma^2 / (\nu - 2)$; and if $\nu \rightarrow \infty$, $t(\mu, \sigma^2, \nu) \rightarrow \mathcal{N}(\mu, \sigma^2)$.¹⁵

Given a latent weight variable τ distributed as the Gamma distribution $\text{Gam}(\nu/2, \nu/2)$, y can also be represented hierarchically as a latent variable model as follows¹⁵:

$$y|\tau \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{\tau}\right), \quad \tau \sim \text{Gam}\left(\frac{\nu}{2}, \frac{\nu}{2}\right). \tag{2}$$

Under model (2), it is easy to obtain the marginal distribution $y \sim t(\mu, \sigma^2, \nu)$ by $f(y; \mu, \sigma^2, \nu) = \int_0^\infty f(y|\tau)f(\tau)d\tau$ ¹⁶ and the posterior distribution of τ given y

$$\tau|y \sim \text{Gam}\left(\frac{\nu+1}{2}, \frac{\nu+\delta^2(\mu, \sigma^2)}{2}\right).$$

3. Novel robust meta-analysis model

In this section, we propose a novel robust meta-analysis model called *tMeta*. In Section 3.1, we present the model. In Section 3.2, we develop an algorithm for parameter estimation. In Section 3.3 and Section 3.4, we give the details for outlier accommodation and detection in *tMeta*.

3.1. The proposed *tMeta* model

Based on the hierarchical representation of the *t* distribution in Section 2.2, we propose a novel robust random effects meta-analysis model, denoted by *tMeta*. Its latent variable model can be expressed by

$$\begin{cases} y_i = \mu + b_i + e_i, & i = 1, \dots, N, \\ b_i|\tau_i \sim \mathcal{N}(0, \sigma^2/\tau_i), & e_i|\tau_i \sim \mathcal{N}(0, s_i^2/\tau_i) \\ \tau_i \sim \text{Gam}(\nu/2, \nu/2), \end{cases} \quad (3)$$

where, unlike *nMeta*, the random effects b_i and the within-study error e_i under *tMeta* are only conditionally independent; that is, b_i and e_i are mutually independent given the latent weight τ_i ; μ is the overall effect size, σ^2 is the unknown between-study variance, s_i^2 is the known within-study variance, and the degrees of freedom $\nu > 0$.

According to (3), integrating out the latent weight τ_i yields the marginal distributions $b_i \sim t(0, \sigma^2, \nu)$ and $e_i \sim t(0, s_i^2, \nu)$. Furthermore, using the property of the normal distribution, it is easy to obtain the conditional distribution of y_i given τ_i

$$y_i|\tau_i \sim \mathcal{N}\left(\mu, \frac{1}{\tau_i}(\sigma^2 + s_i^2)\right). \quad (4)$$

Integrating out the latent weight τ_i , we obtain an important result that the marginal distribution y_i follows a *t* distribution, that is,

$$y_i \sim t(\mu, \sigma^2 + s_i^2, \nu). \quad (5)$$

Note that this result is not available under *tRE-Meta* model, where the marginal distribution of y_i is mathematically intractable. This difference arises because *tMeta* and *tRE-Meta* model outliers in distinct ways. In *tRE-Meta*, outliers are assumed to result solely from extreme variation within studies. By contrast, as shown in (4), *tMeta* models the importance of a study i at the y_i -level by incorporating a latent weight τ_i associated with y_i to reflect the study's significance. The same τ_i is then naturally applied to both the between-study effect b_i and the within-study error e_i , as shown in (3). In other words, outliers in *tMeta* are assumed to result from extreme variation across both the within-study and between-study levels. This hierarchical modeling framework enables a tractable marginal model for the effect y_i .

As a result, the degrees of freedom ν in *tMeta* can be interpreted as an overall measure of deviation from the *nMeta* model across both within-study and between-study levels. The two models differ significantly when ν is small but become similar as ν becomes large. Similar overall measures have appeared in the literature; for example, a total correlation parameter has been used to capture overall correlation across both levels in the normal random-effects model.¹⁷ Notably, *nMeta* emerges as a

special case of *tMeta* in the limit, as the *t* distribution $t(\mu, \sigma^2 + s_i^2, \nu)$ approaches the normal distribution $\mathcal{N}(\mu, \sigma^2 + s_i^2)$ as $\nu \rightarrow \infty$.

3.1.1. Probability distributions

From *tMeta* model (3), it is easy to obtain the following probability distributions

$$\begin{aligned}
 y_i | b_i, \tau_i &\sim \mathcal{N}\left(\mu + b_i, \frac{s_i^2}{\tau_i}\right), \\
 b_i | y_i, \tau_i &\sim \mathcal{N}\left(\frac{\sigma^2(y_i - \mu)}{\sigma^2 + s_i^2}, \frac{\sigma^2 s_i^2}{\tau_i(\sigma^2 + s_i^2)}\right), \\
 b_i | y_i &\sim t\left(\frac{\sigma^2(y_i - \mu)}{\sigma^2 + s_i^2}, \frac{\sigma^2 s_i^2}{(\sigma^2 + s_i^2)}, \nu\right), \\
 \tau_i | y_i &\sim \text{Gam}\left(\frac{\nu + 1}{2}, \frac{\nu + \delta_i^2(\mu, \sigma^2)}{2}\right),
 \end{aligned}
 \tag{6}$$

where

$$\delta_i^2(\mu, \sigma^2) = \frac{(y_i - \mu)^2}{\sigma^2 + s_i^2},
 \tag{7}$$

is the squared Mahalanobis distance of y_i from the overall effect size μ . It is clear that all the probability distributions under *tMeta*, including the marginal distributions of b_i , e_i and y_i given in Section 3.1, are well-known and tractable.

3.1.2. Robust meta-regression with covariates

When several covariates are involved, the model (3) can be extended to a more general model,

$$y_i = \mathbf{x}'_i \boldsymbol{\beta} + b_i + e_i, \quad i = 1, \dots, N,$$

where \mathbf{x}_i represents p -dimensional vector of covariates, $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)'$ is the p -dimensional regression coefficients; the random variables b_i and e_i and the other parameters μ , σ^2 and ν are similar as those in *tMeta* (3). Under this model, we have $y_i \sim t(\mathbf{x}'_i \boldsymbol{\beta}, \sigma^2 + s_i^2, \nu)$.

3.2. Maximum likelihood estimation

In this section, we develop estimation algorithms for obtaining the ML estimates of the parameters $\boldsymbol{\theta} = (\mu, \sigma^2, \nu)$ in the *tMeta* model. Given the effect size vector $\mathbf{y} = (y_1, \dots, y_N)$, from (3) the observed data log-likelihood function \mathcal{L} is (up to a constant),

$$\begin{aligned}
 \mathcal{L}(\boldsymbol{\theta} | \mathbf{y}) &= -\frac{1}{2} \sum_{i=1}^N \{(\nu + 1)(\nu + \delta_i^2(\mu, \sigma^2)) + \ln(\sigma^2 + s_i^2)\} \\
 &\quad + N \left\{ \ln \Gamma\left(\frac{\nu + 1}{2}\right) - \ln \Gamma\left(\frac{\nu}{2}\right) + \frac{\nu}{2} \ln \nu \right\}.
 \end{aligned}
 \tag{8}$$

The maximization of \mathcal{L} in (8) can be obtained by standard numerical optimizers. However, we shall propose an EM-type algorithm to obtain the ML estimate $\hat{\boldsymbol{\theta}}$ because of its simplicity and stability.¹⁵ From (6), the required conditional expectation in the E-step can be obtained as

$$\tilde{\tau}_i \triangleq \mathbb{E}[\tau_i | y_i] = \frac{\nu + 1}{\nu + \delta_i^2(\mu, \sigma^2)}.
 \tag{9}$$

The details about the development of this algorithm can be found in Section A.1 of the Appendix.

