Mathematical and Computational Applications

MDPI

Article

Testing Homogeneity of Proportion Ratios for Stratified Bilateral Correlated Data

Wanqing Tian and Changxing Ma *

Department of Biostatistics, University at Buffalo, Buffalo, NY 14214, USA; wanqingt@buffalo.edu
* Correspondence: cxma@buffalo.edu

Abstract: Intraclass correlation in bilateral data has been investigated in recent decades with various statistical methods. In practice, stratifying bilateral data by some control variables will provide more sophisticated statistical results to satisfy different research proposed in randomized clinical trials. In this article, we propose three test statistics (the likelihood ratio test, score test, and Wald-type test statistics) to evaluate the homogeneity of proportion ratios for stratified bilateral correlated data under an equal correlation assumption. Monte Carlo simulations of Type I error and power are performed, and the score test yields a robust outcome based on empirical Type I error and power. Lastly, two real data examples are conducted to illustrate the proposed three tests.

Keywords: likelihood ratio test; score test; Wald-type test; type I error; power; stratified bilateral correlated data

check for updates

Citation: Tian, W.; Ma, C. Testing Homogeneity of Proportion Ratios for Stratified Bilateral Correlated Data. *Math. Comput. Appl.* **2024**, *29*, 26. https://doi.org/10.3390/mca29020026

Academic Editor: Sandra Ferreira

Received: 9 January 2024
Revised: 19 March 2024
Accepted: 21 March 2024
Published: 22 March 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In randomized clinical trials, bilateral data frequently occurs in patients who receive a treatment based on paired body parts or organs (such as eyes, ears, kidneys, and so on). Since the outcome of bilateral data has been naturally split into three types (no response, unilateral response, and bilateral responses), considering the intraclass correlation in the bilateral data is a natural way to avoid misleading results [1–5]. Intraclass correlation in bilateral data has been investigated in recent decades with various statistical methods [1–3,5–8]. Rosner proposes that the conditional probability of a response occurring at one side of paired body parts or organs that gives a response at the other body parts or organs is a positive constant R times the response rate [1]. Tang et al. provide a statistical inference for correlated data in binary paired data under the R model and also evaluate the asymptotic test with Type I error and power [9]. Ma, Shan, and Liu develop an asymptotic testing method under the homogeneity assumption for Rosner's R model [5]. Donner assumes that all treatment groups share one intraclass correlation coefficient ρ [2], and Thompson evaluates the robustness of Donnar's ρ model in pair data by adopting simulations [10]. Liu et al. test the equality of correlation coefficients based on Donner's ρ model for paired binary data with multiple groups [11]. Later, Liu et al. explore the exact methods of testing the homogeneity of prevalence for correlated binary data under Donnar's ρ model [8]. However, Dallal criticizes Rosner's assumption and points out that "the constant R model will give a poor fit if the characteristic is almost certain to occur bilaterally with widely varying group-specific prevalence" [3]. He proposes that it is more advantageous to assume the characteristic emerges through a triggering mechanism, in which the probability of a subsequent occurrence is unaffected by the probability of initiating the trigger [3]. Dallal believes the conditional probability of a response at one side of paired body parts or organs giving a response at the other body parts or organs is a constant γ [3]. Li et al. develop asymptotic and exact methods following Dallal's model [12]. Then, Chen et al. propose multiple test statistics of response rates in the different groups under Dallal's model [13].

The homogeneity test for the appropriate effect size measure between different groups equates to testing the common value of the measure of effect. Moreover, there are three popular methods to evaluate the effect size in randomized clinical trials: the odds ratio, risk difference, and relative risk [14]. Indeed, relative risk is more informative than risk difference in some cases [15]. Compared with the odds ratio, the relative risk can process sparse data better. The stratifying of bilateral data by some control variables (e.g., disease phases, age, etc.) will provide more sophisticated statistical results to satisfy different research proposed in randomized clinical trials [16,17]. For example, evaluating the appropriate effect size across strata of disease phases between treatment and control groups equates to testing the effect size of homogeneity across the strata. Zhuang et al. investigate the homogeneity test of the ratio of two proportions in the stratified bilateral data based on Donner's ρ model [18]. Shen et al. test the homogeneity of the difference between two proportions for stratified correlated paired binary data under Donner's ρ model [19]. Xue and Ma propose interval estimation of proportion ratios for stratified bilateral correlated binary data under Rosner's constant R model [20].

Rosner's R and Donner's ρ models have limitations and restrictions. Rosner's R model is a conditional probability that a response at one side of paired body parts or organs that gives a response at the other body parts or organs is a positive constant R times the response rate [1]. However, the constant R can not reach one unless each response rate is equal, and the model is not appropriate if the patient's body parts or organs are all responding and different groups have different response rates [3,20]. In Donner's ρ model, it is assumed that all treatment groups share one intraclass correlation coefficient ρ [2]. However, when the correlations between the two groups are significantly different, Donner's ρ model cannot be used for data analysis. In practice, it is first verified that the correlation coefficients are equal to decide whether the Donner model is appropriate. Compared to Rosner's R and Donner's ρ models, Dallal's model does not require concern over whether all of a patient's body parts or organs have responded, nor does it need to worry about whether different groups have varying response rates. Additionally, it eliminates the need to verify that the correlation coefficients are equal. Instead, it provides a straightforward intraclass correlation constant, γ, for the conditional probability of a response in one part of a pair of body parts or organs given a response in the other.

To avoid the limitations of Rosner's R and Donner's ρ models, we investigate the proportion ratios in clinical trial design with stratified bilateral data under Dallal's model. The remainder of this paper is organized as follows: Section 2 presents the data structure and hypotheses, and Section 3 introduces the maximum-likelihood estimation under homogeneity. We propose three tests to examine the homogeneity of proportions across strata in Dallal's model in Section 4. Accordingly, we investigate the performance and robustness of three tests by using simulation studies in Section 5. In Section 6, we use two real data examples to illustrate our proposed methods. Conclusions and future works are in Section 7.

2. Data Structure and Hypotheses

2.1. Notation

Let m_{lij} be the number of patients in the ith group of the jth stratum with lth responses, where $i = 1, 2$, $j = 1, \ldots, J$, and $l = 0, 1, 2$. $m_{+ij} = m_{0ij} + m_{1ij} + m_{2ij}$ represents the number of patients in the ith group of the jth stratum; p_{lij} corresponds to the probability of patients in the ith group of the jth stratum with lth responses. Define Z_{hijk} as the indicator of the response of the kth body parts or organs of the hth patient in the ith group from the jth stratum, where $k = 1, 2$ and $h = 1, 2, \ldots, m_{+ij}$. If $Z_{hijk} = 1$, then the improvement response occurs; otherwise, $Z_{hijk} = 0$. Therefore, denote $\Pr\left(Z_{hijk} = 1\right) = \pi_{ij}$, $\left(0 \leq \pi_{ij} \leq 1\right)$ as the probability of having response on one site. The corresponding data structure is shown in Table 1.

Table 1. Data structure.

Number of Responses (l)	Group		Total
	1	2	
0	$m_{01j}(p_{01j})$	$m_{02j}(p_{02j})$	m_{0+j}
1	$m_{11j}(p_{11j})$	$m_{12j}(p_{12j})$	m_{1+j}
2	$m_{21j}(p_{21j})$	$m_{22j}(p_{22j})$	m_{2+j}
Total	m_{+1j}	m_{+2j}	m_{++j}

We explore the intraclass correlation based on Dallal's model, where the conditional probability of a response at one side of paired body parts or organs, given a response at the other body parts or organs, is a constant γ. Therefore, we assume $P\left(Z_{hijk} = 1 \mid Z_{hij(3-k)} = 1\right) = \gamma_j$, and the probabilities of a body part or organ with none, one, or both can be expressed as follows:

$$p_{0ij} = 1 - (2 - \gamma_j)\pi_{ij},$$

$$p_{1ij} = 2\pi_{ij}(1 - \gamma_j),$$

$$p_{2ij} = \pi_{ij}\gamma_j,$$

and $p_{0ij} + p_{1ij} + p_{2ij} = 1$ for any fixed i and j. The joint likelihood function for the observed data $\boldsymbol{m_{ij}} = (m_{0ij}, m_{1ij}, m_{2ij})^T$ is given by

$$f(\boldsymbol{m_{ij}} \mid \boldsymbol{p_{ij}}) = \prod_{j=1}^{J}\prod_{i=1}^{2} \frac{m_{+ij}!}{m_{0ij}!m_{1ij}!m_{2ij}!} p_{0ij}^{m_{0ij}} p_{1ij}^{m_{1ij}} p_{2ij}^{m_{2ij}},$$

where $\boldsymbol{p_{ij}} = (p_{0ij}, p_{1ij}, p_{2ij})^T$. The corresponding log-likelihood function can be expressed as:

$$l(\boldsymbol{\pi_i}, \gamma) = \sum_{j=1}^{J}\sum_{i=1}^{2}\{m_{0ij}\log[1 - (2-\gamma_j)\pi_{ij}] + m_{1ij}\log[2\pi_{ij}(1-\gamma_j)] + m_{2ij}\log[\pi_{ij}\gamma_j]\} + \log C,$$

where $\boldsymbol{\pi_i} = (\pi_{i1}, \ldots, \pi_{iJ})^T$, $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_J)^T$ and $C = \prod_{j=1}^{J}\prod_{i=1}^{2} \frac{m_{+ij}!}{m_{0ij}!m_{1ij}!m_{2ij}!}$ is a constant.

2.2. Hypotheses for the Proportion Ratio Across Strata

Assuming there is a common ratio of proportions between two groups across J strata, i.e., the ratio of proportions between two groups in the jth stratum is $\delta_j = \pi_{2j}/\pi_{1j}$ for $j = 1, \ldots, J$, the hypotheses are given as

$$H_0 : \delta_1 = \delta_2 = ... = \delta_J \quad \text{versus} \quad H_a : \delta_r \neq \delta_s, \quad \text{for at least one pair of } (r,s), r \neq s.$$

3. Maximum-Likelihood Estimation (MLE) under Homogeneity

3.1. The Constrained MLEs

Since we assume $\delta_j = \pi_{2j}/\pi_{1j}$ for $j = 1, \ldots, J$, then $\pi_{2j} = \delta_j\pi_{1j}$. Under the null hypothesis with a common δ, the log-likelihood function can be expressed as:

$$l(\boldsymbol{\pi_1}, \delta, \gamma) = \sum_{j=1}^{J} l_j(\pi_{1j}, \delta, \gamma_j),$$

where

$$\begin{aligned} l_j(\pi_{1j}, \delta, \gamma_j) = \{&m_{01j}\log[1 - (2-\gamma_j)\pi_{1j}] + m_{11j}\log[2\pi_{1j}(1-\gamma_j)] + m_{21j}\log[\pi_{1j}\gamma_j] \\ &+ m_{02j}\log[1 - (2-\gamma_j)\pi_{1j}\delta] + m_{12j}\log[2\pi_{1j}\delta(1-\gamma_j)] + m_{22j}\log[\pi_{1j}\delta\gamma_j]\} \\ &+ \log C. \end{aligned}$$

