J Med Sci 2020;40(3):119-126
DOI: 10.4103/jmedsci.jmedsci_133_19

ORIGINAL ARTICLE

A Superior Odds Ratio Compared to the Risk Ratio when Estimating Moderator Effects in Meta-Regression Analyses of Randomized Controlled Trials

Chih-Chien Chiu[1], Chien-Fu Chen[2], Po-Jen Hsiao[3], Dung-Jang Tsai[4], Hsueh-Lu Chang[5], Wen-Hui Fang[6], Wei-Teing Chen[7], Jenq-Shyong Chan[3], Min-Tser Liao[8,9], Yi-Jung Ho[4,10], Wen Su[11], Ying-Kai Chen[12], Hui-Han Hu[5], Zheng-Zong Lai[13], Chin Lin[4]

[1]Division of Infectious Diseases, Department of Internal Medicine, Taoyuan Armed Forces General Hospital, Taoyuan, [2]Department of Orthopedics, Taichung Armed Forces General Hospital, Taichung, [3]Renal Division, Department of Internal Medicine, Taoyuan Armed Forces General Hospital, Taoyuan, [4]Graduate Institute of Life Sciences, National Defense Medical Center, Taipei, [5]School of Public Health, National Defense Medical Center, Taipei, [6]Department of Family and Community Medicine, Tri-Service General Hospital, National Defense Medical Center, Taipei, [7]Division of Thoracic Medicine, Department of Medicine, Cheng Hsin General Hospital, Taipei, [8]Department of Pediatrics, Taoyuan Armed Forces General Hospital, Taoyuan, [9]Division of Pediatrics, Department of Medicine, Tri-Service General Hospital, National Defense Medical Center, Taipei, [10]School of Pharmacy, National Defense Medical Center, Taipei, [11]Department of Nursing, Tri-Service General Hospital, Taipei, [12]Division of Nephrology, Department of Medicine, Zuoying Branch of Kaohsiung Armed Forces General Hospital, Kaohsiung, [13]Department and Gradute Institute of Pharmacology, National Defense Medical Center, Taipei, Taiwan, ROC

Background: Moderator effect assessment is important in personalized medicine. We mathematically prove that the average summary value is actually nonlinearly to logRR, and we assess the bias from linear meta-regression on logRR via simulation. **Methods:** In the meta-analysis of randomized controlled trials, the moderator effect is generally evaluated by the linear meta-regression of the logarithmic risk ratio (RR) versus the average summary value of the entire study population. **Conclusions:** We recommend using linear meta-regression on logarithmic odds ratio (logOR) since it has been shown that the average summary value is actually linear to logOR.

Key words: Moderator effect, interaction, personalized medicine, odds ratio, risk ratio, participant-level variable, study-level variable, meta-analysis, meta-regression, randomized controlled trial

INTRODUCTION

Detecting a moderator effect (or interaction) is important in personalized medicine. In randomized controlled trials (RCTs), the moderator effect is typically tested in *post hoc* subgroup analyses.[1] The list of prespecified subgroups could be selected from previous mechanistic studies, epidemiological studies, or RCTs. To explore potential unknown moderate factors was also important; they might help us to generate new hypotheses. When exploring unknown factors in multivariable studies, the significance levels in moderator effect detection must be reduced by adjusting *P* value. However, the power of a single trial is often insufficient after multicomparison correction.[2,3] Meta-analysis is a common method for increasing this power, and subgroup analysis may also help detect the moderator effect. The biggest challenge in detecting the moderator effect in a meta-analysis is data collection. Analyses of the subsets of participants within studies are rarely reviewed in the literature because most reports lack sufficient details to

Received: July 17, 2019; Revised: September 22, 2019; Accepted: October 25, 2019; Published: December 03, 2019
Corresponding Author: Dr. Chin Lin, National Defense Medical Center, Graduate Institute of Life Sciences, No. 161, Min-Chun E. Road, Section 6, Neihu, Taipei 114, Taiwan. Tel: 886-2-87923100#18448; Fax: 886-2-87923147. E-mail: xup6fup@gmail.com

How to cite this article: Chiu CC, Chen CF, Hsiao PJ, Tsai DJ, Chang HL, Fang WH, *et al.* A superior odds ratio compared to the risk ratio when estimating moderator effects in meta-regression analyses of randomized controlled trials. J Med Sci 2020;40:119-26.

extract the data on participant-level variables, such as gender and comorbidity.[4] Thus, most meta-analyses are limited to the moderator effects of the study-level variables, such as geographical location. However, most of potential moderator effects were hidden in participant-level variables, such as diabetes in folic acid treatment effect.[5] Researches needed to explore these potential issues for reducing heterogeneity.

Because participant-level variables are generally reported as their average summary values, there are two common methods for investigating their moderator effects. The first method selects a cutoff point that dichotomizes the included studies into "higher group" and "lower group." However, this method increases the likelihood of Type 1 errors[6,7] and yields imprecise risk estimates. The second method investigates the relation between the average summary value and the logarithmic risk ratio (*RR*) by meta-regression.[8-10] As shown in Figures 1 and S1, this relation is nonlinear; thus, this method may bias the estimates.

In a previous study, we proved a linear relationship between the average summary value of a case group and the logarithmic odds ratio (*OR*).[11] Thus, we inferred that the proposed method would improve moderator effect detection in the meta-analysis of RCTs. However, acknowledging the difficulty of accessing the average summary value of a case group, we also considered replacing this variable with the average summary value of the entire population. Therefore, this study compares the Type 1 error, power and 95% confidence interval coverage rate (CICR) of the following methods: (1) meta-regression of the logarithmic *RR* versus the average summary value of the entire population (the most common method); (2) meta-regression of the logarithmic *OR* versus the average summary value of the entire population; and (3) meta-regression of the logarithmic *OR* versus the average summary value of a case group.

METHODS

Derivations

The moderator effect of participant-level variables in the meta-analyses of RCTs is usually detected by using the average summary value, in which the association between the aggregated summary values of the factor and RR was based on multiple factors. For better understanding of this principle, we hereby describe an example.

When the independent variable (x) is the intervention encoded with value 0 for control group and 1 for treatment group, and the moderator (m) is diabetes status encoded with values 0 and 1 for without and with diabetes, respectively. The dependent variable is a binary outcome event (y) (with 0 and 1 signifying nonoccurrence and occurrence, respectively). Probabilities p_1, p_2, p_3, and p_4 are then defined as follows:

$$\begin{cases} p_1 = p\left(y=1 \mid x=0 \cap m=0\right) \\ p_2 = p\left(y=1 \mid x=1 \cap m=0\right) \\ p_3 = p\left(y=1 \mid x=0 \cap m=1\right) \\ p_4 = p\left(y=1 \mid x=1 \cap m=1\right) \end{cases}$$

which represents the outcome incidence of control group without diabetes, treatment group without diabetes, control group with diabetes, and treatment group with diabetes. According to above setting, the RRs of interest for patient without diabetes (RR_0 [m = 0]) and patient with diabetes (RR_1 [m = 1]) are calculated as follows, respectively:

$$\begin{cases} RR_0 = p_2 / p_1 \\ RR_1 = p_4 / p_3 \end{cases}$$

In an RCT, the pooled or combined *RR* (RR_{combine}) is a function of the diabetes prevalence in the entire study population, $q_{total} = p\left(m=1\right)$ and can be expressed as:

$$RR_{combine} = \frac{p_2\left(1-q_{total}\right)+p_4 q_{total}}{p_1\left(1-q_{total}\right)+p_3 q_{total}}$$

Let the π denote the *RR* of diabetes in the untreated individuals ($\pi = p_3/p_1$). We can express p_1, p_2, p_3, and p_4 in terms of p_1 and π alone including *RR*s as follows:

$$\begin{cases} p_1 = p_1 \\ p_2 = p_1 \times RR_0 \\ p_3 = p_1 \times \pi \\ p_4 = p_1 \times \pi \times RR_1 \end{cases}$$

This simplifies the calculation of RR_{combine}. Let the *ME* denote the moderator effect ($ME = RR_1/RR_0$), RR_{combine} can be calculated as follows [for details, Text S1]:

$$RR_{combine} = RR_0 \times \left(1 + \frac{\pi \times q_{total}}{\left(1-q_{total}\right)+\pi \times q_{total}} \times \left(ME-1\right)\right)$$

and the logarithm of RR_{combine} can be expressed as follows [for details, Text S1]:

$$\log\left(RR_{combine}\right) = \log\left(RR_0\right) + \log\left[\left(1-q_{total}\right)+\pi \times q_{total} \times ME\right] - \frac{\log\left[\left(1-q_{total}\right)+\pi \times q_{total}\right]}{\log\left(ME\right)} \times \log\left(ME\right) \quad (2.1\text{-}1)$$

Let log (RR_i) denotes the observed effect size of ith individual RCT, $q_{\text{total,i}}$ the proportion of the population with moderator status, η_i the random effects, and ε_i the residuals. In general, awareness studies, β_0 is the expected logarithmic *RR* of individuals without diabetes (log[RR_0]), and β_1 is the expected

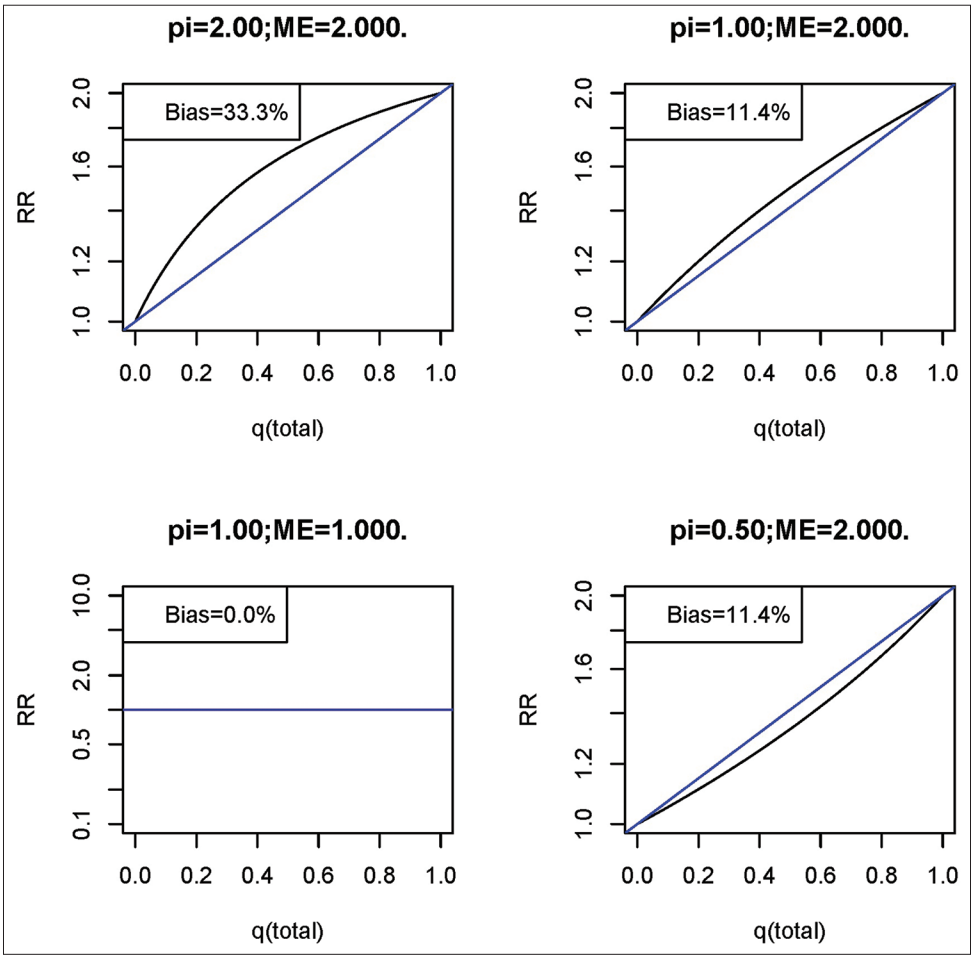

Figure 1: Relation between the average summary value and the logarithmic *RR*. Based on Equation 2.1-3, the true and linear relations differ only with respect to the *RR* of the moderator (π) and the *ME*. This figure plots the true and linear relations in several selected scenarios (π = 2.0 and *ME* = 2.0,π = 1.0 and *ME* = 2.0, π = 1.0 and *ME* = 1.0, and π = 0.5 and *ME* = 2.0). The X axis represents the average summary value throughout the study (*q*). Black and blue lines plot the expected RR and the linear relation, respectively. The bias is defined as the triangular area in the window straddling the two lines. The larger bias appears in that π/*ME* deviate null effect and in the same direction (π = 2.0 and *ME* = 2.0). The only scenarios without bias are that π/*ME* are null effects. There are bias of varying sizes when π or *ME* is not equal to 1. The details scenarios are shown in Figure S1. The intercept RR_0 is set equal to 1, and the Y axis is transformed into the logarithmic scale. *ME*: Moderator effect, *RR*: Risk ratio

logarithmic moderator effect (log[*ME*]). Thus, the traditional meta-regression (here called Method 1) can be written as.

$$\log\left(RR_i\right)=\beta_0+\beta_1\times q_{total,i}+\eta_i+\varepsilon_i \quad (2.1\text{-}2)$$

However, following Equation 2.1-1, the relation between log (RR_i) and $q_{total,i}$ is actually nonlinear:

$$\log\left(RR_i\right)=\beta_0+\beta_1\times\kappa_i+\eta_i+\varepsilon_i \quad (2.1\text{-}3)$$

$$\text{where },\ \kappa_i=\frac{\log\left[\left(1-q_{total,i}\right)+\pi\times q_{total,i}\times ME\right]-\log\left[\left(1-q_{total,i}\right)+\pi\times q_{total,i}\right]}{\log\left(ME\right)}$$

and π is the *RR* of diabetes in the untreated individuals. Note that π is an unknown population parameter not provided by most papers. That is why the traditional studies often used $q_{total,\ i}$ to replace κ_i. Obviously, this may cause bias since $\kappa_i = q_{total,\ i}$ only when π = 1 and *ME* = 1 [Figures 1 and S1]. Plot the true relation between the average summary value of the entire population (prevalence of diabetes in example) and the logarithmic *RR*. We observe that the bias is larger when the directions of π and *ME* are the same, and their effects leave 1 (null effect). The nonlinear relation will cause biased estimates of β_0 and β_1, and the bias extent in different situations is shown in simulation part.