3.3. Outlier accommodation

3.3.1. Adaptive outlier accommodation

Looking at (9), (A.2), and (A.3), the following can be observed.

- (i) When the data contain no outliers and the y_i 's come from nMeta, ν is expected to take on large values. This causes all the weights $\tilde{\tau}_i$ in (9) to be close to 1. Consequently, (A.2) and (A.3) would degenerate to those of nMeta, and hence tMeta adaptively degenerates to nMeta in this case.
- (ii) In the presence of outliers, ν is expected to take on small values, and the outlying study y_i would have a much greater squared Mahalanobis distance $\delta_i^2(\mu, \sigma^2)$ compared with non-outliers, causing the outlier's $\tilde{\tau}_i$ in (9) to be much smaller than those of non-outliers. Consequently, the impact of outliers on the estimators in (A.2) and (A.3) is substantially reduced, allowing tMeta to yield robust estimates.

In summary, the degrees of freedom ν is a robustness tuning parameter that adapts according to the presence of outliers in the data.

3.3.2. Breakdown point

In statistics, the robustness of estimators is assessed by breakdown points, which are the proportion of arbitrarily large outlying observations an estimator can tolerate before giving an incorrect result. The following Proposition 1 gives the breakdown point of tMeta.

Proposition 1. *The upper bound of the breakdown point of tMeta is $1/(\nu + 1)$.*

Proof. As proved by Dümbgen and Tyler,¹⁸ the upper bound of the breakdown point of the d -dimensional multivariate t distribution is $1/(\nu + d)$. For tMeta, the dimension of t -distributed y_i is $d = 1$ and hence the upper bound of tMeta is given by $1/(\nu + 1)$. This completes the proof. \square

In our implementation, we restrict $\nu \geq 1$. Proposition 1 shows that tMeta is a highly robust method as its breakdown point could be close to 50% under this restriction.

3.4. Outlier detection

Similar to that in multivariate t and matrix-variate t distributions,^{19–21} the expected weight $\tilde{\tau}_i$ in tMeta given by (9) can be used as outlier indicator. Let

$$u_i = \frac{N}{\hat{\sigma}^2 + s_i^2} \left/ \sum_{i=1}^N \frac{1}{\hat{\sigma}^2 + s_i^2} \right. . \quad (10)$$

The following Proposition 2 gives the details.

Proposition 2. *Assume that the study $\{y_i\}_{i=1}^N$ follow tMeta model (3). Given the ML estimate $\hat{\theta}$, we have, when the estimate $\hat{\sigma}^2 > 0$,*

$$\frac{1}{N} \sum_{i=1}^N u_i \tilde{\tau}_i = 1,$$

and when $\hat{\sigma}^2 = 0$,

$$\frac{1}{N} \sum_{i=1}^N u_i \tilde{\tau}_i \geq 1,$$

Proof. The proof can be found in Section A.2 of the Appendix. \square

Proposition 2 shows that when the estimate $\hat{\sigma}^2 > 0$, the average of all $u_i \tilde{\tau}_i$'s equals to 1. In other words, the study with $u_i \tilde{\tau}_i$ much smaller than 1 (i.e., $\tilde{\tau}_i$ much smaller than $1/u_i$) or close to 0 can be considered as an outlier. When $\hat{\sigma}^2 = 0$, our experience reveals that $\sum_{i=1}^N u_i \tilde{\tau}_i / N$ may be slightly greater than 1.

In practice, a critical value is needed to judge whether a study is an outlier or not. The following Proposition 3 does this task. Let $F(a, b)$ and $\text{Beta}(a, b)$ stand for the F distribution and Beta distribution with parameters a and b , respectively. The α quantile of $\text{Beta}(a, b)$ is denoted by $\text{Beta}_\alpha(a, b)$.

Proposition 3. *Suppose that the study set $\{y_i\}_{i=1}^N$ follow $t\text{Meta}$ model (3). Then we have that the Mahalanobis distance $\delta_i^2(\mu, \sigma^2) \sim F(1, \nu)$. Given the ML estimate $\hat{\theta}$, the weights $\tilde{\tau}_i$, $i = 1, \dots, N$ converge in distribution to $(1 + 1/\nu)\text{Beta}(\nu/2, 1/2)$ as the study sample size N approaches infinity. Therefore, at a significance level of α , the i -th study with $\tilde{\tau}_i < (1 + 1/\nu)\text{Beta}_\alpha(\nu/2, 1/2)$ could be identified as an outlier.*

Proof. This is a special case with dimension $d = 1$ of the result on the d -dimensional multivariate t distribution proved by Wang and Fun.¹⁹ This completes the proof. \square

4. Results

In this section, we compare the performance of our proposed $t\text{Meta}$ with five closely related methods: $n\text{Meta}$, $t\text{RE-Meta}$, MIX-Meta , SYM-Meta and SKM-Meta using four benchmark real-world datasets. For $t\text{Meta}$, the iteration stops when the relative change in the objective function \mathcal{L} ($|1 - \mathcal{L}^{(t)} / \mathcal{L}^{(t+1)}|$) is smaller than the given threshold $\text{tol} = 10^{-8}$ or the number of iterations exceeds $t_{\max} = 100$. For $n\text{Meta}$, $t\text{RE-Meta}$, and MIX-Meta , we use the R codes available from <https://cran.r-project.org/web/packages/metaplus/>. In addition, we use the default setting for MIX-Meta , i.e., 20 initializations. The code for SYM-Meta and SKM-Meta can be found from the supplementary materials by Baker and Jackson.¹³

To perform outlier detection for $t\text{Meta}$, we utilize the critical value provided in Proposition 3 and set the significance level $\alpha = 0.05$. For better visualization, we equivalently plot the inverse of $\tilde{\tau}_i$. That is, the study with $1/\tilde{\tau}_i > 1/((1 + 1/\nu)\text{Beta}_\alpha(\nu/2, 1/2))$ is identified as an outlier for $t\text{Meta}$. For MIX-Meta , we use the empirical threshold 0.9 as suggested by Beath,¹ which represents the posterior probability that a study belongs to the outlying component. For SYM-Meta and SKM-Meta , we adopt the p -value method specially developed for both models by Baker and Jackson.¹³ Since $t\text{RE-Meta}$ lacks guidelines for setting the threshold, we follow the empirical approach by Baker and Jackson,¹² treating studies with very small values of the relative weight ω_i / ω_i^0 , or equivalently, very large values of ω_i^0 / ω_i as outliers, where ω_i and ω_i^0 are the weights under $t\text{RE-Meta}$ and $n\text{Meta}$, respectively.

