RANDOM-EFFECTS APPROACHES COMPARISONS IN META-ANALYSIS: DATA SIMULATION WITH BINOMIAL OUTCOMES

By

DONGWEN GUO

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2011

To my mom and dad

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincerest gratitude to my supervisor, Dr. Abraham, G. Hartzema, for his guidance, encouragement, support, and faith in me. This dissertation would not have been possible without his support in providing me with an excellent atmosphere for doing research.

It is a great honor to thank Dr. Jonathan J. Shuster. This dissertation would not exist without his innovative research findings. He has made available his support in many ways with extreme patience, enabling me to develop a clear understanding of the subject.

I am very grateful to Dr. Richard Segal, who is always willing to act as a sounding board. Without his constant direction and support, I could not have achieved my goals. I would also like to thank Dr. Teresa Kauf for her generous encouragement when I needed it most. I am thankful to Dr. Xiaomin Lu, who often made time to give her invaluable advice. Special thanks go to Dr. Chris Delaney for making himself available to participate in my final defense and offer his excellent suggestions.

I am indebted to my teachers and professors for their valuable advice and special support, staff members for their timely support, graduate colleagues and friends for their friendship. My sincere regards and blessings reach out to all of those who supported me in any respect during the completion of my dissertation.

Finally, I owe my deepest gratitude to my mother and father, as well as other family members for their unconditional love and unwavering support.

TABLE OF CONTENTS

page

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CI	Confidence interval
D	Diversity between studies
M	Number of studies combined for a meta-analysis study
NHigh	Highest number of subjects in each arm
NLow	Lowest number of subjects in each arm
OR	Odds Ratio
PC (Πc)	Global true event probability in the control arm
PT (Πt)	Global true event probability in the treatment arm
RD	Risk difference
RR	Relative risk
STD	Standard deviation
UW_MM	Unweighted mean of means meta-analysis random-effects approach in terms of odds ratio
UW_OR	Unweighted summary proportions meta-analysis random-effects approach in terms of odds ratio
UW_RR	Unweighted summary proportions meta-analysis random-effects approach in terms of relative risk
W_DL	DerSimonian-Laird model or empirical meta-analysis random-effects model
W_RR	Sample size weighted meta-analysis random-effects approach in terms of relative risk

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

RANDOM-EFFECTS APPROACHES COMPARISONS IN META-ANALYSIS: DATA SIMULATION WITH BINOMIAL OUTCOMES

By

Dongwen Guo

August 2011

Chair: Abraham G. Hartzema
Major: Pharmaceutical Sciences

The empirical (classical) random-effects models provide biased summary effect sizes and generate too narrow 95% CI due to three factors: (a) treating random weights as non-random, (b) a probable association between effect size and weights at low event incidence, (c) use of an asymptotic normal distribution when a small number of studies are combined for a meta-analysis. Two unweighted random-effects approaches , the UW_OR and the UW_MM, were proposed to estimate the unweighted summary effect size by overcoming the issues with the empirical random-effects approach but efficiency of those two approaches are unknown. Additionally, a sample size weighted approach (W_RR) was proposed to estimate the weighted summary effect size. However, the relative performance of the unweighted summary effect size from the UW_RR versus the weighted summary effect size from the W_RR is unknown when they are forced to estimate the same summary effect size.

The study aims to evaluate the statistical properties of the UW_OR and the UW_MM, and the performance of the UW_RR and the W_RR using two major data simulations. This simulation approach is taken because it is impossible at present to

contrast the methods using mathematical means in that the calculation of the parameter dimensions are well beyond our capabilities.

Our findings suggest that the efficiency of the UW_OR and the UW_MM is close when a large number of studies are combined for a meta-analysis and the low event rates are unlikely. Further, the UW_RR generates a similar summary effect size as the W_RR does when the effect size does not depend upon the study sample size. In addition, we demonstrate that a t distribution is more appropriate compared with a normal distribution for all proposed methods, when a small number of studies are combined for a meta-analysis. It might be a good statistical practice to report results from both the sample size weighted (W_RR) and the unweighted (UW_RR) analysis (one as primary and one as secondary). If the two agree qualitatively, the consistent reports add strength to make any inference. If not, since the UW_RR estimates the unweighted summary effect size and the W_RR estimates the weighted summary effect size, the disagreement is not a contradiction, but suggests further investigation of inherent differences between larger and smaller studies.

CHAPTER 1
INTRODUCTION

Statement of Problem

Meta-analyses are systematic reviews which quantitatively combine results of previous research [1]. The Federal Drug and Food Administration (FDA) requires reviews of clinical trials on safety and efficacy in marketing applications [2]. In safety assessment, meta-analysis studies are viewed as primary evidence in support of regulatory decision making [3, 4].

In regards to meta-analytical methods, the most widely used models are the empirical (classical) fixed-effect model and the empirical (classical) random-effects models [5]. Under these two models, each study within a meta-analysis is weighted by the inverse of that study's estimated variance [5]. The empirically weighted random-effects approach suffers from three issues that have been identified in recent years [6]. These issues lead to biased summary effect size estimates and the corresponding smaller estimates of standard errors [6].

In a meta-analysis with a binomial outcome, Shuster et al. [7-9] proposed two unweighted random-effects approaches (UW_MM, UW_OR/UW_RR) and one sample size weighted random-effects approach (W_RR) to overcome issues that the empirical random-effects model has.

Theoretically, the above proposed approaches both provide unbiased summary effect size estimates [6-9]. However, we are uncertain which one is more efficient. Our research goal is to compare the efficiency of the two unweighted random-effects approaches (UW_MM vs. UW_OR) using the odds ratio data simulation, and to compare the efficiency of the unweighted random-effects and the sample size weighted

random-effects approach (UW_RR vs. W_RR) when they estimate the same summary effect size using the relative risk data simulation.

Specific Aims

The following specific aims are proposed:

Specific Aim #1: Evaluate efficiency for the UW_OR and the UW_MM in a meta-analysis of binomial trials when synthesizing a large number of studies.

Hypothesis 1: It is hypothesized that there is no precision difference between the UW_OR and the UW_MM when synthesizing a large number of studies.

Specific Aim #2: Evaluate efficiency of t approximation vs. normal approximation in a meta-analysis of binomial trials when synthesizing a small number of studies.

Hypothesis 2-1: It is hypothesized that a t approximation is more accurate than a normal approximation in terms of closeness to the 95% coverage for the UW_OR when synthesizing a small number of studies.

Hypothesis 2-2: It is hypothesized that a t approximation is more accurate than a normal approximation in terms of closeness to the 95% coverage for the UW_MM when synthesizing a small number of studies.

Specific Aim #3: Evaluate efficiency for the UW_RR and the W_RR when they estimate same summary effect size in a meta-analysis of binomial trials.

Hypothesis 3: It is hypothesized that efficiency of the UW_RR and the W_RR is similar when they estimate the same summary effect size.

Specific Aim #4: Evaluate efficiency of t approximation vs. normal approximation in a meta-analysis of binomial trials when synthesizing a small number of studies.

Hypothesis 4-1: It is hypothesized that a t approximation is more accurate than a normal approximation in terms of closeness to the 95% coverage for the UW_RR when synthesizing a small number of studies.

Hypothesis 4-2: It is hypothesized that a t approximation is more accurate than a normal approximation in terms of closeness to the 95% coverage for the W_RR when synthesizing a small number of studies.

Significance

Evidence from meta-analyses represents the highest authority in the development of guidelines and helps reach consensus when conflicting evidence occurs [1]. In addition to selecting the appropriate source studies for a meta-analysis, the appropriate selection of the meta-analysis method is a very important step in ensuring the validity of synthesized evidence, especially when a meta-analysis is adopted to summarize the low incidence of adverse drug events studies.

At the completion of this series of experiments, we expect to provide more solid supporting evidence on the efficiency of the UW_OR and the UW_MM, and on the efficiency of the UW_RR and the W_RR. The successful completion of this study will help medical researchers select appropriate meta-analysis methods in the future. Moreover, the evidence from an appropriate meta-analysis will help the FDA to evaluate the efficacy and safety of new drugs. Furthermore, formulary managers and healthcare officials will be able to use the evidence from appropriate meta-analyses to develop treatment guidelines or decide upon the inclusion of a new drug in a formulary [1] in a more rigorous manner than previously.

Definition of Terms

EFFICIENCY. In statistics, efficiency refers to the degree to which a statistic is stable from one sample to another sample. In other words, a statistic is more efficient when it is less subject to sampling variation and bias. The efficiency of a statistic can be treated as a synonym for the precision of an estimate. The more efficient a statistic is, the more precise the statistic is as an estimator for a parameter [10].

PRECISION OF METHODS. The precision of methods is a general term for the following statistics of variance, standard error and confidence interval [5].

RANDOM VARIABLE. A random variable denotes a variable whose values may be generated from a random experiment and the values cannot be predicted with certainty before a specific sample is selected [11].

LARGE-SAMPLE THEORY. As a branch of statistics, the large sample theory states that it is relatively easy to obtain good approximate results if the sample size is large [12].

CENTRAL LIMIT THEOREM. As a probability theory, the theorem states conditions under which the means of a sufficiently large number of independent random variables are approximately normally distributed [13].

DELTA METHOD. The delta method is a technique which is used to approximate expected values of functions of random variables when it is not feasible to obtain the expected value directly, such as compute the moments of an approximating asymptotic distribution [14].

PROBABILITY DISTRIBUTION. The probability distribution of a random variable provides the frequencies of each value of the random variable in a population [11].

NORMAL DISTRIBUTION. The normal distribution is characterized as a symmetrically bell-shaped curve. The parameters of μ (the mean describing the center) and σ (the standard deviation describing the spread) are used to describe the distribution of a population (Figure 1-1) [11].

STUDENT'S T DISTRIBUTION. The student's t distribution is a symmetrically bell-shaped curve. It appears similar to the normal distribution but it differs because it has an additional parameter of degrees of freedom (df) which change the shape of the curve. The degrees of freedom are equal to the sample size minus one. The smaller the degrees of freedom, the larger the area under the tail of the curve is. In contrast, the larger the degrees of freedom, the closer it is to the curve of standard normal distribution. The normal distribution is a special case of the student's t distribution with df=∞ (Figure 1-1) [15].

BINOMIAL DISTRIBUTION. The binomial distribution describes the number of occurrences of a particular event in a series of n trials. It needs to meet four conditions. First, the trials are identical. Second, the outcome of each trial is independent, i.e. the outcome is not associated with the outcome from other trials. Third, the outcome has only two values of 'success' or 'failure.' The last condition is that the probability of success (π) stays the same for all trials. The binomial distribution is described by the sample size (n) and the probability of success (π) (Figure 1-2) [11].

Organization of Study

The remainder of the study is organized into four chapters. Chapter two presents a review of the related literature covering the concepts proposed in this chapter. Chapter three describes the research design and statistical analysis plan of the study. Chapter four delineates the results of the study. Chapter five contains discussion and conclusions of this study.

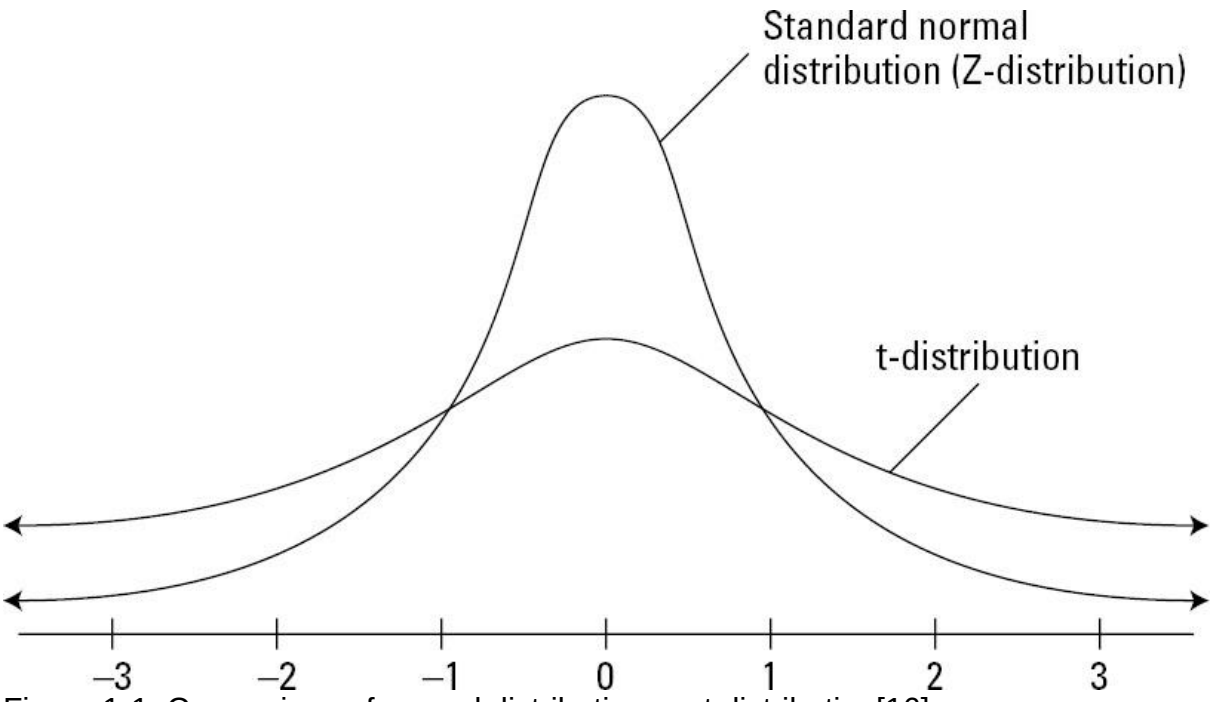

Figure 1-1. Comparison of normal distribution vs. t distribution[16].

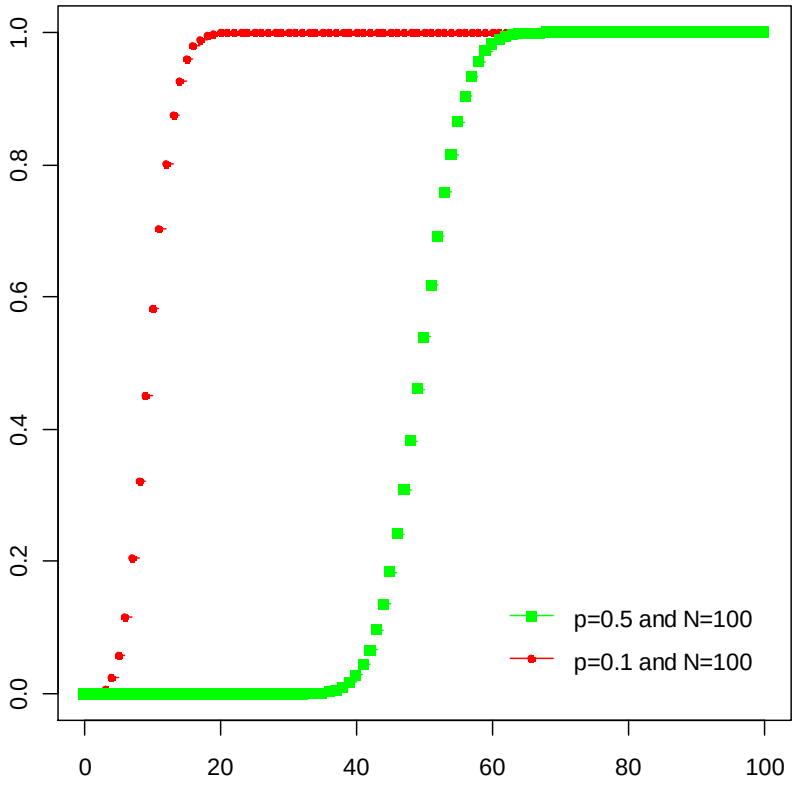

Figure 1-2. Binomial distribution (Y=Probability, X=N) by R software V2.1

CHAPTER 2
REVIEW OF THE LITERATURE

This chapter provides the background and the rationale for conducting this study. We first define meta-analysis. Next, we review the most popularly used statistical meta-analytical models for this study. Third, we explain the selection of the models. Fourth, we summarize the criticism with the empirical random-effects model, as stated by DerSimonian-Laird (W_DL) [17]. Fifth, we introduce the solutions that overcome the issues with the W_DL model. Finally, we propose questions that remain unknown.

Definition of Meta-analysis

Meta-analysis is described as 'any systematic method that uses statistical analyses for combining data from independent studies to obtain a numerical estimate of the overall effect of a particular procedure or variable on a defined outcome.' [18]. In 1976, the term of 'meta-analysis' was coined by Glass [1, 19]. 'meta-analysis' became an accepted term in Medline In 1989 [1]. The evidence from systematic reviews and meta-analyses is often placed on the top of a pyramid in terms of authority (Figure 2-1, [20]), although there is no single, universally-accepted hierarchy of evidence. There are two main steps in conducting a meta-analysis: a critical review of the literature and synthesizing data from each individual study into a summary (pooled or combined) effect size [1] (Figure 2-2).