The average summary value of a case group is approximately linearly related to the logarithmic *OR* under two assumptions: (1) rare disease and (2) independence between the independent variable and moderator, as in an earlier study.[11,12]

The independence assumption holds in the meta-analysis of RCTs, and the proposed method remains robust when the rare disease assumption is violated.[11] Thus, we expect that the meta-regression of the logarithmic *OR* versus the average summary value of the case group in study i ($q_{case,i}$, i.e., Method 3) will best detect the moderator effect in the meta-analysis of RCTs.

As the average summary value of a case group may be difficult to access, we investigate the viability of replacing the value with the average summary value of the entire population. Let RR_0 denote the *RR* of the treatment in individuals without the moderator; then, the variables $q_{total,i}$ and $q_{case,i}$ are related through Equation 2.1-4 [for this derivation, Text S2]:

$$q_{case,i} = \frac{q_{total,i}\left(\pi + RR_0 \times \pi \times ME\right)}{\left(1 - q_{total,i}\right)\left(1 + RR_0\right) + q_{total,i}\left(\pi + RR_0 \times \pi \times ME\right)} \quad (2.1\text{-}4)$$

The difference between $q_{total,\,i}$ and $q_{case,\,i}$ is affected by RR_0, *ME,* and π. The Pearson correlation between q_{total} and q_{case} exceeds 0.95 when RR_0, *ME,* and π are between 0.5 and 2.0 [Table S1]; thus, we considered that q_{total} can effectively represent q_{case}. Thus, the meta-regression of the logarithmic *OR* versus the average summary value of the entire population (i.e., Method 2) is a useful alternative when the average summary value of a case group is unavailable.

Simulations

In this subsection, we simulate a meta-analysis of RCTs. The simulation code is written in the R programming language; this code is provided in Appendix S1. As the meta-analysis data are summarized from individual data, we generated individual simulation studies. We generated 20 studies with each simulation using randomly generated sample sizes from a uniform distribution of (200, 1000). The probability of receiving treatment was set equal to 0.5 (the simple randomized design). The proportions of individuals with moderators were randomly generated from a uniform distribution (0, 1), and the treatment (*t*) was assumed to be independent of the moderator (*m*). The above steps were used to generate the information of the treatment and moderator.

The second step was to generate the disease information. The disease incidence of each individual was based on the relative risk model with four parameters: (1) the disease incidence of the entire population (*p*), which is not the incidence in the patients without treatment or the moderator; (2) the *RR* of treatment in individuals without the moderator (RR_0); (3) the *RR* of the moderator in individuals without treatment (π); and (4) the moderator effect (*ME*). Using these parameters, we calculated the disease incidence of untreated individuals without the moderator (*Incidence* [$t = 0 \cap m = 0$]) as follows (the detailed derivation was shown in the previous study):[11]

$$Incidence\left(t = 0 \cap m = 0\right) = \frac{4 \times p}{1 + RR_0 + \pi + RR_0 \times \pi \times ME}$$

We also calculated the disease incidences of individuals with other conditions as follows:

$$Incidence\left(t = 1 \cap m = 0\right) = Incidence\left(t = 0 \cap m = 0\right) \times RR_0$$

$$Incidence\left(t = 0 \cap m = 1\right) = Incidence\left(t = 0 \cap m = 0\right) \times \pi$$

$$Incidence\left(t = 1 \cap m = 1\right) = Incidence\left(t = 0 \cap m = 0\right) \times RR_0 \times \pi \times ME$$

The disease statuses of the individuals were randomly generated based on the above incidences. The values of *p*, RR_0, π, and *ME* used in the simulations are listed in Table 1.

The third step was to summarize the individual data into meta-analysis data. Each study provided four pieces of information for the following meta-regression analysis: (1) the logarithmic *RR* of treatment without stratification and its variance (i.e., log[RR_i] and var [log[RR_i]], respectively); (2) the logarithmic *OR* of treatment without stratification and its variance (i.e., log [OR_i] and var [log [OR_i]], respectively); (3) the proportion of individuals with the moderator throughout the study (q_{total}); and (4) the proportion of individuals with the moderator in the case group (q_{case}). The meta-regression analyses proceeded as follows, and we used the random effect model for the simulation. We used the "rma" command in the "metafor" package[13] in *R* to calculate the following meta-regression.

1. Meta-regression of the logarithmic *RR* versus the average summary value of the entire population

Regression formula: $\log\left(RR_i\right) = \beta_0 + \beta_1 q_{total,i} + \eta_i + \varepsilon_i$;

Table 1: Simulation conditions of this study

P	RR_0	*Π*	*ME*
0.1	1.5	2.0	2.0
0.2	1.0	1.5	1.5
	0.667	1.0	1.2
		0.667	1.0
		0.5	0.833
			0.667
			0.5

P: Disease incidence throughout the population, RR_0: Risk ratio of treatment in individuals without the moderator, *π*: Risk ratio of the moderator in individuals without treatment, ME: Moderator effect

$$Weight_i = \frac{1}{\mathrm{var}\left(\log\left(RR_i\right)\right) + \tau^2}$$

2. Meta-regression of the logarithmic OR versus the average summary value of the entire population

Regression formula: $\log\left(OR_i\right) = \beta_0 + \beta_1 q_{total,i} + \eta_i + \varepsilon_i$;

$$Weight_i = \frac{1}{\mathrm{var}\left(\log\left(OR_i\right)\right) + \tau^2}$$

3. Meta-regression of the logarithmic OR versus the average summary value of the case group

Regression formula: $\log\left(OR_i\right) = \beta_0 + \beta_1 q_{total,i} + \eta_i + \varepsilon_i$;

$$Weight_i = \frac{1}{\mathrm{var}\left(\log\left(OR_i\right)\right) + \tau^2}$$

where log (RR_i) is the observed logarithmic *RR* in study i; $q_{total,i}$ is the proportion of the population with moderator status in study i; $q_{case,i}$ is the proportion of the population with moderator status of case group in study i; η_i is the random effects; ε_i is the residuals; i is an identifier of an individual RCT ranging from 1 to 20; β_0 is the expected logarithmic *RR* of individuals without moderator status; β_1 is the expected logarithmic moderator effect; and τ^2 is the variance of random effects, which is estimated by the restricted maximum likelihood method.

The primary outcomes were the 95% CICRs of the moderator effect (β_1) and the intercept (β_0). The CICR defines the proportion of the 95% CIs that include the real parameter. The appropriate CI coverage was 95%. In addition, Type 1 errors were assessed in the null moderator effect model (*ME* = 1). As nonsignificant results are often ignored, we also computed the power of the moderator-effect assessment as the secondary outcome. Data under each condition were acquired from 10,000 simulations.

RESULTS

Figure 2 shows selected simulations relating to Equation 2.1-3 and Figure 1. Conditions A, B, and C differ only by the value of π. Method 3 is shown to produce the most robust results, and the 95% CICRs of the slope and intercept approximate 0.95 under all conditions. Methods 1 and 2 introduce varying degrees of bias, which is higher in the intercept than in the slope. However, Method 2 is more robust than Method 1. As expected, the conditions with larger bias (π/ME = 2.0/0.5) provide the same results as the earlier derivation [Figures 1 and S1]. In all methods, the 95% CICRs approximate 0.95 when *ME* = 1.0, which indicates that the false-positive rates of all methods are acceptable. However, the methods differ in their statistical power. Overall, Methods 2 and 3 exhibit higher statistical powers than Method 1. As described above, the parameters π and *ME* affect the bias in Method 1 because they damage the linear relation between the logarithmic *RR* and the average summary value throughout the study.

Figures S2 and S3 in the supplementary material present the detailed 95% CICRs of the slope and intercept, respectively, in the three meta-regression methods. Figure S4 shows the false-positive rates and the powers of the moderator effect at the 0.05 significance level in the meta-regression methods. The 95% CICR of the slope is marginally reduced for *P* = 0.2 and π/ME = 2.0/0.5 in Method 3; however, the bias nearly disappears when *P* = 0.1. This may correspond to the rare disease assumption of Method 3. Method 2 shows higher bias than Method 3 under all conditions but remains more robust than Method 1. The powers of Methods 2 and 3 are similar under all conditions and frequently exceed that of Method 1. Method 1 is insensitive to the values of RR_0 and *p*, which is also consistent with the earlier derivation. The marginal difference introduced by these parameters might be due to the weights in each study. The effects of the weights are described in the following section.

DISCUSSION

Meta-analysis is a powerful tool to determine the consistency of evidence. However, the pooled results often have high heterogeneity because the treatment effect is different in different patients.[14] The mission of personalized medicine is to advocate for the practice of personalized health care; therefore, investigating the source of heterogeneity is important. Meta-regression is a common method to explain heterogeneity, but the previous study suggested that meta-regression was only suitable on study-level variables.[4] However, personalized medicine is focused on effects at the individual patient level rather than the study level. Thus, investigating the source of heterogeneity of participant-level variables is important. Although certain studies investigated the relation between the average summary participant-level value and the logarithmic *RR* by meta-regression,[8-10] this relation is nonlinear, as shown in Figures 1 and S1. Thus, recent research has criticized this method.[14]

As an earlier study proved a linear relationship between the average summary value of a case group and the logarithmic OR,[11] we expected that the performance of the meta-regression using OR would exceed that when using RR. Consistent with the expected results, Method 1 exhibited higher bias than Methods 2 and 3 in all conditions investigated in this study. Thus, although certain conditions improved the power of Method 1 above those of Methods 2 and 3, we inferred that Method 1 is not a robust method. Meta-analysis attempts