To compare the computational efficiency, we will report their total CPU time consumed by various methods, which is sum of the time used for training model and that for detecting outliers. For $t\text{Meta}$ and MIX-Meta , outlier detection is a byproduct of the model training and incurs no additional time cost. However, $t\text{RE-Meta}$, SYM-Meta and SKM-Meta require additional time cost for outlier detection. To be specific, $t\text{RE-Meta}$ requires numerical methods to compute ω_i^0 / ω_i while SYM-Meta and SKM-Meta necessitate additional efforts to implement the p -value method.

4.1. Intravenous magnesium

The Mag dataset²² comprises 16 studies. As can be seen from the forest plot shown in Figure 1(a), it looks difficult to visually identify which study is an outlier except that study 16 seems different from others due to its relatively large y_i value and low s_i^2 . Previous researches^{1,6} have analyzed this dataset and found no outliers. Below we perform outlier detection with various methods.

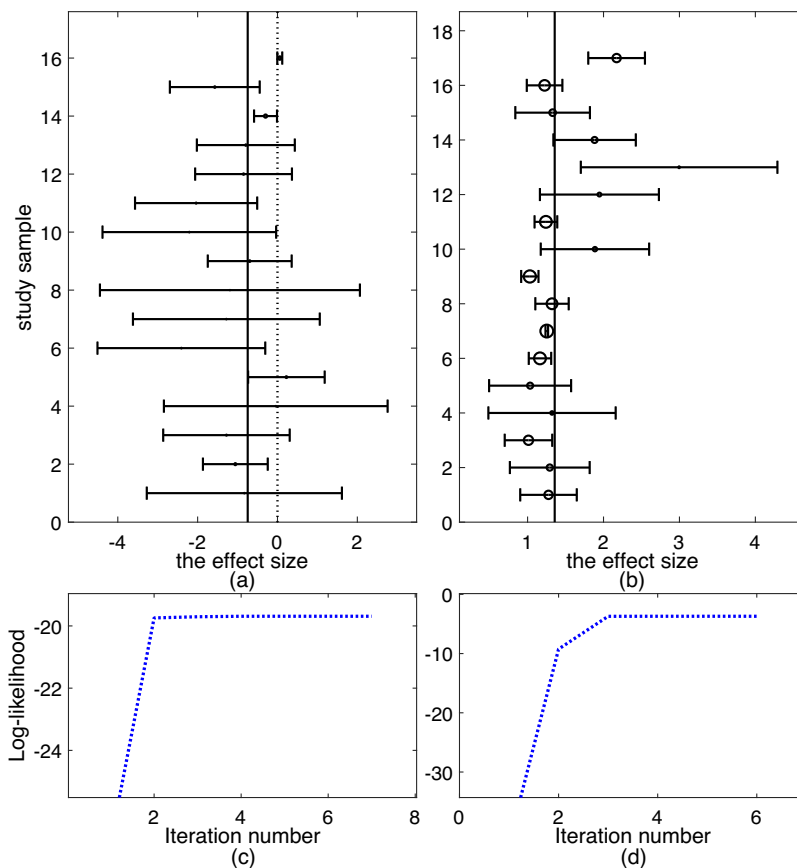


Figure 1. Top row: forest plots on two datasets: (a) Mag and (b) Hipfrac, where each effect size γ_i and 95% confidence interval are shown as circle and solid line, respectively. Bottom row: evolvement of log-likelihood of \mathcal{L} versus number of iterations: (c) Mag and (d) Hipfrac.

Table 1. Results of parameter estimates, negative log-likelihood, and CPU time (in seconds) by various methods on Mag dataset.

Methods	μ	σ	ν	$-\mathcal{L}$	Time
nMeta	-0.746	0.504	—	19.685	—
<i>t</i> RE-Meta	-0.746	0.504	inf	19.685	3.0
MIX-Meta	-0.746	0.504	—	19.685	32.4
SYM-Meta	-0.746	0.504	—	19.685	0.3
SKM-Meta	-0.746	0.504	—	19.685	0.3
<i>t</i> Meta	-0.746	0.504	inf	19.685	0.05

Note: The best method is shown in boldface. ‘—’ indicates that a method does not have corresponding results.

We fit all the six methods on the Mag dataset. Table 1 collects the results. The results in Table 1 show all the six methods yield similar performance. This means that all the five methods *t*RE-Meta, MIX-Meta, SYM-Meta, SKM-Meta and *t*Meta could degrade to nMeta. Nevertheless, among the five robust methods, *t*Meta is computationally the most efficient while *t*RE-Meta and MIX-Meta require much more time. Figure 1(c) shows the evolvement of log-likelihood \mathcal{L} versus number of iterations when fitting *t*Meta. It can be seen from Figure 1(c) that *t*Meta converges within 7 iterations on this dataset.

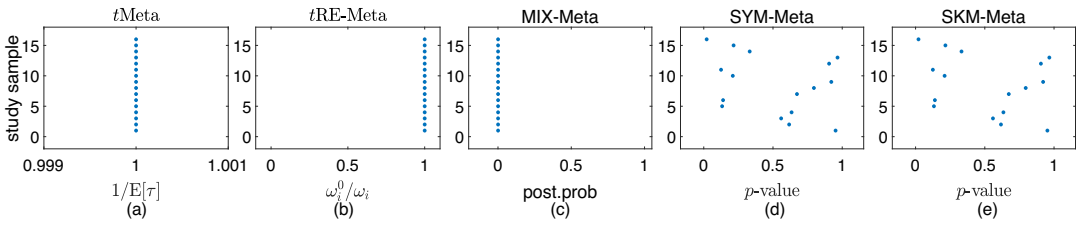


Figure 2. Results on outlier detection by various methods on Mag dataset: (a) *tMeta*; (b) *tRE-Meta*; (c) *MIX-Meta*; (d) *SYM-Meta*; (e) *SKM-Meta*. The marker solid point in blue represents normal studies judged by a method.

Table 2. Results of parameter estimates, negative log-likelihood, BIC, and CPU time (in seconds) by various methods on Hipfrac dataset.

Methods	μ	σ	ν	$-\mathcal{L}$	BIC	Time
nMeta	1.357	0.260	—	8.498	22.661	—
<i>tRE-Meta</i>	1.251	0.013	0.582	6.575	21.649	79.2
<i>MIX-Meta</i>	1.252	0.000	—	4.507	20.347	364.1
<i>SYM-Meta</i>	1.220	0.074	—	5.670	19.840	0.222
<i>SKM-Meta</i>	1.202	0.063	—	1.439	14.212	0.2
<i>tMeta</i>	1.252	0.000	1.871	3.700	15.899	0.03

Note: The best method is shown in boldface. ‘—’ indicates that a method does not have corresponding results.

Figure 2 shows the results of detecting outliers by the five methods. It can be seen from Figure 2 that all the five methods suggest no outliers for Mag dataset. This finding is consistent with that by Beath.¹

4.2. Hipfrac dataset

The Hipfrac dataset²³ contains 17 studies, collected from an investigation on the magnitude and duration of excess mortality after hip fracture among older men. Figure 1(b) shows the forest plot, from which it seems hard to identify which study is an outlier. Below we perform outlier analysis with various methods.

We fit all the six methods on the Hipfrac dataset. Table 2 summarizes the results. The results in Table 2 show that *tMeta* and *SKM-Meta* obtain significantly better BIC than the other methods and *SKM-Meta* wins by a narrow margin. In terms of computational efficiency among the five robust methods, *tMeta* is the fastest while *tRE-Meta* and *MIX-Meta* are the slowest runners. Figure 1(d) shows the evolution of log-likelihood \mathcal{L} versus number of iterations when fitting *tMeta*. It can be seen from Figure 1(d) that *tMeta* converges within 6 iterations on this dataset.