Models of Meta-analysis

In regards to meta-analysis models, researchers usually have two choices, a fixed-effect model and a random-effects model [5] (Figure 2-3). The common underlying assumptions are that each individual study is independent and each one produces an unbiased estimate of that study's effect size [6]. As for a meta-analysis with binomial

outcomes, most meta-analysis models presume that individual effect sizes in the form of odds ratios or relative risks are an independent, identically normal distribution on a log scale [7, 17, 21, 22], while some work with other prior distribution, i.e. the Bayesian approach [23-25]. We elaborate next upon the most popular used models: fixed-effect and random-effects models [5].

Fixed-effect Model

The fixed-effect model was proposed by Mantel and Haenszel in 1959 [1, 26]. The specific underlying assumption for the fixed-effect model is that only one common true value of effect size exists for any study included in a meta-analysis [5]. The variation in the meta-analysis comes from sampling error within each individual primary study. The weight is usually the inverse of that study's estimated variance.

Random-effects Model

Another meta-analysis model is the random-effects model developed by William Cochran in 1954 [27], but was not used in the medical field until 1986 by DerSimonian and Laird [1, 17]. The specific underlying assumption for the random-effects model is that the effect size of each individual study is normally distributed as opposed to sharing a common true value for the fixed-effect model. Under this model, the source of variation in the meta-analysis comes from the sampling error within each study and the difference across studies combined for the meta-analysis [28].

In summary, the commonly used models for a meta-analysis are the fixed-effect and the random-effects models, which have different specific underlying assumptions. The fixed-effect model is a special case of the empirical random-effects model. Researchers need to choose an appropriate model before conducting a meta-analysis.

Selection of Meta-analysis Models

Two ways can help a researcher to select the fixed-effect model or the random-effects model. One is using a Cochran Q test. The other one is based upon clinical judgment.

Method 1-Cochran Q test

One method of selecting a meta-analysis model is to use the Cochran Q test and corresponding p value. The Cochran Q test is a diagnostic tool that is used to determine which model is appropriate, the fixed-effect model or the random-effects model. If the Q statistic is statistically significant with the p value less than 0.05, the random-effects model is the choice. However, if the number of studies is small the Q statistic may not have sufficient power to detect the between-study variance. Therefore, after the test, we only have two possible conclusions: (a) study effect sizes are statistically diverse ($p<0.05$), suggesting choosing the random-effects model or (b) it is inconclusive as to whether study effect sizes are diverse, i.e. we do not know which model is appropriate. Therefore, the Cochran Q test is not recommended as a diagnostic tool.

Method 2-Clinical Judgment

Another means of selecting a meta-analysis model is based on the distribution of the effect sizes of individual studies and the relevant sources of errors [5]. In other words, we should make our decision based on the nature of the studies being combined for a meta-analysis instead of using a data driven test, such as the Cochran Q test.

Fixed-effect model. The fixed-effect model is the choice when (a) we believe that all the studies examine the same hypothesis, and (b) the research goal is to calculate the common effect size, which intends to be generalized to the same population. Under these conditions, the fixed-effect model is a better choice for a multicenter clinical trial.

This is because when a drug company conducts several studies with the same way of recruiting patients, using the same researchers, and using the same dosing regimens [28].

Random-effects model. In contrast, the random-effects model is more easily justified than the fixed-effect model when a researcher combines data from a series of studies performed by different investigators [28]. First, it is inappropriate to assume that all the studies examine the common effect size in that these studies would differ in many ways that impact the results, such as different study designs, different study populations, different comparison groups, different dosing regimens, different statistical methods etc. Additionally, the goal of the meta-analysis is often to generalize to a range of populations.

In summary, although the two models are popularly used, the random-effects model is a more appropriate approach when a meta-analysis is conducted at a study level [28]. Among available random-effects models, the empirically weighted random-effects approach (W_DL) is usually used [6].

Issues Regarding Empirically Weighted Random-effects Approach

In this section, we describe the empirically weighted random-effects approach (W_DL) first, and then identify the issues with this approach.

As for the most popularly used random-effects model, the empirical approach is a weighted approach. Each study is weighted by the inverse of the each study's estimated variance, including the within-study variance and the between-study variance. The rationale behind the weights is that it intends to emphases the impact from larger sample size studies on the summary effect size because the larger the sample size, the lower within-study random error.

The conceptual formulas for the summary weighted mean (θ_w) and corresponding 95% CI are computed below [28]. The outcome metrics could be odds ratio, relative risk or risk difference. If the outcome is measured in the forms of odds ratio or relative risk, they need to be transformed to a log scale then apply the following formulas. An illustration diagram in the form of odds ratio is displayed in Figure 2-4.

The conceptual formula for the summary weighted mean (θ_w)

$$\theta_w = \frac{\sum_j W_j \ \theta_j}{\sum_j W_j}$$

Where

θ_w refers to the weighted summary mean

θ_j refers to the j th study's effect size

j refers to j th study under a meta-analysis

$$W_j = \frac{1}{V_j^*}$$

W_j refers a weight for the j th study

V_j^* refers to the variance of the j th study, including the within–study variance for the j th study plus the between-studies variance (τ^2).

$$V_j^* = V_j + \tau^2$$

Where

$$\tau^2 = \frac{Q - df}{C}$$

Where

τ^2 (Tau squared) refers to the between study variance estimate

Q refers to the Q statistic.

$$Q = \sum_j W_j \theta_j^2 - \frac{\left(\sum_j W_j \theta_j\right)^2}{\sum_j W_j}$$

$$C = \sum W_j - \frac{W_j^2}{W_j}$$

df refers to the degrees of freedom, which is equal to M-1

Where M refers to the number of studies for a meta-analysis

95% CI is computed blow:

$$LL = \theta_w - 1.96 * SE_{\theta_w}$$
$$UL = \theta_w + 1.96 * SE_{\theta_w}$$

Where

LL refers to the low limit of 95% confidence interval

UL refers to the upper limit of 95% confidence interval

1.96 refers to the z score for a normal distribution

$$SE_{\theta_w} = \sqrt{V_{\theta_w}}$$

Where $V_{\theta_w} = \frac{\sum\left[W_j \ SE_{\theta_j}\right]^2}{\sum W_j^2}$

However, the empirical random-effects model (W_DL) has three issues (Table 2-1). First, the W_DL is widely used with weights [28]. As we mentioned earlier, the

weight refers to the inverse of the estimated variance of each study, including both variance from within study and between studies. Empirically, the weighted random-effects analysis treats its weights as non-random against its random nature, which is shown above in the conceptual formula for the weighted summary mean (θ_w). Therefore, the estimate of the summary effect size is not only biased but also does not estimate what it claims to estimate of the unweighted summary effect size. In addition, the confidence intervals derived from empirical weights are narrower than they should be [29].

The second, and often the case, is that the low event rates make the estimate of the effect size for that individual study highly biased, including its standard error. The reason is that the estimates of the effect size and the standard error for an individual study are calculated according to a large sample theory [12]. This situation is similar to that of the low expected cell number problems in two-by-two contingency tables, which explains why the Fisher's Exact Test is popular. Moreover, at low event rates, the weights are generally associated with the effect sizes of studies [6].

The third issue occurs when the number of studies combined for a meta-analysis is small. This is because the large sample theory plays a role not only in each individual study but also in the number of studies combined for a meta-analysis. The estimator of between-study variations is based upon study's τ^2. In other words, when a small number of studies are combined for a meta-analysis, the large sample approach is unreliable. Therefore, the z score of 1.96 that is based upon the standard normal distribution applying to calculate the 95% CI is not appropriate. Instead a t distribution with a degree of freedom (i.e. the number of studies minus one) might be used. But as

the number of studies increases, the summary effect size with a t distribution is closer to the summary effect size with a normal distribution.

In conclusion, although the empirical weighted random-effects model (W_DL) often suffers from three issues of the empirical weighting, the low event rates, and the asymptotic normal distribution assumption, it has been widely used [6].

Solutions

To overcome the issues using the empirically weighted random-effects approach, Shuster et al. proposed the two unweighted random-effects approaches [7, 8] and one sample size weighted approach [6]. The advantages for each method are listed in Table 2-3. Now we introduce how a summary effect size is computed in the following order: UW_MM, UW_OR/UW_RR and W_RR.

Solution 1: Unweighted Random-effects Approach I

Shuster et al. proposed an unweighted random-effects approach (UW_MM) in 2010 [8]. By using this approach, we compute the odds ratio for each study first, and then transform the odds ratio into its log form. In the log scale, we assume that each log odds ratio is normally distributed. Therefore, we derive the mean log odds ratio by adding each log odds ratio together and dividing the summation by the number of studies. Last, we take the antilog on the mean log odds ratio back to the original odds ratio scale. We illustrate the steps in Figure 2-5. The detailed formulas are displayed as follows.

The operational unweighted estimate for a summary odds ratio for a sample is:

$$OR = e^{(\sum_j LOR_j)/M}$$

Where

M=number of studies

$LOR_j = In(OR_j)$

$OR_j = \left\{ \pi_{2j}(1-\pi_{1j}) / \pi_{1j}(1-\pi_{2j}) \right\}$

$\pi_{ij} = F_{ij} / N_{ij}$

F_{ij} and N_{ij} denote the number of events and sample size for treatment i and trial j

i =1 represents control arm, i =2 represents treatment arm

95% CI for LOR in a log scale is

$$LOR \pm t_{1-\alpha/2} SE_{LOR}$$

Where

$Z_{LOR} = \left\{ LOR - LOR \right\} / SE_{LOR}$ has an asymptotic t-distribution with M-1 degrees of freedom

$$SE_{LOR} = SQRT \left\{ \left(\frac{\sum_j \left(LOR_j \right)^2 - M \left(\frac{\sum_j LOR_j}{M} \right)^2}{M-1} \right) / M \right\}$$

95% CI for OR in the original scale is

$$OR * e^{\pm t_{1-\alpha/2} SE_{LOR}}$$

Solution 2: Unweighted Random-effects Approach II

Shuster et al. [7, 9] proposed another unweighted random-effects approach that completely avoids the low event issue with the weighted random-effects approach by taking ratios only after combining the proportions of each study (Table 2-2, Figure 2-6).

The only place that the large sample theory is ever employed is at the number of studies being combined. This method provides an unbiased summary effect size by avoiding taking the log of the effect size for each individual study and the antilog of the summary effect size back to the original scale. Moreover, this method avoids the chance of not being able to calculate individual effect size because of zero event in one or both arms.

Odds Ratio Estimation:

The operational unweighted estimate for a summary odds ratio for a sample is:

$$OR = \left\{ \pi_2 (1-\pi_1) / \pi_1 (1-\pi_2) \right\}$$

Where

$$\pi_i = \sum_j P_{ij} / M$$

$$P_{ij} = F_{ij} / N_{ij}$$

F_{ij} and N_{ij} denote the number of events and sample size for treatment $\boldsymbol{i}$ and trial j

$\boldsymbol{i}$ =1 represents control arm, $\boldsymbol{i}$ =2 represents treatment arm

M=number of studies

95% CI for $\log OR$ in a log scale is

$$LOR \pm t_{1-\alpha/2} SE_{LOR}$$

Where

$Z_{LOR} = \left\{ LOR - LOR \right\} / SE_{LOR}$ has an asymptotic t-distribution with M-1 degrees of freedom

$$SE_{LOR} = SQRT \left\{ (Q + R - 2S) / M \right\}$$

$$Q = C_{11} / \left[\pi_1 (1-\pi_1) \right]^2,$$

$R = C_{22} / \left[\pi_2 (1-\pi_2) \right]^2$,

$S = C_{12} / \left[\pi_1 (1-\pi_1) \pi_2 (1-\pi_2) \right]$

95% CI for OR in the original scale is

$$OR * e^{\pm t_{1-\alpha/2} SE_{LOR}}$$

Relative Risk Estimation:

The operational unweighted estimate for a summary relative risk for a sample is:

$$RR = \pi_2 / \pi_1$$

Where

$\pi_{ij} = \sum_j P_{ij} / M$

$P_{ij} = F_{ij} / N_{ij}$

F_{ij} and N_{ij} denote the number of events and sample size for treatment i and trial j

i =1 represents control arm, i =2 represents treatment arm

M=number of studies

95% CI for $\log RR$ in a log scale is

$$LRR \pm t_{1-\alpha/2} SE_{LRR}$$

$Z_{LRR} = \left\{ Log(RR) - Log(RR) \right\} / SE_{LRR}$ has an asymptotic t-distribution with M-1 degrees of freedom.

Where

$SE_{LRR} = SQRT \left\{ (Q_1 + R_1 - 2S_1) / M \right\}$

Where

$Q_1 = C_{11} / \left[\pi_1 \right]^2$,

$R_1 = C_{22} / \left[\pi_2 \right]^2$,

$$S_1 = C_{12} / \left[\pi_1 \pi_2 \right]$$

$$C_{kl} = \sum_j \left\{ P_{kj} - \pi_k \right\} \left\{ P_{lj} - \pi_l \right\} / \left(M - 1 \right)$$

Where k=1,2; l=1,2

95% CI for RR in the original scale is

$$RR * e^{\pm t_{1-\alpha/2} SE_{L\,RR}}$$

In all, theoretically the above two unweighted random-effects approaches (UW_MM and UW_OR), which are both in odds ratio fashion provide valid summary effect size estimates for a meta-analysis. However, the efficiency of both methods remains unknown.

Solution 3: Sample-size Weighted Random-effects Approach

This method was proposed by Shuster et al. in 2011 [6]. The basic idea is that we conceptually draw a subject at random from the target population of past, present, and future subjects. The probability of selecting a given study to which this subject belongs is proportional to the total sample size of the study. The inferential framework differs from the solution 2 of the unweighted approach (UW_RR) in that this sample size weighted approach is at a patient level, not at a study level. This approach is illustrated in Figure 2-7.

The conceptual formula for the weighted summary effect size is

$$\theta_w = \sum_j W_j \theta_j$$

Where

W_j represents the fraction of subjects in the universe, which belongs to trial j.

θ_j is effect size estimate for j th study

θ_W is completely different from the θ defined for solution 1-2, unless there is no association between the sample size and the effect size.

The formula for the weighted summary effect size for a sample of studies is

$$\theta_W = \sum_j U_j \theta_j / \sum_j U_j$$

Where

θ_j denotes the estimated effect size for the j th study

$$Uj = (N_{1j} + N_{2j}) / 2$$

The operational formula for the weighted summary effect size for a sample is

$$RR = \frac{\overline{A_2}}{\overline{A_1}}$$

Where

$$\overline{A_i} = \sum_j A_{ij} / M$$

$A_{ij} = U_j P_{ij}$ is the adjusted number of events

i=1,2 , 1 indicating the control arm, 2 indicating the treatment arm

j=1,2,…M indicating the number of studies

P_{ij} denotes the estimate of risk for i th arm and j th study

95% CI for $\log RR$ in a log scale is

$$LRR \pm t_{1-\alpha/2} SE_{LRR}$$

Where

$Z_{LRR} = \left\{ LRR - LRR \right\} / SE_{LRR}$ is asymptotically t distribution with (M-1) degrees of freedom

$$SE_{LRR} = SQRT\left[\begin{array}{l}\left\{S(A_{1j})/\overline{A}_1\right\}^2 + \left\{S(A_{2j})/\overline{A}_2\right\}^2 - \\ 2\left\{C(A_{1j}, A_{2j})/(\overline{A}_1\overline{A}_2)\right\}\end{array}\right]/M$$

Where S() represents the sample standard deviation and

C(,) represents the sample covariance.

95% CI for RR in the original scale is

$$RR * e^{\pm t_{1-\alpha/2} SE_{LRR}}$$

In all, although the two unweighted random-effects approaches (UW_OR/UW_RR) have been proven to be valid, we do not know their efficiency. In addition, another solution of the sample size weighted random-effects approach (W_RR) was proposed, but it estimates the weighted summary effect size, which is different from the unweighted summary effect size. It is unknown if we force them to estimate the same effect size, which one would be more efficient.