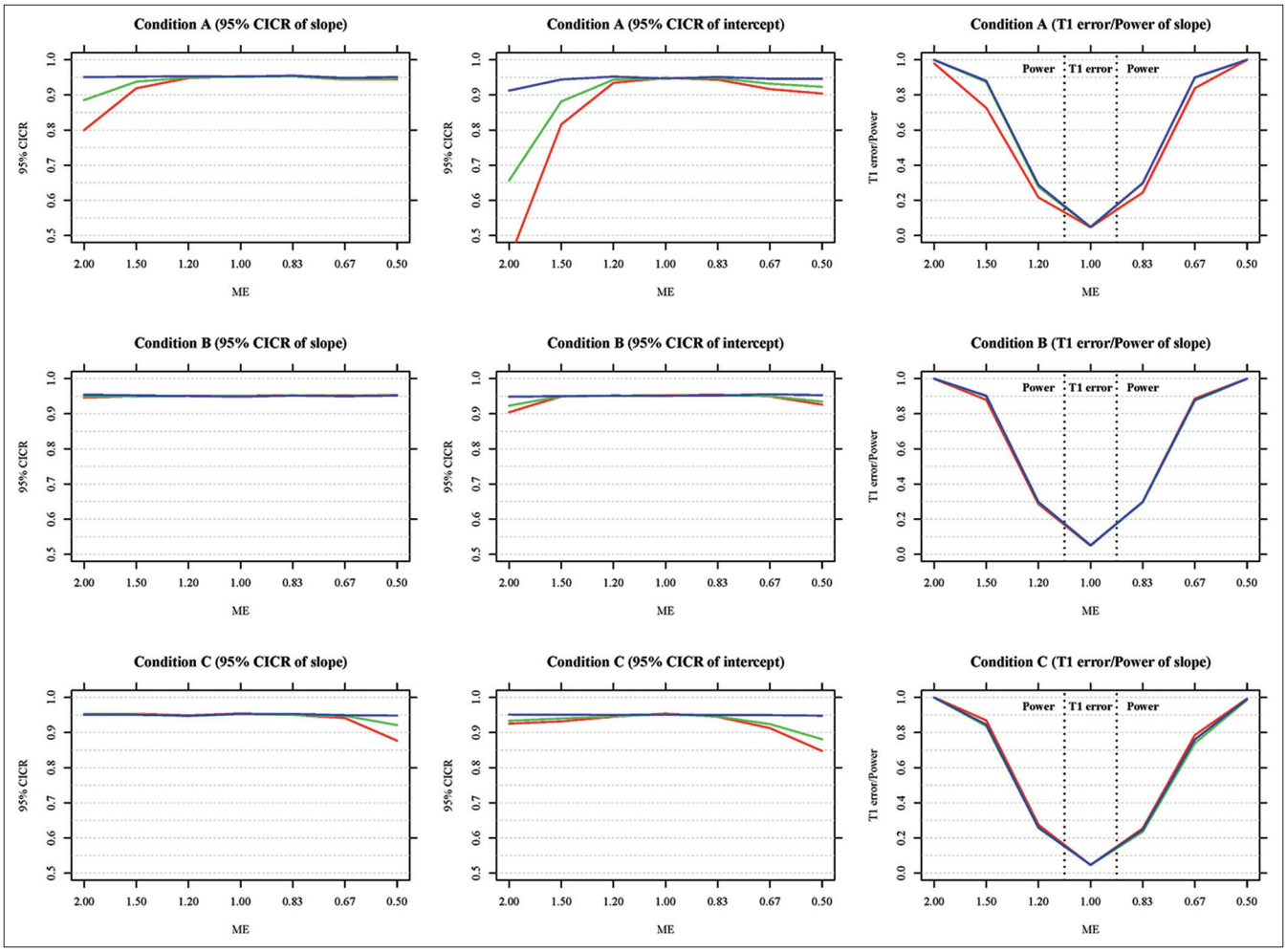

Figure 2: Results from selected simulations. Based on Equation 2.1-3, Figures 1 and S1, the largest bias occurs when $\pi/ME = 2.0/0.5$. Thus, under conditions A and C, π is set equal to 2.0 and 0.5, respectively. We also select a condition that produces the smallest expected bias (condition B). The parameters P and RR_0 are set equal to 0.2 and 1.0, respectively. Condition A: $P = 0.2$; $RR_0 = 1.0$; $\pi = 2.0$; ME = 2.0-0.5; Condition B: $P = 0.2$; $RR_0 = 1.0$; $\pi = 1.0$; ME = 2.0-0.5; Condition C: $P = 0.2$; $RR_0 = 1.0$; $\pi = 0.5$; ME = 2.0-0.5. Red line: Method 1; Green line: Method 2; Blue line: Method 3. *ME*: Moderator effect

to determine the P value and to estimate the accuracy of a treatment effect.[4] Thus, a low-bias method is required and is achieved by Methods 2 and 3.

The larger bias of Method 2 compared to Method 3 might arise from the attenuation bias because we replace the average summary value of the case group with the average summary value of the entire population. [15,16] However, the Pearson correlations between q_{total} and q_{case} exceeded 0.95 for appropriate values of RR_0, *ME* and π (i.e., between 0.5 and 2.0) [Table S1]. The Pearson correlations between q_{total} and q_{case} decrease as *ME* increases, but a large *ME* is easily detected in a single study.[17] The meta-analysis should focus on a smaller *ME*, which is difficult to detect in a single study. Thus, we considered that the average summary value of the entire population is an acceptable substitute for the average summary value of the case group. The lower bias in Method 2 compared to Method 1 indicates the superiority of *OR* over *RR* in the meta-regression.

The results with different simulation conditions were not entirely consistent with the expected bias. In Figures 1 and S1, the biases are equal when $\pi/ME = 2.0$ and $\pi/ME = 0.5$; however, differences emerge in Figure 2. These differences are introduced by the weights in each study, which depend on several factors: the sample size (N), the random effect variance (τ^2), the disease incidence of untreated individuals without the moderator (p_1), the *RR* of treatment in individuals without the moderator (RR_0), the *RR* of the moderator in untreated individuals (π), the moderator effect (*ME*), and the proportions of individuals with the moderator throughout the study population in study i ($q_{total, i}$). The dependences of the study weights ($Weight_i$) in study i on these parameters are given by Equations 4-1.1 and 4-1.2 (for detailed derivations), [Text S1 and S3]. Thus, the $Weight_i$

changes when π/ME decreases from 2.0 to 0.5 [Figure S5]. This phenomenon also explains the marginal dependence of the simulation results on P and RR_0. Specifically, we have:

$$Weight_i = \frac{1}{\mathrm{var}(RR_i) + \tau^2} = \frac{p_1 \times RR_i \times N}{2 - 4 \times p_1 \times RR_i + RR_i + p_1 \times RR_i \times N \times \tau^2} \quad (4\text{-}1.1)$$

$$RR_i = RR_0 \times \left(1 + \frac{\pi \times q_{total,i}}{(1 - q_{total,i}) + \pi \times q_{total,i}} \times (ME - 1) \right) \quad (4\text{-}1.2)$$

The bias introduced by the weights is not easily simplified because it involves the distribution of q_{total} among the included studies. However, if the average summary value is a linear function of *OR*/*RR*, the studies' weights will not bias the estimates. Thus, the linear relation is important and again highlights the superior mathematical behavior of *OR*, whose linear relation with the average summary value has been proven.[11]

The selection of summary statistics for the meta-analysis of RCTs has been addressed in previous studies.[15,16] The authors concluded that *OR* and *RR* were both acceptable selections due to their heterogeneity under different risks.[18,19] *RR* is typically preferred because it is simpler to interpret than *OR*.[20-22] This preference might underlie the common use of *RR* to detect the moderator effect in meta-regression. Although *RR* might be bias-free when the candidate moderator is a study-level variable, it may not be suitable for participant-level variables. We found that the mathematical properties of *OR* are superior to those of *RR*; *RR* might also introduce serious bias. Although *OR* is more difficult to interpret,[20-22] it can be converted into *RR* via Equation 4-2,[4] where *ACR* is the assumed control risk. Thus, we consider *OR* as the better selection in the meta-regression analysis of RCTs:

$$RR = \frac{OR}{1 - ACR \times (1 - OR)} \quad (4\text{-}2)$$

Although *OR* reduces bias in the meta-regression analysis, it retains the nature of the meta-regression. The limitations of meta-regression analysis[14,16] are important to understand. First, the results of RCT meta-regression must be considered to be epidemiological data, which cannot be randomized to underpin causality.[16] Thus, the associations found in a meta-regression should be considered to be hypothesis-generating rather than proof of causality.[14] Second, individual patient data provide better results than summary data;[11,23] thus, the average data must be considered to be a suboptimal choice. Third, meta-regression must frequently include more than ten studies, and the moderator of interest should be preproposed and backed by an adequate theoretical basis.[4,14] Multiple comparisons are also problematic in meta-regression analysis.[24] Fourth, the average summary values in each included study are calculated from a small sample size and thus may include significant random errors. The resulting attenuation bias[15,16] will steer the result toward a null association. In this study, we confirmed only that *OR* exhibits stronger mathematical properties than *RR* in meta-regression analyses; the inherent limitations of meta-regression remain.

CONCLUSIONS

This study shows the superior linear properties of *OR*-based meta-regression compared to *RR*-based meta-regression. If most of the included studies report the average summary values of case groups, the accuracy and power of the *OR*-based meta-regression increase. However, we suggested researchers can present *OR*- and *RR*-based results simultaneously because the explanation of *RR* is better than *OR*. The *OR* and *RR*-based results were similar in most of the situation; the authors will not do different conclusions based on these two methods. Moreover, this mathematical improvement does not reduce the inherent limitations of meta-regression. Researchers need to understand that meta-regression results are useful for hypothesis-generating but not for causality inference.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. Pocock SJ, Hughes MD, Lee RJ. Statistical problems in the reporting of clinical trials. A survey of three medical journals. N Engl J Med 1987;317:426-32.
2. Brookes ST, Whitley E, Peters TJ, Mulheran PA, Egger M, Davey Smith G, *et al.* Subgroup analyses in randomised controlled trials: Quantifying the risks of false-positives and false-negatives. Health Technol Assess 2001;5:1-56.
3. Brookes ST, Whitely E, Egger M, Smith GD, Mulheran PA, Peters TJ, *et al.* Subgroup analyses in randomized trials: Risks of subgroup-specific analyses; power and sample size for the interaction test. J Clin Epidemiol 2004;57:229-36.
4. Higgins JP, Green S. Cochrane Handbook for Systematic Reviews of Interventions. Ver. 5.1.0. The Cochrane

Collaboration; 2011. http://handbook.cochrane.org.
5. Qin X, Huo Y, Xie D, Hou F, Xu X, Wang X, *et al.* Homocysteine-lowering therapy with folic acid is effective in cardiovascular disease prevention in patients with kidney disease: A meta-analysis of randomized controlled trials. Clin Nutr 2013;32:722-7.
6. Altman DG, Lausen B, Sauerbrei W, Schumacher M. Dangers of using "optimal" cutpoints in the evaluation of prognostic factors. J Natl Cancer Inst 1994;86:829-35.
7. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: A bad idea. Stat Med 2006;25:127-41.
8. Simmonds MC, Higgins JP, Stewart LA, Tierney JF, Clarke MJ, Thompson SG, *et al.* Meta-analysis of individual patient data from randomized trials: A review of methods used in practice. Clin Trials 2005;2:209-17.
9. Lyman GH, Kuderer NM. The strengths and limitations of meta-analyses based on aggregate data. BMC Med Res Methodol 2005;5:14.
10. Riley RD, Simmonds MC, Look MP. Evidence synthesis combining individual patient data and aggregate data: A systematic review identified current practice and possible methods. J Clin Epidemiol 2007;60:431-9.
11. Lin C, Chu CM, Lin J, Yang HY, Su SL. Gene-gene and gene-environment interactions in meta-analysis of genetic association studies. PLoS One 2015;10:e0124967.
12. Lin C, Yang HY, Wu CC, Lee HS, Lin YF, Lu KC, *et al.* Angiotensin-converting enzyme insertion/deletion polymorphism contributes high risk for chronic kidney disease in asian male with hypertension – A meta-regression analysis of 98 observational studies. PLoS One 2014;9:e87604.
13. Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw 2010;36:1-48.
14. Baker WL, White CM, Cappelleri JC, Kluger J, Coleman CI; Health Outcomes, Policy, and Economics (HOPE) Collaborative Group, *et al.* Understanding heterogeneity in meta-analysis: The role of meta-regression. Int J Clin Pract 2009;63:1426-34.
15. Daniels MJ, Hughes MD. Meta-analysis for the evaluation of potential surrogate markers. Stat Med 1997;16:1965-82.
16. Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? Stat Med 2002;21:1559-73.
17. Demidenko E. Sample size and optimal design for logistic regression with binary interaction. Stat Med 2008;27:36-46.
18. Furukawa TA, Guyatt GH, Griffith LE. Can we individualize the 'number needed to treat'? An empirical study of summary effect measures in meta-analyses. Int J Epidemiol 2002;31:72-6.
19. Deeks JJ. Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. Stat Med 2002;21:1575-600.
20. Sinclair JC, Bracken MB. Clinically useful measures of effect in binary analyses of randomized trials. J Clin Epidemiol 1994;47:881-9.
21. Straus SE, Sackett DL. Using research findings in clinical practice. BMJ 1998;317:339-42.
22. Akobeng AK. Understanding randomised controlled trials. Arch Dis Child 2005;90:840-4.
23. Lambert PC, Sutton AJ, Abrams KR, Jones DR. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. J Clin Epidemiol 2002;55:86-94.
24. Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Statistics in medicine – Reporting of subgroup analyses in clinical trials. N Engl J Med 2007;357:2189-94.

SUPPORTING INFORMATION

Text S1: Derivation of the bias in the linear relation between the logarithmic risk ratio and the average summary value.