Figure 3 shows the results of detecting outliers by the five methods. It can be seen from Figure 3 that both *tMeta* and *tRE-Meta* identify study 17 as an outlier. This result is consistent with that obtained by Lin et al.² In contrast, *MIX-Meta* identifies one more outlier: study 9, while *SYM-Meta* and *SKM-Meta* fail completely.

It is interesting to make a further comparison between *tMeta* and *tRE-Meta*. From Figure 3(a), it can be seen that *tMeta* detects that study 9 as being close to an outlier, while it is not the case by *tRE-Meta* from Figure 3(b). In fact, Lin et al.² have considered study 9 as a potential outlier and perform a sensitivity analysis by removing this study. As a result, they found that study 9 is not influential. Therefore, the result by *tMeta* is well match that obtained by Lin et al.²

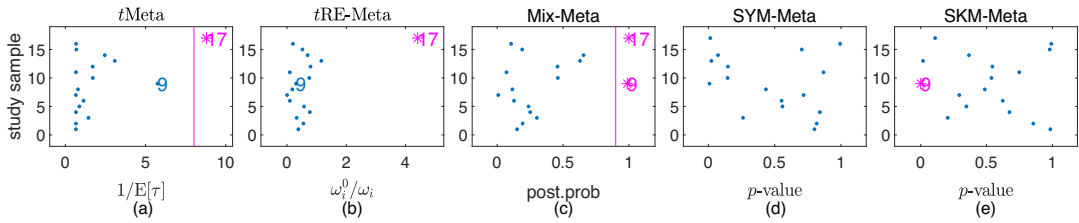


Figure 3. Results on outlier detection by various methods on Hipfrac dataset: (a) *tMeta*; (b) *tRE-Meta*; (c) *MIX-Meta*; (d) *SYM-Meta*; (e) *SKM-Meta*. The vertical line indicates the critical value for *tMeta* (red) and the threshold 0.9 (magenta) for *MIX-Meta*. The marker solid point in blue represents normal studies judged by a method. Star ‘*’ signals outlying studies, with red for *tMeta* and magenta for the other methods.

4.3. Fluoride toothpaste

This dataset contains 70 studies, obtained from an evaluation of fluoride’s efficacy in preventing childhood dental caries.²⁴ The effect size y_i denotes the difference between control and treatment groups, with negative values signifying significant therapeutic effects.

Previous works^{1,6,12} have concluded that there exist three outliers in this dataset: study 63, study 50 and study 38. Contrarily, the analysis with *SKM-Meta* suggests no outliers in the dataset.¹³ To better examine the outlier detection performance by various methods, we shall perform two experiments in this section. In the first experiment of Section 4.3, we use the original dataset (Flu). In the second experiment of Section 4.3, we add the original dataset with one more artificial outlier. The resulting dataset is called modified Flu for clarity.

4.3.1. Original Flu

Figure 4(a) shows the forest plot of the original dataset Flu. It can be observed from Figure 4(a) that studies 38, 50, and 63 look like abnormal. We then perform further analysis to identify outliers.

We fit all the six methods on Flu. Table 3 summarizes the results. The results in Table 3 show that *tMeta*, *SYM-Meta* and *SKM-Meta* obtain substantially better BIC than the other methods and *SKM-Meta* is again the best. Among the five robust methods, *tMeta* is the most computationally efficient while *tRE-Meta* and *MIX-Meta* are the most inefficient. Figure 4(c) shows the evolution of log-likelihood \mathcal{L} versus number of iterations when fitting *tMeta*. It can be seen from Figure 4(c) that *tMeta* converges within 18 iterations on this dataset.

The top row in Figure 5 shows the results of detecting outliers by the five methods. It can be seen from Figure 5 that *tMeta*, *tRE-Meta* and *MIX-Meta* all identify three studies: 63, 50, 38. This means that the result by *tMeta* is consistent with those in previous works.^{1,6,12} In contrast, *SYM-Meta* only detects the most abnormal study 63 as one outlier while *SYM-Meta* identify no outlier.

4.3.2. Modified Flu

In the modified Flu, the outlier (study 71) is introduced as follows. The effect size y_{71} is generated from the uniform distribution U on the interval $[1, 2]$, i.e., $y_{71} \sim U(1, 2)$. We set its within-study variance $s_{71}^2 = 1/12$. Figure 4(b) shows the forest plot of modified Flu, from which it can be seen that the newly added study 71 looks like a mild outlier as it is very different from all the other studies.

Table 3 summarizes the results by six methods. The results in Table 3 show that *tMeta* yields the best BIC on this dataset, which is then followed by *SYM-Meta*, and *SKM-Meta* is the third best. Among the five robust methods, *tMeta* is again the best performer in computational efficiency while *tRE-Meta* and *MIX-Meta* are still the most inefficient. Figure 4(d) shows the evolution of log-likelihood \mathcal{L} versus number of iterations when fitting *tMeta*. It can be seen from Figure 4(d) that *tMeta* converges within 19 iterations on this dataset.

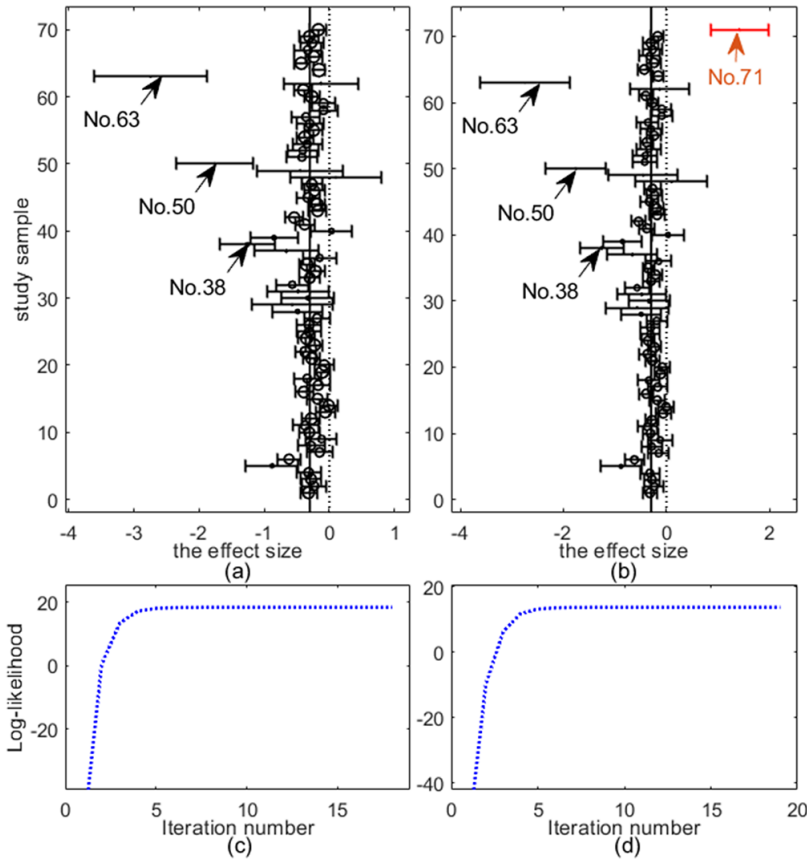


Figure 4. Top row: forest plots on the fluoride toothpaste dataset: (a) Flu and (b) modified Flu, where each effect size γ_i and 95% confidence interval are shown as circle and solid line, respectively. Bottom row: evolution of log-likelihood of \mathcal{L} versus number of iterations: (c) Flu and (d) modified Flu.