Figure 2-1. Evidence Hierarchy

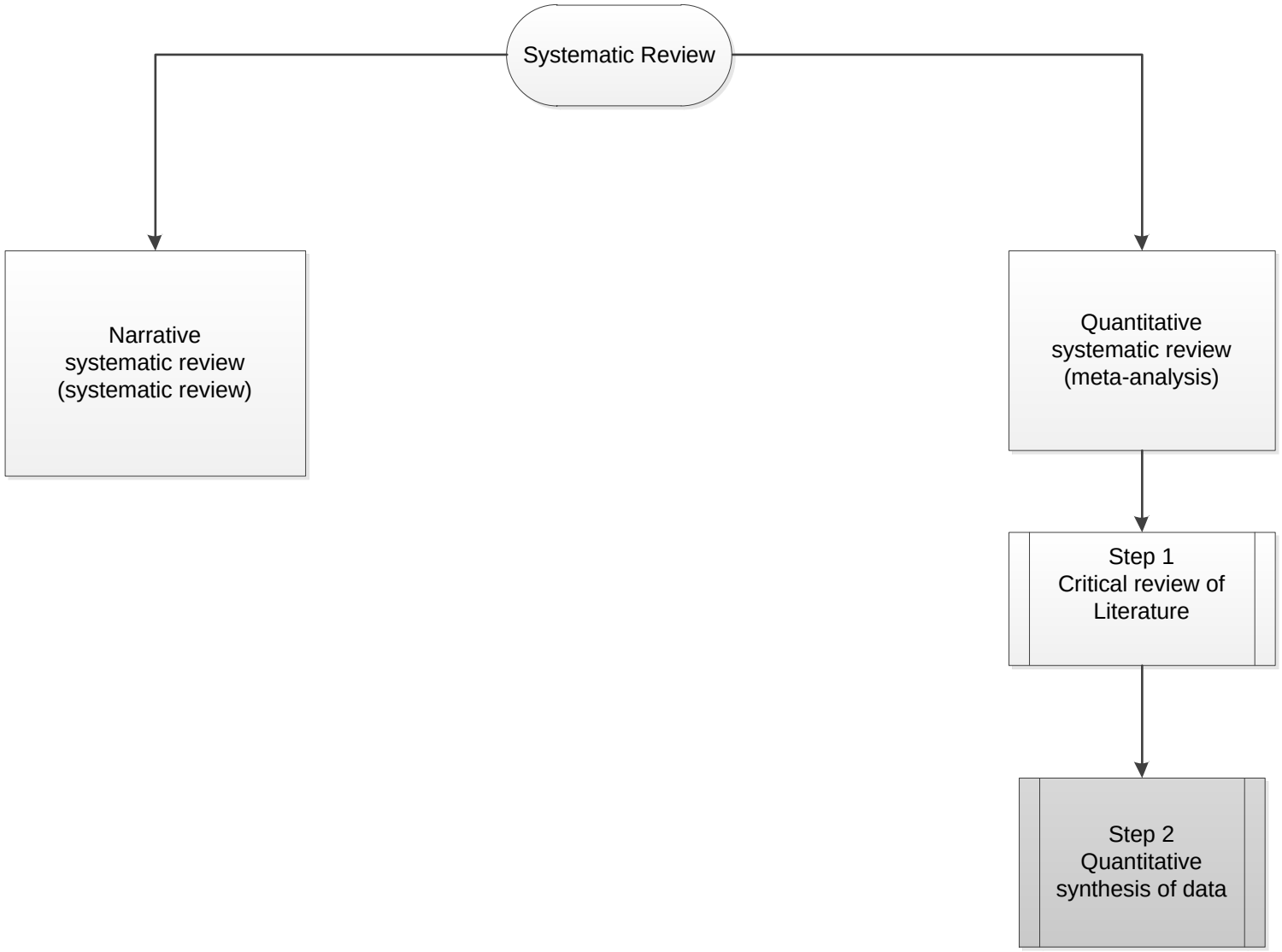

Figure 2-2. Overview of systematic review

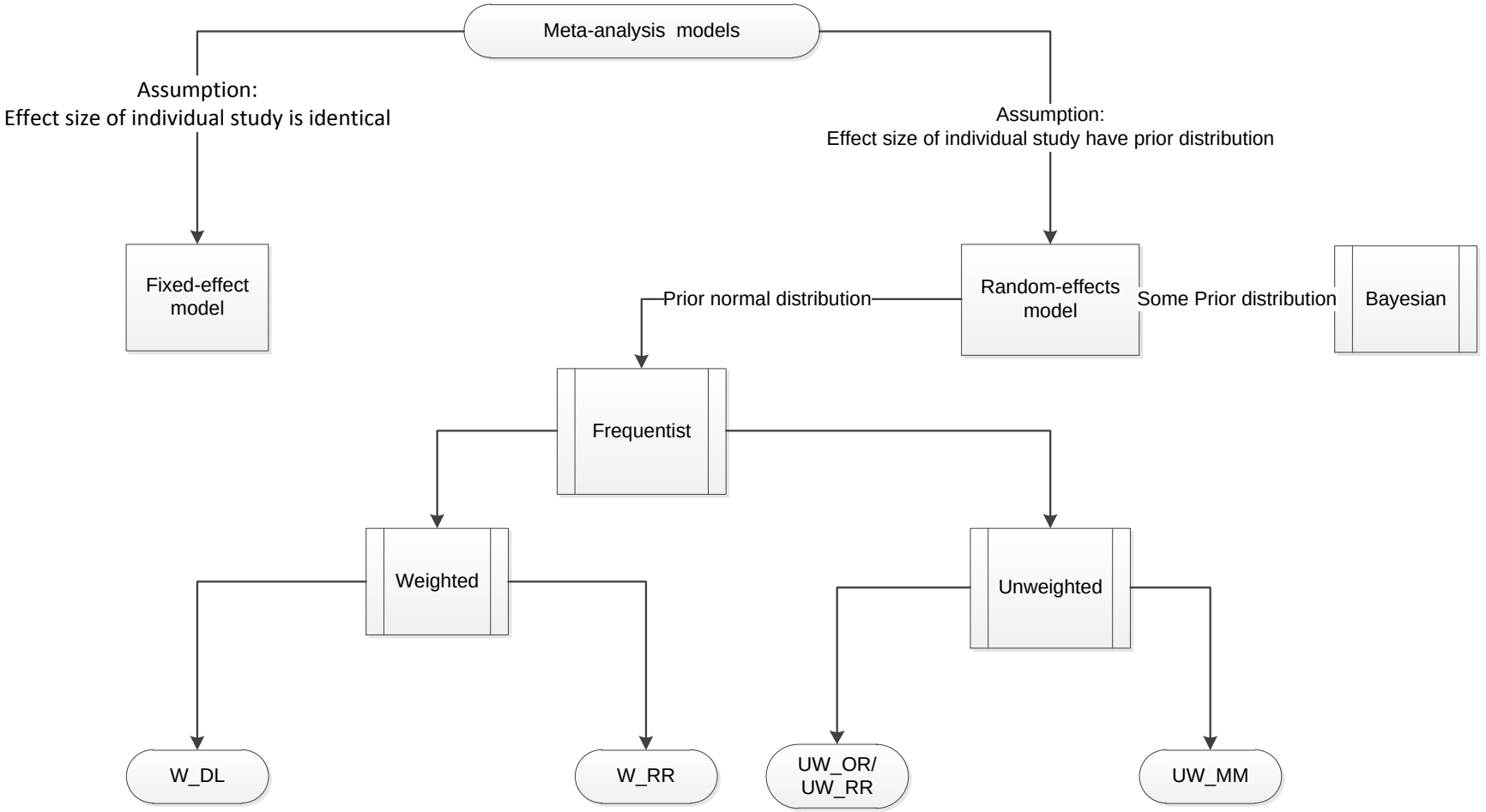

Figure 2-3. Overview of meta-analysis

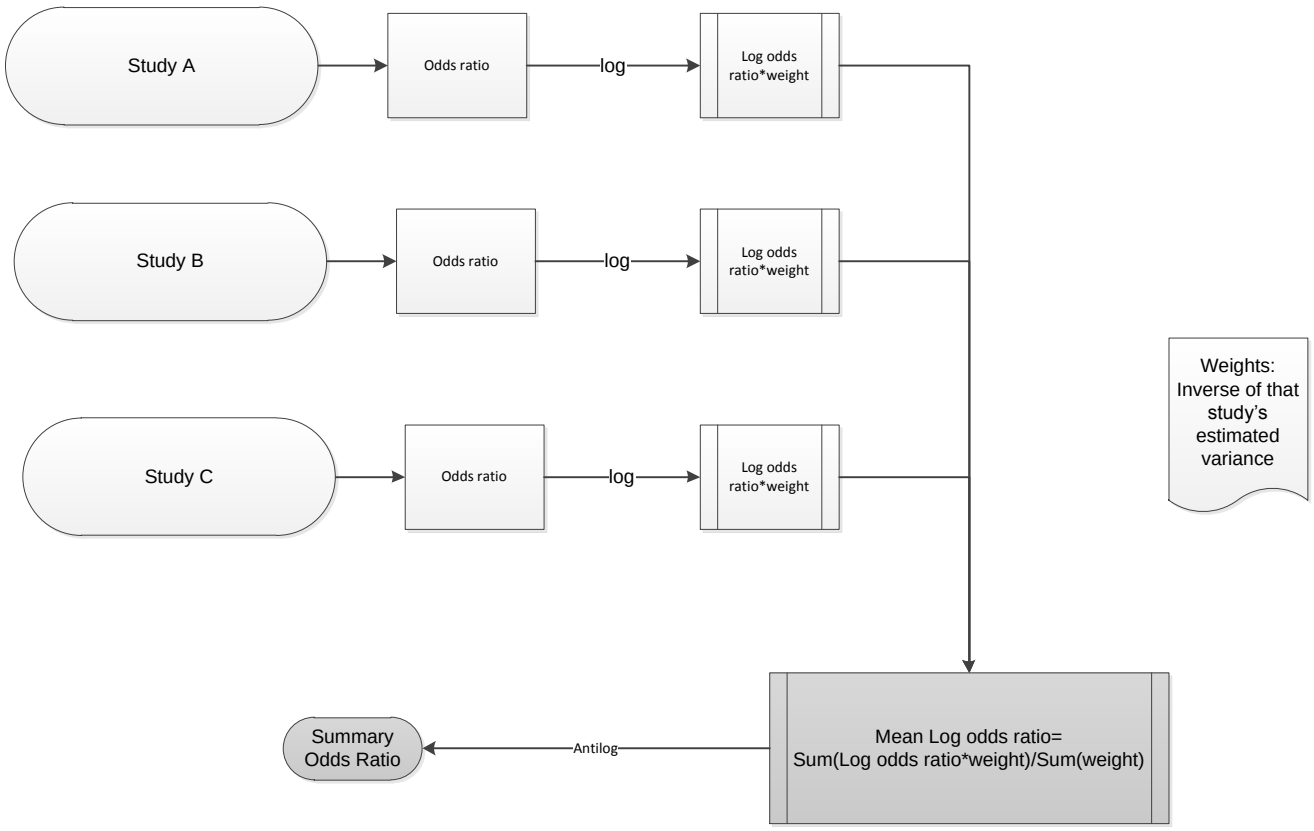

Figure 2-4. Illustration of empirically weighted random-effects approach (W_DL)

Table 2-1. Illustration of issues with empirically weighted random-effects approach

No. of Issues	Issues	Summary effect size	SE	95% CI
1	Empirical weights are random but treated as non-random	• Does not estimate the claimed unweighted effect size • Biased summary effect size	Inaccurate SE	Too narrow
2	Estimate of point effect size for an individual study is computed based upon the large sample theory. This becomes a problem when low event occurs.	• Biased effect size for individual study • Bias from log transformation on each individual effect size	Inaccurate SE	Skewed
3	Low number of studies is combined for a meta-analysis but with a normal distribution assumption (z=1.96) to calculate 95% CI	N/A	N/A	Too narrow

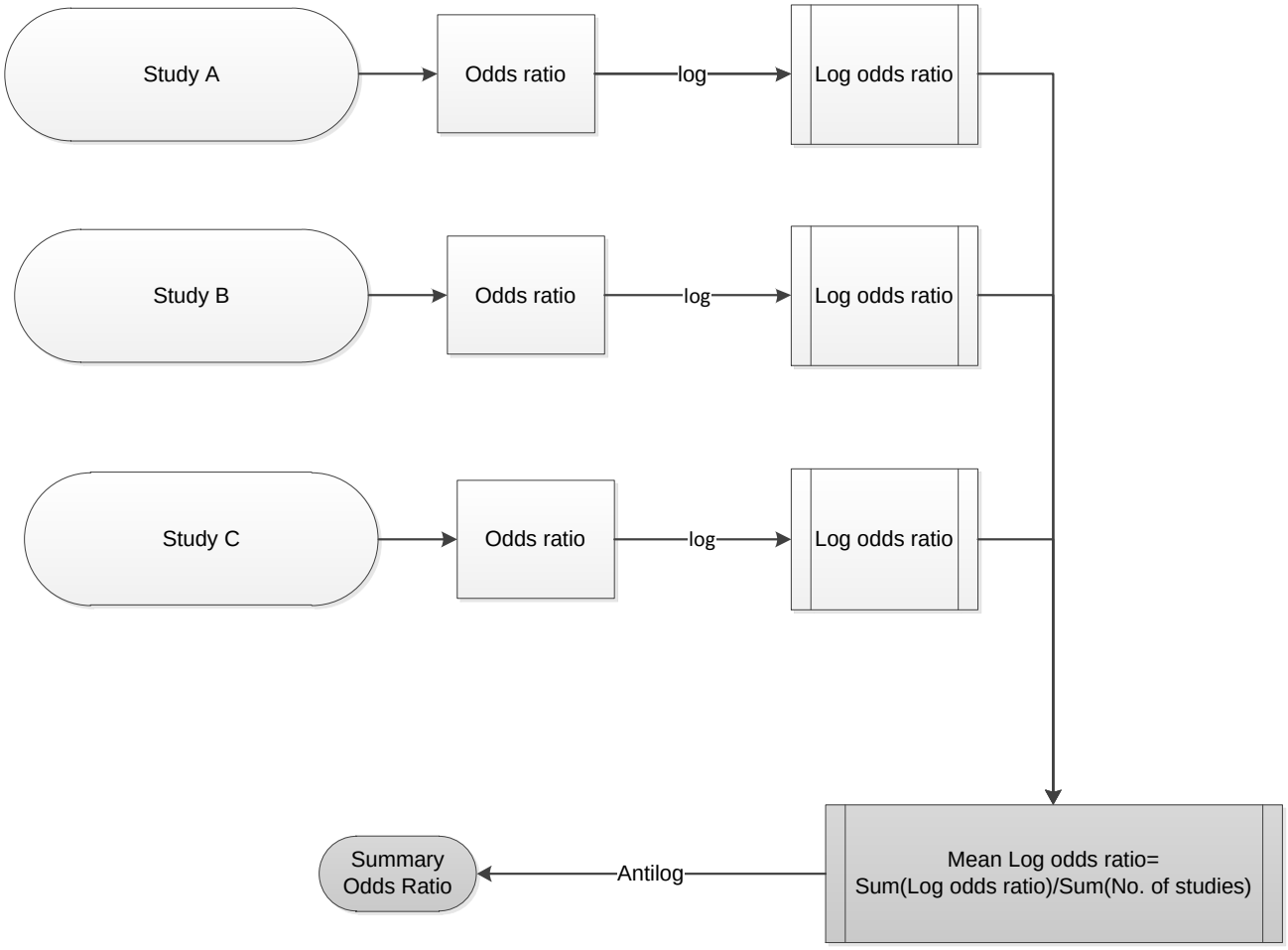

Figure 2-5. Illustration of unweighted random-effects approach (UW_MM)

Table 2-2. Nomenclature for 2 by 2 table of outcome by treatment

	Outcomes (+)	Outcomes (-)	N
Treatment	a	b	n 1
Control	c	d	n 2

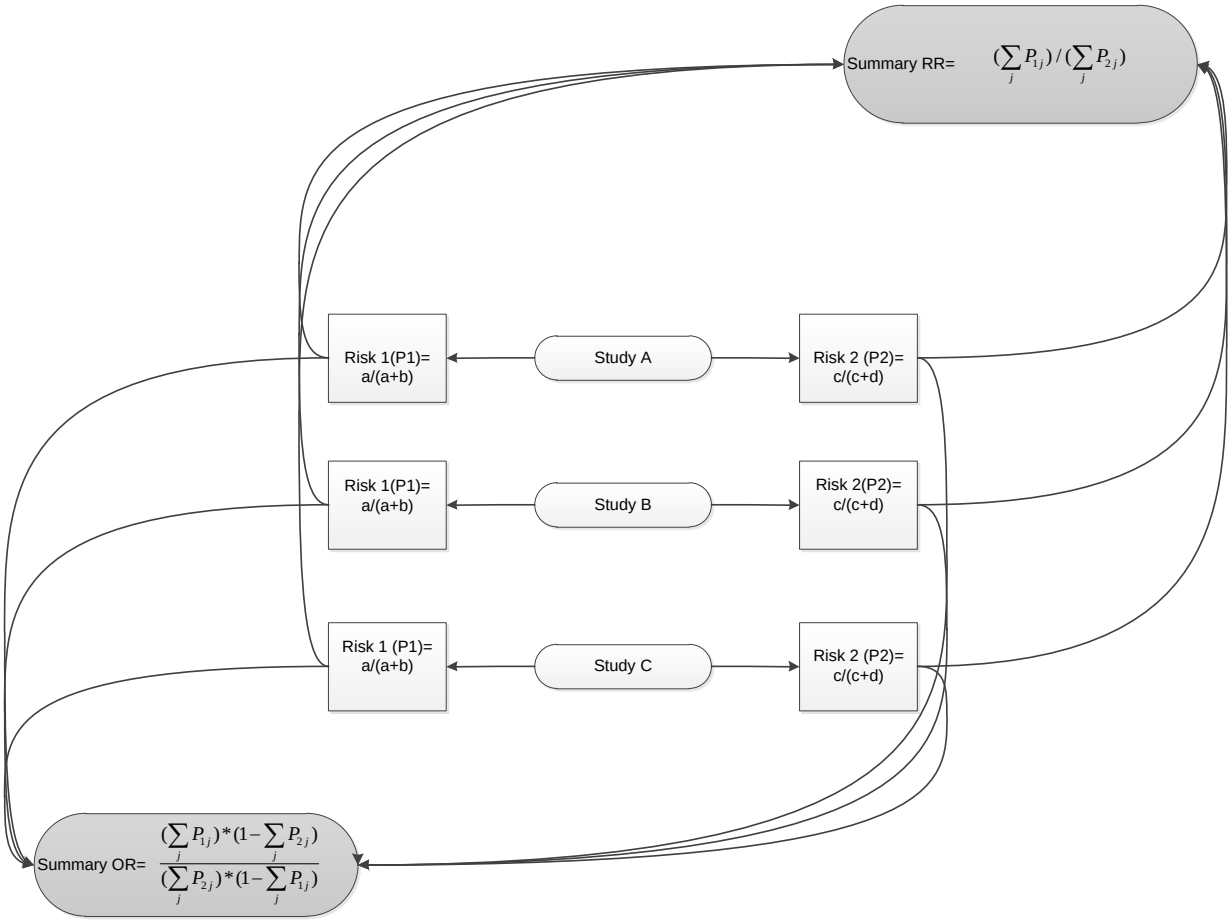

Figure 2-6. Illustration of unweighted random-effects approach (UW_OR/UW_RR)