Differentiating $l(\boldsymbol{\pi_1}, \delta, \boldsymbol{\gamma})$ with respect to $\boldsymbol{\pi_1}$ and $\boldsymbol{\gamma}$, we have

$$\frac{\partial l}{\partial \pi_{1j}} = \frac{m_{1+j}}{\pi_{1j}} + \frac{m_{2+j}}{\pi_{1j}} + \frac{m_{01j}\left(\gamma_j - 2\right)}{\pi_{1j}\left(\gamma_j - 2\right) + 1} + \frac{\delta\, m_{02j}\left(\gamma_j - 2\right)}{\delta\, \pi_{1j}\left(\gamma_j - 2\right) + 1},$$

$$\frac{\partial l}{\partial \gamma_j} = \frac{m_{21j}}{\gamma_j} + \frac{m_{22j}}{\gamma_j} + \frac{m_{11j}}{\gamma_j - 1} + \frac{m_{12j}}{\gamma_j - 1} + \frac{m_{01j}\, \pi_{1j}}{\pi_{1j}\left(\gamma_j - 2\right) + 1} + \frac{\delta\, m_{02j}\, \pi_{1j}}{\delta\, \pi_{1j}\left(\gamma_j - 2\right) + 1}.$$

We can set $\frac{\partial l}{\partial \pi_{1j}} = 0$ with $\frac{\partial l}{\partial \gamma_j} = 0$ to obtain the MLEs $\hat{\pi}_{1j}$ and $\hat{\gamma}_j$. Indeed, the MLE of γ_j has a closed-form solution. Meanwhile, the MLE of π_{1j} is a function of m_{lij} and δ. For the MLE of δ, we will update by using Fisher scoring iterative algorithm by consider the initial value of $\delta^{(0)}$ as mean of unconstrained MLE $\tilde{\delta}_j$ ($\frac{\sum_{j=1}^{J} \tilde{\delta}_j}{J}$), and

$$\hat{\delta} = \delta^{(t+1)} = \delta^{(t)} + I^{-1}\left(\delta^{(t)}\right)\left(\frac{\partial l}{\partial \delta}\right)\Big|_{\delta=\delta(t), \hat{\pi}_{1j}, \hat{\gamma}_j},$$

where

$$\frac{\partial l}{\partial \delta} = \sum_{j=1}^{J}\left(\frac{m_{12j}}{\delta} + \frac{m_{22j}}{\delta} + \frac{m_{02j}\, \pi_{1j}\left(\gamma_j - 2\right)}{\delta\, \pi_{1j}\left(\gamma_j - 2\right) + 1}\right),$$

$$\frac{\partial^2 l}{\partial \delta^2} = \sum_{j=1}^{J}\left(-\frac{m_{12j}}{\delta^2} - \frac{m_{22j}}{\delta^2} - \frac{m_{02j}\, {\pi_{1j}}^2\left(\gamma_j - 2\right)^2}{\left(\delta\, \pi_{1j}\left(\gamma_j - 2\right) + 1\right)^2}\right),$$

and

$$\begin{aligned} I = -E\left(\frac{\partial^2 l}{\partial \delta^2}\right) &= -E\left(\sum_{j=1}^{J}\left(-\frac{m_{12j}}{\delta^2} - \frac{m_{22j}}{\delta^2} - \frac{m_{02j}\, {\pi_{1j}}^2\left(\gamma_j - 2\right)^2}{\left(\delta\, \pi_{1j}\left(\gamma_j - 2\right) + 1\right)^2}\right)\right) \\ &= \sum_{j=1}^{J}\left(\frac{m_{+2j} p_{12j}}{\delta^2} + \frac{m_{+2j} p_{22j}}{\delta^2} + \frac{m_{+2j} p_{02j}\, {\pi_{1j}}^2\left(\gamma_j - 2\right)^2}{\left(\delta\, \pi_{1j}\left(\gamma_j - 2\right) + 1\right)^2}\right). \end{aligned}$$

3.2. The Unconstrained MLEs

Under the alternative hypothesis, the log-likelihood can be presented as:

$$l(\boldsymbol{\pi_1}, \boldsymbol{\delta}, \boldsymbol{\gamma}) = \sum_{j=1}^{J} l_j\left(\pi_{1j}, \delta_j, \gamma_j\right),$$

where

$$\begin{aligned} l_j\left(\pi_{1j}, \delta_j, \gamma_j\right) = \sum_{j=1}^{J}\{ & m_{01j}\log\left[1 - \left(2 - \gamma_j\right)\pi_{1j}\right] + m_{11j}\log\left[2\pi_{1j}\left(1 - \gamma_j\right)\right] + m_{21j}\log\left[\pi_{1j}\gamma_j\right] \\ & + m_{02j}\log\left[1 - \left(2 - \gamma_j\right)\pi_{1j}\delta_j\right] + m_{12j}\log\left[2\pi_{1j}\delta_j\left(1 - \gamma_j\right)\right] + m_{22j}\log\left[\pi_{1j}\delta_j\gamma_j\right]\} + \log C, \end{aligned}$$

and the MLEs of three parameters, π_{1j}, δ_j, and γ_j, can be derived by setting the partial differentiation equal to zero and then the MLEs as follows:

$$\tilde{\pi}_{1j} = \frac{\left(m_{11j} + m_{21j}\right)\left(m_{11j} + m_{12j} + 2\, m_{21j} + 2\, m_{22j}\right)}{2\left(m_{01j} + m_{11j} + m_{21j}\right)\left(m_{11j} + m_{12j} + m_{21j} + m_{22j}\right)},$$

$$\tilde{\gamma}_j = \frac{2\, m_{2+j}}{m_{1+j} + 2\, m_{2+j}},$$

$$\tilde{\delta}_j = -\frac{\left(m_{12j} + m_{22j}\right)\left(m_{01j} - m_{12j} + m_{1+j} - m_{22j} + m_{2+j}\right)}{\left(m_{02j} + m_{12j} + m_{22j}\right)\left(m_{12j} - m_{1+j} + m_{22j} - m_{2+j}\right)}.$$

4. Testing Methods

4.1. Likelihood Ratio Test (T_L)

The likelihood ratio test statistic is given by

$$T_L = 2\sum_{j=1}^{J}\left[l_j\left(\tilde{\delta}_j, \tilde{\pi}_{1j}, \tilde{\gamma}_j\right) - l_j\left(\hat{\delta}, \hat{\pi}_{1j}, \hat{\gamma}_j\right)\right],$$

Moreover, T_L is asymptotically distributed as a Chi-square distribution with $J-1$ degrees of freedom under the null hypothesis. Then, reject the null hypothesis if $T_L > \chi^2_{J-1,1-\alpha}$ at a significant level α.

4.2. Score Test (T_{SC})

Under the assumption that each stratum has the same ratio of proportions δ_j ($\delta = \delta_j$), the score test statistic is given by

$$\begin{aligned} T_{SC} &= \sum_{j=1}^{J} U_j I_j^{-1} U_j^T \mid \delta = \hat{\delta}, \pi_{1j} = \hat{\pi}_{1j}, \gamma_j = \hat{\gamma}_j \\ &= \sum_{j=1}^{J}\left(\frac{\partial l_j}{\partial \delta_j}\right)^2 \frac{1}{D_j} I_j^{-1}(1,1) \mid \delta = \hat{\delta}, \pi_{1j} = \hat{\pi}_{1j}, \gamma_j = \hat{\gamma}_j, \end{aligned}$$

$$I_j^{-1} = \left(\frac{1}{D_j}\right)\begin{pmatrix} \left(I_{22}^{(j)} I_{33}^{(j)} - I_{23}^{(j)} I_{32}^{(j)}\right) & -\left(I_{12}^{(j)} I_{33}^{(j)} - I_{13}^{(j)} I_{32}^{(j)}\right) & \left(I_{12}^{(j)} I_{23}^{(j)} - I_{13}^{(j)} I_{22}^{(j)}\right) \\ -\left(I_{21}^{(j)} I_{33}^{(j)} - I_{23}^{(j)} I_{31}^{(j)}\right) & \left(I_{11}^{(j)} I_{33}^{(j)} - I_{13}^{(j)} I_{31}^{(j)}\right) & -\left(I_{11}^{(j)} I_{23}^{(j)} - I_{13}^{(j)} I_{21}^{(j)}\right) \\ \left(I_{21}^{(j)} I_{32}^{(j)} - I_{22}^{(j)} I_{31}^{(j)}\right) & -\left(I_{11}^{(j)} I_{32}^{(j)} - I_{12}^{(j)} I_{31}^{(j)}\right) & \left(I_{11}^{(j)} I_{22}^{(j)} - I_{12}^{(j)} I_{21}^{(j)}\right) \end{pmatrix},$$

$$D_j = I_{11}^{(j)} I_{22}^{(j)} I_{33}^{(j)} - I_{11}^{(j)} I_{23}^{(j)} I_{32}^{(j)} - I_{12}^{(j)} I_{21}^{(j)} I_{33}^{(j)} + I_{12}^{(j)} I_{23}^{(j)} I_{31}^{(j)} + I_{13}^{(j)} I_{21}^{(j)} I_{32}^{(j)} - I_{13}^{(j)} I_{22}^{(j)} I_{31}^{(j)},$$

where the score function for the j th stratum $U_j = \left(\frac{\partial l_j}{\partial \delta}, 0, 0\right)$ and $I_j^{-1}(1,1) = \left(I_{22}^{(j)} I_{33}^{(j)} - I_{23}^{(j)} I_{32}^{(j)}\right)$ is the first diagonal element of the inverse of the Fisher information matrix (see Appendix A for more detail). Therefore, T_{SC} is asymptotically distributed as a Chi-square distribution with $J-1$ degrees of freedom under the null hypothesis. Then, reject the null hypothesis if $T_{SC} > \chi^2_{J-1,1-\alpha}$ at a significant level α.

4.3. Wald-Type Test (T_W)

An alternative way to express the null hypothesis is in matrix form as $\boldsymbol{C}\boldsymbol{\beta}^T = 0$, where

$$\boldsymbol{\beta} = \left(\delta_1, \pi_{11}, \gamma_1, \delta_2, \pi_{12}, \gamma_2, \cdots, \delta_J, \pi_{1J}, \gamma_J\right)$$

and

$$\boldsymbol{C} = \begin{pmatrix} 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & 0 & 0 & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & \cdots & -1 & 0 & 0 \end{pmatrix}.$$

The information matrix for this Wald-type test I_W is a $3J \times 3J$ block diagonal matrix, with each block being a 3×3 information matrix of a parameter vector $\left(\delta_j, \pi_{1j}, \gamma_j\right)$ within each stratum. We used the same derivation approach that we previously used to derive the information matrix for the score test.