The bottom row in Figure 5 shows the results of detecting outliers by the five methods. It can be seen from the bottom row of Figure 5 that *t*Meta, *r*RE-Meta and MIX-Meta successfully identify four outliers: 71, 63, 50, 38. In contrast, SYM-Meta fails to detect any outlier, while SKM-Meta can detect the newly added study 71.

4.4. CDP-choline

The CDP-choline dataset²⁵ is obtained by exploring the cytidinediphosphocholine analysis in cognitive and behavioural disorders associated with chronic brain diseases in the elderly. The sample size is $N = 10$.

Previous analyses^{1,6,12} have concluded that there is one outlier in this dataset: study 8. Like Section 4.3, we perform two experiments. In the first experiment of Section 4.4, we use the original dataset (CDP). In the second experiment of Section 4.4, we modify CDP so that it contains more outliers, which is denoted by modified CDP for clarity.

4.4.1. Original CDP

Figure 6(a) shows the forest plot of the original CDP. It can be observed from Figure 6(a) that study 8 looks like abnormal. We then perform further analysis to identify outliers.

We fit all the six methods on CDP. Table 4 summarizes the results. The results in Table 4 show that *t*Meta, SYM-Meta and SKM-Meta obtain significantly better BIC than the other methods and

Table 3. Results by various methods on the original and modified fluoride toothpaste dataset, including parameter estimates, negative log-likelihood, BIC, and CPU time (in seconds).

Methods	μ	σ	ν	$-\mathcal{L}$	BIC	Time
Original Flu						
nMeta	-0.300	0.119	—	1.233	10.963	—
tRE-Meta	-0.280	0.049	1.158	-13.121	-13.497	64.0
MIX-Meta	-0.281	0.090	—	-14.636	-12.277	27.0
SYM-Meta	-0.282	0.092	—	-17.148	-21.551	0.4
SKM-Meta	-0.273	0.081	—	-21.914	-26.834	0.6
tMeta	-0.282	0.051	2.754	-18.283	-23.820	0.05
Modified Flu						
nMeta	-0.297	0.139	—	15.760	40.046	—
tRE-Meta	-0.279	0.047	1.023	-7.774	-2.761	59.2
MIX-Meta	-0.280	0.088	—	-10.062	-3.072	26.9
SYM-Meta	-0.282	0.092	—	-12.399	-12.010	0.7
SKM-Meta	-0.277	0.088	—	-13.144	-9.238	0.6
tMeta	-0.281	0.047	2.367	-13.791	-14.794	0.05

Note: The best method is shown in boldface. ‘—’ indicates that a method does not have corresponding results.

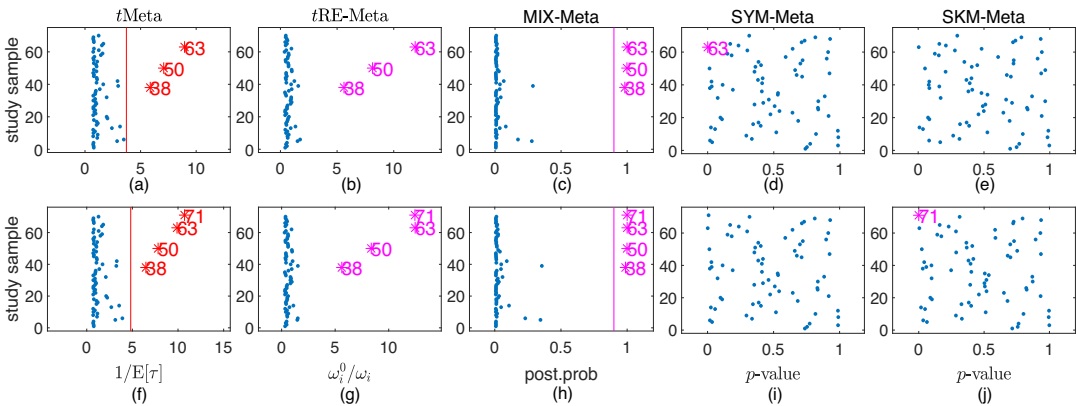


Figure 5. Results on outlier detection by various methods on fluoride toothpaste dataset. Top row: the original dataset; Bottom row: the modified dataset. (a), (f) tMeta; (b), (g) tRE-Meta; (c), (h) MIX-Meta; (d), (i) SYM-Meta; (e), (j) SKM-Meta. The vertical line indicates the critical value for tMeta and the threshold 0.9 for MIX-Meta. The marker solid point in blue represents normal studies judged by a method. Star ‘*’ signals outlying studies, with red for tMeta and magenta for the other methods.

SKM-Meta is again the best. In terms of computational efficiency among the five robust methods, tMeta is the most efficient while tRE-Meta and MIX-Meta are the slowest runners. Figure 6(c) shows the evolution of log-likelihood \mathcal{L} versus number of iterations when fitting tMeta. It can be seen from Figure 6(c) that tMeta converges within 10 iterations on this dataset.

The top row in Figure 7 shows the results of detecting outliers by the five methods. It can be seen from Figure 5 that all the five methods successfully identify study 8 as an outlier. This means that the results by tMeta, SYM-Meta and SKM-Meta are consistent with those in previous works.^{1,6,12}

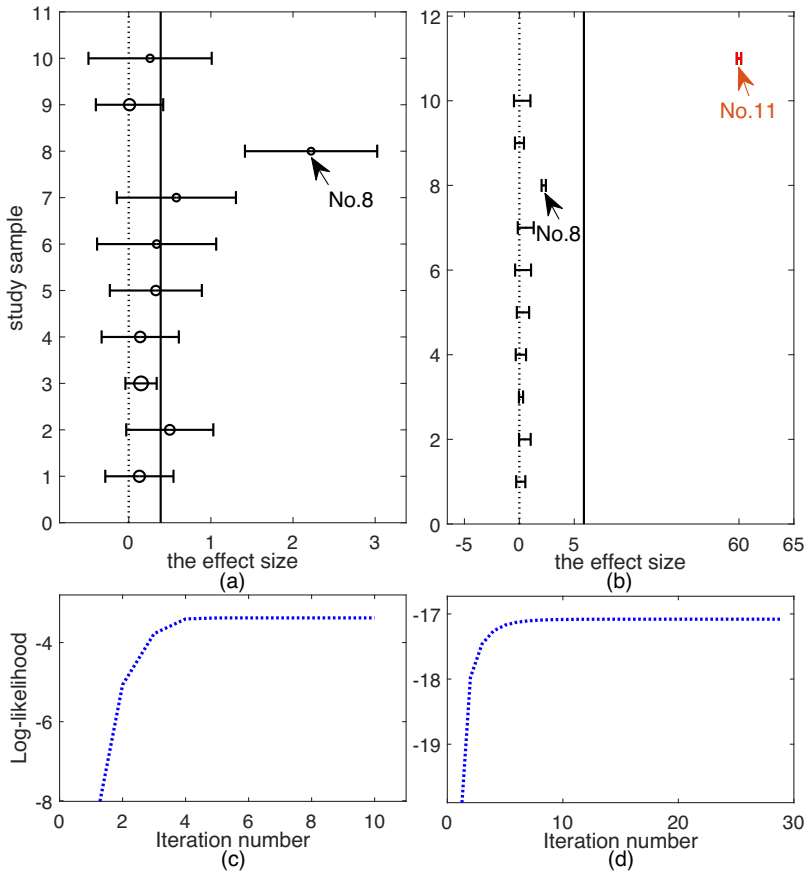


Figure 6. Top row: forest plots on CDP-choline dataset: (a) original dataset; (b) modified dataset. Bottom row: evolution of log-likelihood of \mathcal{L} versus number of iterations: (c) original dataset and (d) modified dataset.