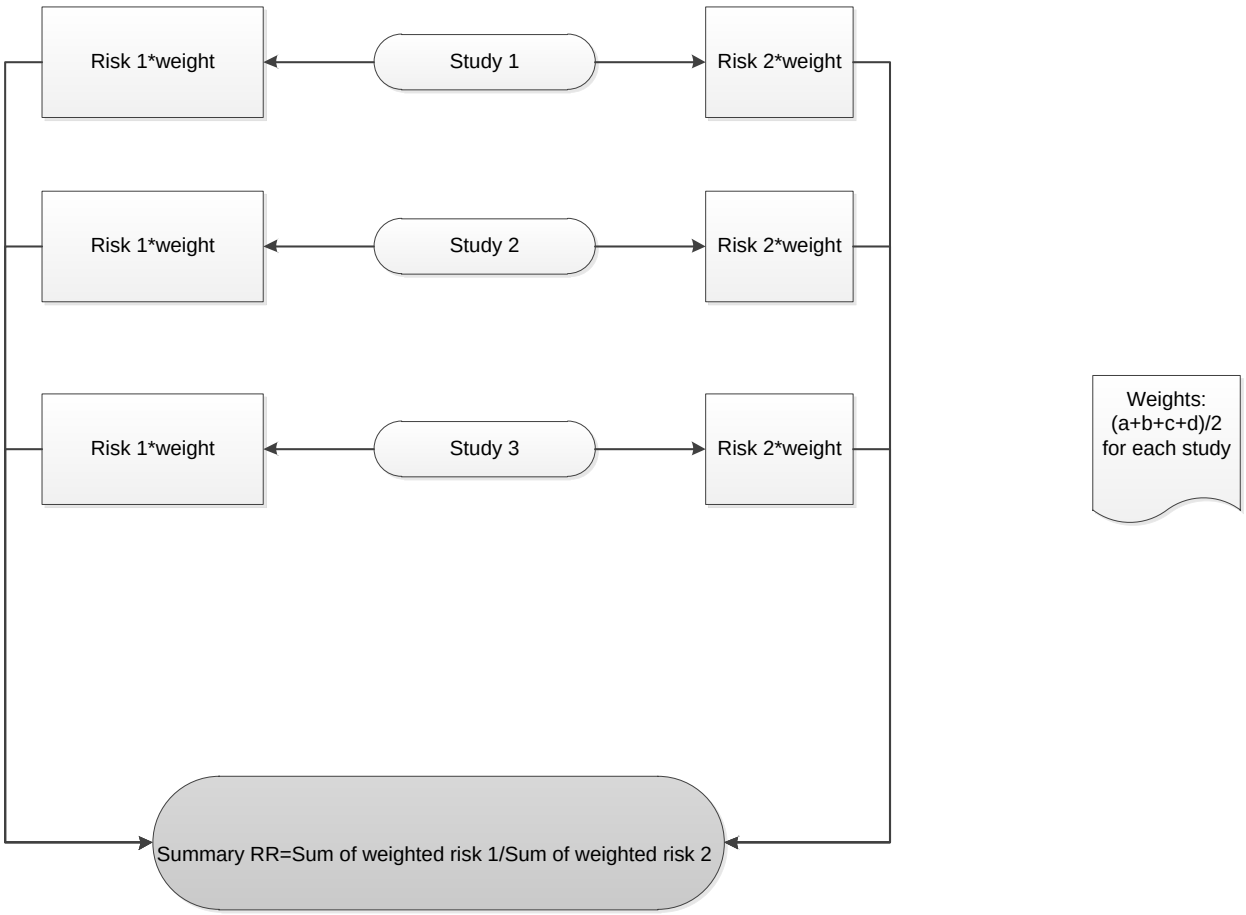

Figure 2-7. Illustration of weighted random-effects approach (W_RR)

Table 2-3. Illustration of advantages of three proposed random-effects approaches

No. of Issues	Issues w/DL	UW_MM	UW_OR/UW_RR	W_RR
1	Empirical weights are random but treated as non-random	Using no weight	Using no weight	• Using sample size as weights and treating them as random, plus
2	Estimate of point effect size for an individual study is computed based upon the large sample theory. This becomes a problem when low event occurs.	N/A	• Using summation of proportions for each individual study and doing ratio at summary level, plus • Avoiding log transformation for each point effect size	• Using summation of weighted proportions for each individual study, plus • Avoiding log transformation for each point effect size
3	Low number of studies is combined for a meta-analysis but with a normal distribution assumption (z=1.96) to calculate 95% CI	A t distribution is close to a normal distribution when a large number of studies are combined	Same as UW_MM	Same as UW_MM

CHAPTER 3
METHODS

Given the issues of the empirical random-effects model (W_DL), the two unweighted random-effects approaches and one weighted random-effects approach were proposed [7, 8]. Aim 1 and aim 2 of this study compare the efficiency of the two unweighted random-effects approaches (UW_OR vs. UW_MM) using an odds ratio data simulation. Aim 3 and aim 4 of this study compare the efficiency of the UW_RR and the W_RR when they estimate the same effect size using a relative risk data simulation.

Two sections are addressed below. First, we delineate the data simulation process, including the odds ratio data simulation and the relative risk data simulation. Second, we elaborate statistical analyses for each aim.

Odds Ratio Data Simulation

In this section, two components are included. First, we define scenarios by using predetermined parameters. Second, we repeatedly simulate random samples 10,000 times [30-32] under each scenario to conduct corresponding series of meta-analyses (Figure 3-1). The concept of a scenario can be thought as a source target population in terms of individual studies.

Definition of a Scenario

We use six parameters to define a scenario, including the global true probability of an event on the treatment arm (Πt), the global true probability of an event on the control arm (Πc), the minimum number of subjects in each arm (NLow), the maximum number of subjects in each arm (NHigh), the diversity across studies (D), and the number of studies (M) requested for conducting a meta-analysis under the scenario.

A series of scenarios are generated from the combination of these six parameters of interest (Figure 3-1). Additionally, each scenario needs to meet the following three criteria: (1) Πt is less or equal to Πc and (2) Πc (1+D) is less than or equal to the value of one and (3) if NLow is not equal to NHigh. The reasons for the first two criteria are two folds as follows. a) These ensure the probability of the event of interest is not higher than the value of one on each arm. b) Limiting the values for the metric of odds ratio to be less than one is for reducing the redundancy of scenarios, because the failure rate and success rate are symmetric for the metric of odds ratio. The third criterion is to ensure variation in the sample sizes across studies in the predefined range between the lowest number of subjects and the highest number of subjects in an individual study. In this study, we assume that the number of subjects is equally distributed in both arms. The values for each parameter are listed in Table 3-1. The considerations of selecting these values are based upon empirical clinical trials combined for a meta-analysis and the efficiency of data simulations. We also provide three examples of scenarios with different parameters in Table 3-2.

Random Sampling for Meta-analyses

Once a scenario is defined, i.e. the pool of studies is established, we start doing random sampling of studies from the pool. Each selected study is structured with four random variables, the number of events in the treatment arm (treatment event), the number of events in the control arm (control event), and the number of subjects (treatment N / Control N) in each arm.

Under a scenario, in order to generate an individual study, the true event rate in the treatment arm (ET) for an individual study is uniformly distributed from (1-D) Πt to (1+D) Πt. The true event rate in the control arm (EC) for the same study is uniformly

distributed from (1-D) Пc to (1+D) Пc. The number of subjects (treatment N/control N) in each arm is uniformly distributed between NLow and NHigh. The number of events in the treatment arm is binomially distributed based upon the two factors of the sample size in the treatment arm and the true event rate in the treatment arm (ET) for the study, while the number of events in the control arm is binomially distributed based upon the two factors of the sample size in the control arm and the true event rate in the control arm (EC) for the same study. The attributes of the random variables are listed in Table 3-3.

According to the definition of a scenario, in order to conduct a meta-analysis, M studies are needed to be chosen randomly under the scenario (i.e. the source of studies). Table 3-4 provides an example with five randomly selected studies based upon scenario 1 in Table 3-2. Once we have a set of M studies, i.e. we finish a round of random sampling, then we are ready to apply each proposed meta-analysis approach to calculate outcomes of interest, which we describe in the next section. Under a scenario, we randomly select 10,000 times M number of studies to prepare for further analyses.

Relative Risk Data Simulation

The relative risk data simulation is similar to the odds ratio simulation (Figure 3-2). The only difference lies in that we use a different outcome metric of relative risk (RR) instead of odds ratio (OR). We use the RR simulation due to the precision issue with the calculation of standard error for the sample size weighted random-effects approach (W_RR) . The reason is that the parameters used to calculate the standard error for the odds ratio are 14 (4 means, 4 variances, and 6 covariances), whereas 5 parameters are needed for the W_RR (2 means, 2 variances, and 1 covariance) [6].

Definition of a Scenario

The scenarios are generated from the combination of the six parameters of interest (Table 3-5). A scenario needs to meet the same three criteria as we described in the odds ratio data simulation.

Random Sampling for Meta-analyses

The detailed random sampling process is the same as it is for the odds ratio data simulation (Figure 3-2).

In sum, to ensure the quality of two data simulations, we randomly selected a couple of samplings and conducted corresponding meta-analyses for each proposed approach manually to ensure the results from data simulations and from manual calculations reach consensus. The odds ratio data simulation is used for aim 1 and aim 2 of the study, while the relative risk data simulation is used aim 3 and aim 4 of the study.

Statistical Analyses

The outcomes of interest are measured in two ways in this simulation study. One is the mean coverage of the global true value of odds ratio or relative risk for the 95% confidence intervals (95% CI) for each scenario. Each scenario has one global true OR value or one global true RR value. The other way is the mean length of the 95% CI in the log scale, which is derived from an upper limit minus a lower limit of the 95% CI for each scenario. The closer the mean coverage probability to 95% and the narrower the mean length of the 95% CI indicate a better efficiency of an approach under a scenario.

Aim 1

Evaluate efficiency for the UW_OR and the UW_MM in a meta-analysis of binomial trials when synthesizing a large number of studies.

Under each scenario, we carried out the following meta-analysis using both approaches (UW_OR vs. UW_MM) as shown in Figure 3-3.

Step 1, for each random sample of M studies, we conducted a meta-analysis to calculate the coverage score of the global true odds ratio (OR). The coverage score is a dummy variable with the value of 1 or 0. If the 95% CI contains the global true OR, then the score is equal to 1, otherwise it is scored as 0. Meanwhile we calculated the length of the 95% CI for each approach.

Step 2, after 10,000 random samplings, i.e. 10,000 times meta-analysis, we calculated the mean coverage of the global true OR and the mean length of the 95% CI for the scenario.

In all, we followed the same two steps described above for each scenario. At the end, for each scenario, we summarized two measurements for each of the two approaches of interest (UW_OR vs. UW_MM).

Aim 2

Evaluate efficiency of t approximation vs. normal approximation in a meta-analysis of binomial trials when synthesizing a small number of studies.

As we stated in Chapter II, when a small number of studies are combined for a meta-analysis, it is suggested that a t distribution is used instead of a normal distribution for the summary effect size in a log scale for binomial outcomes. However, the impact of the different distributions on the efficiency of unweighted random-effects approaches (UW_OR vs. UW_MM) remains unknown when a meta-analysis contains a small number of studies with binomial outcomes.

Therefore, we modified the two proposed unweighted random-effects approaches [7, 8] using a normal distribution instead of a t distribution. Next, we followed the same

two steps outlined in aim 1 to compare the UW_OR with a t distribution (original version) vs. the UW_OR with a normal distribution (modified version), and to compare the UW_MM with a t distribution (original version) vs. the UW_MM with a normal distribution (modified version) (see Figure 3-4).

Aim 3

Evaluate efficiency for the UW_RR and the W_RR when they estimate same summary effect size in a meta-analysis of binomial trials.

Under each scenario, we carried out the following meta-analysis using both approaches (UW_RR vs. W_RR) as shown in Figure 3-5. The UW_RR selected in the comparison of the W_RR is based upon the following two reasons. First, although the UW_RR and the W_RR do not estimate the same summary effect size (unweighted vs. weighted), we forced them to estimate the same effect size by independently defining ET and EC for individual study and assumed that there is no association between the effect size and the sample size for each individual study. Second, they both avoid the log transformation on the individual effect size (Figure 2-6, 2-7). The steps are similar to those in aim 1.

At the end, for each scenario, we have two summarized measurements for each of the two approaches of interest (UW_RR vs. W_RR).

Aim 4

Evaluate efficiency of t approximation vs. normal approximation in a meta-analysis of binomial trials when synthesizing a small number of studies.

As we stated in Chapter II, when a small number of studies are combined in a meta-analysis, it is suggested that we assume a t distribution over a normal distribution for the summary effect size in a log scale for binomial outcomes. However, the impact of

the different distributions on the efficiency of random-effects approaches (UW_RR vs. W_RR) remains unknown when a meta-analysis contains a small number of studies with binomial outcomes.

Therefore, we modified the two random-effects approaches, the UW_RR and the W_RR, from a t distribution to a normal distribution. Next, we followed the same two steps in aim 1 to compare the UW_RR with a t distribution (original version) vs. the UW_RR with a normal distribution (modified version), and to compare the W_RR with a t distribution (original version) vs. the W_RR with a normal distribution (modified version) (see Figure 3-6).

In summary, the data simulations and statistical analyses are performed using the statistical SAS software (9.2, Cary). We expect that our research aims of interest can be achieved through this study. The results are reported in Chapter IV in the order of specific aims.

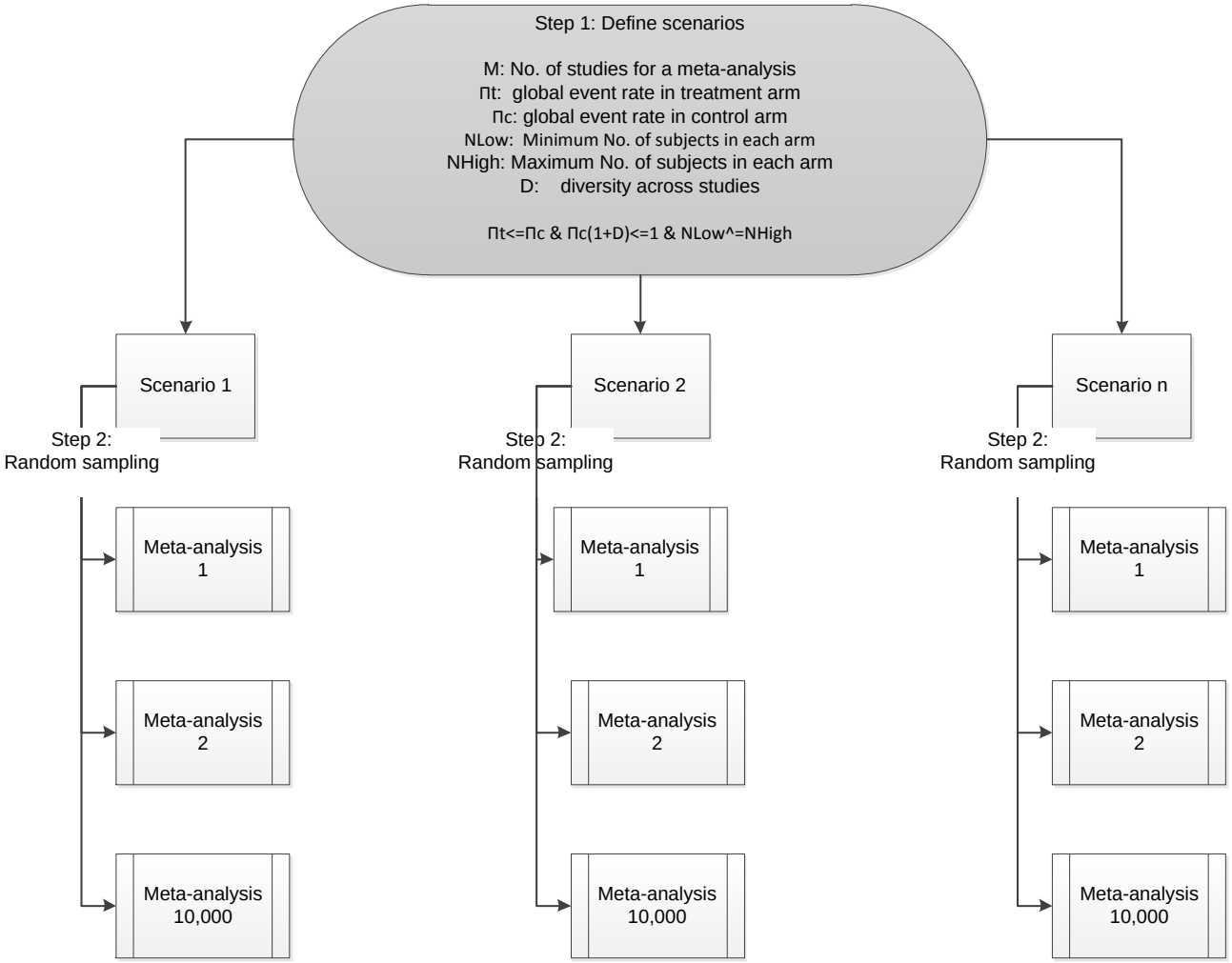

Figure 3-1. Odds ratio data simulation flow chart.