$$I_W = \begin{bmatrix} I_W^{(1)} & 0 & \cdots & 0 \\ 0 & I_W^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I_W^{(J)} \end{bmatrix}_{3J \times 3J},$$

with

$$I_W^{(j)} = \begin{bmatrix} I_{W11}^{(j)} & I_{W12}^{(j)} & I_{W13}^{(j)} \\ I_{W21}^{(j)} & I_{W22}^{(j)} & I_{W23}^{(j)} \\ I_{W31}^{(j)} & I_{W32}^{(j)} & I_{W33}^{(j)} \end{bmatrix}_{3 \times 3},$$

where the explicit form of each element $\left(I_{Wik}^{(j)}\right)$ is exactly the same as $\left(I_{ik}^{j}\right)$, derived in the score test. Then, the Wald-type test statistic can be expressed as:

$$T_W = \left(\boldsymbol{\beta}\boldsymbol{C}^T\right)\left(\boldsymbol{C}I_W^{-1}\boldsymbol{C}^T\right)^{-1}\left(\boldsymbol{C}\boldsymbol{\beta}^T\right)\Big|_{\boldsymbol{\beta}=\left(\tilde{\delta}_1,\ldots,\tilde{\delta}_J,\tilde{\pi}_{11},\ldots,\tilde{\pi}_{1J},\tilde{\gamma}_1,\ldots,\tilde{\gamma}_J\right)}$$

$$I_W^{-1} = \begin{bmatrix} \left(I_W^{(1)}\right)^{-1} & 0 & \cdots & 0 \\ 0 & \left(I_W^{(2)}\right)^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \left(I_W^{(J)}\right)^{-1} \end{bmatrix}_{J \times J},$$

where

$$(I_W^{(j)})^{-1} = \left(\frac{1}{D_W^{(j)}}\right)\begin{pmatrix} \left(I_{W22}^{(j)} I_{W33}^{(j)} - I_{W23}^{(j)} I_{W32}^{(j)}\right) & -\left(I_{W12}^{(j)} I_{W33}^{(j)} - I_{W13}^{(j)} I_{W32}^{(j)}\right) & \left(I_{W12}^{(j)} I_{W23}^{(j)} - I_{W13}^{(j)} I_{W22}^{(j)}\right) \\ -\left(I_{W21}^{(j)} I_{W33}^{(j)} - I_{W23}^{(j)} I_{W31}^{(j)}\right) & \left(I_{W11}^{(j)} I_{W33}^{(j)} - I_{W13}^{(j)} I_{W31}^{(j)}\right) & -\left(I_{W11}^{(j)} I_{W23}^{(j)} - I_{W13}^{(j)} I_{W21}^{(j)}\right) \\ \left(I_{W21}^{(j)} I_{W32}^{(j)} - I_{W22}^{(j)} I_{W31}^{(j)}\right) & -\left(I_{W11}^{(j)} I_{W32}^{(j)} - I_{W12}^{(j)} I_{W31}^{(j)}\right) & \left(I_{W11}^{(j)} I_{W22}^{(j)} - I_{W12}^{(j)} I_{W21}^{(j)}\right) \end{pmatrix},$$

$$D_W^{(j)} = I_{W11}^{(j)} I_{W22}^{(j)} I_{W33}^{(j)} - I_{W11}^{(j)} I_{W23}^{(j)} I_{W32}^{(j)} - I_{W12}^{(j)} I_{W21}^{(j)} I_{W33}^{(j)} + I_{W12}^{(j)} I_{W23}^{(j)} I_{W31}^{(j)} + I_{W13}^{(j)} I_{W21}^{(j)} I_{W32}^{(j)} - I_{W13}^{(j)} I_{W22}^{(j)} I_{W31}^{(j)}.$$

Meanwhile, T_W is asymptotically distributed as a Chi-square distribution with $J-1$ degrees of freedom under the null hypothesis. Then, reject the null hypothesis if $T_W > \chi^2_{J-1,1-\alpha}$ at a significant level α.

5. Simulation Studies

This section investigates the empirical performances of three proposed test statistics in the previous section by three Monte Carlo simulation studies to evaluate the quality of relative risk in terms of the empirical Type I error rate and the power.

5.1. Empirical Type I Error Rates

In the first Monte Carlo simulation study, we investigate the behavior of the empirical Type I rate for three proposed tests under various procedures, where m = 25, 50, or 100 in strata J = 2, 4, 6, or 8. By considering that π_{1j} and γ_j are either common or different cross strata, we provide the parameter settings under different sample sizes and various sets of parameters in Table 2. For each configuration, 50,000 replications are randomly generated under the null hypothesis $H_0 : \delta$ = 1, 1.2, and 0.8, and the empirical Type I error rates are calculated by dividing the number of rejections by 50,000, and the nominal significant level α = 0.05. According to the previous study [9,20], the robustness of empirical Type I error rates is in the range from 0.04 to 0.06. In Table 3 and Appendix A.3 Tables A1–A3, we observe that the Wald-type test has poor performance, and the likelihood ratio test does not work well under a small sample with multiple strata. However, the score test is more robust than the likelihood ratio and Wald-type test.

Table 2. Parameter setups for computing empirical Type I error rates and powers.

		Number of Strata			
Parameter	**Cases**	$J = 2$	$J = 4$	$J = 6$	$J = 8$
	I	(0.2, 0.4)	(0.2, 0.4, 0.2, 0.4)	(0.2, 0.4, 0.2, 0.4, 0.2, 0.4)	(0.2, 0.4, 0.2, 0.4, 0.2, 0.4, 0.2, 0.4)
γ	II	(0.3, 0.3)	(0.3, 0.3, 0.3, 0.3)	(0.3, 0.3, 0.3, 0.3, 0.3, 0.3)	(0.3, 0.3, 0.3, 0.3, 0.3, 0.3, 0.3, 0.3)
	III	(0.3, 0.5)	(0.3, 0.5, 0.3, 0.5)	(0.3, 0.5, 0.3, 0.5, 0.3, 0.5)	(0.3, 0.5, 0.3, 0.5, 0.3, 0.5, 0.3, 0.5)
	IV	(0.6, 0.6)	(0.6, 0.6, 0.6, 0.6)	(0.6, 0.6, 0.6, 0.6, 0.6, 0.6)	(0.6, 0.6, 0.6, 0.6, 0.6, 0.6, 0.6, 0.6)
π_1	a	(0.2, 0.4)	(0.2, 0.4, 0.2, 0.4)	(0.2, 0.4, 0.2, 0.4, 0.2, 0.4)	(0.2, 0.4, 0.2, 0.4, 0.2, 0.4, 0.2, 0.4)
	b	(0.3, 0.3)	(0.3, 0.3, 0.3, 0.3)	(0.3, 0.3, 0.3, 0.3, 0.3, 0.3)	(0.3, 0.3, 0.3, 0.3, 0.3, 0.3, 0.3, 0.3)
	c	(0.2, 0.3)	(0.2, 0.3, 0.2, 0.3)	(0.2, 0.3, 0.2, 0.3, 0.2, 0.3)	(0.2, 0.3, 0.2, 0.3, 0.2, 0.3, 0.2, 0.3)
δ	H_{a1}:	(0.5, 1)	(0.5, 1, 0.5, 1)	(0.5, 1, 0.5, 1, 0.5, 1)	(0.5, 1, 0.5, 1, 0.5, 1, 0.5, 1)
	H_{a2}	(0.5, 1.2)	(0.5, 1.2, 0.5, 1.2)	(0.5, 1.2, 0.5, 1.2, 0.5, 1.2)	(0.5, 1.2, 0.5, 1.2, 0.5, 1.2, 0.5, 1.2)
	H_{a3}	(0.5, 1.4)	(0.5, 1.4, 0.5, 1.4)	(0.5, 1.4, 0.5, 1.4, 0.5, 1.4)	(0.5, 1.4, 0.5, 1.4, 0.5, 1.4, 0.5, 1.4)

Table 3. Simulation results of the empirical sizes (percentage) for 2 strata.

			$m = 25$			$m = 50$			$m = 100$		
δ	γ	π_1	T_L	T_{SC}	T_W	T_L	T_{SC}	T_W	T_L	T_{SC}	T_W
1.0	I	a	5.46	5.28	3.76	5.28	5.17	3.97	5.06	5.03	4.41
		b	5.39	5.37	2.04	5.28	5.23	3.40	5.09	5.07	4.20
		c	5.31	5.25	1.71	5.19	5.15	2.99	5.13	5.11	4.02
	II	a	5.65	5.43	4.68	5.35	5.23	4.45	5.27	5.22	4.73
		b	5.42	5.42	1.92	5.39	5.33	3.33	4.94	4.92	4.05
		c	5.54	5.46	2.39	5.22	5.16	3.23	5.10	5.08	4.00
	III	a	5.36	5.21	3.42	5.33	5.24	3.92	5.14	5.10	4.27
		b	5.50	5.47	1.83	5.21	5.15	3.14	5.07	5.04	4.12
		c	5.41	5.29	1.52	5.14	5.11	2.79	5.11	5.09	3.87
	IV	a	5.51	5.21	3.95	5.12	5.01	3.77	5.03	4.97	4.23
		b	5.49	5.44	1.20	5.22	5.20	2.79	4.94	4.90	3.78
		c	5.48	5.34	1.70	5.32	5.20	2.71	5.24	5.18	3.73
1.2	I	a	5.37	5.22	3.66	5.24	5.15	4.00	4.98	4.95	4.29
		b	5.54	5.48	1.70	5.16	5.10	3.15	5.06	5.04	4.08
		c	5.56	5.47	1.48	5.25	5.18	2.76	5.12	5.11	3.87
	II	a	5.42	5.29	4.47	5.11	5.06	4.28	5.14	5.10	4.70
		b	5.19	5.15	1.51	5.05	5.00	3.03	5.00	4.98	4.01
		c	5.49	5.44	2.17	5.11	5.07	3.15	5.07	5.04	3.91
	III	a	5.25	5.13	3.32	5.08	4.99	3.63	5.09	5.05	4.27
		b	5.28	5.24	1.46	5.14	5.09	2.99	5.13	5.10	4.00
		c	5.53	5.46	1.31	5.29	5.25	2.56	4.88	4.86	3.65
	IV	a	5.37	5.12	3.76	5.17	5.05	3.87	5.20	5.12	4.29
		b	5.62	5.57	1.05	5.12	5.09	2.48	5.26	5.24	3.84
		c	5.64	5.50	1.44	5.14	5.09	2.46	5.17	5.14	3.55
0.8	I	a	5.31	5.13	3.81	5.16	5.07	4.14	5.15	5.09	4.47
		b	5.36	5.33	2.49	5.15	5.10	3.60	5.10	5.06	4.34
		c	5.54	5.39	2.08	5.17	5.12	3.12	5.08	5.04	4.10
	II	a	5.40	5.14	4.63	5.20	5.08	4.34	5.09	5.03	4.63
		b	5.41	5.36	2.30	5.30	5.24	3.68	5.10	5.08	4.29
		c	5.29	5.14	2.45	5.25	5.18	3.39	4.99	4.95	4.00
	III	a	5.45	5.28	3.77	5.20	5.11	3.95	5.12	5.06	4.42
		b	5.31	5.28	2.02	5.25	5.22	3.49	5.00	4.99	4.14
		c	5.29	5.18	1.75	5.26	5.17	2.95	5.19	5.12	3.98
	IV	a	5.72	5.43	4.07	5.23	5.12	3.99	5.12	5.08	4.39
		b	5.43	5.34	1.55	5.30	5.26	3.02	5.01	4.98	3.84
		c	5.51	5.19	1.95	5.23	5.12	2.82	5.34	5.29	3.86

To obtain completed and robust empirical performances of the three proposed tests, we propose the second simulation by randomly choosing parameter configurations. We randomly generate 1000 parameter configurations within a parameter space for different sample sizes m = 15, 25, 50, 100 and a variety of strata J = 2, 4, 6, and 8 for 50,000 replications. Just as for the first simulation, the empirical Type I error rates can be calculated as the number of rejections divided by 50,000. Furthermore, the corresponding box plots and violin plots are shown in Figures 1–4. Comparing empirical Type I error rates among the three tests, the score and likelihood ratio tests are stable and robust under all conditions. Still, the Wald-type test performs poorly compared to the other two tests. Indeed, the accuracy of the Wald-type test improves with larger sample sizes. The Wald-type test may perform poorly in small samples, leading to inaccurate p-values.