When the independent variable (x) is the intervention encoded with value 0 for control group and 1 for treatment group, and the moderator (m) is diabetes status encoded with values 0 and 1 for without and with diabetes. The dependent variable is a binary outcome event (y) (with 0 and 1 signifying non-occurrence and occurrence, respectively). Probabilities p_1, p_2, p_3 and p_4 are then defined as follows:

$$\begin{cases} p_1 = p(y=1 \mid x=0 \cap m=0) \\ p_2 = p(y=1 \mid x=1 \cap m=0) \\ p_3 = p(y=1 \mid x=0 \cap m=1) \\ p_4 = p(y=1 \mid x=1 \cap m=1) \end{cases}$$

which represent the outcome incidence of control group without diabetes, treatment group without diabetes, control group with diabetes and treatment group with diabetes. According to above setting, the risk ratios (RR) of interest for patient without diabetes [RR_0 ($m = 0$)] and patient with diabetes [RR_1 ($m = 1$)] are calculated as follows:

$$\begin{cases} RR_0 = p_2 / p_1 \\ RR_1 = p_4 / p_3 \end{cases}$$

In an RCT, the pooled or combined RR (RR_{combine}) is a function of the diabetes prevalence in the entire studied population, $q_{total} = p(m=1)$ and can be expressed as

$$RR_{combine} = \frac{p_2(1-q_{total}) + p_4 q_{total}}{p_1(1-q_{total}) + p_3 q_{total}}$$

Let the π denote the RR of diabetes in the untreated individuals ($\pi = p_3/p_1$). We can express p_1, p_2, p_3 and p_4 in terms of p_1 and π alone including RRs as follows:

$$\begin{cases} p_1 = p_1 \\ p_2 = p_1 \times RR_0 \\ p_3 = p_1 \times \pi \\ p_4 = p_1 \times \pi \times RR_1 \end{cases}$$

This simplifies the calculation of RR_{combine}. Let the ME denote the moderator effect ($ME = RR_1/RR_0$), RR_{combine} can be calculated as follows:

$$\begin{aligned} RR_{combine} &= \frac{p_2(1-q_{total}) + p_4 q_{total}}{p_1(1-q_{total}) + p_3 q_{total}} \\ &= \frac{RR_0 \times (1-q_{total}) + RR_1 \times \pi \times q_{total}}{(1-q_{total}) + \pi \times q_{total}} \\ &= RR_0 \times \frac{(1-q_{total})}{(1-q_{total}) + \pi \times q_{total}} + RR_1 \times \frac{\pi \times q_{total}}{(1-q_{total}) + \pi \times q_{total}} \\ &= RR_0 + \frac{\pi \times q_{total}}{(1-q_{total}) + \pi \times q_{total}} \times (RR_1 - RR_0) \\ &= RR_0 + \frac{\pi \times q_{total}}{(1-q_{total}) + \pi \times q_{total}} \times (RR_0 \times ME - RR_0) \\ &= RR_0 + RR_0 \times \frac{\pi \times q_{total}}{(1-q_{total}) + \pi \times q_{total}} \times (ME - 1) \\ &= RR_0 \times \left(1 + \frac{\pi \times q_{total}}{(1-q_{total}) + \pi \times q_{total}} \times (ME - 1)\right) \end{aligned}$$

and the logarithm of $RR_{combine}$ can be expressed as follows:

$$\log\left(RR_{combine}\right)=\log\left(RR_0\right)+\log\left(1+\frac{\pi\times q_{total}}{\left(1-q_{total}\right)+\pi\times q_{total}}\times\left(ME-1\right)\right) \quad (2.1\text{-}1)$$

$$=\log\left(RR_0\right)+\frac{\log\left(1+\frac{\pi\times q_{total}}{\left(1-q_{total}\right)+\pi\times q_{total}}\times\left(ME-1\right)\right)}{\log\left(ME\right)}\times\log\left(ME\right)$$

$$=\log\left(RR_0\right)+\frac{\log\left(\frac{\left(1-q_{total}\right)+\pi\times q_{total}+\pi\times q_{total}\times\left(ME-1\right)}{\left(1-q_{total}\right)+\pi\times q_{total}}\right)}{\log\left(ME\right)}\times\log\left(ME\right)$$

$$=\log\left(RR_0\right)+\frac{\log\left(\frac{\left(1-q_{total}\right)+\pi\times q_{total}\times ME}{\left(1-q\right)q_{total}+\pi\times q_{total}}\right)}{\log\left(ME\right)}\times\log\left(ME\right)$$

$$=\log\left(RR_0\right)+\frac{\log\left[\left(1-q_{total}\right)+\pi\times q_{total}\times ME\right]-\log\left[\left(1-q_{total}\right)+\pi\times q_{total}\right]}{\log\left(ME\right)}\times\log\left(ME\right)$$

Let log (RR_i) denotes the observed effect size of ith individual RCT, $q_{total,\,i}$ the proportion of the population with moderator status, η_i the random effects, and ε_i the residuals. In general awareness studies, β_0 is the expected logarithmic RR of individuals without diabetes [log (RR_0)], and β_1 is the expected logarithmic moderator effect [log (ME)]. Thus, the traditional meta-regression (here called Method 1) can be written as.

$$\log\left(RR_i\right)=\beta_0+\beta_1\times q_{total,i}+\eta_i+\varepsilon_i \quad (2.1\text{-}2)$$

However, following Equation 2.1-1, the relation between log (RR_i) and $q_{total,\,i}$ is actually nonlinear:

$$\log\left(RR_i\right)=\beta_0+\beta_1\times\kappa_i+\eta_i+\varepsilon_i \quad (2.1\text{-}3)$$

where , $\kappa_i=\frac{\log\left[\left(1-q_{total,i}\right)+\pi\times q_{total,i}\times ME\right]-\log\left[\left(1-q_{total,i}\right)+\pi\times q_{total,i}\right]}{\log\left(ME\right)}$

and π is the RR of diabetes in the untreated individuals. Note that π is an unknown population parameter not provided by most papers. That is why the traditional studies often used $q_{total,\,i}$ to replace κ_i. Obviously, this may cause bias since $\kappa_i = q_{total,\,i}$ only when $\pi = 1$ and $ME = 1$.

Text S2: Relation between the average summary values of the case group (q_{case}) and the entire population (q_{total}).

The independent variable (x) and the moderator (m) are assumed to be binary variables, where 0 and 1 signify nonexposure and exposure, respectively. The dependent variable is a binary outcome event (y), where 0 and 1 signify nonoccurrence and occurrence, respectively. p_1, p_2, $p_{3,}$ and p_4 are then defined as follows:

$$\begin{cases} p_1=p\left(y=1\,|\,x=0\cap m=0\right)\\ p_2=p\left(y=1\,|\,x=1\cap m=0\right)\\ p_3=p\left(y=1\,|\,x=0\cap m=1\right)\\ p_4=p\left(y=1\,|\,x=1\cap m=1\right)\end{cases}$$

Defining RR_0 as the risk ratio of the independent variable when $m = 0$; π as the risk ratio of the moderator when $x = 0$; and ME as the moderator effect, the parameters p_1, p_2, $p_{3,}$ and p_4 can be rewritten in terms of p_1 and π alone:

$$\begin{cases} p_1=p_1\\ p_2=p_1\times RR_0\\ p_3=p_1\times\pi\\ p_4=p_1\times RR_0\times\pi\times ME\end{cases}$$

Setting the expected proportion of $x = 0$ in an RCT equal to 0.5 and assuming that the independent variable does not depend on the moderator because the process is randomized, we can relate the average summary value of the case group (q_{case}) to the average summary value of the entire population (q_{total}) as follows:

$$
\begin{aligned}
q_{case} &= p(m=1 \mid y=1) \\
&= \frac{p(m=1) \times p(y=1 \mid m=1)}{p(y=1)} \\
&= \frac{q_{total} \times \frac{p_3 + p_4}{2}}{(1-q_{total}) \times \frac{p_1 + p_2}{2} + q_{total} \times \frac{p_3 + p_4}{2}} \\
&= \frac{q_{total}(p_3 + p_4)}{(1-q_{total})(p_1 + p_2) + q_{total}(p_3 + p_4)} \\
&= \frac{q_{total}(p_1 \times \pi + p_1 \times RR_0 \times \pi \times ME)}{(1-q_{total})(p_1 + p_1 \times RR_0) + q_{total}(p_1 \times \pi + p_1 \times RR_0 \times \pi \times ME)} \\
&= \frac{q_{total}(\pi + RR_0 \times \pi \times ME)}{(1-q_{total})(1 + RR_0) + q_{total}(\pi + RR_0 \times \pi \times ME)}
\end{aligned}
$$

Text S3: Weights of the included RCTs.

The independent variable (x) is assumed to be a binary variable, where 0 and 1 signify nonexposure and exposure, respectively. The dependent variable is a binary outcome event (y), where 0 and 1 signify non-occurrence and occurrence, respectively. The parameters p_1 and p_2 are then defined as follows:

$$
\begin{cases}
p_1 = p(y=1 \mid x=0) \\
p_2 = p(y=1 \mid x=1)
\end{cases}
$$

Because the probability of the treatment group is often set equal to 0.5 in randomized controlled trails, the expected sample size in each group is $N/2$, where N is the total sample size. Thus, the variance of a specific risk ratio (RR) can be expressed as follows:

$$
\begin{aligned}
\operatorname{var}(RR) &= \frac{1-p_1}{p_1 \times N/2} + \frac{1-p_2}{p_2 \times N/2} \\
&= \frac{p_2 \times (1-p_1) + p_1 \times (1-p_2)}{p_1 \times p_2 \times N/2} \\
&= \frac{p_1 \times RR \times (1-p_1) + p_1 \times (1-p_1 \times RR)}{p_1 \times p_1 \times RR \times N/2} \\
&= \frac{RR \times (1-p_1) + (1-p_1 \times RR)}{p_1 \times RR \times N/2} \\
&= \frac{2 - 4 \times p_1 \times RR + RR}{p_1 \times RR \times N}
\end{aligned}
$$

Finally, the weights of each study are computed by the inverse variance method (i.e. the most common weighting method):

$$
Weight = \frac{1}{\operatorname{var}(RR) + \tau^2} = \frac{p_1 \times RR \times N}{2 - 4 \times p_1 \times RR + RR + p_1 \times RR \times N \times \tau^2}
$$

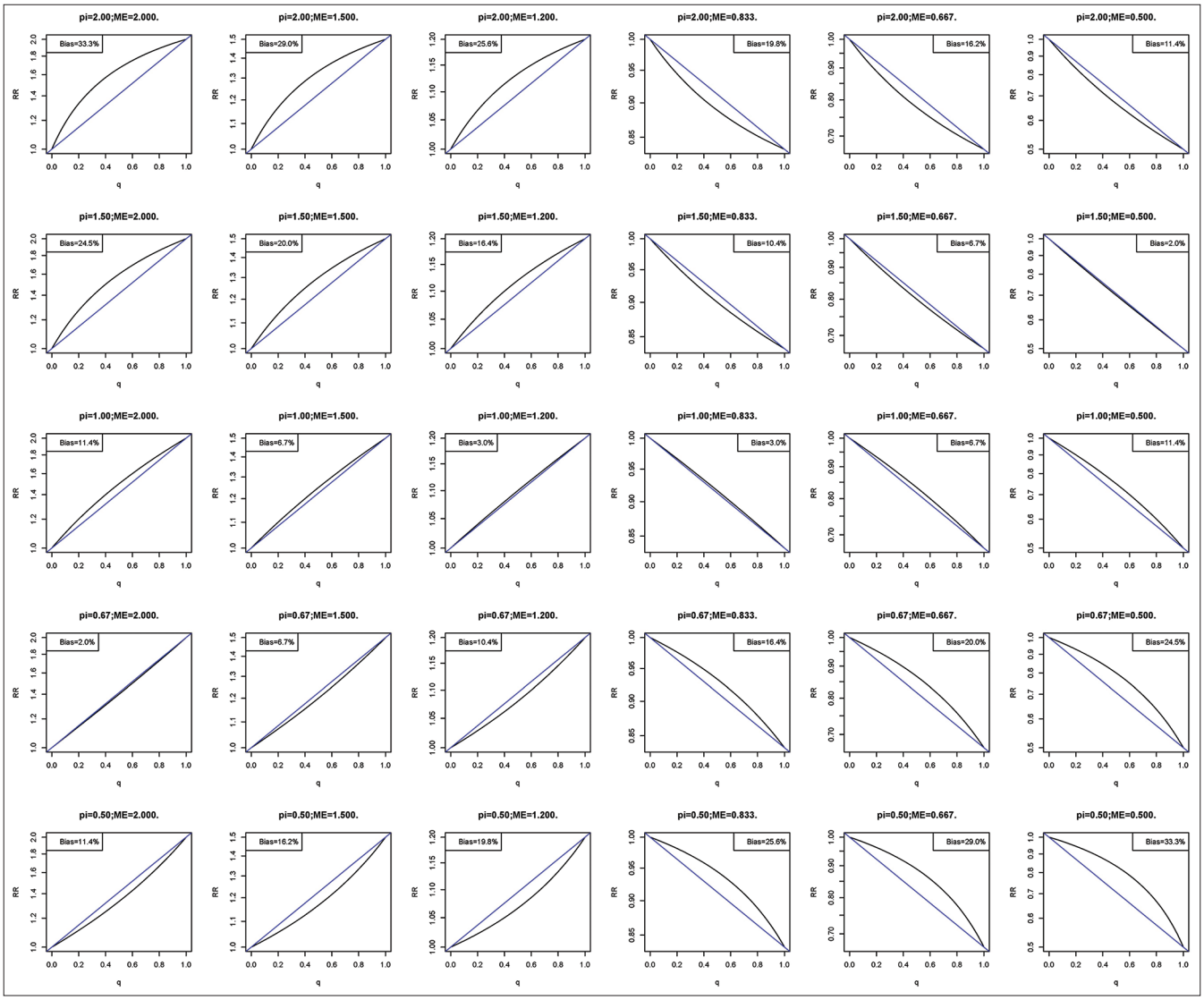

Figure S1: Relation between the average summary value and the logarithmic RR. Based on Equation 2.1-3, the true and linear relations differ only with respect to the RR of the moderator (π) and the *ME*. This figure plots the true and linear relations for different π (2.0, 1.5, 1.0, 0.67, and 0.5) and *ME* (2.0, 1.5, 1.2, 0.83, 0.67, and 0.5) values. The X axis represents the average summary value throughout the study (*q*). Black and blue lines plot the expected *RR* and the linear relation, respectively. The bias is defined as the triangular area in the window straddling the two lines. The bias is larger when the directions of π and *ME* are the same and their effects leave l (null effect). The intercept RR_0 is set equal to 1, and the Y axis is transformed into the logarithmic scale. *ME*: Moderator effect, *RR*: Risk ratio