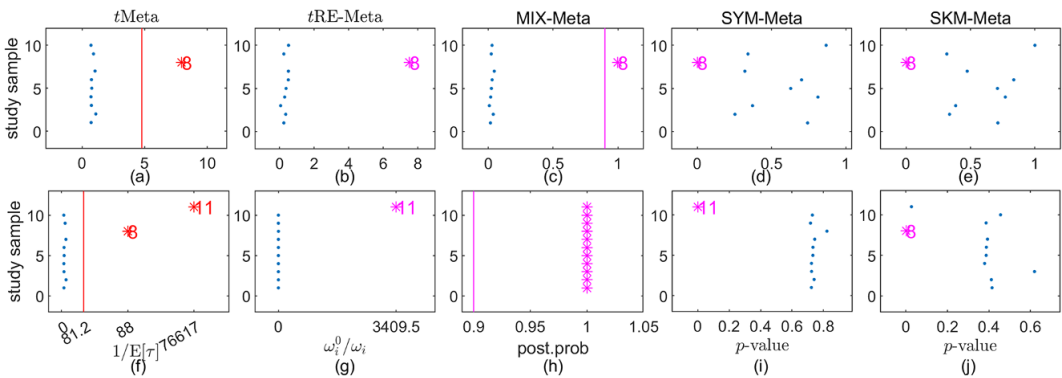


Figure 7. Results on outlier detection by various methods on CDP-choline dataset. Top row: the original dataset; Bottom row: the modified dataset. (a), (f) tMeta; (b), (g) tRE-Meta; (c), (h) MIX-Meta; (d), (i) SYM-Meta; (e), (j) SKM-Meta. The vertical line indicates the critical value for tMeta and the threshold 0.9 for MIX-Meta. The marker solid point \bullet in blue represents normal studies judged by a method. Star ‘*’ signals outlying studies, with red for tMeta and magenta for the other methods.

Table 4. Results by various methods on the CDP-choline dataset, including parameter estimates, negative log-likelihood, BIC, and CPU time (in seconds).

Methods	μ	σ	ν	$-\mathcal{L}$	BIC	Time
Original CDP						
nMeta	0.389	0.383	—	8.199	21.002	—
<i>t</i> RE-Meta	0.195	0.006	0.494	4.058	15.024	24.7
MIX-Meta	0.191	1.777	—	3.007	15.225	47
SYM-Meta	0.194	0.000	—	2.847	12.602	0.14
SKM-Meta	0.193	0.000	—	1.403	12.016	0.2
<i>t</i> Meta	0.187	0.000	2.380	3.377	13.662	0.03
Modified CDP						
nMeta	5.879	17.126	—	46.855	98.506	—
<i>t</i> RE-Meta	0.193	0.002	0.273	13.768	34.729	65.9
MIX-Meta	5.879	2.455	—	46.855	103.302	19.5
SYM-Meta	5.880	17.117	—	46.855	100.904	0.3
SKM-Meta	0.484	0.711	—	21.622	52.836	0.4
<i>t</i> Meta	0.200	0.115	1.000	17.081	41.355	0.03

Note: The best method is shown in boldface. ‘—’ indicates that a method does not have corresponding results.

4.4.2. Modified CDP

In the modified CDP, we make two modifications: (i) add one outlier, namely study 11, which is set as $y_{11} = 60, s_{11}^2 = 0.01$; (ii) set $s_8^2 = 0.01$. Figure 6(b) shows the forest plot of modified CDP, from which it can be seen that the newly added study 11 is a gross outlier as it is extremely different from all the other studies and study 8 is a mild outlier but now it has a smaller within-study variance than that in the original CDP.

Table 4 summarizes the results by six methods. The results in Table 4 show that *t*RE-Meta and *t*Meta have better BIC than the other methods on this dataset and *t*RE-Meta is the best. In terms of computational efficiency among the five robust methods, *t*Meta is still the most efficient while *t*RE-Meta and MIX-Meta demand the most time. Figure 6(d) shows the evolution of log-likelihood \mathcal{L} versus number of iterations when fitting *t*Meta. It can be seen from Figure 6(d) that *t*Meta requires 29 iterations to converge on this dataset.

The bottom row in Figure 7 shows the results of detecting outliers by the five methods. It can be seen that *t*Meta performs reliably as it successfully detects the two outliers: study 11, 8. In contrast, MIX-Meta and SKM-Meta fail to detect the most extreme study 11. *t*RE-Meta and SYM-Meta can identify study 11 but they fail to detect study 8.

5. Conclusion

For outlier accommodation and detection simultaneously, in this article we propose a novel robust meta-analysis model using student’s *t* distribution, namely *t*Meta. *t*Meta can be expressed as a hierarchical latent variable model while the marginal distribution of the effect size y_i follows a tractable *t* distribution. To obtain the ML estimates of the parameters, we develop an ECME algorithm, which is computationally much more efficient than related methods as shown in our experiments. Empirical results on real datasets show that *t*Meta not only improves the robustness of nMeta as expected but also is compared favorably with closely related competitors in that it can provide the best performance for outlier accommodation and detection simultaneously, for both mild and gross outliers.

The experiment results show that SKM-Meta on some datasets yields better performance in outlier accommodation. For future work it would be interesting to extend *t*Meta using the skew-*t* distribution for further accommodating skewed data.

Acknowledgements. The authors thank the Associate Editor and the anonymous reviewers for their insightful comments, which have contributed to an improved version of the manuscript.

Author contributions. Y.W.: Methodology; data curation; formal analysis; investigation; validation; writing—original draft. J.Z.: Conceptualization; methodology; supervision; funding acquisition; writing—review and editing. F.J.: Methodology, investigation, validation, writing—original draft. L.S.: Methodology; funding acquisition; writing—review and editing. J.P.: Resources; funding acquisition; writing – review and editing.

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

Data availability statement. The code to reproduce the results in our experiments is provided in the online supplementary materials and is also available on GitHub: the R code is available at <https://github.com/wangyue4127/tmeta-R-code.git>, and the MATLAB code can be found at <https://github.com/wangyue4127/tmeta.git>.