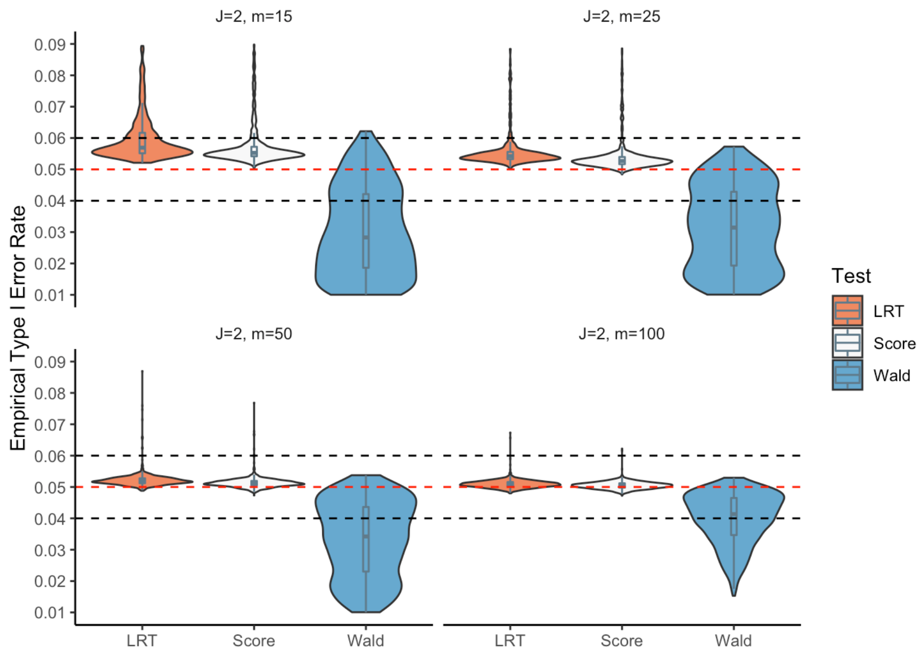

Figure 1. Violin plots and box plots of empirical sizes ($J = 2$).

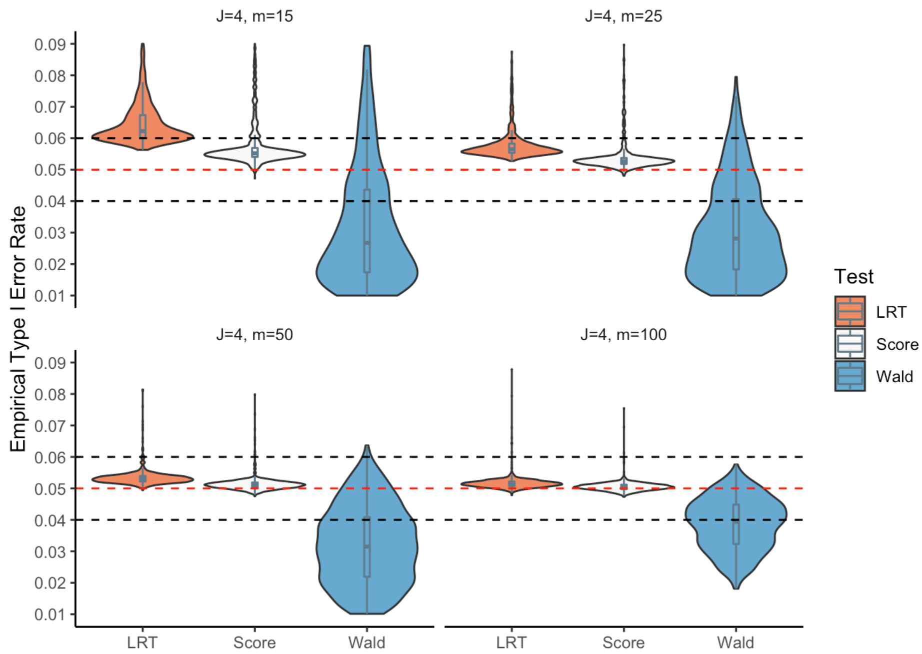

Figure 2. Violin plots and box plots of empirical sizes ($J = 4$).

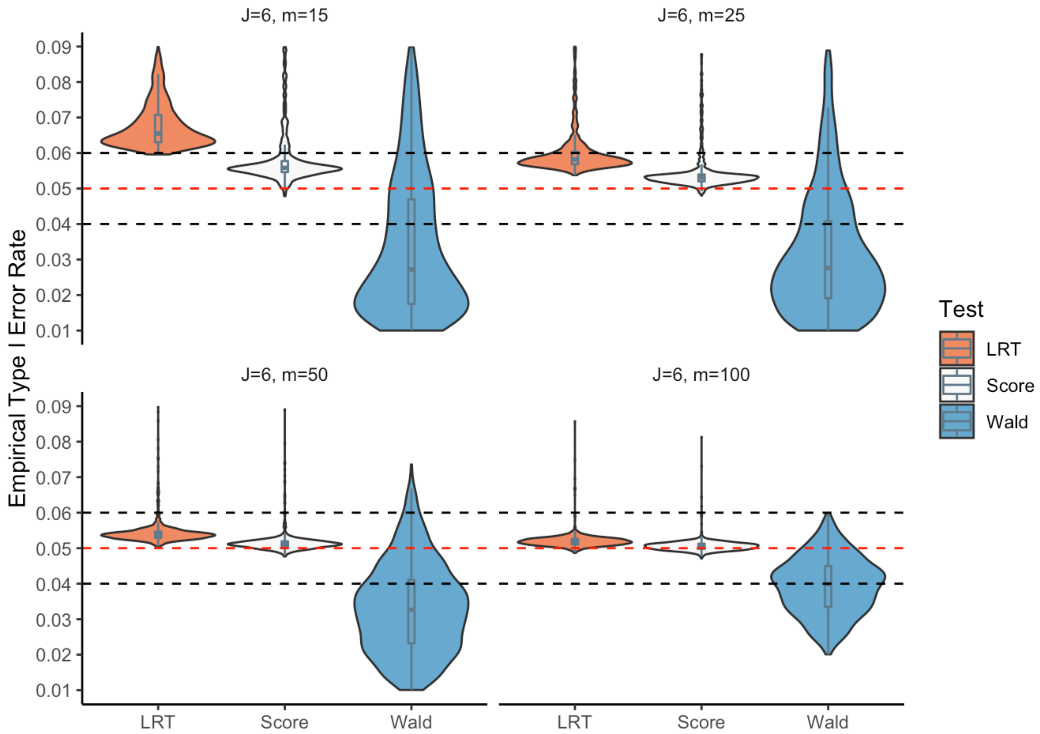

Figure 3. Violin plots and box plots of empirical sizes ($J = 6$).

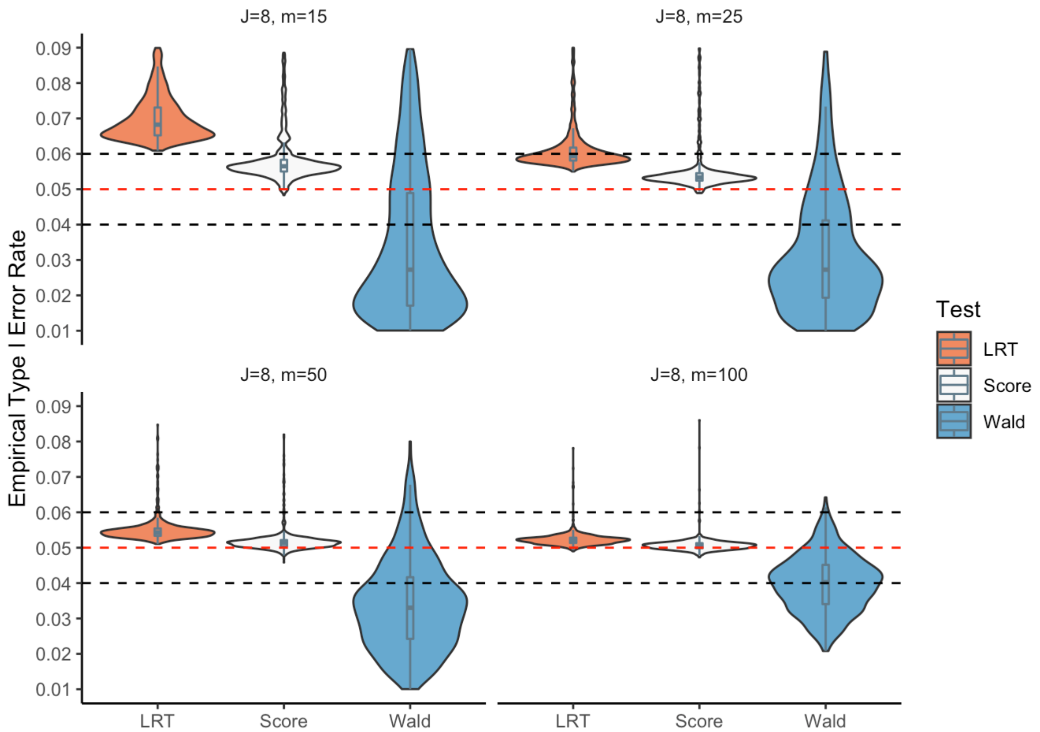

Figure 4. Violin plots and box plots of empirical sizes ($J = 8$).

5.2. Power

The third Monte Carlo simulation aims to investigate the power performance among the three proposed tests. Based on the alternative hypothesis with respect to δ = H_{a1}, H_{a2}, and H_{a3}, we consider the same strata, sample size, and parameter setting (Table 2) as in the first Monte Carlo simulation of the empirical Type I error. As the sample size increases, the powers of the three proposed tests increase for all strata settings (Figures 5–8). In addition, as the strata increase, the powers increase for the three proposed tests. Considering the empirical Type I error, using the score test is highly recommended.

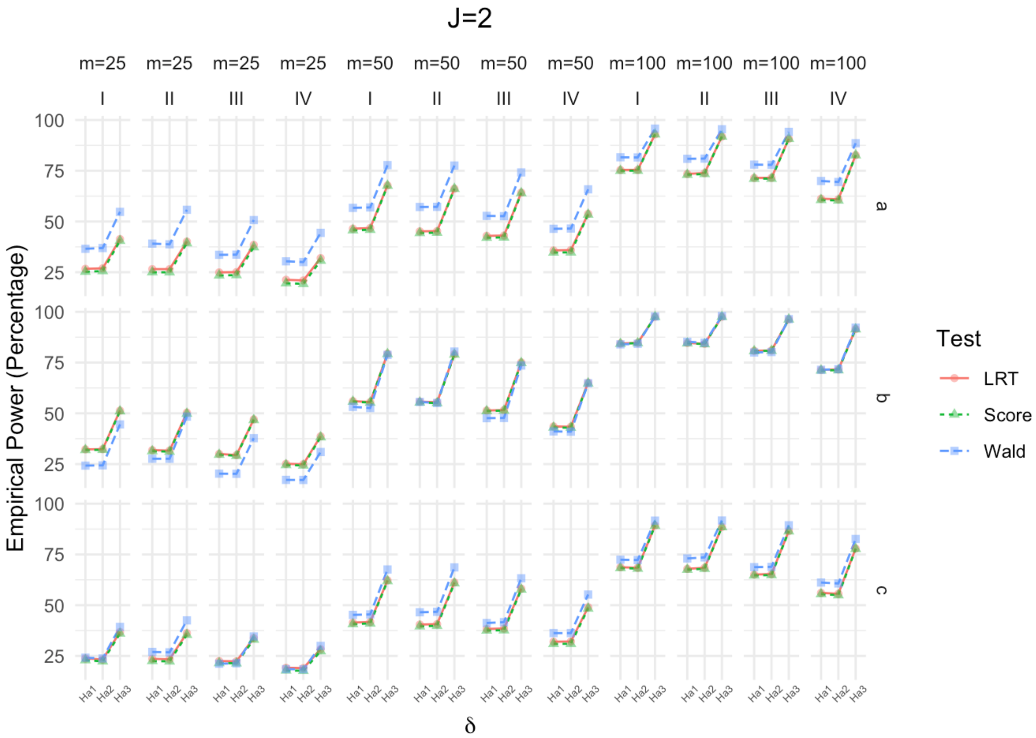

Figure 5. Line plots of empirical power ($J = 2$).