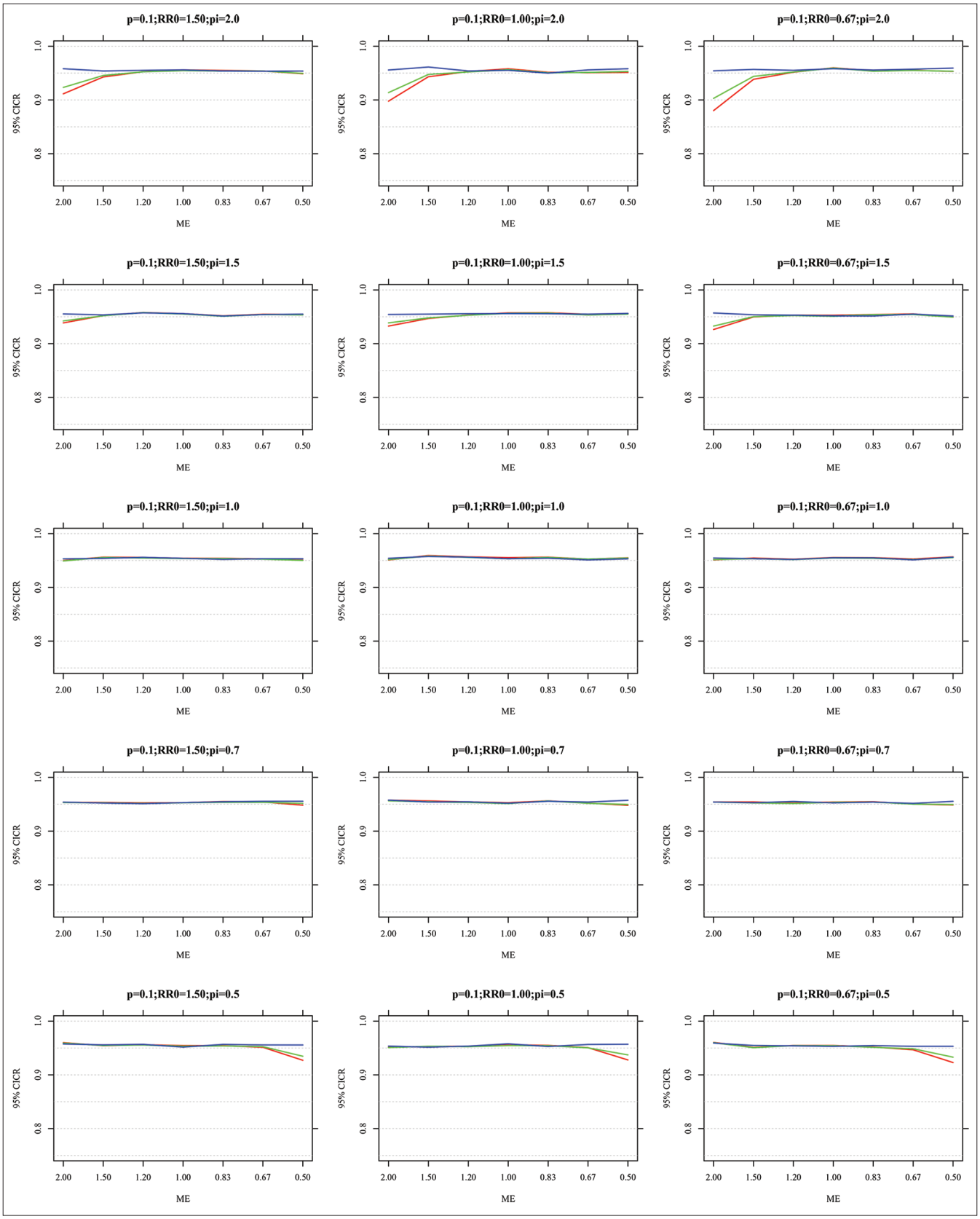

Figure S2: 95% Confidence interval coverage rate of the slope from different meta-regression methods. Red line: Method 1; Green line: Method 2; Blue line: Method 3

Contd...

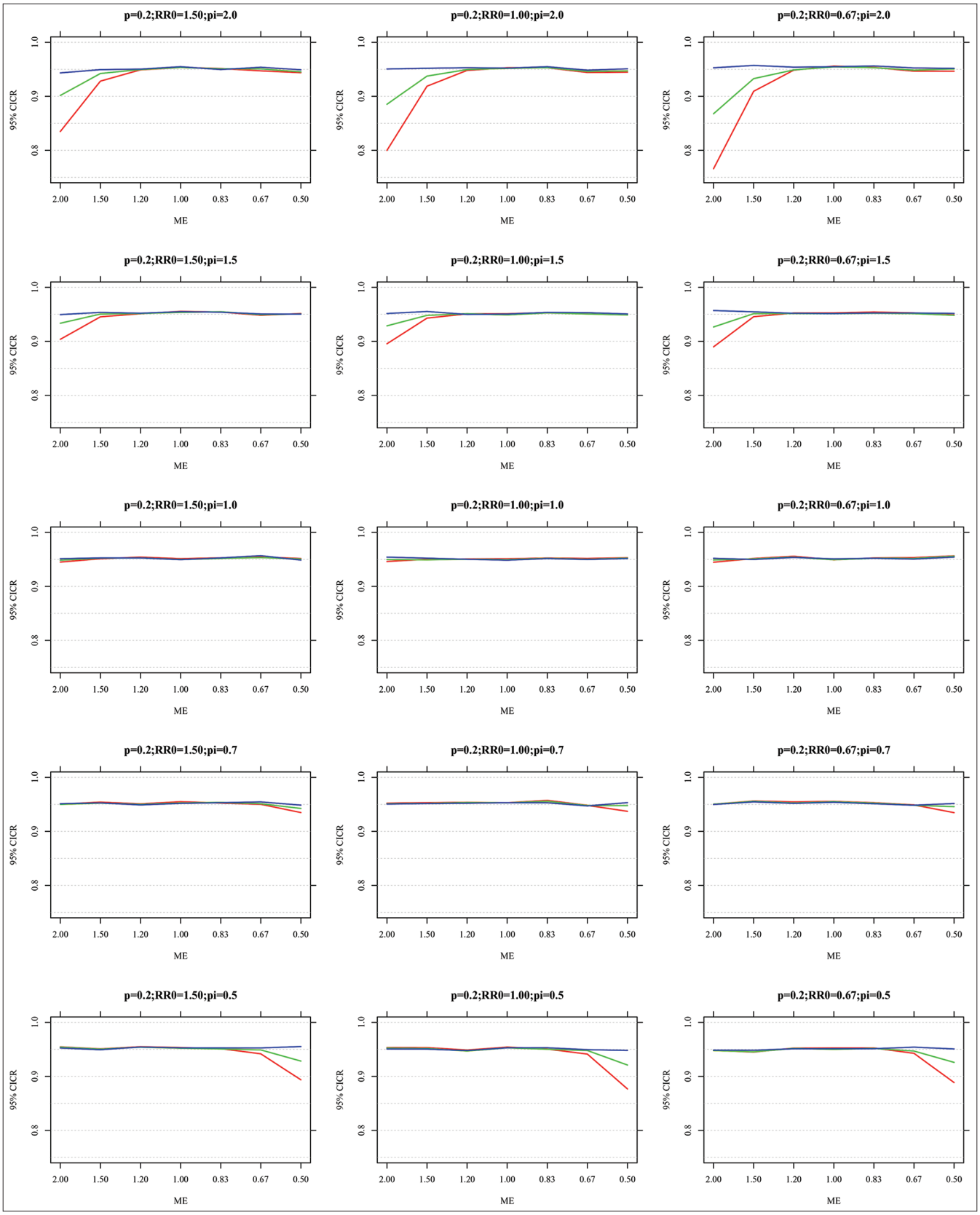

Figure S2: Contd...

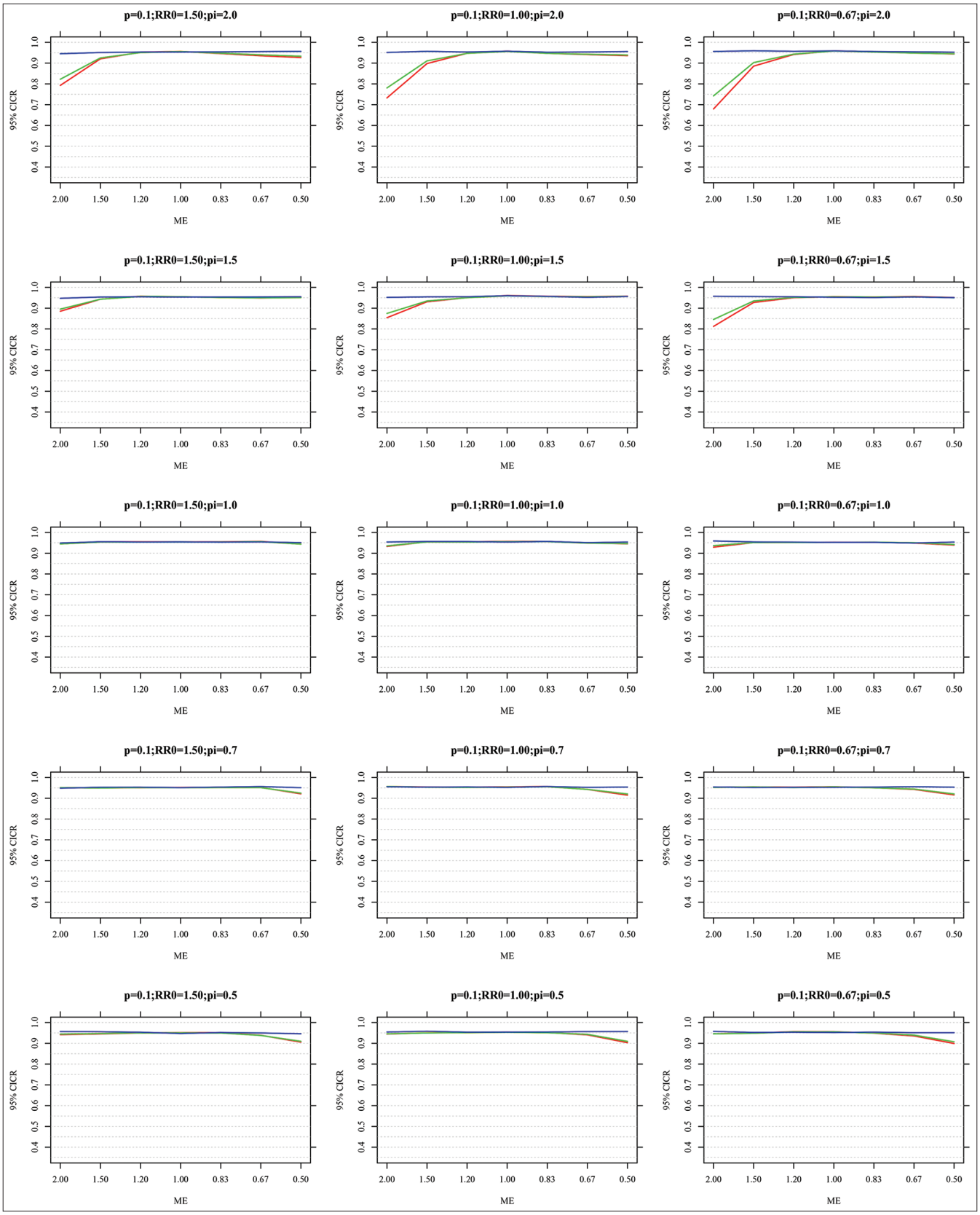

Figure S3: 95% Confidence interval coverage rate of the intercept from different meta-regression methods. Red line: Method 1; Green line: Method 2; Blue line: Method 3

Contd...