Funding statement. This work was supported partly by the National Natural Science Foundation of China under Grant 12161089, Grant 11931015, Grant 12271471 and Grant 12271047; partly by Yunnan Province ‘Xingdian Talent Support Program’ under Grant of J.H. Zhao and Grant of L. Shi; partly by Yunnan Fundamental Research Projects under Grant 202201AU070105; partly by the Science Foundation of Yunnan University of Finance and Economics under Grant 2021D10 and Grant 2024YUFEYC085; partly by the Science Foundation of Yunnan Educational Department under Grant 2025Y0755 and Grant 2025Y0802; partly by the National Philosophy and Social Science Fund under Grant 22&ZD158 and Grant 22VRCO49; partly by the Guangdong Provincial Key Laboratory of Interdisciplinary Research and Application for Data Science, BNU-HKBU United International College under Grant 2022B1212010006.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/rsm.2025.8>.

References

- [1] Beath KJ. A finite mixture method for outlier detection and robustness in meta-analysis. *Res. Synth. Methods* 2014;5(4): 285–293. <https://doi.org/10.1002/jrsm.1114>.
- [2] Lin L, Chu H, Hodges JS. Alternative measures of between-study heterogeneity in meta-analysis: Reducing the impact of outlying studies. *Biometrics* 2017;73(1): 156–166. <https://doi.org/10.1111/biom.12543>.
- [3] Noma H, Nagashima K, Kato S, Teramukai S, Furukawa TA. Meta-analysis using flexible random-effects distribution models. *J. Epidemiol.* 2022;32(10): 441–448. <https://doi.org/10.2188/jea.JE20200376>.
- [4] Hardy RJ, Thompson SG. Detecting and describing heterogeneity in meta-analysis. *Stat. Med.* 1998;17(8): 841–856. [https://doi.org/10.1002/\(SICI\)1097-0258\(19980430\)17:8<841::AID-SIM781>3.0.CO;2-D](https://doi.org/10.1002/(SICI)1097-0258(19980430)17:8<841::AID-SIM781>3.0.CO;2-D).
- [5] Wang Y, Lin L, Thompson CG, Chu H. A penalization approach to random-effects meta-analysis. *Stat. Med.* 2022;41(3): 500–516. <https://doi.org/10.1002/sim.9261>.
- [6] Gumedze FN, Jackson D. A random effects variance shift model for detecting and accommodating outliers in meta-analysis. *BMC Med. Res. Methodol.* 2011;11: 1–9. <https://doi.org/10.5812/ircmj.17648>.
- [7] Mavridis D, Moustaki I, Wall M, Salanti G. Detecting outlying studies in meta-regression models using a forward search algorithm. *Res. Synth. Methods* 2017;8(2): 199–211. <https://doi.org/10.1002/jrsm.1197>.
- [8] Viechtbauer W, Cheung MWL. Outlier and influence diagnostics for meta-analysis. *Res. Synth. Methods* 2010;1(2): 112–125. <https://doi.org/10.1002/jrsm.11>.
- [9] Shi L, Zuo S, Yu D, Zhou X. Influence diagnostics in meta-regression model. *Res. Synth. Methods* 2017;8(3): 343–354. <https://doi.org/10.1002/jrsm.1247>.
- [10] Zhang J, Fu H, Carlin BP. Detecting outlying trials in network meta-analysis. *Stat. Med.* 2015;34(19): 2695–2707. <https://doi.org/10.1002/sim.6509>.
- [11] Yu D, Ding C, He N, Wang R, Zhou X, Shi L. Robust estimation and confidence interval in meta-regression models. *Comput. Stat. Data Anal.* 2019;129: 93–118. <https://doi.org/10.1016/j.csda.2018.08.010>.
- [12] Baker R, Jackson D. A new approach to outliers in meta-analysis. *Health Care Manag. Sci.* 2008;11(3): 121–131. <https://doi.org/10.1007/s10729-007-9041-8>.
- [13] Baker R, Jackson D. New models for describing outliers in meta-analysis. *Res. Synth. Methods* 2016;7(3): 314–328. <https://doi.org/10.1002/jrsm.1191>.
- [14] Hardy RJ, Thompson SG. A likelihood approach to meta-analysis with random effects. *Stat. Med.* 1996;15(6): 619–629. [https://doi.org/10.1002/\(SICI\)1097-0258\(19960330\)15:6<619::AID-SIM188>3.0.CO;2-A](https://doi.org/10.1002/(SICI)1097-0258(19960330)15:6<619::AID-SIM188>3.0.CO;2-A).
- [15] Liu C, Rubin DB. ML estimation of the *t* distribution using EM and its extensions, ECM and ECME. *Stat. Sin.* 1995;5(1): 19–39.
- [16] Zhao J, Jiang Q. Probabilistic PCA for *t* distributions. *Neurocomputing* 2006;69(16–18): 2217–2226. <https://doi.org/10.1016/j.neucom.2005.07.011>.

[17] Riley RD, Thompson JR, Abrams KR. An alternative model for bivariate random-effects meta-analysis when the within-study correlations are unknown. *Biostatistics* 2008;9(1): 172–186.

[18] Dümmbgen L, Tyler DE. On the breakdown properties of some multivariate M-functionals. *Scand. J. Stat.* 2005;32(2): 247–264. <https://doi.org/10.1111/j.1467-9469.2005.00425.x>.

[19] Wang WL, Fan TH. Estimation in multivariate t linear mixed models for multiple longitudinal data. *Stat. Sin.* 2011;21: 1857–1880. <https://doi.org/10.5705/ss.2009.306>.

[20] Ma X, Zhao J, Wang Y, Shang C, Jiang F. Robust factored principal component analysis for matrix-valued outlier accommodation and detection. *Comput. Stat. Data Anal.* 2023;179: 107657. <https://doi.org/10.1016/j.csda.2022.107657>.

[21] Zhao J, Ma X, Shi L, Wang Z. Robust bilinear probabilistic PCA using a matrix variate t distribution. *IEEE Trans. Neural Netw. Learn. Syst.* 2023;34(12): 10683–10697. <https://doi.org/10.1109/TNNLS.2022.3170797>.

[22] Sterne JAC, Bradburn MJ, Egger M. Meta-analysis in Stata. In: Egger M, Smith GD, Altman DG. eds. *Systematic Reviews in Health Care: Meta-Analysis in Context*. 2nd ed. BMJ Publishing Group; 2001.

[23] Haentjens P, Magaziner J, Colón-Emeric CS, et al. Meta-analysis: Excess mortality after hip fracture among older women and men. *Ann. Internal Med.* 2010;152(6): 380–390. <https://doi.org/10.7326/0003-4819-152-6-201003160-00008>.

[24] Marinho VC, Higgins J, Logan S, Sheiham A. Fluoride toothpastes for preventing dental caries in children and adolescents. 2002. <http://www.cochrane.org/reviews/en/ab002278.html>.

[25] Fioravanti M, Yanagi M. Cytidinediphosphocholine (CDP-choline) for cognitive and behavioural disturbances associated with chronic cerebral disorders in the elderly(review). 2005. <http://www.cochrane.org/reviews/en/ab000269.html>.

A. Appendix

A.1. The proposed ECME algorithm for tMeta

In this subsection, we develop an Expectation Conditional Maximization of Either (ECME) algorithm, a variant of the EM algorithm with faster monotone convergence. [15] Our ECME consists of an E-step followed by three conditional maximization (CM)-steps. In each CM step, a parameter in $\theta = (\mu, \sigma^2, \nu)$ is maximized while keeping the others fixed.