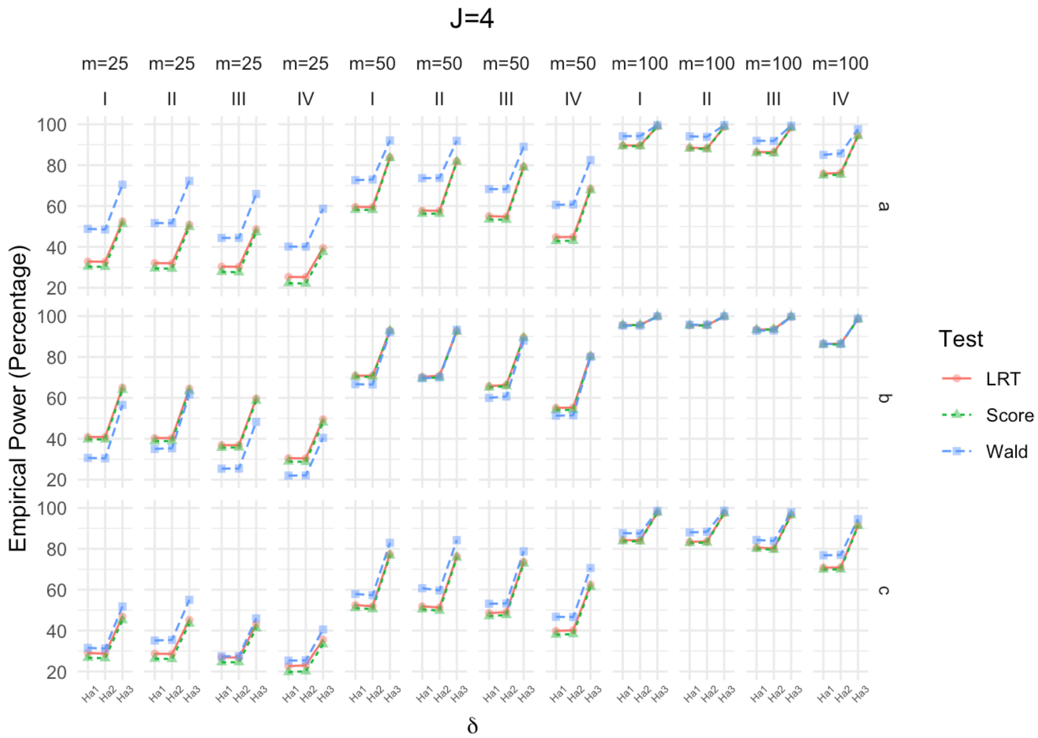

Figure 6. Line plots of empirical power ($J = 4$).

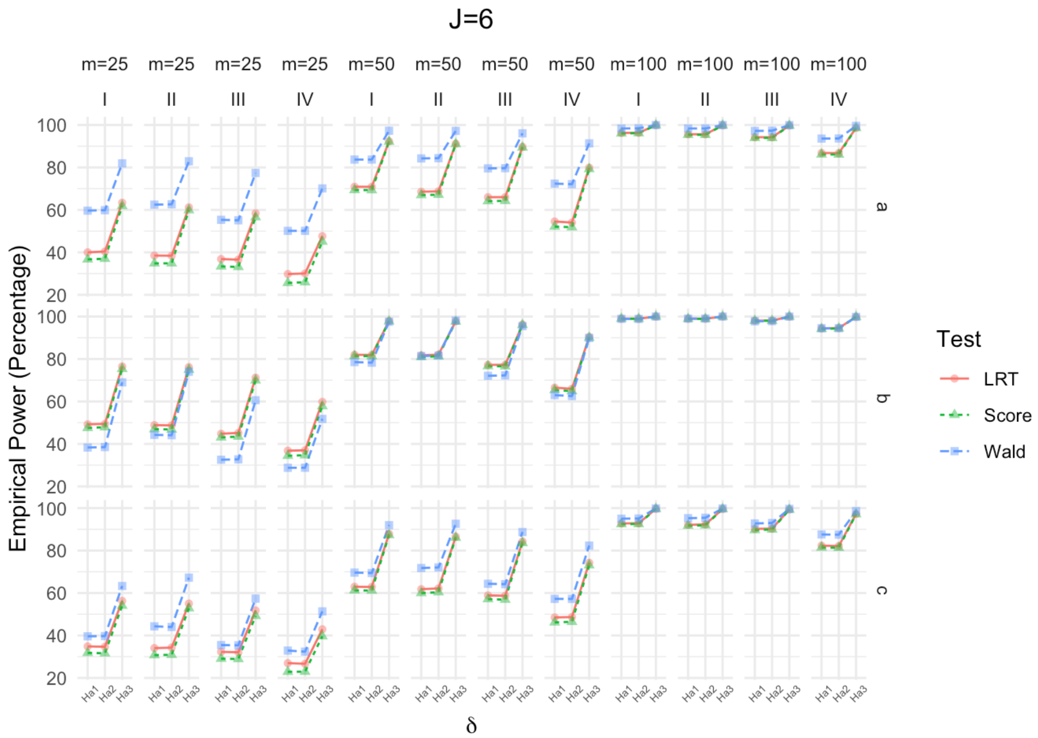

Figure 7. Line plots of empirical power ($J = 6$).

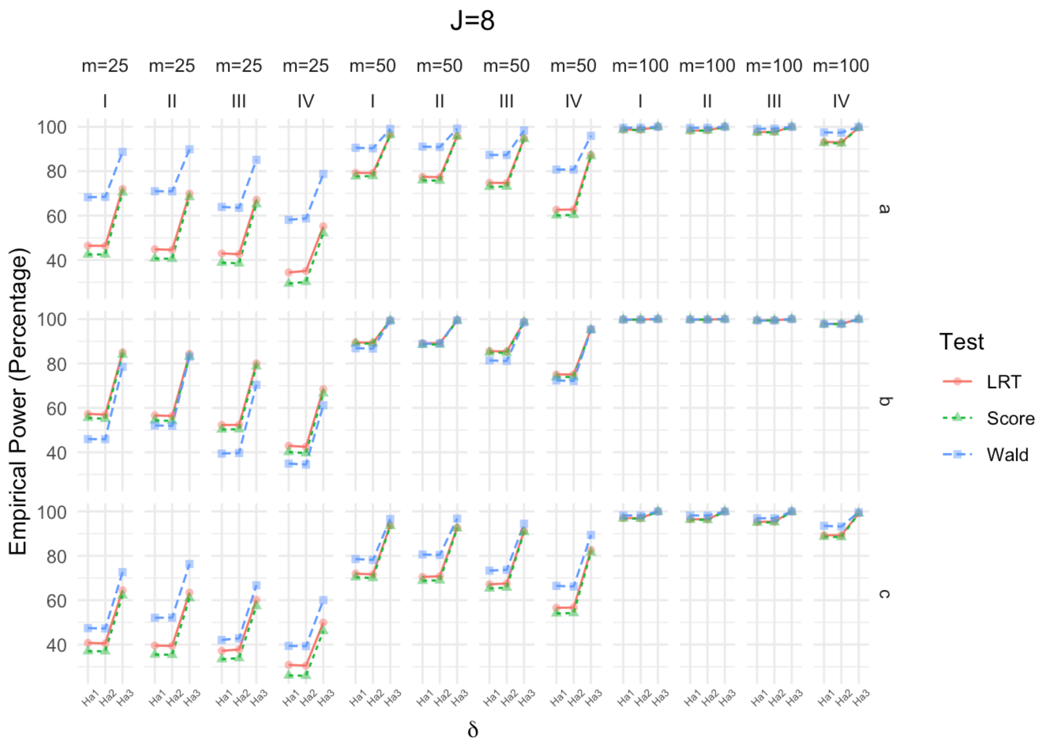

Figure 8. Line plots of empirical power ($J = 8$).

6. Real Data Examples

Two randomized clinical trial data are used as real data examples to illustrate the proposed three tests. The first real data example is a double-blinded randomized clinical trial data proposed by Mandel et al. [21]. In this clinical trial, children who suffer from otitis media with effusion (OME) and simultaneously have bilateral tympanocentesis are randomized into two groups: receiving 14-day treatments of amoxicillin or cefaclor [21].

After the treatment, the number of cured ears is summarized in Table 4 with the age group as strata. To explore whether the cured rates between the two groups (amoxicillin and cefaclor) among age strata are clinically equivalent, we test the homogeneity based on the three proposed tests. The MLEs of the parameters, based on observed data, are listed in Table 5, and the three test statistics and p-values are summarized in Table 6. However, all of the p-values are greater than 0.05, which means there is no statistical evidence to reject the null hypothesis $H_0 : \delta_1 = \delta_2 = \delta_3$.

Table 4. Frequency of number of OME-free ears after treatment.

	Age Group					
	<2 years		**2–5 years**		**≥6 years**	
Number of OME-Free Ears	**Cefactor**	**Amoxicillin**	**Cefactor**	**Amoxicillin**	**Cefactor**	**Amoxicillin**
0	8	11	6	3	0	1
1	2	2	6	1	1	0
2	8	2	10	5	3	6

Table 5. MLEs of parameters based on observed data (OME-free ears).

	Unconstrained MLEs			**Constrained MLEs**		
Age Groups	$\tilde{\pi}_1$	$\tilde{\gamma}$	$\tilde{\delta}$	$\hat{\pi}_1$	$\hat{\gamma}$	$\hat{\delta}$
<2 years	0.4762	0.8333	0.4800	0.4036	0.8333	0.8174
2–5 years	0.6116	0.8108	0.9167	0.6249	0.8108	-
≥6 years	0.9500	0.9474	0.8572	0.9500	0.9474	-

Table 6. The values of statistics and p-values for three different tests (OME-free ears).

	T_L	T_{SC}	T_W
Statistic	1.6918	1.6392	2.3520
P	0.4292	0.4406	0.3085

Another example used to illustrate the three proposed tests is a randomized double-blinded placebo-controlled trial presented by Postlethwaite et al. [22]. In this trial, the primary measurement is the modified Rodnan Skin Score (MRSS), and patients diagnosed with diffuse scleroderma were randomly divided into the treatment and the control group. The duration of the disease (early or late phase) was considered two separate strata. Meanwhile, the number of improved hands is summarized by groups and disease phases in Table 7. We obtain the MLEs of the parameters in Table 8, and the statistics and p-values are summarized in Table 9. All of the p-values are greater than 0.05, which means there is no statistical evidence to reject the null hypothesis $H_0 : \delta_1 = \delta_2$.

Table 7. Number of patients with hand MRSS decreased.

	Early		**Late**	
Number of Hands with Improvement	**Collagen**	**Placebo**	**Collagen**	**Placebo**
0	20	23	9	22
1	2	3	3	2
2	5	4	3	2

Table 8. MLEs of parameters based on observed data (MRSS).

	Unconstrained MLEs			Constrained MLEs		
Phase	$\tilde{\pi}_1$	$\tilde{\gamma}$	$\tilde{\delta}$	$\hat{\pi}_1$	$\hat{\gamma}$	$\hat{\delta}$
Early	0.213	0.783	0.900	0.248	0.783	0.626
Late	0.300	0.667	0.385	0.245	0.667	-

Table 9. The values of statistics and *p*-values for three different tests (MRSS).