Table 3-1. Parameter list and values for OR data simulation

	Minimum Value	Maximum Value	Incremental Value	No. of Values
M	5	55	10	6
Пt	0.02	0.22	0.05	5
Пc	0.02	0.22	0.05	5
NLow in each arm	100	400	100	4
NHigh in each arm	100	500	100	5
D	0.1	0.2	0.3	3

Table 3-2. Examples with combination of parameters for scenarios

	M	Пt	Пc	NLow	NHigh	D
Scenario 1	5	0.2	0.2	100	200	0.1
Scenario 2	15	0.2	0.2	200	300	0.1
Scenario n	55	0.2	0.2	400	500	0.3

Table 3-3. Random sampling based upon parameters under a scenario

	Treatment event	Control event	Event rate in treatment arm	Event rate in control arm	Treatment N (=Control N)
Attributes	Random	Random	Random	Random	Random
Distribution	Binomial	Binomial	Uniform	Uniform	Uniform

Table 3-4. Example of studies structure for a random sampling in a meta-analysis

M	Treatment event	Control event	Treatment N	Control N
1	4	3	101	101
2	0	2	123	123
3	4	4	143	143
4	6	0	156	156
5	2	1	178	178

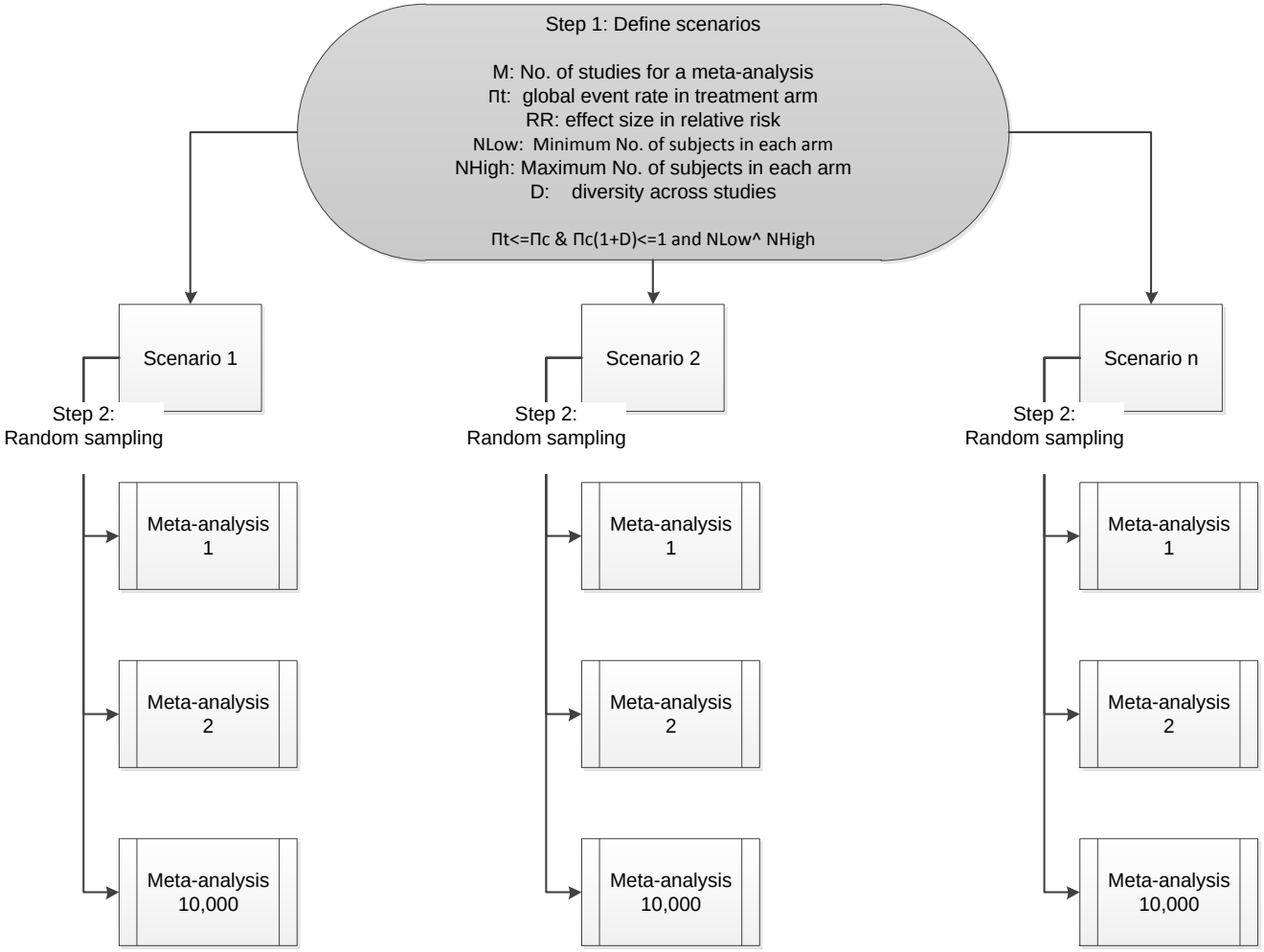

Figure 3-2. Relative risk data simulation flow chart.

Table 3-5. Parameter list and values for RR data simulation

	Minimum Value	Maximum Value	Incremental Value	No. of Values
M	5	55	10	6
Πt	0.02	0.22	0.05	5
RR	0.1	1	0.2	5
NLow in each arm	100	400	100	4
NHigh in each arm	100	500	100	5
D	0.1	0.2	0.3	3