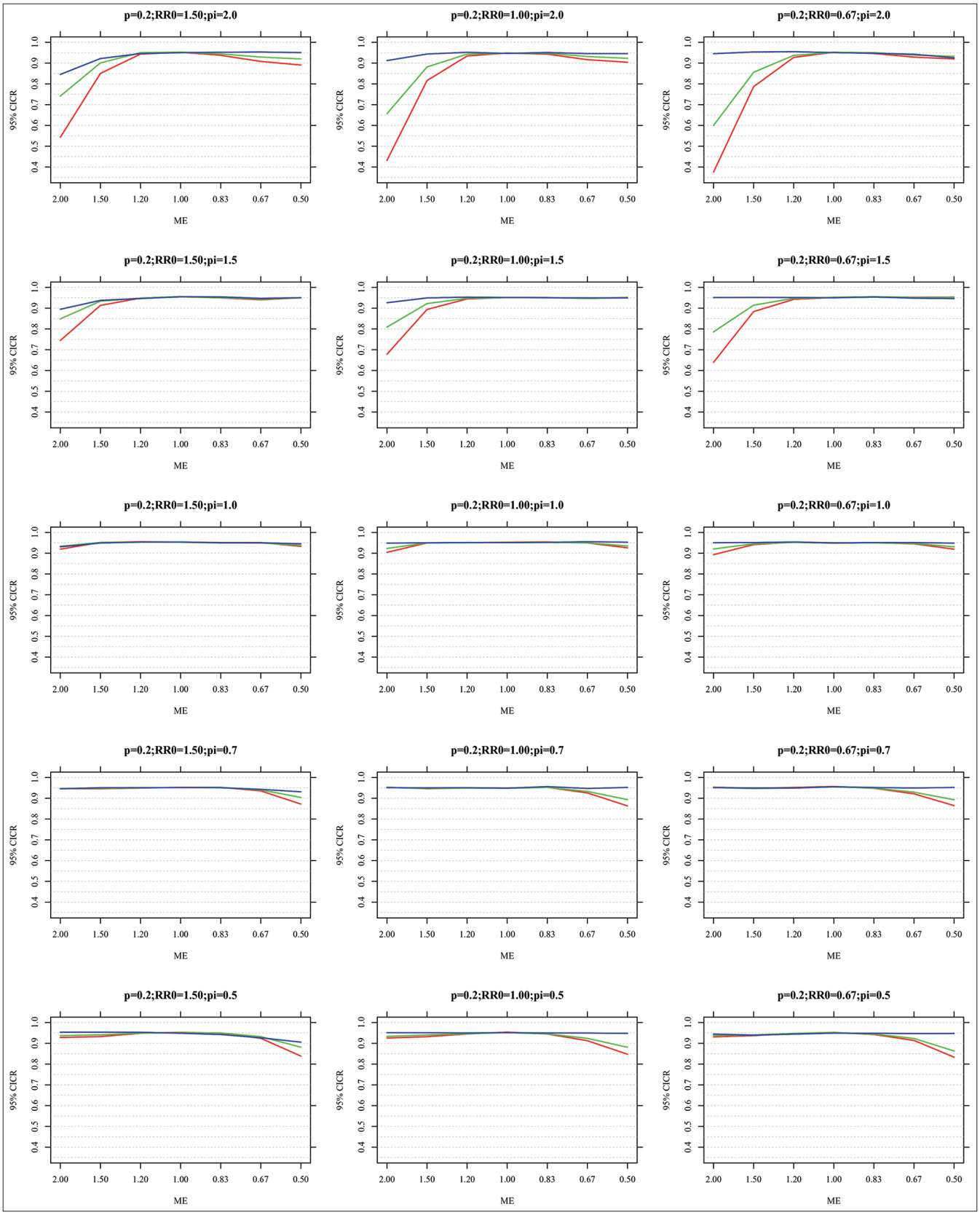

Figure S3: Contd...

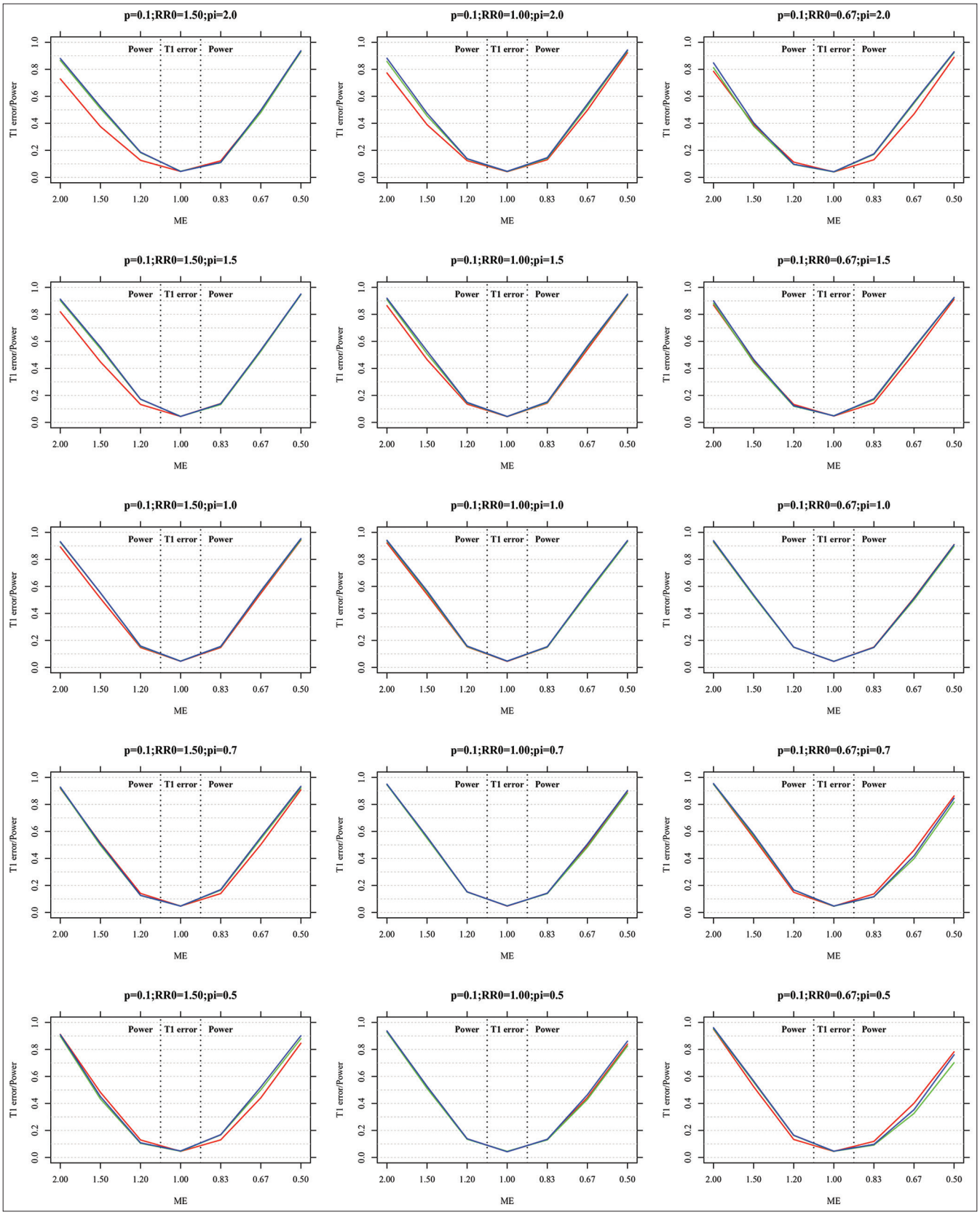

Figure S4: False-positive rate and power of the moderator effect (%) at the 0.05 significance level from different meta-regression methods. Red line: Method 1; Green line: Method 2; Blue line: Method 3

Contd...

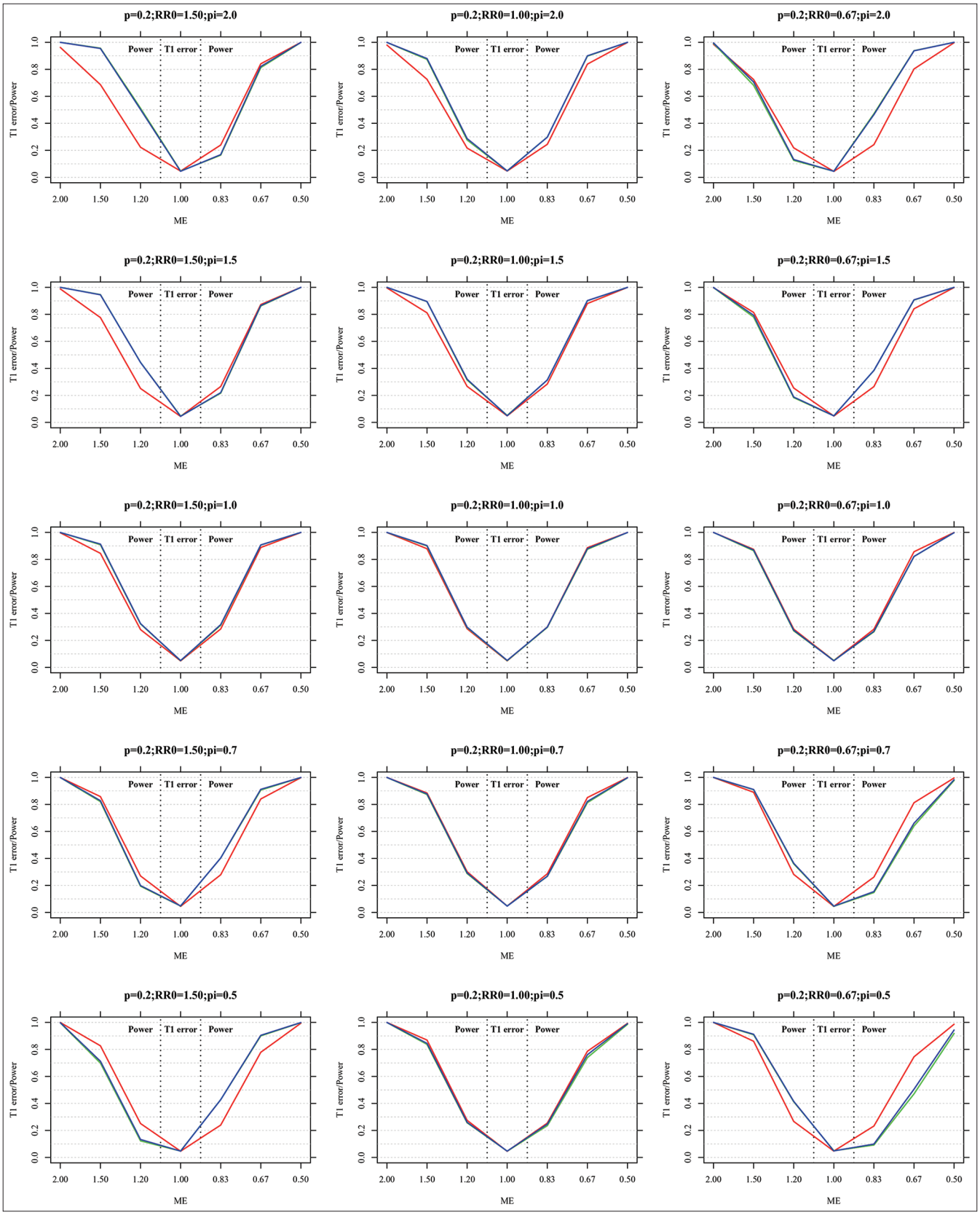

Figure S4: Contd...

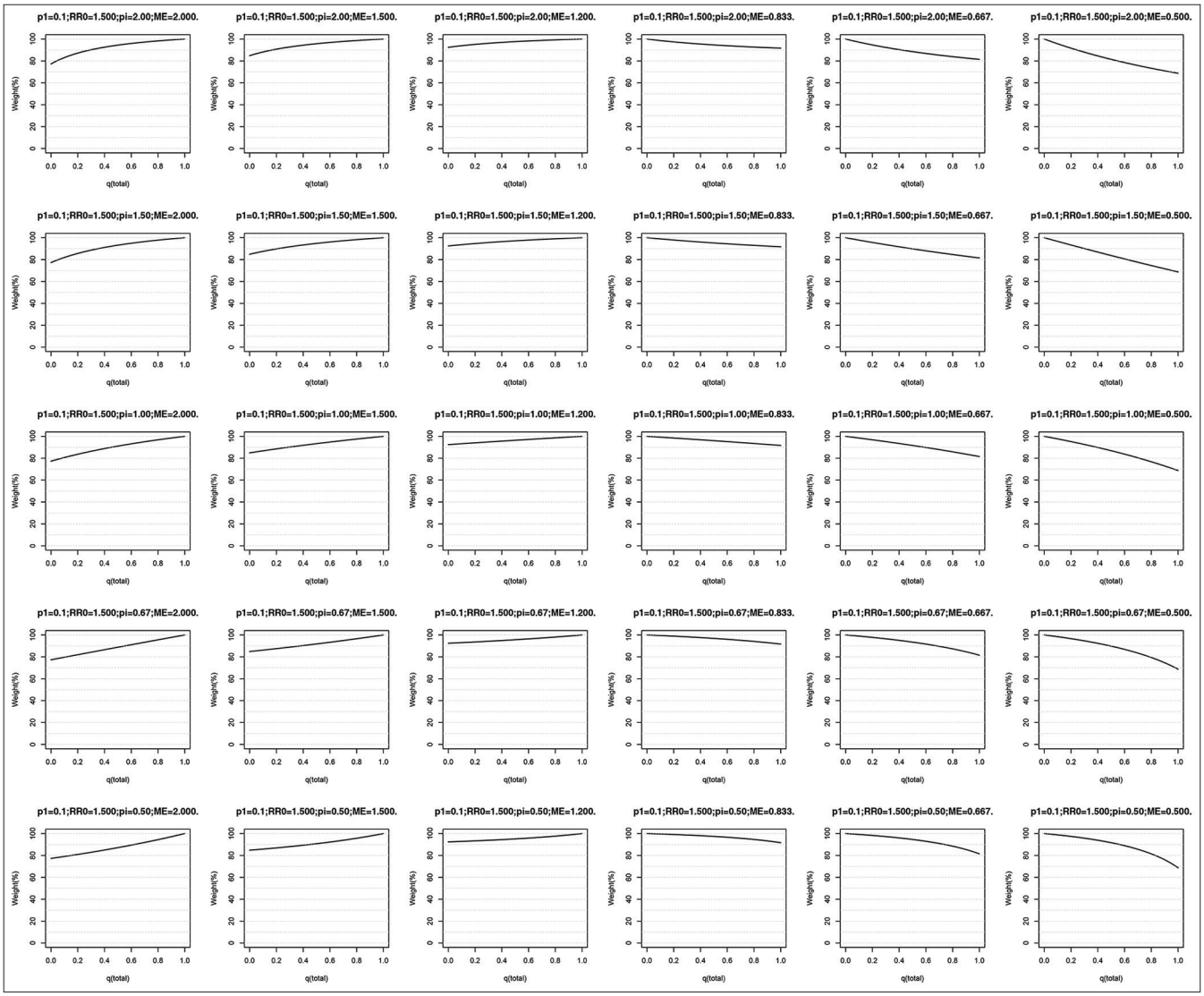

Figure S5: Relation between the average summary value and study weight. By varying the disease incidence in individuals without treatment or the moderator (p_1); the *RR* of treatment in individuals without the moderator (RR_0); the *RR* of the moderator in untreated individuals (π); and the moderator effect (*ME*), we can determine how the average summary value throughout the study (q_{total}) depends on the study weight. The weights, which are normalized by their maximum value, were calculated using Equations 4-1.1 and 4-1.2. The sample size (*N*) was excluded from these calculations because it is frequently unrelated to the relevant factors

Contd...