	T_L	T_{SC}	T_W
Statistic	1.3979	1.3955	1.2046
P	0.2371	0.2375	0.2724

7. Conclusions

This article utilizes three MLE-based tests (LRT, Wald-type test, and score test) to test the homogeneity relative risk of two proportions on stratified bilateral correlated data.

The three Monte Carlo simulation results show that the score test yields a robust performance with the empirical Type I error and power. Even with a small sample size and multiple strata, the score test still generates a stable empirical Type I error and satisfactory power. Meanwhile, the likelihood ratio test and Wald-type test offer reasonable power but come with unstable empirical Type I error performances. By incorporating a small sample size, the Wald-type test shows unsatisfactory performance of the empirical Type I error, but the performance tends to be reasonable by increasing the sample size. Under a small sample size with multiple strata, the score test is slightly unstable, but the performance also tends to be robust by increasing the sample size.

However, asymptotic methods may have some limitations due to poor performance under a small sample size with multiple strata. Future work might consider exact tests to investigate related issues.

Author Contributions: Conceptualization, W.T. and C.M.; Methodology, W.T. and C.M.; Writing—Original Draft Preparation, W.T.; Writing—Review & Editing, W.T. and C.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data supporting this study are included within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Appendix A.1. Derivation of Score Statistic

The first-order differential equations of stratum j are:

$$\frac{\partial l_j}{\partial \delta} = \frac{m_{12j}}{\delta} + \frac{m_{22j}}{\delta} + \frac{m_{02j}\,\pi_{1j}\,(\gamma_j - 2)}{\delta\,\pi_{1j}\,(\gamma_j - 2) + 1},$$

$$\frac{\partial l_j}{\partial \pi_{1j}} = \frac{m_{12j}}{\pi_{1j}} + \frac{m_{21j}}{\pi_{1j}} + \frac{m_{22j}}{\pi_{1j}} + \frac{m_{01j}\,(\gamma_j - 2)}{\pi_{1j}\,(\gamma_j - 2) + 1} + \frac{\delta\, m_{02j}\,(\gamma_j - 2)}{\delta\,\pi_{1j}\,(\gamma_j - 2) + 1} + \frac{m_{11j}\,(2\,\gamma_j - 2)}{2\,\pi_{1j}\,(\gamma_j - 1)},$$

$$\frac{\partial l_j}{\partial \gamma_j} = \frac{m_{21j}}{\gamma_j} + \frac{m_{22j}}{\gamma_j} + \frac{m_{11j}}{\gamma_j - 1} + \frac{m_{12j}}{\gamma_j - 1} + \frac{m_{01j}\,\pi_{1j}}{\pi_{1j}\,(\gamma_j - 2) + 1} + \frac{\delta\, m_{02j}\,\pi_{1j}}{\delta\,\pi_{1j}\,(\gamma_j - 2) + 1}.$$

The second-order differential equations of stratum j are:

$$\frac{\partial^2 l_j}{\partial \delta^2} = -\frac{m_{12j}}{\delta^2} - \frac{m_{22j}}{\delta^2} - \frac{m_{02j}\,{\pi_{1j}}^2\,(\gamma_j - 2)^2}{\left(\delta\,\pi_{1j}\,(\gamma_j - 2) + 1\right)^2},$$

$$\frac{\partial^2 l_j}{\partial\delta\partial\pi_{1j}} = \frac{m_{02j}\,(\gamma_j-2)}{\delta\,\pi_{1j}\,(\gamma_j-2)+1} - \frac{\delta\,m_{02j}\,\pi_{1j}\,(\gamma_j-2)^2}{(\delta\,\pi_{1j}\,(\gamma_j-2)+1)^2},$$

$$\frac{\partial^2 l_j}{\partial\delta\partial\gamma_j} = \frac{m_{02j}\,\pi_{1j}}{\delta\,\pi_{1j}\,(\gamma_j-2)+1} - \frac{\delta\,m_{02j}\,{\pi_{1j}}^2\,(\gamma_j-2)}{(\delta\,\pi_{1j}\,(\gamma_j-2)+1)^2},$$

$$\frac{\partial^2 l_j}{\partial\pi_{1j}\partial\delta} = \frac{m_{02j}\,(\gamma_j-2)}{\delta\,\pi_{1j}\,(\gamma_j-2)+1} - \frac{\delta\,m_{02j}\,\pi_{1j}\,(\gamma_j-2)^2}{(\delta\,\pi_{1j}\,(\gamma_j-2)+1)^2},$$

$$\frac{\partial^2 l_j}{\partial\pi_{1j}^2} = -\frac{m_{12j}}{{\pi_{1j}}^2} - \frac{m_{21j}}{{\pi_{1j}}^2} - \frac{m_{22j}}{{\pi_{1j}}^2} - \frac{m_{01j}\,(\gamma_j-2)^2}{(\pi_{1j}\,(\gamma_j-2)+1)^2} - \frac{\delta^2\,m_{02j}\,(\gamma_j-2)^2}{(\delta\,\pi_{1j}\,(\gamma_j-2)+1)^2} - \frac{m_{11j}\,(2\,\gamma_j-2)}{2\,{\pi_{1j}}^2\,(\gamma_j-1)},$$

$$\frac{\partial^2 l_j}{\partial\pi_{1j}\partial\gamma_j} = \frac{m_{01j}}{(\pi_{1j}\,(\gamma_j-2)+1)^2} + \frac{\delta\,m_{02j}}{(\delta\,\pi_{1j}\,(\gamma_j-2)+1)^2},$$

$$\frac{\partial^2 l_j}{\partial\gamma_j\partial\delta} = \frac{m_{02j}\,\pi_{1j}}{\delta\,\pi_{1j}\,(\gamma_j-2)+1} - \frac{\delta\,m_{02j}\,{\pi_{1j}}^2\,(\gamma_j-2)}{(\delta\,\pi_{1j}\,(\gamma_j-2)+1)^2},$$

$$\frac{\partial^2 l_j}{\partial\gamma_j\partial\pi_{1j}} \frac{m_{01j}}{(\pi_{1j}\,(\gamma_j-2)+1)^2} + \frac{\delta\,m_{02j}}{(\pi_{1j}\,(2\,\delta-\delta\,\gamma_j)-1)^2},$$

$$\frac{\partial^2 l_j}{\partial\gamma_j^2} = -\frac{m_{21j}}{{\gamma_j}^2} - \frac{m_{22j}}{{\gamma_j}^2} - \frac{m_{11j}}{(\gamma_j-1)^2} - \frac{m_{12j}}{(\gamma_j-1)^2} - \frac{m_{01j}\,{\pi_{1j}}^2}{(\pi_{1j}\,(\gamma_j-2)+1)^2} - \frac{\delta^2\,m_{02j}\,{\pi_{1j}}^2}{(\delta\,\pi_{1j}\,(\gamma_j-2)+1)^2}.$$

The Fisher information of stratum j is:

$$I_{11}^{(j)} = E\left(-\frac{\partial^2 l_j}{\partial\delta^2}\right) = \frac{m_{+2j}}{\delta^2\,(\delta\,\pi_{1j}\,(\gamma_j-2)+1)} - \frac{m_{+2j}}{\delta^2},$$

$$I_{12}^{(j)} = I_{21}^{(j)} = E\left(-\frac{\partial^2 l_j}{\partial\delta\partial\pi_{1j}}\right) = \frac{m_{+2j}\,(\gamma_j-2)}{\delta\,\pi_{1j}\,(\gamma_j-2)+1},$$

$$I_{13}^{(j)} = I_{31}^{(j)} = E\left(-\frac{\partial^2 l_j}{\partial\delta\partial\gamma_j}\right) = -\frac{m_{+2j}\,\pi_{1j}}{\delta\,\pi_{1j}\,(\gamma_j-2)+1},$$

$$I_{22}^{(j)} = E\left(-\frac{\partial^2 l_j}{\partial\pi_{1j}^2}\right) = \frac{m_{+1j}\,(\gamma_j-2)^2}{\pi_{1j}\,(\gamma_j-2)+1} - \frac{(\gamma_j-2)\,(m_{+1j}+\delta\,m_{+2j})}{\pi_{1j}} - \frac{\delta^2\,m_{+2j}\,(\gamma_j-2)^2}{\pi_{1j}\,(2\,\delta-\delta\,\gamma_j)-1},$$

$$I_{23}^{(j)} = I_{32}^{(j)} = E\left(-\frac{\partial^2 l_j}{\partial\pi_{1j}\gamma_j}\right) = \frac{\delta\,m_{+2j}}{\pi_{1j}\,(2\,\delta-\delta\,\gamma_j)-1} - \frac{m_{+1j}}{\pi_{1j}\,(\gamma_j-2)+1},$$

$$\begin{aligned} I_{33}^{(j)} = E\left(-\frac{\partial^2 l_j}{\partial\gamma_j^2}\right) &= \frac{m_{+1j}\,\pi_{1j}}{\gamma_j} - \frac{2\,m_{+1j}\,\pi_{1j}}{\gamma_j-1} + \frac{m_{+1j}\,{\pi_{1j}}^2}{\pi_{1j}\,(\gamma_j-2)+1} \\ &+ \frac{\delta^2\,m_{+2j}\,{\pi_{1j}}^2}{\delta\,\pi_{1j}\,(\gamma_j-2)+1} + \frac{\delta\,m_{+2j}\,\pi_{1j}}{\gamma_j} - \frac{2\,\delta\,m_{+2j}\,\pi_{1j}}{\gamma_j-1}. \end{aligned}$$

Appendix A.2. Derivation of Wald-Type Statistic

The first-order differential equations of stratum j are:

$$\frac{\partial l_j}{\partial\delta_j} = \frac{m_{12j}}{\delta_j} + \frac{m_{22j}}{\delta_j} + \frac{m_{02j}\,\pi_{1j}\,(\gamma_j-2)}{\delta_j\,\pi_{1j}\,(\gamma_j-2)+1},$$

$$\frac{\partial l_j}{\partial \pi_{1j}} = \frac{m_{12j}}{\pi_{1j}} + \frac{m_{21j}}{\pi_{1j}} + \frac{m_{22j}}{\pi_{1j}} + \frac{m_{01j}\,(\gamma_j-2)}{\pi_{1j}\,(\gamma_j-2)+1} + \frac{\delta_j\, m_{02j}\,(\gamma_j-2)}{\delta_j\,\pi_{1j}\,(\gamma_j-2)+1} + \frac{m_{11j}\,(2\,\gamma_j-2)}{2\,\pi_{1j}\,(\gamma_j-1)},$$

$$\frac{\partial l_j}{\partial \gamma_j} = \frac{m_{21j}}{\gamma_j} + \frac{m_{22j}}{\gamma_j} + \frac{m_{11j}}{\gamma_j-1} + \frac{m_{12j}}{\gamma_j-1} + \frac{m_{01j}\,\pi_{1j}}{\pi_{1j}\,(\gamma_j-2)+1} + \frac{\delta_j\, m_{02j}\,\pi_{1j}}{\delta_j\,\pi_{1j}\,(\gamma_j-2)+1}.$$