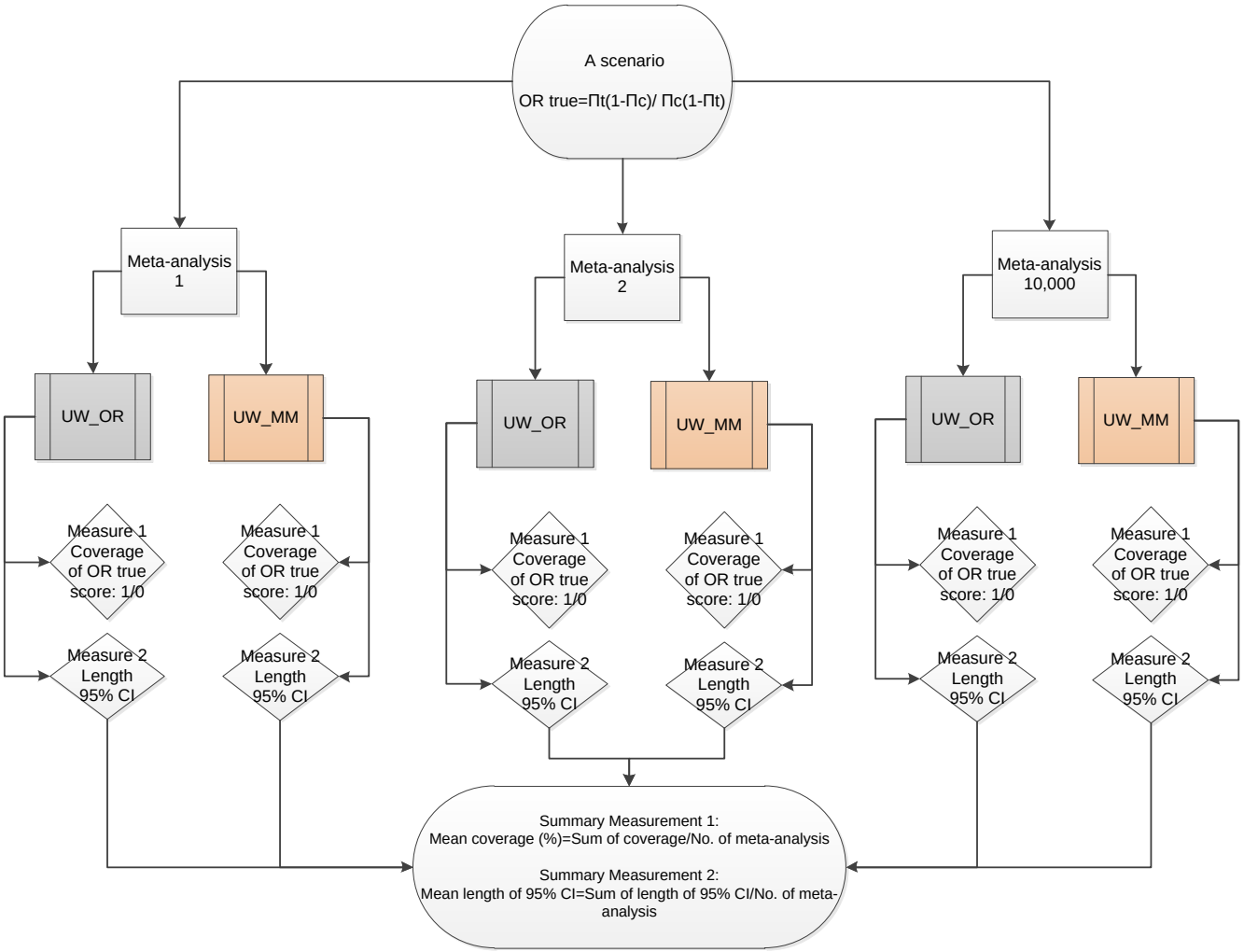

Figure 3-3. Illustration of analysis steps for aim 1

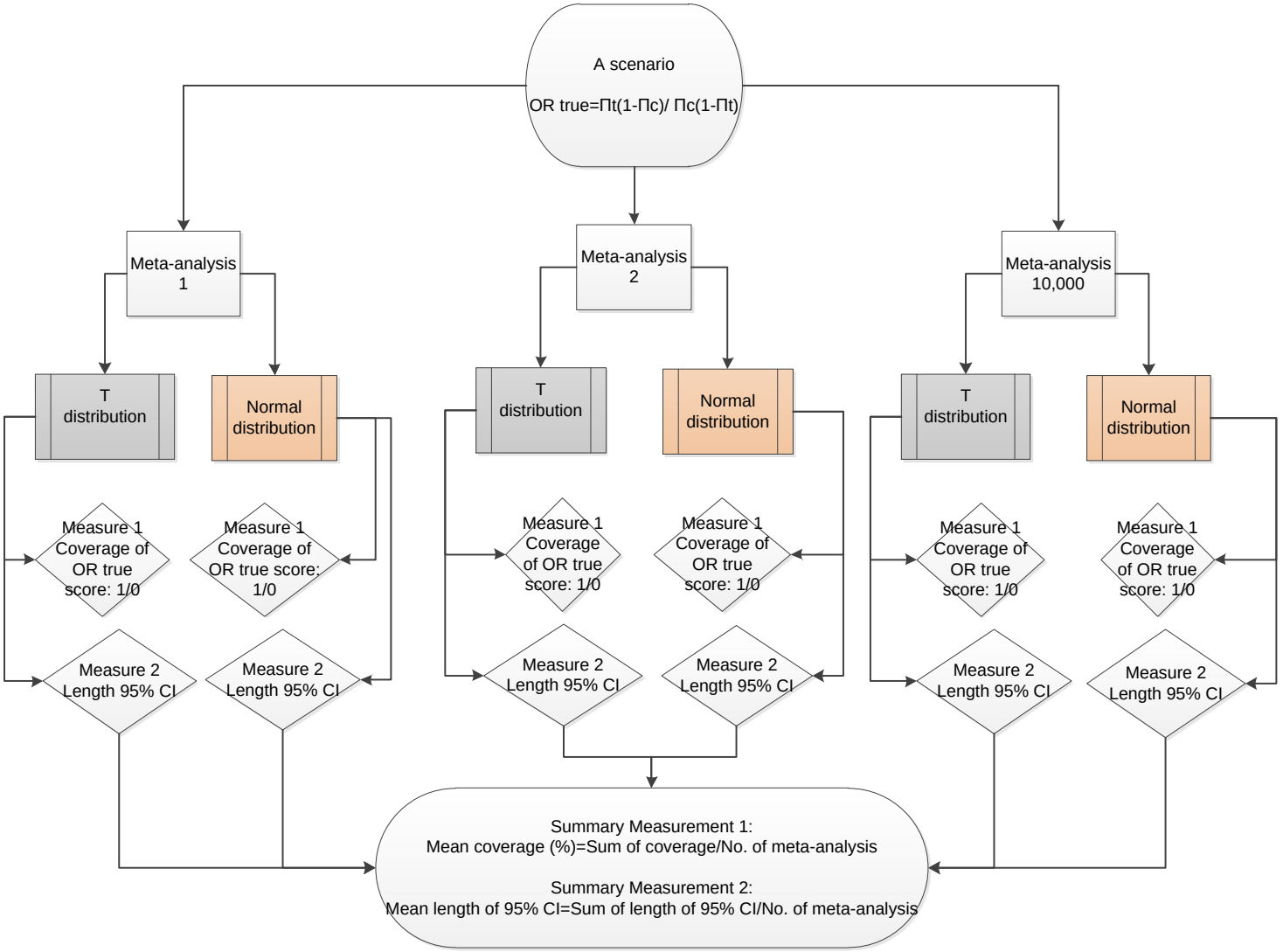

Figure 3-4. Illustration of analysis steps for aim 2

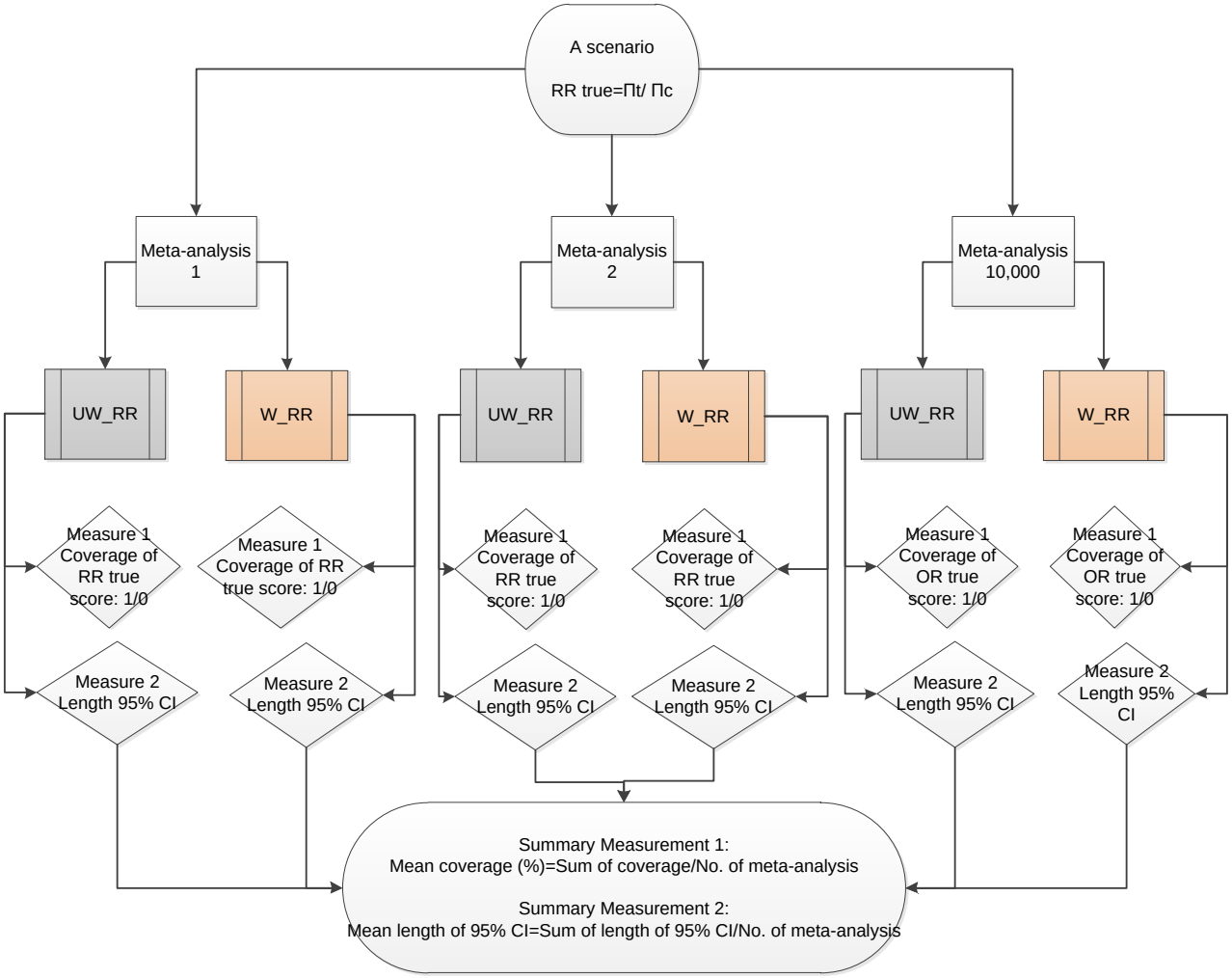

Figure 3-5. Illustration of analysis steps for aim 3

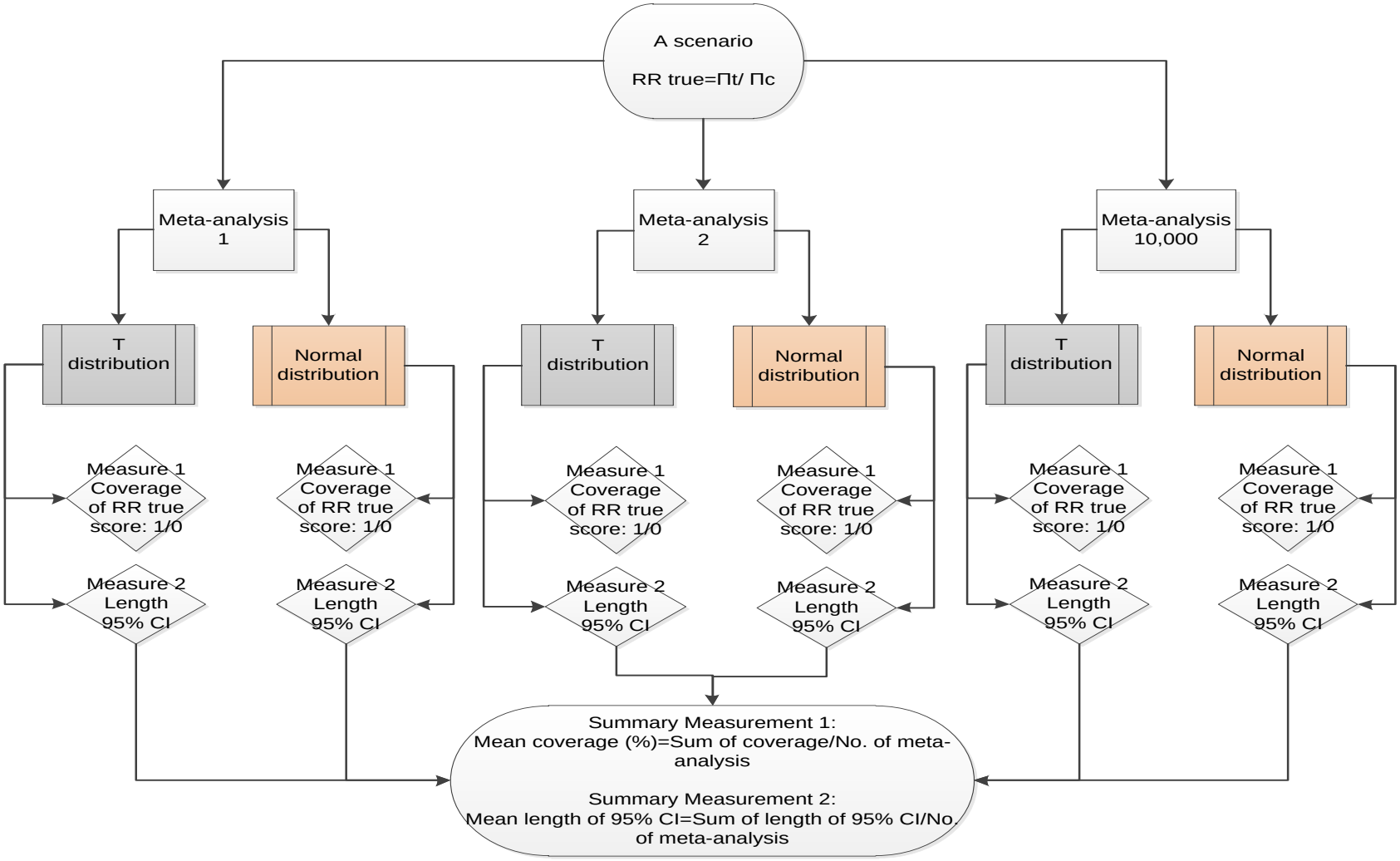

Figure 3-6. Illustration of analysis steps for aim 4

CHAPTER 4
RESULTS

Results are reported for each hypothesis under each aim. For each hypothesis, the results are presented separately as a primary analysis, a subgroup analysis stratified by the global true event rate in the treatment arm (PT/ Πt), and the number of studies (M). Furthermore, under each primary analysis and subgroup analysis, the results are presented in terms of the two forms of assessment as we explain in the next paragraph.

Under each scenario, the efficiency of a method is measured using two metrics: the mean coverage of the global true value by the purported 95% CI in percentage and the mean length of the 95% CI in a log scale. The combination of these two metrics helps to determine which method is more efficient. The closer the mean coverage is to 95% and the lower mean length of the 95% CI indicate a superior efficiency. If one method has a more accurate mean coverage and a wider mean length of the 95% CI or if one method has a less accurate mean coverage and a narrower mean length of the 95% CI, we consider first whether the mean coverage for the method is close to 95% coverage. If the mean coverage for both approaches is all close to 95%, then the narrower the mean length of the 95% CI, the better that approach is.

Two forms of figures are adopted to compare the efficiency of any two methods of interest. One form of figures is use the difference of the mean coverage for two comparators on the y axis and the difference of the mean length of the 95% CI for two comparators on the x axis. One red dot indicates one scenario. Red dots falling into the top left or bottom right quadrant indicate the superior efficiency of a method. The other figure is use the mean coverage on the y axis and the mean length of the 95% CI on the

x axis. Therefore, the blue and red dots differentiate the comparators of interest for each scenario.

In this section, concepts of the mean coverage or the mean length of the 95% CI are used within a scenario with a divisor of 10,000 instances of meta-analysis. Otherwise the concepts of the average of the mean coverage or the average of the mean length 95% CI are adopted for all scenarios with a divisor of the number of scenarios of interest.

As we mentioned in Chapter III, the methods section, the odds ratio simulation is performed for aim 1 and aim 2. A total of 6,300 scenarios are generated based upon the combination of six predetermined parameters, 2,700 of which meet the predefined criteria in the odds ratio simulation. Results of aim 1 and aim 2 are generated based upon these qualified scenarios.

At the same time, the relative risk simulation is performed for aim 3 and aim 4. A total of 6,300 scenarios are generated based upon the combination of six predetermined parameters, 4,200 of which meet the same predefined criteria as in the odds ratio simulation. The reason that the more scenarios, 4,200, meet the predefined criteria is that there are 24 values for the global true event probability in the control arm (Πc) in the relative risk simulation, but 5 values in Πc in the odds ratio simulation. The results of aim 3 and aim 4 are generated based upon these qualified scenarios.

Aim 1

Hypothesis 1: It is hypothesized that there is no precision difference between the UW_OR and the UW_MM when synthesizing a large number of studies.

The difference in the mean coverage is derived from the mean coverage of the UW_OR minus that of the UW_MM, which is represented on the y axis (Figure 4-1, 4-2). The difference of the mean length of the 95% CI is derived from the mean length of the 95% CI of the UW_OR minus that of the UW_MM, which is represented on the x axis (Figure 4-1, 4-2).

Primary analysis. We note that the majority of scenarios in the form of dots, 1,850 out of 2700 (68.52%), fall into the top left quadrant. It means that the UW_OR approach has a higher mean coverage and a narrower mean length of 95 % CI, indicating that the UW_OR is more efficient than the UW_MM (Figure 4-1). The average of the mean coverage of the UW_OR for scenarios is 95.1% (range 94.4% to 95.9%), whereas that of the UW_MM is 93.1% (range 62.0% to 96.2%) which is away to the expected 95% coverage (Table 4-1). Moreover, the average of the mean length of the 95% CI of the UW_OR is 0.369, which is shorter than that of the UW_MM, 0.400.

Subgroup analysis. When the number of studies is as small as five, the efficiency of both methods is close regardless of the event rate in the treatment arm (Figure 4-2, 4-4).

At low event rates in the treatment arm (PT<=0.12), as the number of studies (M) increases, the UW_OR is more efficient than the UW_MM in terms of the higher mean coverage difference between the UW_OR and the UW_MM (Figure 4-2), while the mean length of 95% CI for both approaches is close (Figure 4-2, Table 4-2). The reason for the higher mean coverage difference is that the mean coverage of the UW_MM drops as the number of studies increases, while that of the UW_OR is robust and close to the expected 95% coverage (Figure 4-4).

At high event rates in the treatment arm (PT>0.12), the UW_OR is as efficient as the UW_MM (Figure 4-2) regardless of the number of studies (M).

For both methods, the mean length of the 95% CI is shorter as the event rate in the treatment arm is increased or the number of studies is increased. The average of mean length of the 95% CI for the UW_OR and the UW_MM under all scenarios is dropped from 1.232 and 1.382 respectively at the low event of 0.02 in the treatment arm and the number of studies of five to 0.129 and 0.131 respectively at the high event rate of 0.22 in the treatment arm and the number of studies of 55 (Table 4-2).

Aim 2

Hypothesis 2-1: It is hypothesized that a t approximation is more accurate than a normal approximation in terms of closeness to the 95% coverage for the UW_OR when synthesizing a small number of studies.

The difference in the mean coverage is derived from the mean coverage of the UW_OR with a t distribution minus that of the UW_OR with a normal distribution, which is represented on the y axis (Figure 4-5, 4-6). The difference in the mean length of the 95% CI is derived from the mean length of the 95% CI from the UW_OR with a t distribution minus that of the UW_OR with a normal distribution, as represented on the x axis (Figure 4-5, 4-6).

Primary analysis. All 2,700 scenarios fall into the top right quadrant (Figure4-5). The average of the mean coverage of the UW_OR with a t distribution is 95.1% (range 94.4% to 95.9%) (Table 4-3), whereas the average of the mean coverage of the UW_OR with a normal distribution is 93.1% (range 87.1% to 95.2%) which is away from the expected 95% coverage (Table 4-3). The findings indicate that the assumption of a t

distribution is more accurate than the assumption of a normal distribution for the UW_OR.

Subgroup analysis. When the number of studies is as low as five, the average of the mean coverage of the UW_OR with a t distribution is 95.5% (range 95.1% to 95.9%), whereas the average of the mean coverage of the UW_OR with a normal distribution is as low as 88.6% (range 87.9% to 89.3%) (Figure 4-8, Table 4-4), which is far away from the expected 95% coverage.

As the number of studies (M) increases, the mean coverage for the UW_OR with a normal distribution is close to that of the UW_OR with a t distribution regardless of the event rate in the treatment arm (Figure 4-8), whereas the mean coverage of a t distribution is robust within the range of 94.4% to 95.9% regardless of the number of studies or the event rate in the treatment arm (Table 4-3, Figure 4-8).

The mean length of the 95% CI for both a t distribution and a normal distribution for the UW_OR is shorter when the number of studies increases or the event rate in the treatment arm increases (Figure 4-8). The average of the mean length of the 95% CI for the UW_OR with a t distribution and the UW_OR with a normal distribution is shorter from 1.232 and 0.870 respectively at the low event of 0.02 in the treatment arm and the number of studies of five to 0.129 and 0.126 respectively at the high event rate of 0.22 in the treatment arm and the number of studies of 55 (Table 4-4).

Hypothesis 2-2: It is hypothesized that a t approximation is more accurate than a normal approximation in terms of closeness to the 95% coverage for the UW_MM when synthesizing a small number of studies.

The difference of the mean coverage is derived from the mean coverage of the UW_MM with a t distribution minus that of the UW_MM with a normal distribution, which is represented on the y axis (Figure 4-9, 4-10). The difference of the mean length of the 95% CI is derived from the mean length of the 95% CI of the UW_MM with a t distribution minus that of the UW_MM with a normal distribution, which is represented on the x axis (Figure 4-9, 4-10).

Primary analysis. All 2,700 scenarios fall into the top right quadrant (Figure4-9). The average of the mean coverage of the UW_MM with a t distribution is 93.1% (range 62.0% to 96.2%), which is closer the expected 95% coverage, whereas that of the UW_MM with a normal distribution is 90.8% (range 60.4% to 95.2%), which is lower than the expected 95% coverage (Table 4-5). The findings indicate that the assumption of a t distribution is more accurate than the assumption of a normal distribution for the UW_MM.

Subgroup analysis. When the number of studies is as low as five, the average of the mean coverage for the UW_MM with a t distribution is 95.5% (range 94.9% to 96.2%), whereas that of the UW_MM with a normal distribution is 88.4% (range 87.4% to 89.5%) that is far away from the expected 95% coverage (Figure 4-12, Table 4-6).

As the number of studies (M) increases, the UW_MM with a normal distribution is close to the UW_MM with a t distribution regardless of the event rate in the treatment arm (Figure 4-12). At the lower event rate of 0.02 in the treatment arm, as the number of studies increases from 5 to 55, the average of the mean coverage of the UW_MM drops dramatically from 95.5% to 82.6% for the one with a t distribution and from 88.4% to 81.3% for the one with a normal distribution (Table 4-6). Whereas at the high event

rate of 0.22 in the treatment arm, as the number of studies increases from 5 to 55, the average of the mean coverage of the UW_MM stay stably from 95.3% to 95.1% for a t distribution and from 88.1% to 94.6% for a normal distribution (Table 4-6).

The mean length of the 95% CI for the UW_MM with a t distribution and the UW_MM with a normal distribution is shorter when the number of studies increases or the event rate in the treatment arm increases (Figure 4-12). The average of the mean length of the 95% CI of the UW_MM from 1.382 for a t distribution and 0.976 for a normal distribution at the low event of 0.02 in the treatment arm and the number of studies of five drop to 0.131 for a t distribution and 0.128 for a normal distribution at the high event rate of 0.22 in the treatment arm and number of studies of 55 (Table 4-6).

Aim 3

Hypothesis 3: It is hypothesized that efficiency of the UW_RR and the W_RR is similar when they estimate the same summary effect size.

The difference in the mean coverage is derived from the mean coverage of the UW_RR minus that of the W_RR, which is represented on the y axis (Figure 4-13, 4-14). The difference of the mean length of the 95% CI is derived from the mean length of the 95% CI from the UW_RR minus that of the W_RR, as is represented on the x axis (Figure 4-13, 4-14).

Primary analysis. The average of the mean coverage for both the UW_RR and the W_RR is closer to the expected 95% coverage, which is 95.1% (range 94.4% to 95.9%) vs. 95.0% (range 94.0% and 95.9%) (Table 4-7). However, the W_RR provides a narrower average of mean length of the 95% CI of 0.291 (range 0.058 to 1.976) as compared to 0.299 for the UW_RR (range 0.058 to 2.014) (Table 4-7). As we note, 3,062 out of 4,200 (72.90%) fall into the top right quadrant while 1,110 out of 4,200

(26.43%) fall into the bottom right quadrant (Figure 4-13), indicating that the UW_RR and the W_RR have similar efficiency in terms of the mean coverage, but the W_RR is more accurate as compared with the UW_RR in some scenarios.

Subgroup analysis. In regards to the mean coverage, as the number of studies increases, the mean coverage difference between the UW_RR and the W_RR decreases regardless of the event rate in the treatment arm (Figure 4-14), indicating that the efficiency of both approaches is closer as the number of studies increases. In addition, the mean length of the 95% CI for both the UW_RR and the W_RR is getting narrower as the number of studies increases or the event rate in the treatment arm increases (Figure 4-16).

Aim 4

Hypothesis 4-1: It is hypothesized that a t approximation is more accurate than a normal approximation in terms of closeness to the 95% coverage for the UW_RR when synthesizing a small number of studies.

The difference of the mean coverage is derived from the mean coverage of the UW_RR with a t distribution minus that of the UW_RR with a normal distribution, as represented on the y axis (Figure 4-17, 4-18). The difference of the mean length of the 95% CI is derived from the mean length of the 95% CI from the UW_RR with a t distribution minus that of the UW_RR with a normal distribution, which is represented on the x axis (Figure 4-17, 4-18).

Primary analysis. All 4,200 scenarios fall into the top right quadrant (Figure4-17). The average of the mean coverage of the UW_RR with a t distribution is 95.1% (range 94.4% to 95.9%), whereas that of the UW_RR with a normal distribution is 93.1% (range 87.0% to 95.2%), which is lower than the expected 95% coverage (Table 4-9).

The findings indicate that the assumption of a t distribution is more accurate than the assumption of a normal distribution for the UW_RR.

Subgroup analysis. When the number of studies is as low as five, the difference of the mean coverage between the UW_RR with a t distribution and the UW_RR with a normal distribution is large (Figure 4-18). The average of the mean coverage of the UW_RR with a t distribution is 95.5% (range 95.1% to 95.9%), whereas the average of the mean coverage of the UW_RR with a normal distribution is 88.6% (range 87.8% to 89.4%), which is far from the expected 95% coverage (Table 4-10).

As the number of studies (M) increases, the mean coverage for the UW_RR with a normal distribution is close to that of the UW_RR with a t distribution regardless of the event rate in the treatment arm (Figure 4-20). Whereas the mean coverage of the UW_RR with a t distribution is robust within the range of 94.4% to 95.9% regardless of the number of studies or the event rate in the treatment arm (Table 4-9, Figure 4-20).