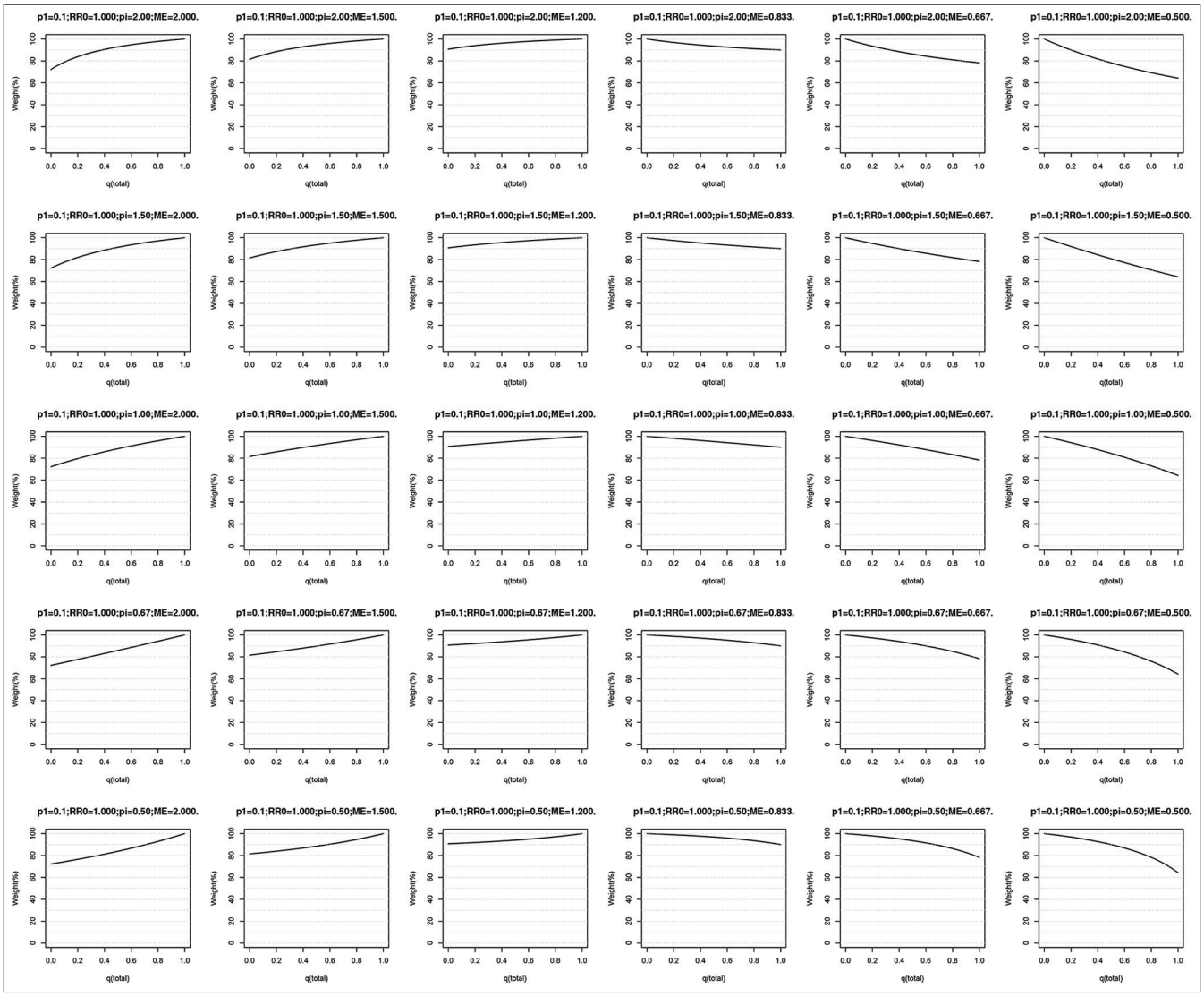

Figure S5: Contd...

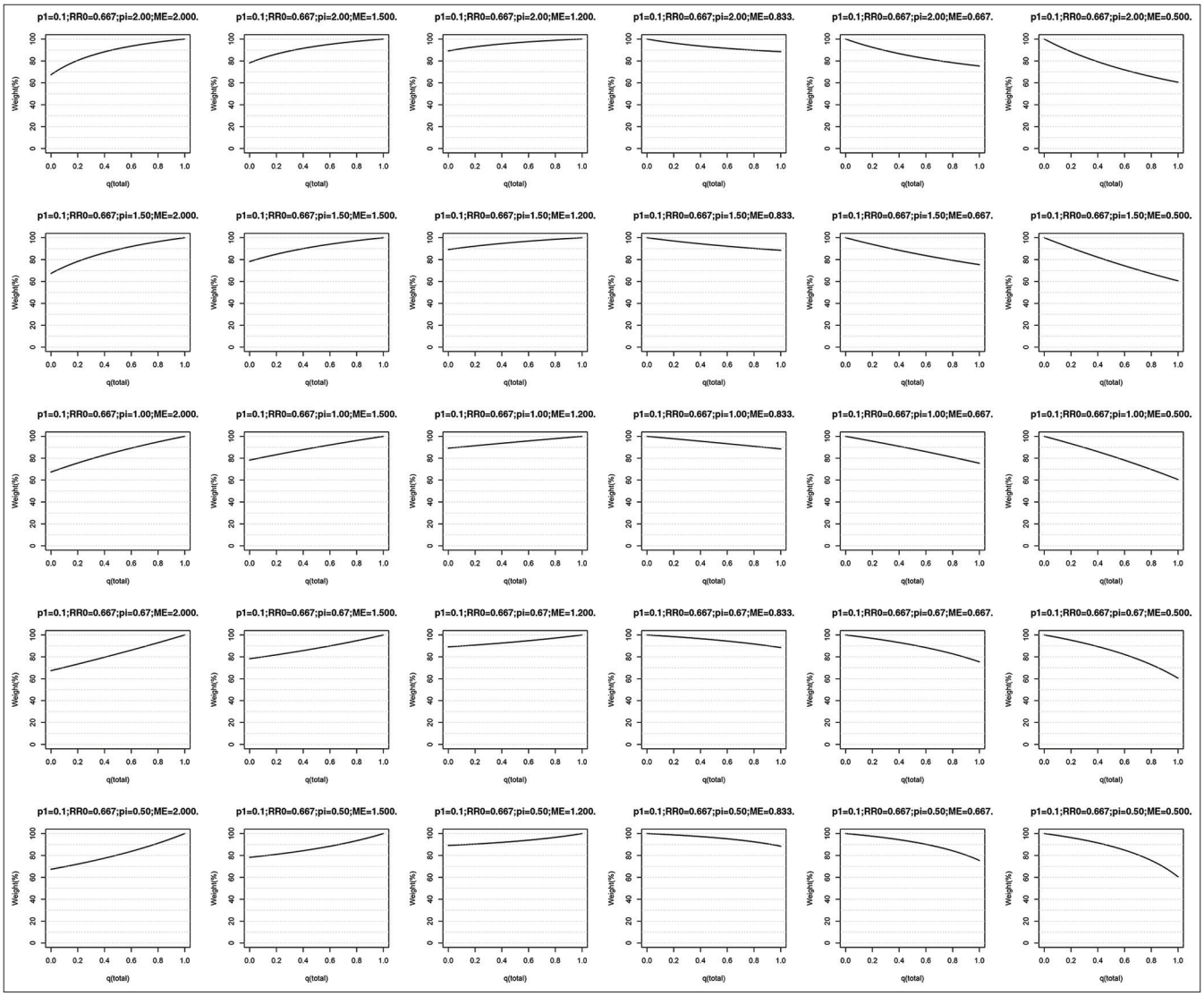

Figure S5: Contd...

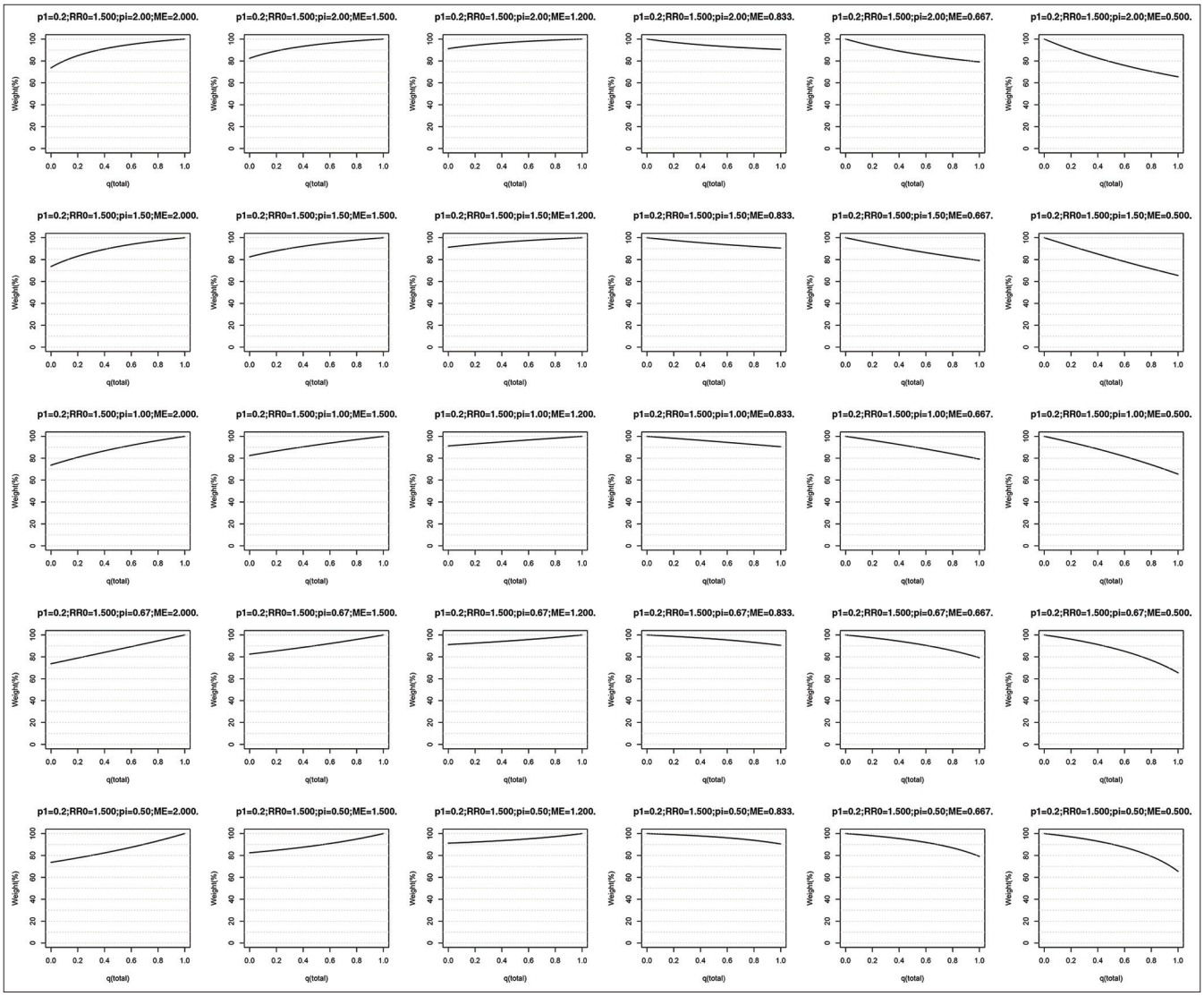

Figure S5: Contd...