The second-order differential equations of stratum j are:

$$\frac{\partial^2 l_j}{\partial \delta_j^2} = -\frac{m_{12j}}{\delta_j^2} - \frac{m_{22j}}{\delta_j^2} - \frac{m_{02j}\,{\pi_{1j}}^2\,(\gamma_j-2)^2}{(\delta_j\,\pi_{1j}\,(\gamma_j-2)+1)^2},$$

$$\frac{\partial^2 l_j}{\partial \delta_j \partial \pi_{1j}} = \frac{m_{02j}\,(\gamma_j-2)}{\delta_j\,\pi_{1j}\,(\gamma_j-2)+1} - \frac{\delta_j\, m_{02j}\,\pi_{1j}\,(\gamma_j-2)^2}{(\delta_j\,\pi_{1j}\,(\gamma_j-2)+1)^2},$$

$$\frac{\partial^2 l_j}{\partial \delta_j \partial \gamma_j} = \frac{m_{02j}\,\pi_{1j}}{\delta_j\,\pi_{1j}\,(\gamma_j-2)+1} - \frac{\delta_j\, m_{02j}\,{\pi_{1j}}^2\,(\gamma_j-2)}{(\delta_j\,\pi_{1j}\,(\gamma_j-2)+1)^2},$$

$$\frac{\partial^2 l_j}{\partial \pi_{1j} \partial \delta_j} = \frac{m_{02j}\,(\gamma_j-2)}{\delta_j\,\pi_{1j}\,(\gamma_j-2)+1} - \frac{\delta_j\, m_{02j}\,\pi_{1j}\,(\gamma_j-2)^2}{(\delta_j\,\pi_{1j}\,(\gamma_j-2)+1)^2},$$

$$\frac{\partial^2 l_j}{\partial \pi_{1j}^2} = -\frac{m_{12j}}{{\pi_{1j}}^2} - \frac{m_{21j}}{{\pi_{1j}}^2} - \frac{m_{22j}}{{\pi_{1j}}^2} - \frac{m_{01j}\,(\gamma_j-2)^2}{(\pi_{1j}\,(\gamma_j-2)+1)^2} - \frac{\delta^2\, m_{02j}\,(\gamma_j-2)^2}{(\delta_j\,\pi_{1j}\,(\gamma_j-2)+1)^2} - \frac{m_{11j}\,(2\,\gamma_j-2)}{2\,{\pi_{1j}}^2\,(\gamma_j-1)},$$

$$\frac{\partial^2 l_j}{\partial \pi_{1j} \partial \gamma_j} = \frac{m_{01j}}{(\pi_{1j}\,(\gamma_j-2)+1)^2} + \frac{\delta_j\, m_{02j}}{(\delta_j\,\pi_{1j}\,(\gamma_j-2)+1)^2},$$

$$\frac{\partial^2 l_j}{\partial \gamma_j \partial \delta_j} = \frac{m_{02j}\,\pi_{1j}}{\delta_j\,\pi_{1j}\,(\gamma_j-2)+1} - \frac{\delta_j\, m_{02j}\,{\pi_{1j}}^2\,(\gamma_j-2)}{(\delta_j\,\pi_{1j}\,(\gamma_j-2)+1)^2},$$

$$\frac{\partial^2 l_j}{\partial \gamma_j \partial \pi_{1j}} \frac{m_{01j}}{(\pi_{1j}\,(\gamma_j-2)+1)^2} + \frac{\delta_j\, m_{02j}}{(\pi_{1j}\,(2\,\delta_j - \delta_j\,\gamma_j)-1)^2},$$

$$\frac{\partial^2 l_j}{\partial \gamma_j^2} = -\frac{m_{21j}}{{\gamma_j}^2} - \frac{m_{22j}}{{\gamma_j}^2} - \frac{m_{11j}}{(\gamma_j-1)^2} - \frac{m_{12j}}{(\gamma_j-1)^2} - \frac{m_{01j}\,{\pi_{1j}}^2}{(\pi_{1j}\,(\gamma_j-2)+1)^2} - \frac{\delta^2\, m_{02j}\,{\pi_{1j}}^2}{(\delta_j\,\pi_{1j}\,(\gamma_j-2)+1)^2}.$$

The Fisher information of stratum j is:

$$I_{W11}^{(j)} = E\left(-\frac{\partial^2 l_j}{\partial \delta_j^2}\right) = \frac{m_{+2j}}{\delta_j^2\,(\delta_j\,\pi_{1j}\,(\gamma_j-2)+1)} - \frac{m_{+2j}}{\delta_j^2},$$

$$I_{W12}^{(j)} = I_{W21}^{(j)} = E\left(-\frac{\partial^2 l_j}{\partial \delta_j \partial \pi_{1j}}\right) = \frac{m_{+2j}\,(\gamma_j-2)}{\delta_j\,\pi_{1j}\,(\gamma_j-2)+1},$$

$$I_{W13}^{(j)} = I_{W31}^{(j)} = E\left(-\frac{\partial^2 l_j}{\partial \delta_j \partial \gamma_j}\right) = -\frac{m_{+2j}\,\pi_{1j}}{\delta_j\,\pi_{1j}\,(\gamma_j-2)+1},$$

$$I_{W22}^{(j)} = E\left(-\frac{\partial^2 l_j}{\partial \pi_{1j}^2}\right) = \frac{m_{+1j}\,(\gamma_j-2)^2}{\pi_{1j}\,(\gamma_j-2)+1} - \frac{(\gamma_j-2)\,(m_{+1j}+\delta_j\, m_{+2j})}{\pi_{1j}} - \frac{\delta_j^2\, m_{+2j}\,(\gamma_j-2)^2}{\pi_{1j}\,(2\,\delta_j-\delta_j\,\gamma_j)-1},$$

$$I_{W23}^{(j)} = I_{W32}^{(j)} = E\left(-\frac{\partial^2 l_j}{\partial \pi_{1j} \gamma_j}\right) = \frac{\delta_j\, m_{+2j}}{\pi_{1j}\,(2\,\delta_j-\delta_j\,\gamma_j)-1} - \frac{m_{+1j}}{\pi_{1j}\,(\gamma_j-2)+1},$$

$$I_{W33}^{(j)} = E\left(-\frac{\partial^2 l_j}{\partial \gamma_j^2}\right) = \frac{m_{+1j}\,\pi_{1j}}{\gamma_j} - \frac{2\,m_{+1j}\,\pi_{1j}}{\gamma_j - 1} + \frac{m_{+1j}\,{\pi_{1j}}^2}{\pi_{1j}\,(\gamma_j - 2) + 1} + \frac{\delta_j^2\,m_{+2j}\,{\pi_{1j}}^2}{\delta_j\,\pi_{1j}\,(\gamma_j - 2) + 1} + \frac{\delta_j\,m_{+2j}\,\pi_{1j}}{\gamma_j} - \frac{2\,\delta_j\,m_{+2j}\,\pi_{1j}}{\gamma_j - 1}.$$

Appendix A.3. Empirical Type I Error Rates

Table A1. Simulation results of the empirical sizes (percentage) for 4 strata.

			***m* = 25**			***m* = 50**			***m* = 100**		
δ	γ	π_1	T_L	T_{SC}	T_W	T_L	T_{SC}	T_W	T_L	T_{SC}	T_W
1.0	I	a	5.82	5.45	4.33	5.33	5.17	4.35	5.07	5.03	4.59
		b	5.41	5.20	2.44	5.21	5.09	3.29	5.13	5.08	4.04
		c	5.73	5.46	2.49	5.31	5.20	3.19	5.11	5.05	4.00
	II	a	5.52	5.13	4.96	5.23	5.10	4.78	5.12	5.03	4.92
		b	5.54	5.28	2.41	5.23	5.12	3.27	5.12	5.07	4.06
		c	5.63	5.31	2.89	5.29	5.15	3.30	5.13	5.04	4.09
	III	a	5.75	5.41	3.79	5.26	5.08	4.11	5.06	5.01	4.53
		b	5.52	5.30	2.39	5.05	4.93	3.14	5.05	5.01	3.91
		c	5.69	5.35	2.22	5.35	5.20	3.20	5.15	5.10	3.81
	IV	a	5.59	5.16	4.05	5.23	5.00	4.55	5.19	5.10	4.64
		b	5.55	5.30	1.84	5.23	5.12	2.85	5.16	5.09	3.72
		c	5.77	5.33	2.43	5.36	5.18	3.01	5.15	5.08	3.75
1.2	I	a	5.67	5.34	3.67	5.26	5.10	4.16	5.17	5.08	4.51
		b	5.65	5.40	2.05	5.23	5.13	3.18	5.12	5.07	3.90
		c	5.65	5.35	1.96	5.26	5.13	2.90	5.21	5.14	3.74
	II	a	5.77	5.41	4.60	5.19	5.09	4.55	4.97	4.89	4.70
		b	5.54	5.32	1.92	5.34	5.19	2.96	5.20	5.17	4.02
		c	5.71	5.37	2.37	5.32	5.18	3.33	5.04	4.95	3.96
	III	a	5.56	5.24	3.56	5.21	5.03	3.85	5.27	5.17	4.41
		b	5.45	5.21	1.86	5.41	5.26	3.07	5.17	5.11	3.85
		c	5.53	5.22	1.82	5.32	5.20	2.74	5.38	5.32	3.81
	IV	a	5.72	5.26	4.00	5.30	5.13	4.11	4.96	4.88	4.38
		b	5.47	5.19	1.56	5.15	5.04	2.48	5.23	5.18	3.61
		c	5.78	5.37	1.92	5.29	5.13	2.84	5.41	5.30	3.65
0.8	I	a	5.56	5.23	4.86	5.24	5.03	4.66	5.16	5.07	4.71
		b	5.50	5.21	2.98	5.03	4.92	3.59	5.09	5.01	4.25
		c	5.77	5.34	2.98	5.24	5.06	3.37	5.24	5.16	4.18
	II	a	5.60	5.14	5.68	5.21	5.04	5.18	4.94	4.88	4.84
		b	5.53	5.23	2.87	5.37	5.23	3.64	5.14	5.08	4.22
		c	5.73	5.33	3.47	5.20	5.01	3.73	5.24	5.17	4.37
	III	a	5.64	5.19	4.57	5.42	5.25	4.60	5.04	4.95	4.62
		b	5.63	5.34	2.96	5.43	5.27	3.76	5.09	5.04	4.19
		c	5.80	5.36	2.82	5.39	5.20	3.46	5.19	5.10	4.04
	IV	a	5.83	5.24	5.10	5.38	5.15	4.90	5.15	5.06	4.78
		b	5.66	5.30	2.35	5.41	5.23	3.24	5.13	5.06	3.96
		c	5.75	5.12	3.02	5.33	5.07	3.36	5.02	4.89	3.79

Table A2. Simulation results of the empirical sizes (percentage) for 6 strata.