The mean length of the 95% CI for the UW_RR with both a t distribution and a normal distribution is shorter when the number of studies increases or the event rate in the treatment arm increases (Figure 4-20). The average of the mean length of the 95% CI for the UW_RR with a t distribution and with a normal distribution is shorter from 1.312 and 0.926 respectively at the low event of 0.02 in the treatment arm and the number of studies of five to 0.092 and 0.090 respectively at the high event rate of 0.22 in the treatment arm and the number of studies of 55 (Table 4-10).

Hypothesis 4-2: It is hypothesized that a t approximation is more accurate than a normal approximation in terms of closeness to the 95% coverage for the W_RR when synthesizing a small number of studies.

The difference in the mean coverage is derived from the mean coverage of the W_RR with a t distribution minus that of the W_RR with a normal distribution, as represented on the y axis (Figure 4-21, 4-22). The difference in the mean length of the 95% CI is derived from the mean length of the 95% CI from the W_RR with a t distribution minus that of the W_RR with a normal distribution, as represented on the x axis (Figure 4-21, 4-22).

Primary analysis. All 4,200 scenarios fall into the top right quadrant (Figure4-21). The average of the mean coverage of the W_RR with a t distribution is 95.0% (range 94.0% to 95.9%), whereas that of the W_RR with a normal distribution is 92.9% (range 86.0% to 95.2%), which is lower than the expected 95% coverage (Table 4-11). The findings indicate that the assumption of a t distribution is more accurate than the assumption of a normal distribution for the W_RR.

Subgroup analysis. When the number of studies is as low as five, the difference in the mean coverage between the W_RR with a t distribution and the W_RR with a normal distribution is large (Figure 4-22). The average of the mean coverage of the W_RR with a t distribution is 95.2% (range 94.4% to 95.9%), whereas the average of the mean coverage of the W_RR with a normal distribution is 88.2% (range 86.8% to 89.2%), which is far away from the expected 95% coverage (Table 4-12).

As the number of studies (M) increases, the W_RR with a normal distribution is close to the W_RR with a t distribution regardless of the event rate in the treatment arm (Figure 4-24). Whereas the mean coverage for the W_RR with a t distribution is robust within the range of 94.0% to 95.9% regardless of the number of studies or the event rate in the treatment arm (Table 4-11, Figure 4-24).

The mean length of the 95% CI for the W_RR with both a t distribution and a normal distribution is shorter when the number of studies increases or the event rate in the treatment arm increases (Figure 4-24). The average of the mean length of the 95% CI for the W_RR with a t distribution and with a normal distribution decreases from 1.273 and 0.899, respectively at the low event of 0.02 in the treatment arm and the number of studies of five to 0.091 and 0.089, respectively at the high event rate of 0.22 in the treatment arm with the number of studies of 55 (Table 4-12).

In summary, through the odds ratio simulation, we find that the efficiency of the UW_OR and the UW_MM is close as long as no rare event occurs, suggesting that we cannot reject the null hypothesis for aim 1. In addition, when a small number of studies are combined for a meta-analysis, both the UW_OR and the UW_MM show that the assumption with a t distribution is more accurate than the assumption with a normal distribution in terms of closeness to the 95% coverage, which suggests that we cannot reject the null hypotheses for aim 2.

In regards to the relative risk simulation, we find that when the UW_RR and the W_RR estimate the same summary effect size, both of which are efficient in terms of closeness to the 95% coverage, but the W_RR is marginally precise in terms of the mean length of the 95% CI, suggesting that we cannot reject the null hypothesis for aim 3. Moreover, when a small number of studies are combined for a meta-analysis, both the UW_RR and the W_RR show that the assumption with a t distribution that is more accurate than the assumption with a normal distribution in terms of closeness to the 95% coverage. It suggests that we cannot reject the null hypotheses for aim 4.

In next Chapter V, we discuss our findings and related issues in more depth.

Table 4-1. Aim 1: Comparisons of UW_OR vs. UW_MM

Outcomes	Mean	STD	Minimum	Maximum
Mean coverage for UW_OR	0.951	0.002	0.944	0.959
Mean coverage for UW_MM	0.931	0.047	0.62	0.962
Mean 95% CI length for UW_OR	0.369	0.279	0.091	2.053
Mean 95% CI length for UW_MM	0.400	0.312	0.092	2.289

Table 4-2. Aim 1: Comparisons of UW_OR vs. UW_MM by event rate in treatment arm and number of studies

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.02	5	Mean coverage for UW_OR	150	0.955	0.002	0.951	0.959
		Mean coverage for UW_MM	150	0.955	0.002	0.949	0.962
		Mean 95% CI length for UW_OR	150	1.232	0.251	0.856	2.053
		Mean 95% CI length for UW_MM	150	1.382	0.278	0.955	2.289
	15	Mean coverage for UW_OR	150	0.952	0.002	0.947	0.957
		Mean coverage for UW_MM	150	0.939	0.013	0.902	0.957
		Mean 95% CI length for UW_OR	150	0.567	0.115	0.394	0.937
		Mean 95% CI length for UW_MM	150	0.656	0.129	0.457	1.071
	25	Mean coverage for UW_OR	150	0.950	0.002	0.946	0.955
		Mean coverage for UW_MM	150	0.911	0.031	0.826	0.957
		Mean 95% CI length for UW_OR	150	0.425	0.086	0.295	0.701
		Mean 95% CI length for UW_MM	150	0.495	0.096	0.347	0.803
	35	Mean coverage for UW_OR	150	0.950	0.002	0.946	0.956
		Mean coverage for UW_MM	150	0.885	0.049	0.760	0.958
		Mean 95% CI length for UW_OR	150	0.354	0.072	0.246	0.583
		Mean 95% CI length for UW_MM	150	0.413	0.079	0.290	0.669
	45	Mean coverage for UW_OR	150	0.950	0.002	0.945	0.955
		Mean coverage for UW_MM	150	0.856	0.067	0.689	0.955
		Mean 95% CI length for UW_OR	150	0.310	0.063	0.215	0.510
		Mean 95% CI length for UW_MM	150	0.363	0.069	0.255	0.586
	55	Mean coverage for UW_OR	150	0.951	0.002	0.944	0.957
		Mean coverage for UW_MM	150	0.826	0.087	0.620	0.957
		Mean 95% CI length for UW_OR	150	0.279	0.056	0.194	0.459
		Mean 95% CI length for UW_MM	150	0.327	0.062	0.230	0.528

Table 4-2. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.07	5	Mean coverage for UW_OR	120	0.953	0.003	0.946	0.958
		Mean coverage for UW_MM	120	0.954	0.003	0.945	0.959
		Mean 95% CI length for UW_OR	120	0.757	0.130	0.521	1.129
		Mean 95% CI length for UW_MM	120	0.784	0.144	0.531	1.208
	15	Mean coverage for UW_OR	120	0.952	0.002	0.947	0.956
		Mean coverage for UW_MM	120	0.950	0.002	0.946	0.956
		Mean 95% CI length for UW_OR	120	0.352	0.061	0.243	0.525
		Mean 95% CI length for UW_MM	120	0.368	0.068	0.249	0.570
	25	Mean coverage for UW_OR	120	0.951	0.001	0.948	0.954
		Mean coverage for UW_MM	120	0.947	0.004	0.936	0.952
		Mean 95% CI length for UW_OR	120	0.264	0.045	0.183	0.394
		Mean 95% CI length for UW_MM	120	0.277	0.051	0.188	0.428
	35	Mean coverage for UW_OR	120	0.950	0.001	0.948	0.955
		Mean coverage for UW_MM	120	0.944	0.007	0.924	0.953
		Mean 95% CI length for UW_OR	120	0.221	0.038	0.153	0.328
		Mean 95% CI length for UW_MM	120	0.231	0.043	0.157	0.357
	45	Mean coverage for UW_OR	120	0.950	0.003	0.944	0.955
		Mean coverage for UW_MM	120	0.942	0.009	0.911	0.955
		Mean 95% CI length for UW_OR	120	0.193	0.033	0.134	0.287
		Mean 95% CI length for UW_MM	120	0.203	0.038	0.137	0.313
	55	Mean coverage for UW_OR	120	0.951	0.002	0.945	0.953
		Mean coverage for UW_MM	120	0.939	0.012	0.902	0.954
		Mean 95% CI length for UW_OR	120	0.174	0.030	0.120	0.259
		Mean 95% CI length for UW_MM	120	0.183	0.034	0.124	0.282

Table 4-2. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.12	5	Mean coverage for UW_OR	90	0.953	0.003	0.945	0.957
		Mean coverage for UW_MM	90	0.953	0.002	0.945	0.957
		Mean 95% CI length for UW_OR	90	0.642	0.103	0.446	0.910
		Mean 95% CI length for UW_MM	90	0.654	0.109	0.451	0.942
	15	Mean coverage for UW_OR	90	0.952	0.002	0.948	0.956
		Mean coverage for UW_MM	90	0.951	0.002	0.948	0.956
		Mean 95% CI length for UW_OR	90	0.299	0.048	0.209	0.424
		Mean 95% CI length for UW_MM	90	0.306	0.051	0.211	0.442
	25	Mean coverage for UW_OR	90	0.951	0.001	0.948	0.953
		Mean coverage for UW_MM	90	0.950	0.001	0.947	0.953
		Mean 95% CI length for UW_OR	90	0.225	0.036	0.157	0.318
		Mean 95% CI length for UW_MM	90	0.230	0.039	0.159	0.331
	35	Mean coverage for UW_OR	90	0.950	0.001	0.948	0.954
		Mean coverage for UW_MM	90	0.949	0.002	0.943	0.953
		Mean 95% CI length for UW_OR	90	0.187	0.030	0.131	0.265
		Mean 95% CI length for UW_MM	90	0.192	0.032	0.133	0.277
	45	Mean coverage for UW_OR	90	0.950	0.003	0.945	0.955
		Mean coverage for UW_MM	90	0.948	0.004	0.939	0.955
		Mean 95% CI length for UW_OR	90	0.164	0.026	0.115	0.232
		Mean 95% CI length for UW_MM	90	0.168	0.028	0.116	0.242
	55	Mean coverage for UW_OR	90	0.951	0.002	0.946	0.955
		Mean coverage for UW_MM	90	0.949	0.003	0.939	0.955
		Mean 95% CI length for UW_OR	90	0.148	0.024	0.103	0.209
		Mean 95% CI length for UW_MM	90	0.152	0.025	0.105	0.218

Table 4-2. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.17	5	Mean coverage for UW_OR	60	0.953	0.002	0.947	0.957
		Mean coverage for UW_MM	60	0.953	0.002	0.947	0.957
		Mean 95% CI length for UW_OR	60	0.588	0.093	0.413	0.811
		Mean 95% CI length for UW_MM	60	0.596	0.096	0.416	0.829
	15	Mean coverage for UW_OR	60	0.951	0.002	0.948	0.956
		Mean coverage for UW_MM	60	0.951	0.002	0.948	0.955
		Mean 95% CI length for UW_OR	60	0.274	0.044	0.193	0.378
		Mean 95% CI length for UW_MM	60	0.279	0.045	0.195	0.388
	25	Mean coverage for UW_OR	60	0.950	0.001	0.947	0.953
		Mean coverage for UW_MM	60	0.950	0.001	0.946	0.953
		Mean 95% CI length for UW_OR	60	0.206	0.033	0.145	0.284
		Mean 95% CI length for UW_MM	60	0.209	0.034	0.146	0.291
	35	Mean coverage for UW_OR	60	0.950	0.001	0.948	0.954
		Mean coverage for UW_MM	60	0.950	0.001	0.947	0.953
		Mean 95% CI length for UW_OR	60	0.172	0.027	0.121	0.236
		Mean 95% CI length for UW_MM	60	0.175	0.028	0.122	0.243
	45	Mean coverage for UW_OR	60	0.950	0.002	0.946	0.954
		Mean coverage for UW_MM	60	0.950	0.003	0.944	0.954
		Mean 95% CI length for UW_OR	60	0.151	0.024	0.106	0.207
		Mean 95% CI length for UW_MM	60	0.153	0.025	0.107	0.213
	55	Mean coverage for UW_OR	60	0.951	0.002	0.946	0.954
		Mean coverage for UW_MM	60	0.951	0.002	0.946	0.955
		Mean 95% CI length for UW_OR	60	0.136	0.021	0.096	0.187
		Mean 95% CI length for UW_MM	60	0.138	0.022	0.096	0.192

Table 4-2. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.22	5	Mean coverage for UW_OR	30	0.953	0.002	0.948	0.956
		Mean coverage for UW_MM	30	0.953	0.002	0.948	0.956
		Mean 95% CI length for UW_OR	30	0.558	0.089	0.395	0.759
		Mean 95% CI length for UW_MM	30	0.564	0.092	0.397	0.772
	15	Mean coverage for UW_OR	30	0.951	0.002	0.948	0.954
		Mean coverage for UW_MM	30	0.951	0.002	0.948	0.954
		Mean 95% CI length for UW_OR	30	0.261	0.042	0.185	0.354
		Mean 95% CI length for UW_MM	30	0.264	0.043	0.186	0.361
	25	Mean coverage for UW_OR	30	0.951	0.002	0.946	0.953
		Mean coverage for UW_MM	30	0.950	0.002	0.946	0.953
		Mean 95% CI length for UW_OR	30	0.196	0.031	0.139	0.266
		Mean 95% CI length for UW_MM	30	0.198	0.032	0.140	0.271
	35	Mean coverage for UW_OR	30	0.950	0.001	0.948	0.953
		Mean coverage for UW_MM	30	0.950	0.001	0.948	0.954
		Mean 95% CI length for UW_OR	30	0.163	0.026	0.116	0.222
		Mean 95% CI length for UW_MM	30	0.165	0.027	0.117	0.226
	45	Mean coverage for UW_OR	30	0.950	0.002	0.946	0.954
		Mean coverage for UW_MM	30	0.950	0.002	0.947	0.954
		Mean 95% CI length for UW_OR	30	0.143	0.023	0.102	0.194
		Mean 95% CI length for UW_MM	30	0.145	0.024	0.102	0.198
	55	Mean coverage for UW_OR	30	0.951	0.002	0.946	0.955
		Mean coverage for UW_MM	30	0.951	0.002	0.948	0.955
		Mean 95% CI length for UW_OR	30	0.129	0.021	0.091	0.175
		Mean 95% CI length for UW_MM	30	0.131	0.021	0.092	0.179

Table 4-3. Aim 2-1: Comparisons of t distribution vs. normal distribution for UW_OR

Outcomes	Mean	STD	Minimum	Maximum
Mean coverage for normal distribution	0.931	0.022	0.871	0.952
Mean coverage for t distribution	0.951	0.002	0.944	0.959
Mean 95% CI length for normal distribution	0.315	0.191	0.089	1.449
Mean 95% CI length for t distribution	0.369	0.279	0.091	2.053

Table 4-4. Aim 2-1: Comparisons of t distribution vs. normal distribution for UW_OR by event rate in treatment arm and number of studies

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.02	5	Mean coverage for normal distribution	150	0.886	0.003	0.879	0.893
		Mean coverage for t distribution	150	0.955	0.002	0.951	0.959
		Mean 95% CI length for normal distribution	150	0.870	0.177	0.604	1.449
		Mean 95% CI length for t distribution	150	1.232	0.251	0.856	2.053
	15	Mean coverage for normal distribution	150	0.932	0.002	0.925	0.937
		Mean coverage for t distribution	150	0.952	0.002	0.947	0.957
		Mean 95% CI length for normal distribution	150	0.518	0.105	0.360	0.856
		Mean 95% CI length for t distribution	150	0.567	0.115	0.394	0.937
	25	Mean coverage for normal distribution	150	0.939	0.002	0.933	0.943
		Mean coverage for t distribution	150	0.950	0.002	0.946	0.955
		Mean 95% CI length for normal distribution	150	0.403	0.082	0.280	0.665
		Mean 95% CI length for t distribution	150	0.425	0.086	0.295	0.701
	35	Mean coverage for normal distribution	150	0.941	0.002	0.938	0.949
		Mean coverage for t distribution	150	0.950	0.002	0.946	0.956
		Mean 95% CI length for normal distribution	150	0.341	0.069	0.237	0.562
		Mean 95% CI length for t distribution	150	0.354	0.072	0.246	0.583
	45	Mean coverage for normal distribution	150	0.943	0.002	0.939	0.949
		Mean coverage for t distribution	150	0.950	0.002	0.945	0.955
		Mean 95% CI length for normal distribution	150	0.301	0.061	0.209	0.496
		Mean 95% CI length for t distribution	150	0.310	0.063	0.215	0.510
	55	Mean coverage for normal distribution	150	0.945	0.002	0.939	0.952
		Mean coverage for t distribution	150	0.951	0.002	0.944	0.957
		Mean 95% CI length for normal distribution	150	0.273	0.055	0.189	0.449
		Mean 95% CI length for t distribution	150	0.279	0.056	0.194	0.459