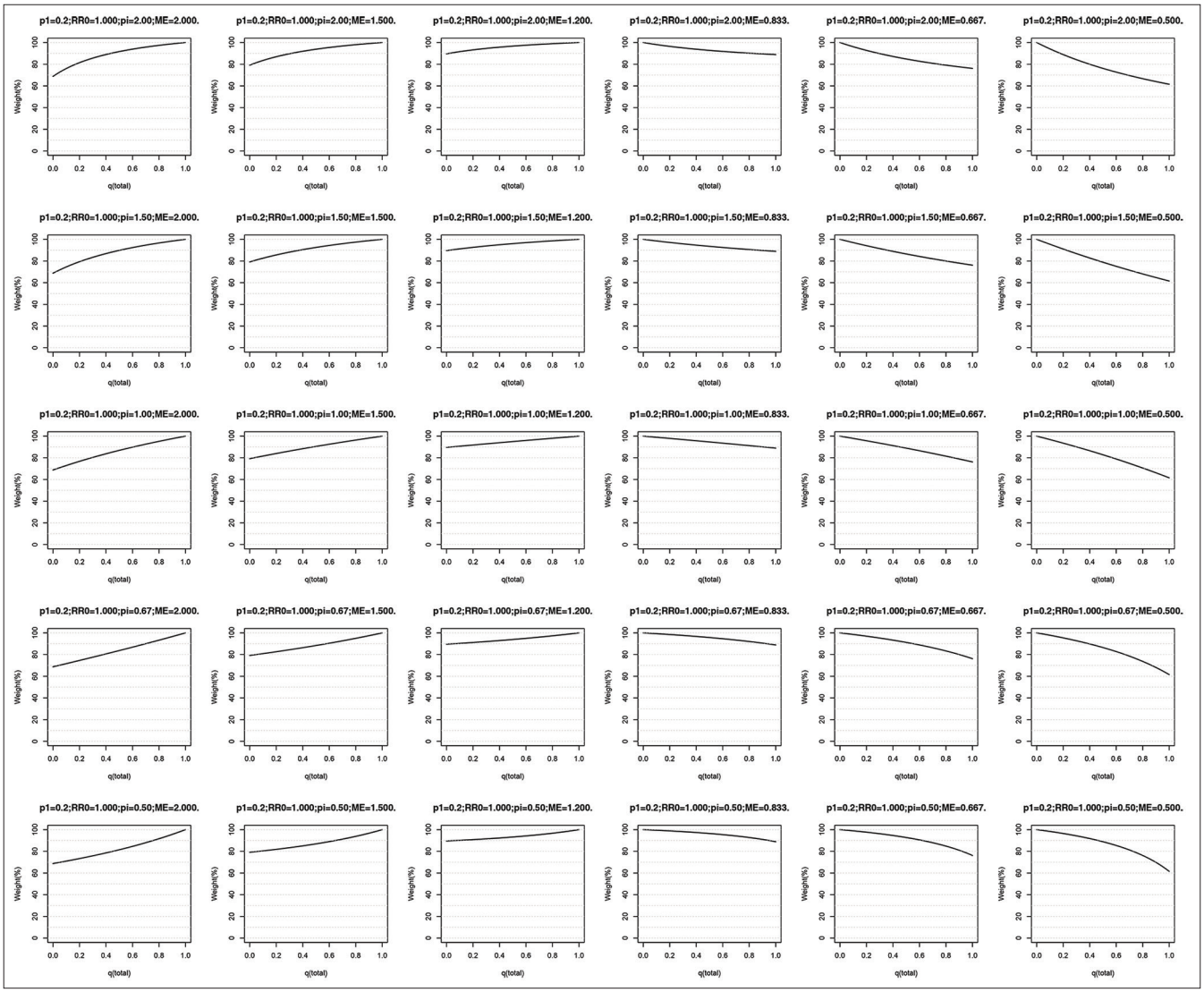

Figure S5: Contd...

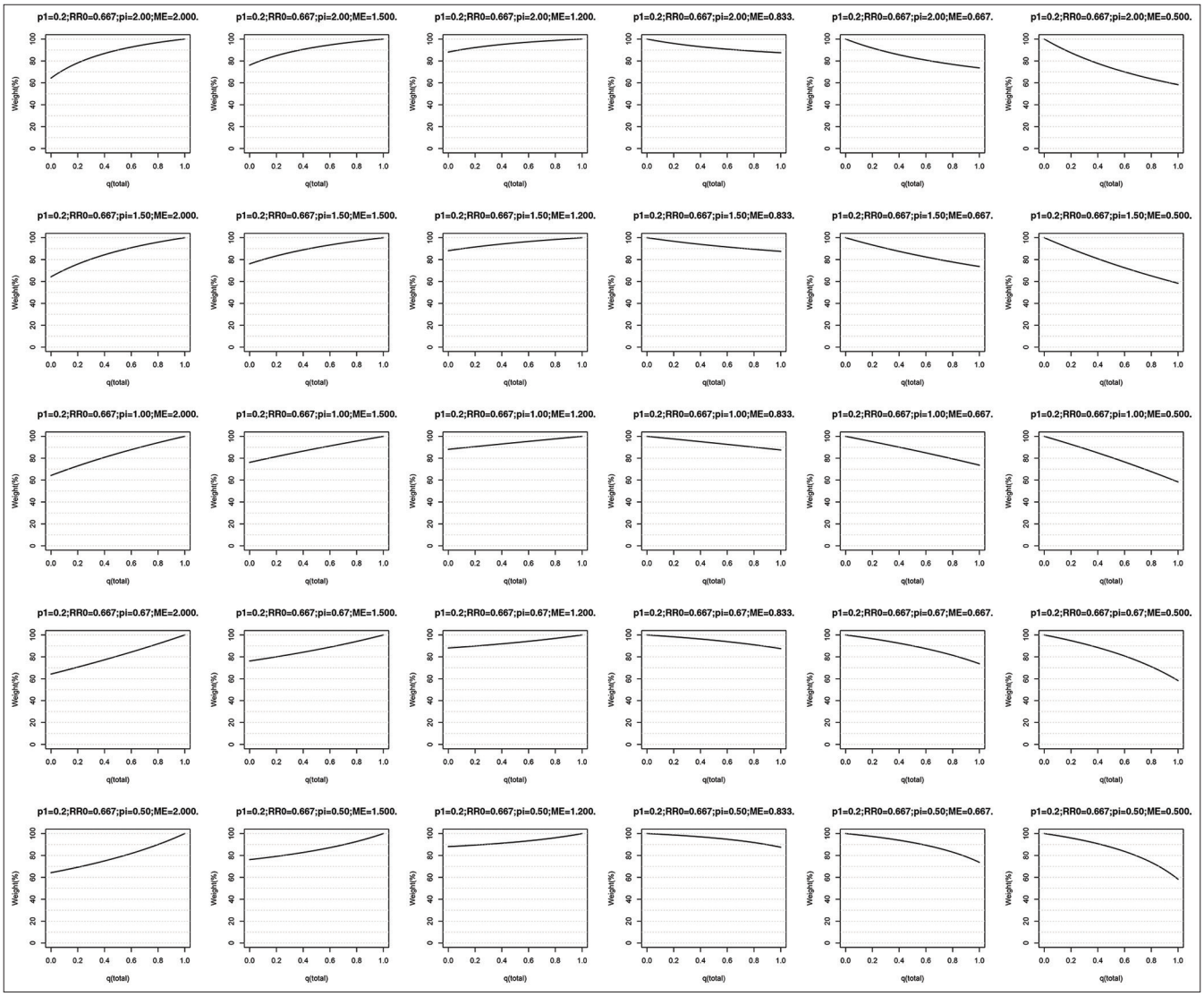

Figure S5: Contd...

Appendix S1: Simulation code

Table S1: The Pearson correlation between the average summary value of case group (q_{case}) and the average summary value of whole population (q_{total})

RR_0=2.000		*ME*						
		2.000	1.500	1.200	1.000	0.833	0.667	0.500
π	2.000	0.9532	0.9686	0.9780	0.9841	0.9890	0.9935	0.9972
	1.500	0.9725	0.9841	0.9907	0.9945	0.9972	0.9992	1.0000
	1.200	0.9841	0.9927	0.9968	0.9989	0.9999	0.9998	0.9983
	1.000	0.9913	0.9972	0.9995	1.0000	0.9995	0.9979	0.9945
	0.833	0.9964	0.9996	0.9999	0.9989	0.9970	0.9937	0.9886
	0.667	0.9996	0.9995	0.9974	0.9945	0.9909	0.9857	0.9784
	0.500	0.9989	0.9945	0.9893	0.9841	0.9784	0.9708	0.9608
RR_0=1.500		*ME*						
		2.000	1.500	1.200	1.000	0.833	0.667	0.500
π	2.000	0.9562	0.9701	0.9786	0.9841	0.9886	0.9927	0.9962
	1.500	0.9749	0.9853	0.9911	0.9945	0.9970	0.9989	0.9999
	1.200	0.9859	0.9934	0.9971	0.9989	0.9998	0.9999	0.9990
	1.000	0.9927	0.9977	0.9996	1.0000	0.9996	0.9983	0.9958
	0.833	0.9972	0.9998	0.9998	0.9989	0.9972	0.9945	0.9904
	0.667	0.9999	0.9993	0.9972	0.9945	0.9913	0.9869	0.9809
	0.500	0.9983	0.9938	0.9889	0.9841	0.9790	0.9725	0.9641
RR_0=1.200		*ME*						
		2.000	1.500	1.200	1.000	0.833	0.667	0.500
π	2.000	0.9587	0.9714	0.9791	0.9841	0.9882	0.9919	0.9953
	1.500	0.9768	0.9862	0.9914	0.9945	0.9968	0.9986	0.9997
	1.200	0.9874	0.9940	0.9973	0.9989	0.9997	1.0000	0.9994
	1.000	0.9937	0.9981	0.9996	1.0000	0.9997	0.9987	0.9966
	0.833	0.9979	0.9999	0.9998	0.9989	0.9974	0.9951	0.9917
	0.667	1.0000	0.9991	0.9970	0.9945	0.9917	0.9878	0.9827
	0.500	0.9978	0.9932	0.9885	0.9841	0.9795	0.9738	0.9666
RR_0=1.000		*ME*						
		2.000	1.500	1.200	1.000	0.833	0.667	0.500
π	2.000	0.9608	0.9725	0.9795	0.9841	0.9878	0.9913	0.9945
	1.500	0.9784	0.9869	0.9917	0.9945	0.9966	0.9983	0.9995
	1.200	0.9886	0.9945	0.9974	0.9989	0.9997	1.0000	0.9996
	1.000	0.9945	0.9983	0.9997	1.0000	0.9997	0.9989	0.9972
	0.833	0.9983	0.9999	0.9997	0.9989	0.9976	0.9956	0.9927
	0.667	1.0000	0.9989	0.9968	0.9945	0.9919	0.9886	0.9841
	0.500	0.9972	0.9927	0.9882	0.9841	0.9800	0.9749	0.9686
RR_0=0.833		*ME*						
		2.000	1.500	1.200	1.000	0.833	0.667	0.500
π	2.000	0.9629	0.9736	0.9800	0.9841	0.9875	0.9907	0.9937
	1.500	0.9800	0.9877	0.9919	0.9945	0.9964	0.9981	0.9993
	1.200	0.9897	0.9950	0.9976	0.9989	0.9996	1.0000	0.9998
	1.000	0.9953	0.9986	0.9997	1.0000	0.9998	0.9991	0.9978
	0.833	0.9988	1.0000	0.9997	0.9989	0.9977	0.9960	0.9936
	0.667	1.0000	0.9987	0.9966	0.9945	0.9922	0.9892	0.9855
	0.500	0.9966	0.9921	0.9878	0.9841	0.9804	0.9759	0.9704
RR_0=0.667		*ME*						
		2.000	1.500	1.200	1.000	0.833	0.667	0.500
π	2.000	0.9655	0.9749	0.9805	0.9841	0.9871	0.9900	0.9927
	1.500	0.9818	0.9886	0.9923	0.9945	0.9962	0.9977	0.9989
	1.200	0.9911	0.9956	0.9978	0.9989	0.9996	0.9999	0.9999
	1.000	0.9962	0.9989	0.9998	1.0000	0.9998	0.9993	0.9983
	0.833	0.9992	1.0000	0.9996	0.9989	0.9979	0.9965	0.9945
	0.667	0.9998	0.9983	0.9964	0.9945	0.9925	0.9900	0.9869
	0.500	0.9958	0.9913	0.9874	0.9841	0.9809	0.9770	0.9725
RR_0=0.500		*ME*						
		2.000	1.500	1.200	1.000	0.833	0.667	0.500
π	2.000	0.9686	0.9764	0.9811	0.9841	0.9866	0.9890	0.9913
	1.500	0.9841	0.9896	0.9927	0.9945	0.9960	0.9972	0.9983
	1.200	0.9927	0.9962	0.9980	0.9989	0.9995	0.9999	1.0000
	1.000	0.9972	0.9992	0.9999	1.0000	0.9999	0.9995	0.9989
	0.833	0.9996	1.0000	0.9995	0.9989	0.9981	0.9970	0.9956
	0.667	0.9995	0.9979	0.9961	0.9945	0.9929	0.9909	0.9886
	0.500	0.9945	0.9904	0.9869	0.9841	0.9814	0.9784	0.9749

According the equation 2.1-3, the relationship between q_{total} and q_{case} is impacted by RR_0, ME, and π. Where the RR_0 is the risk ratio of treatment in individuals without moderator, *ME* is the moderator effect, and the π is the risk ratio of moderator in individuals without treatment. In these tables, we present the Pearson correlation between q_{case} and q_{total} in a series of RR_0 (2.0, 1.5, 1.2, 1, 0.833, 0.667, and 0.5), *ME* (2.0, 1.5, 1.2, 1, 0.833, 0.667, and 0.5), and π (2.0, 1.5, 1.2, 1, 0.833, 0.667, and 0.5). Finally, we can find the Pearson correlations between q_{total} and q_{case} are more than 0.95 when RR_0, *ME*, and π are between 0.5 and 2.0. *ME*: Moderator effect