			***m* = 25**			***m* = 50**			***m* = 100**		
δ	γ	π_1	T_L	T_{SC}	T_W	T_L	T_{SC}	T_W	T_L	T_{SC}	T_W
1.0	I	a	5.90	5.47	4.77	5.30	5.11	4.55	5.06	4.99	4.63
		b	5.56	5.25	2.67	5.32	5.17	3.39	5.23	5.15	4.19
		c	5.89	5.49	2.80	5.37	5.17	3.31	5.17	5.07	3.95
	II	a	5.72	5.20	5.61	5.33	5.10	5.00	5.03	4.92	4.85
		b	5.57	5.27	2.54	5.50	5.32	3.45	5.11	5.05	3.97
		c	5.58	5.15	3.07	5.33	5.12	3.63	5.25	5.16	4.33
	III	a	5.66	5.19	4.50	5.31	5.08	4.37	5.29	5.20	4.71
		b	5.79	5.41	2.57	5.33	5.17	3.40	5.14	5.05	4.00
		c	6.03	5.59	2.73	5.45	5.23	3.30	5.08	4.98	3.88
	IV	a	5.98	5.38	5.21	5.38	5.13	4.72	5.14	5.01	4.73
		b	5.97	5.58	2.38	5.26	5.07	3.06	5.18	5.09	3.87
		c	5.91	5.27	2.93	5.31	5.04	3.40	5.26	5.15	3.91
1.2	I	a	5.79	5.34	4.14	5.28	5.04	4.24	5.22	5.13	4.55
		b	5.75	5.39	2.27	5.40	5.24	3.23	5.13	5.07	4.01
		c	5.56	5.19	2.32	5.60	5.44	3.24	5.24	5.16	3.78

Table A2. *Cont.*

δ	γ	π_1	m = 25 T_L	T_{SC}	T_W	m = 50 T_L	T_{SC}	T_W	m = 100 T_L	T_{SC}	T_W
	II	a	5.93	5.43	4.87	5.27	5.00	4.59	5.12	4.98	4.75
		b	5.83	5.47	2.11	5.14	4.97	3.02	5.11	5.03	3.92
		c	5.79	5.37	2.74	5.54	5.32	3.55	5.30	5.19	4.08
	III	a	5.81	5.32	3.92	5.36	5.14	4.13	5.25	5.10	4.35
		b	5.68	5.33	2.15	5.37	5.23	3.05	5.15	5.10	3.81
		c	5.79	5.36	2.06	5.28	5.08	2.96	5.31	5.20	3.72
	IV	a	5.96	5.42	4.48	5.21	4.92	4.37	5.27	5.17	4.62
		b	5.77	5.42	1.77	5.36	5.16	2.81	5.27	5.19	3.66
		c	5.74	5.21	2.28	5.58	5.32	3.17	5.36	5.26	3.84
0.8	I	a	5.97	5.44	5.71	5.23	4.99	5.01	5.21	5.11	5.17
		b	5.81	5.41	3.51	5.41	5.24	4.04	5.21	5.11	4.32
		c	6.06	5.46	3.69	5.21	4.99	3.65	5.15	5.06	4.29
	II	a	5.88	5.29	6.69	5.29	5.07	5.64	5.00	4.87	5.15
		b	5.69	5.32	3.41	5.28	5.09	3.83	5.02	4.93	4.21
		c	5.93	5.38	4.21	5.36	5.13	4.17	5.18	5.09	4.53
	III	a	6.00	5.39	5.73	5.29	5.09	4.87	5.31	5.22	4.99
		b	5.90	5.52	3.49	5.25	5.07	3.67	5.32	5.21	4.41
		c	6.05	5.44	3.63	5.26	5.06	3.62	5.26	5.16	4.22
	IV	a	6.24	5.46	6.31	5.43	5.09	5.45	5.35	5.19	5.09
		b	5.93	5.44	3.03	5.35	5.15	3.46	5.10	4.98	4.10
		c	6.14	5.34	4.21	5.43	5.11	3.93	5.24	5.10	4.18

Table A3. Simulation results of the empirical sizes (percentage) for 8 strata.

δ	γ	π_1	m = 25 T_L	T_{SC}	T_W	m = 50 T_L	T_{SC}	T_W	m = 100 T_L	T_{SC}	T_W
1.0	I	a	5.87	5.39	5.01	5.40	5.15	4.72	5.02	4.91	4.68
		b	5.66	5.28	2.79	5.37	5.16	3.58	5.17	5.08	4.08
		c	6.04	5.51	3.06	5.40	5.19	3.68	5.16	5.06	4.03
	II	a	5.98	5.33	5.96	5.47	5.16	5.43	5.11	4.97	5.02
		b	5.62	5.19	2.64	5.36	5.19	3.54	5.13	5.04	3.99
		c	5.67	5.08	3.61	5.40	5.16	3.78	5.03	4.94	4.06
	III	a	5.81	5.31	4.85	5.42	5.19	4.65	5.11	4.98	4.52
		b	5.80	5.38	2.83	5.13	4.94	3.24	5.24	5.15	4.06
		c	5.97	5.48	2.96	5.44	5.21	3.44	5.31	5.21	4.11
	IV	a	5.90	5.21	5.68	5.53	5.23	5.07	5.31	5.16	4.90
		b	6.07	5.62	2.38	5.23	5.02	3.10	5.19	5.10	3.90
		c	6.19	5.45	3.45	5.44	5.18	3.63	5.20	5.08	4.04
1.2	I	a	5.95	5.37	4.26	5.37	5.12	4.38	5.27	5.14	4.72
		b	5.76	5.27	2.31	5.50	5.28	3.27	5.31	5.21	3.97
		c	5.84	5.36	2.35	5.39	5.15	3.15	5.43	5.33	4.06
	II	a	6.09	5.40	5.15	5.49	5.20	5.09	5.19	5.06	4.95
		b	5.72	5.29	2.12	5.20	4.98	3.02	5.13	5.03	3.90
		c	6.03	5.49	3.03	5.43	5.23	3.40	5.21	5.11	4.07
	III	a	5.86	5.22	3.99	5.34	5.01	4.19	5.30	5.15	4.52
		b	5.71	5.28	2.18	5.40	5.26	3.11	5.18	5.10	4.00
		c	5.80	5.31	2.37	5.49	5.29	3.02	5.33	5.23	3.93
	IV	a	5.82	5.22	4.93	5.43	5.15	4.62	5.12	4.98	4.63
		b	5.59	5.17	1.80	5.30	5.11	2.80	5.38	5.30	3.75
		c	6.01	5.38	2.58	5.52	5.23	3.30	5.25	5.10	3.83
0.8	I	a	6.10	5.48	6.34	5.39	5.12	5.46	5.16	5.01	5.15
		b	5.85	5.43	3.84	5.40	5.19	4.12	5.14	5.02	4.47
		c	5.83	5.23	4.21	5.35	5.08	4.11	5.04	4.94	4.23
	II	a	5.92	5.28	7.28	5.47	5.24	5.95	5.30	5.19	5.45
		b	5.77	5.32	3.69	5.33	5.14	3.94	5.20	5.11	4.25
		c	6.12	5.41	4.79	5.42	5.17	4.40	5.18	5.07	4.61
	III	a	5.94	5.33	5.89	5.45	5.18	5.44	5.08	4.99	4.92
		b	5.79	5.28	3.77	5.56	5.32	4.04	5.23	5.13	4.43
		c	6.03	5.32	4.22	5.45	5.13	3.99	5.44	5.28	4.36
	IV	a	6.30	5.38	7.31	5.48	5.12	5.70	5.20	5.04	5.29
		b	5.77	5.25	3.46	5.46	5.23	3.73	5.33	5.23	4.25
		c	6.35	5.36	4.69	5.48	5.14	4.36	5.37	5.18	4.51

References

1. Rosner, B. Statistical methods in ophthalmology: An adjustment for the intraclass correlation between eyes. *Biometrics* **1982**, *38*, 105–114. [CrossRef]
2. Donner, A. Statistical methods in ophthalmology: An adjusted chi-square approach. *Biometrics* **1989**, *45*, 605–611. [CrossRef]

3. Dallal, G.E. Paired Bernoulli Trials. *Biometrics* **1988**, *44*, 253–257. [CrossRef] [PubMed]
4. Ma, C.X.; Liu, S. Testing equality of proportions for correlated binary data in ophthalmologic studies. *J. Biopharm. Stat.* **2017**, *27*, 611–619. [CrossRef]
5. Ma, C.X.; Shan, G.; Liu, S. Homogeneity test for binary correlated data. *PLoS ONE* **2015**, *10*, e0124337. [CrossRef]
6. Tang, N.S.; Zhang, B.; Li, H.Q. Homogeneity test of difference between two correlated proportions in stratified matched-pair studies. *J. Biopharm. Stat.* **2013**, *23*, 1261–1280. [CrossRef] [PubMed]
7. Tang, N.S.; Tang, M.L.; Qiu, S.F. Testing the equality of proportions for correlated otolaryngologic data. *Comput. Stat. Data Anal.* **2008**, *52*, 3719–3729. [CrossRef]
8. Liu, X.; Shan, G.; Tian, L.; Ma, C.X. Exact methods for testing homogeneity of proportions for correlated multiple groups paired binary data. *Commun. Stat. Comput.* **2017**, *46*, 6074–6082. [CrossRef]
9. Tang, M.L.; Tang, N.S.; Rosner, B. Statistical inference for correlated data in ophthalmologic studies. *Stat. Med.* **2006**, *25*, 2771–2783. [CrossRef] [PubMed]
10. Thompson, J.R. The chi-square test for data collected on eyes. *Br. J. Ophthalmol.* **1993**, *77*. [CrossRef] [PubMed]
11. Liu, X.; Liu, S.; Ma, C.X. Testing equality of correlation coefficients for paired binary data from multiple groups. *Commun. Stat. Comput.* **2015**, *86*, 1686–1696. [CrossRef]
12. Li, Z.; Ma, C.; Ai, M. Statistical tests under Dallal's model: Asymptotic and exact methods. *PLoS ONE* **2020**, *15*, e0242722. [CrossRef] [PubMed]
13. Chen, Y.; Li, Z.; Ma, C. Further study on testing the equality of response rates under Dallal's model. *Stat. Interface* **2022**, *15*, 115–126. [CrossRef]
14. Newcombe, R.G. *Confidence Intervals for Proportions and Related Measures of Effect Size*; CRC Press: Boca Raton, FL, USA, 2012.
15. Wang, W.; Shan, G. Exact confidence intervals for the relative risk and the odds ratio. *Biometrics* **2015**, *71*, 985–995. [CrossRef] [PubMed]
16. Tang, N.S.; Qiu, S.F.; Tang, M.L.; Pei, Y.B. Asymptotic confidence interval construction for proportion difference in medical studies with bilateral data. *Stat. Methods Med. Res.* **2011**, *20*, 233–259. [CrossRef]
17. Tang, N.S.; Qiu, S.F. Homogeneity test, sample size determination and interval construction of difference of two proportions in stratified bilateral-sample designs. *J. Stat. Plan. Inference* **2012**, *142*. [CrossRef]
18. Zhuang, T.; Tian, G.L.; Ma, C.X. Homogeneity Test of Ratio of Two Proportions in Stratified Bilateral Data. *Stat. Biopharm. Res.* **2018**, *11*, 200–209. [CrossRef]
19. Shen, X.; Ma, C.X. Testing homogeneity of difference of two proportions for stratified correlated paired binary data. *J. Appl. Stat.* **2018**, *45*, 1410–1425. [CrossRef]
20. Xue, Y.; Ma, C.X. Interval estimation of proportion ratios for stratified bilateral correlated binary data. *Stat. Methods Med. Res.* **2020**, *29*, 1987–2014. [CrossRef]
21. Mandel, E.M.; Bluestone, C.D.; Rockette, H.E.; Blatter, M.M.; Reisinger, K.S.; Wucher, F.P.; Harper, J. Duration of effusion after antibiotic treatment for acute otitis media: Comparison of cefaclor and amoxicillin. *Pediatr. Infect. Dis.* **1982**, *1*, 310–316. [CrossRef]
22. Postlethwaite, A.E.; Wong, W.K.; Clements, P.; Chatterjee, S.; Fessler, B.J.; Kang, A.H.; Korn, J.; Mayes, M.; Merkel, P.A.; Molitor, J.A.; et al. A multicenter, randomized, double-blind, placebo-controlled trial of oral type I collagen treatment in patients with diffuse cutaneous systemic sclerosis: I. oral type I collagen does not improve skin in all patients, but may improve skin in late-phase disease. *Arthritis Rheumatol.* **2008**, *58*, 1810–1822.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.