Let $\tau = (\tau_1, \dots, \tau_N)$ be the missing data. From (3), the log-likelihood function of complete data (\mathbf{y}, τ) is given by

$$\mathcal{L}_c(\theta|\mathbf{y}, \tau) = \sum_{i=1}^N \ln\{p(y_i|\tau_i)p(\tau_i)\}.$$

E-step: Compute the expected complete data log-likelihood function \mathcal{L}_c with respect to the conditional distribution $p(\tau|\mathbf{y}, \theta)$,

$$Q(\theta) = \mathbb{E}[\mathcal{L}_c(\theta|\mathbf{y}, \tau)|\mathbf{y}] = Q_1(\mu, \sigma^2) + Q_2(\nu),$$

where, up to a constant

$$Q_1(\mu, \sigma^2) = -\frac{1}{2} \sum_{i=1}^N \{ \ln(\sigma^2 + s_i^2) + \mathbb{E}[\tau_i|y_i] \delta_i^2(\mu, \sigma^2) \}. \tag{A.1}$$

Here, $\delta_i^2(\mu, \sigma^2)$ is given by (7). The required conditional expectation can be obtained by (9).

In our ECME, the first two CM-steps maximize Q while the third CM-step maximize \mathcal{L} . In detail, **CM-step 1:** Given (σ^2, ν) , maximize Q_1 in (A.1) with respect to μ yielding

$$\tilde{\mu} = \sum_{i=1}^N \frac{\tilde{\tau}_i y_i}{\sigma^2 + s_i^2} \bigg/ \sum_{i=1}^N \frac{\tilde{\tau}_i}{\sigma^2 + s_i^2}. \tag{A.2}$$

CM-step 2: Given $(\tilde{\mu}, \nu)$, maximize Q_1 in (A.1) with respect to σ^2 under the same restriction $\tilde{\sigma}^2 \geq 0$ as in nMeta, [9] yielding

$$\begin{aligned} \sigma_t^2 &= \sum_{i=1}^N \frac{\tilde{\tau}_i (y_i - \tilde{\mu})^2 - s_i^2}{(\sigma^2 + s_i^2)^2} \bigg/ \sum_{i=1}^N \frac{1}{(\sigma^2 + s_i^2)^2}, \\ \tilde{\sigma}^2 &= \max \{ \sigma_t^2, 0 \}. \end{aligned} \tag{A.3}$$

CM-step 3: Given $(\tilde{\mu}, \tilde{\sigma}^2)$, maximize the observed data log-likelihood function \mathcal{L} in (8) w.r.t. ν . This is equivalent to finding the root of the following equation

$$\begin{aligned} \mathcal{L}'(\nu) = & -\psi\left(\frac{\nu}{2}\right) + \ln\left(\frac{\nu}{2}\right) + 1 + \psi\left(\frac{\nu+1}{2}\right) - \ln\left(\frac{\nu+1}{2}\right) \\ & + \frac{1}{N} \sum_{i=1}^N \left\{ \ln\left(\frac{\nu+1}{\nu + \tilde{\delta}_i^2}\right) - \left(\frac{\nu+1}{\nu + \tilde{\delta}_i^2}\right) \right\} = 0, \end{aligned} \tag{A.4}$$

where $\tilde{\delta}_i^2 = \delta_i^2(\tilde{\mu}, \tilde{\sigma}^2)$, and $\psi(x) = d \ln(\Gamma(x))/dx$ is the digamma function. Solving (A.4) only requires one-dimensional search, which can be performed, e.g., by the bisection method. [15]

For clarity, the complete ECME algorithm is summarized in Algorithm 1.

Algorithm 1 The ECME algorithm for *tMeta*

Input: Data \mathbf{y} and initialization of $\theta=(\mu, \sigma^2, \nu)$.

1. **repeat**

- 2. *E-step:* Compute $\tilde{\tau}_i$ via (9).
- 3. *CM-step 1:* Update $\tilde{\mu}$ via (A2).
- 4. *CM-step 2:* Update $\tilde{\sigma}^2$ via (A3).
- 5. *CM-step 3:* Update $\tilde{\nu}$ via (A4).
- 6. **until** the relative change of \mathcal{L} in (8) is smaller than a threshold.

Output: $\hat{\theta} = (\tilde{\mu}, \tilde{\sigma}^2, \tilde{\nu})$.

A.2. Proof for Proposition 2

Proof. For ML estimate $\hat{\theta}$, multiplying (9) by $\hat{\nu} + \delta_i^2(\hat{\mu}, \sigma^2)$, we obtain

$$\hat{\nu} + 1 = \hat{\nu}\tilde{\tau}_i + \tilde{\tau}_i\delta_i^2(\hat{\mu}, \hat{\sigma}^2). \tag{A.5}$$

On both sides of (A.3), multiply by $\sum_{i=1}^N 1/(\hat{\sigma}^2 + s_i^2)^2$ and then add $\sum_{i=1}^N s_i^2/(\hat{\sigma}^2 + s_i^2)^2$. On noting (7), when $\hat{\sigma}^2 > 0$, we have

$$\sum_{i=1}^N \frac{\tilde{\tau}_i\delta_i^2(\hat{\mu}, \hat{\sigma}^2)}{\hat{\sigma}^2 + s_i^2} = \sum_{i=1}^N \frac{1}{\hat{\sigma}^2 + s_i^2}, \tag{A.6}$$

and when $\hat{\sigma}^2 = 0$, we have

$$\sum_{i=1}^N \frac{\tilde{\tau}_i\delta_i^2(\hat{\mu}, \hat{\sigma}^2)}{\hat{\sigma}^2 + s_i^2} \leq \sum_{i=1}^N \frac{1}{\hat{\sigma}^2 + s_i^2}. \tag{A.7}$$

On both sides of (A.5), divide by $\hat{\sigma}^2 + s_i^2$ and take the sum over i from 1 to N , yielding

$$\sum_{i=1}^N \frac{\hat{\nu} + 1}{\hat{\sigma}^2 + s_i^2} = \sum_{i=1}^N \frac{\tilde{\tau}_i\delta_i^2(\hat{\mu}, \hat{\sigma}^2)}{\hat{\sigma}^2 + s_i^2} + \sum_{i=1}^N \frac{\hat{\nu}\tilde{\tau}_i}{\hat{\sigma}^2 + s_i^2}. \tag{A.8}$$

Substituting (A.6) and (A.7) into (A.8), respectively, we obtain, when $\hat{\sigma}^2 > 0$,

$$\sum_{i=1}^N \frac{\tilde{\tau}_i}{\hat{\sigma}^2 + s_i^2} = \sum_{i=1}^N \frac{1}{\hat{\sigma}^2 + s_i^2}, \tag{A.9}$$

and when $\hat{\sigma}^2 = 0$,

$$\sum_{i=1}^N \frac{\tilde{\tau}_i}{\hat{\sigma}^2 + s_i^2} \geq \sum_{i=1}^N \frac{1}{\hat{\sigma}^2 + s_i^2}, \quad (\text{A.10})$$

When $\hat{\sigma}^2 > 0$, from (A.9) we have

$$\frac{1}{N} \sum_{i=1}^N u_i \tilde{\tau}_i = 1,$$

and when $\hat{\sigma}^2 = 0$, from (A.10) we have

$$\frac{1}{N} \sum_{i=1}^N u_i \tilde{\tau}_i \geq 1.$$

where u_i is given by (10). This completes the proof. \square