Table 4-4. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.07	5	Mean coverage for normal distribution	120	0.882	0.004	0.873	0.891
		Mean coverage for t distribution	120	0.953	0.003	0.946	0.958
		Mean 95% CI length for normal distribution	120	0.534	0.092	0.368	0.797
		Mean 95% CI length for t distribution	120	0.757	0.130	0.521	1.129
	15	Mean coverage for normal distribution	120	0.931	0.002	0.927	0.936
		Mean coverage for t distribution	120	0.952	0.002	0.947	0.956
		Mean 95% CI length for normal distribution	120	0.322	0.056	0.222	0.480
		Mean 95% CI length for t distribution	120	0.352	0.061	0.243	0.525
	25	Mean coverage for normal distribution	120	0.939	0.001	0.936	0.942
		Mean coverage for t distribution	120	0.951	0.001	0.948	0.954
		Mean 95% CI length for normal distribution	120	0.251	0.043	0.174	0.374
		Mean 95% CI length for t distribution	120	0.264	0.045	0.183	0.394
	35	Mean coverage for normal distribution	120	0.942	0.002	0.938	0.947
		Mean coverage for t distribution	120	0.950	0.001	0.948	0.955
		Mean 95% CI length for normal distribution	120	0.213	0.036	0.147	0.316
		Mean 95% CI length for t distribution	120	0.221	0.038	0.153	0.328
	45	Mean coverage for normal distribution	120	0.944	0.003	0.938	0.949
		Mean coverage for t distribution	120	0.950	0.003	0.944	0.955
		Mean 95% CI length for normal distribution	120	0.188	0.032	0.130	0.279
		Mean 95% CI length for t distribution	120	0.193	0.033	0.134	0.287
	55	Mean coverage for normal distribution	120	0.946	0.002	0.940	0.949
		Mean coverage for t distribution	120	0.951	0.002	0.945	0.953
		Mean 95% CI length for normal distribution	120	0.170	0.029	0.118	0.253
		Mean 95% CI length for t distribution	120	0.174	0.030	0.120	0.259

Table 4-4. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.12	5	Mean coverage for normal distribution	90	0.881	0.003	0.872	0.890
		Mean coverage for t distribution	90	0.953	0.003	0.945	0.957
		Mean 95% CI length for normal distribution	90	0.453	0.073	0.315	0.642
		Mean 95% CI length for t distribution	90	0.642	0.103	0.446	0.910
	15	Mean coverage for normal distribution	90	0.931	0.002	0.927	0.935
		Mean coverage for t distribution	90	0.952	0.002	0.948	0.956
		Mean 95% CI length for normal distribution	90	0.273	0.044	0.191	0.387
		Mean 95% CI length for t distribution	90	0.299	0.048	0.209	0.424
	25	Mean coverage for normal distribution	90	0.939	0.001	0.936	0.942
		Mean coverage for t distribution	90	0.951	0.001	0.948	0.953
		Mean 95% CI length for normal distribution	90	0.213	0.034	0.149	0.302
		Mean 95% CI length for t distribution	90	0.225	0.036	0.157	0.318
	35	Mean coverage for normal distribution	90	0.942	0.002	0.939	0.945
		Mean coverage for t distribution	90	0.950	0.001	0.948	0.954
		Mean 95% CI length for normal distribution	90	0.181	0.029	0.126	0.256
		Mean 95% CI length for t distribution	90	0.187	0.030	0.131	0.265
	45	Mean coverage for normal distribution	90	0.944	0.003	0.938	0.949
		Mean coverage for t distribution	90	0.950	0.003	0.945	0.955
		Mean 95% CI length for normal distribution	90	0.160	0.026	0.111	0.226
		Mean 95% CI length for t distribution	90	0.164	0.026	0.115	0.232
	55	Mean coverage for normal distribution	90	0.946	0.002	0.941	0.949
		Mean coverage for t distribution	90	0.951	0.002	0.946	0.955
		Mean 95% CI length for normal distribution	90	0.145	0.023	0.101	0.204
		Mean 95% CI length for t distribution	90	0.148	0.024	0.103	0.209

Table 4-4. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.12	5	Mean coverage for normal distribution	90	0.881	0.003	0.872	0.890
		Mean coverage for t distribution	90	0.953	0.003	0.945	0.957
		Mean 95% CI length for normal distribution	90	0.453	0.073	0.315	0.642
		Mean 95% CI length for t distribution	90	0.642	0.103	0.446	0.910
	15	Mean coverage for normal distribution	90	0.931	0.002	0.927	0.935
		Mean coverage for t distribution	90	0.952	0.002	0.948	0.956
		Mean 95% CI length for normal distribution	90	0.273	0.044	0.191	0.387
		Mean 95% CI length for t distribution	90	0.299	0.048	0.209	0.424
	25	Mean coverage for normal distribution	90	0.939	0.001	0.936	0.942
		Mean coverage for t distribution	90	0.951	0.001	0.948	0.953
		Mean 95% CI length for normal distribution	90	0.213	0.034	0.149	0.302
		Mean 95% CI length for t distribution	90	0.225	0.036	0.157	0.318
	35	Mean coverage for normal distribution	90	0.942	0.002	0.939	0.945
		Mean coverage for t distribution	90	0.950	0.001	0.948	0.954
		Mean 95% CI length for normal distribution	90	0.181	0.029	0.126	0.256
		Mean 95% CI length for t distribution	90	0.187	0.030	0.131	0.265
	45	Mean coverage for normal distribution	90	0.944	0.003	0.938	0.949
		Mean coverage for t distribution	90	0.950	0.003	0.945	0.955
		Mean 95% CI length for normal distribution	90	0.160	0.026	0.111	0.226
		Mean 95% CI length for t distribution	90	0.164	0.026	0.115	0.232
	55	Mean coverage for normal distribution	90	0.946	0.002	0.941	0.949
		Mean coverage for t distribution	90	0.951	0.002	0.946	0.955
		Mean 95% CI length for normal distribution	90	0.145	0.023	0.101	0.204
		Mean 95% CI length for t distribution	90	0.148	0.024	0.103	0.209

Table 4-4. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.17	5	Mean coverage for normal distribution	60	0.881	0.004	0.871	0.888
		Mean coverage for t distribution	60	0.953	0.002	0.947	0.957
		Mean 95% CI length for normal distribution	60	0.415	0.066	0.292	0.572
		Mean 95% CI length for t distribution	60	0.588	0.093	0.413	0.811
	15	Mean coverage for normal distribution	60	0.931	0.002	0.927	0.934
		Mean coverage for t distribution	60	0.951	0.002	0.948	0.956
		Mean 95% CI length for normal distribution	60	0.251	0.040	0.176	0.346
		Mean 95% CI length for t distribution	60	0.274	0.044	0.193	0.378
	25	Mean coverage for normal distribution	60	0.939	0.001	0.935	0.942
		Mean coverage for t distribution	60	0.950	0.001	0.947	0.953
		Mean 95% CI length for normal distribution	60	0.196	0.031	0.138	0.269
		Mean 95% CI length for t distribution	60	0.206	0.033	0.145	0.284
	35	Mean coverage for normal distribution	60	0.942	0.001	0.939	0.945
		Mean coverage for t distribution	60	0.950	0.001	0.948	0.954
		Mean 95% CI length for normal distribution	60	0.166	0.026	0.117	0.228
		Mean 95% CI length for t distribution	60	0.172	0.027	0.121	0.236
	45	Mean coverage for normal distribution	60	0.944	0.003	0.938	0.949
		Mean coverage for t distribution	60	0.950	0.002	0.946	0.954
		Mean 95% CI length for normal distribution	60	0.146	0.023	0.103	0.202
		Mean 95% CI length for t distribution	60	0.151	0.024	0.106	0.207
	55	Mean coverage for normal distribution	60	0.946	0.002	0.941	0.949
		Mean coverage for t distribution	60	0.951	0.002	0.946	0.954
		Mean 95% CI length for normal distribution	60	0.133	0.021	0.093	0.182
		Mean 95% CI length for t distribution	60	0.136	0.021	0.096	0.187

Table 4-4. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.22	5	Mean coverage for normal distribution	30	0.881	0.003	0.873	0.887
		Mean coverage for t distribution	30	0.953	0.002	0.948	0.956
		Mean 95% CI length for normal distribution	30	0.394	0.063	0.279	0.536
		Mean 95% CI length for t distribution	30	0.558	0.089	0.395	0.759
	15	Mean coverage for normal distribution	30	0.931	0.002	0.927	0.934
		Mean coverage for t distribution	30	0.951	0.002	0.948	0.954
		Mean 95% CI length for normal distribution	30	0.238	0.038	0.169	0.324
		Mean 95% CI length for t distribution	30	0.261	0.042	0.185	0.354
	25	Mean coverage for normal distribution	30	0.939	0.002	0.935	0.942
		Mean coverage for t distribution	30	0.951	0.002	0.946	0.953
		Mean 95% CI length for normal distribution	30	0.186	0.030	0.132	0.252
		Mean 95% CI length for t distribution	30	0.196	0.031	0.139	0.266
	35	Mean coverage for normal distribution	30	0.942	0.001	0.939	0.944
		Mean coverage for t distribution	30	0.950	0.001	0.948	0.953
		Mean 95% CI length for normal distribution	30	0.157	0.025	0.112	0.214
		Mean 95% CI length for t distribution	30	0.163	0.026	0.116	0.222
	45	Mean coverage for normal distribution	30	0.944	0.002	0.939	0.948
		Mean coverage for t distribution	30	0.950	0.002	0.946	0.954
		Mean 95% CI length for normal distribution	30	0.139	0.022	0.099	0.189
		Mean 95% CI length for t distribution	30	0.143	0.023	0.102	0.194
	55	Mean coverage for normal distribution	30	0.946	0.002	0.941	0.950
		Mean coverage for t distribution	30	0.951	0.002	0.946	0.955
		Mean 95% CI length for normal distribution	30	0.126	0.020	0.089	0.171
		Mean 95% CI length for t distribution	30	0.129	0.021	0.091	0.175

Table 4-5. Aim 2-2: Comparisons of t distribution vs. normal distribution for UW_MM

Outcomes	Mean	STD	Minimum	Maximum
Mean coverage for normal distribution	0.908	0.049	0.604	0.952
Mean coverage for t distribution	0.931	0.047	0.620	0.962
Mean 95% CI length for normal distribution	0.343	0.217	0.090	1.616
Mean 95% CI length for t distribution	0.400	0.312	0.092	2.289

Table 4-6. Aim 2-2: Comparisons of t distribution vs. normal distribution for UW_MM by event rate in treatment arm and number of studies

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.02	5	Mean coverage for normal distribution	150	0.884	0.004	0.874	0.895
		Mean coverage for t distribution	150	0.955	0.002	0.949	0.962
		Mean 95% CI length for normal distribution	150	0.976	0.196	0.674	1.616
		Mean 95% CI length for t distribution	150	1.382	0.278	0.955	2.289
	15	Mean coverage for normal distribution	150	0.912	0.017	0.862	0.938
		Mean coverage for t distribution	150	0.939	0.013	0.902	0.957
		Mean 95% CI length for normal distribution	150	0.599	0.118	0.417	0.979
		Mean 95% CI length for t distribution	150	0.656	0.129	0.457	1.071
	25	Mean coverage for normal distribution	150	0.892	0.036	0.795	0.943
		Mean coverage for t distribution	150	0.911	0.031	0.826	0.957
		Mean 95% CI length for normal distribution	150	0.470	0.091	0.329	0.762
		Mean 95% CI length for t distribution	150	0.495	0.096	0.347	0.803
	35	Mean coverage for normal distribution	150	0.869	0.053	0.735	0.951
		Mean coverage for t distribution	150	0.885	0.049	0.760	0.958
		Mean 95% CI length for normal distribution	150	0.399	0.077	0.280	0.646
		Mean 95% CI length for t distribution	150	0.413	0.079	0.290	0.669
	45	Mean coverage for normal distribution	150	0.842	0.072	0.669	0.949
		Mean coverage for t distribution	150	0.856	0.067	0.689	0.955
		Mean 95% CI length for normal distribution	150	0.353	0.067	0.248	0.570
		Mean 95% CI length for t distribution	150	0.363	0.069	0.255	0.586
	55	Mean coverage for normal distribution	150	0.813	0.090	0.604	0.952
		Mean coverage for t distribution	150	0.826	0.087	0.620	0.957
		Mean 95% CI length for normal distribution	150	0.320	0.061	0.225	0.517
		Mean 95% CI length for t distribution	150	0.327	0.062	0.230	0.528

Table 4-6. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.07	5	Mean coverage for normal distribution	120	0.882	0.004	0.873	0.891
		Mean coverage for t distribution	120	0.954	0.003	0.945	0.959
		Mean 95% CI length for normal distribution	120	0.554	0.101	0.375	0.853
		Mean 95% CI length for t distribution	120	0.784	0.144	0.531	1.208
	15	Mean coverage for normal distribution	120	0.929	0.002	0.922	0.934
		Mean coverage for t distribution	120	0.950	0.002	0.946	0.956
		Mean 95% CI length for normal distribution	120	0.337	0.063	0.228	0.521
		Mean 95% CI length for t distribution	120	0.368	0.068	0.249	0.570
	25	Mean coverage for normal distribution	120	0.934	0.004	0.921	0.941
		Mean coverage for t distribution	120	0.947	0.004	0.936	0.952
		Mean 95% CI length for normal distribution	120	0.263	0.049	0.178	0.406
		Mean 95% CI length for t distribution	120	0.277	0.051	0.188	0.428
	35	Mean coverage for normal distribution	120	0.935	0.007	0.911	0.945
		Mean coverage for t distribution	120	0.944	0.007	0.924	0.953
		Mean 95% CI length for normal distribution	120	0.223	0.041	0.151	0.345
		Mean 95% CI length for t distribution	120	0.231	0.043	0.157	0.357
	45	Mean coverage for normal distribution	120	0.935	0.010	0.902	0.949
		Mean coverage for t distribution	120	0.942	0.009	0.911	0.955
		Mean 95% CI length for normal distribution	120	0.197	0.037	0.134	0.305
		Mean 95% CI length for t distribution	120	0.203	0.038	0.137	0.313
	55	Mean coverage for normal distribution	120	0.933	0.012	0.894	0.949
		Mean coverage for t distribution	120	0.939	0.012	0.902	0.954
		Mean 95% CI length for normal distribution	120	0.179	0.033	0.121	0.276
		Mean 95% CI length for t distribution	120	0.183	0.034	0.124	0.282

Table 4-6. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.12	5	Mean coverage for normal distribution	90	0.881	0.004	0.874	0.891
		Mean coverage for t distribution	90	0.953	0.002	0.945	0.957
		Mean 95% CI length for normal distribution	90	0.462	0.077	0.318	0.665
		Mean 95% CI length for t distribution	90	0.654	0.109	0.451	0.942
	15	Mean coverage for normal distribution	90	0.930	0.002	0.926	0.935
		Mean coverage for t distribution	90	0.951	0.002	0.948	0.956
		Mean 95% CI length for normal distribution	90	0.280	0.047	0.193	0.403
		Mean 95% CI length for t distribution	90	0.306	0.051	0.211	0.442
	25	Mean coverage for normal distribution	90	0.938	0.001	0.934	0.941
		Mean coverage for t distribution	90	0.950	0.001	0.947	0.953
		Mean 95% CI length for normal distribution	90	0.219	0.037	0.151	0.315
		Mean 95% CI length for t distribution	90	0.230	0.039	0.159	0.331
	35	Mean coverage for normal distribution	90	0.941	0.002	0.933	0.946
		Mean coverage for t distribution	90	0.949	0.002	0.943	0.953
		Mean 95% CI length for normal distribution	90	0.185	0.031	0.128	0.267
		Mean 95% CI length for t distribution	90	0.192	0.032	0.133	0.277
	45	Mean coverage for normal distribution	90	0.942	0.004	0.932	0.950
		Mean coverage for t distribution	90	0.948	0.004	0.939	0.955
		Mean 95% CI length for normal distribution	90	0.164	0.027	0.113	0.236
		Mean 95% CI length for t distribution	90	0.168	0.028	0.116	0.242
	55	Mean coverage for normal distribution	90	0.944	0.004	0.934	0.950
		Mean coverage for t distribution	90	0.949	0.003	0.939	0.955
		Mean 95% CI length for normal distribution	90	0.148	0.025	0.102	0.214
		Mean 95% CI length for t distribution	90	0.152	0.025	0.105	0.218

Table 4-6. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.17	5	Mean coverage for normal distribution	60	0.881	0.003	0.874	0.888
		Mean coverage for t distribution	60	0.953	0.002	0.947	0.957
		Mean 95% CI length for normal distribution	60	0.420	0.068	0.294	0.585
		Mean 95% CI length for t distribution	60	0.596	0.096	0.416	0.829
	15	Mean coverage for normal distribution	60	0.930	0.001	0.927	0.935
		Mean coverage for t distribution	60	0.951	0.002	0.948	0.955
		Mean 95% CI length for normal distribution	60	0.255	0.041	0.178	0.355
		Mean 95% CI length for t distribution	60	0.279	0.045	0.195	0.388
	25	Mean coverage for normal distribution	60	0.939	0.001	0.935	0.942
		Mean coverage for t distribution	60	0.950	0.001	0.946	0.953
		Mean 95% CI length for normal distribution	60	0.199	0.032	0.139	0.277
		Mean 95% CI length for t distribution	60	0.209	0.034	0.146	0.291
	35	Mean coverage for normal distribution	60	0.942	0.001	0.939	0.945
		Mean coverage for t distribution	60	0.950	0.001	0.947	0.953
		Mean 95% CI length for normal distribution	60	0.169	0.027	0.118	0.234
		Mean 95% CI length for t distribution	60	0.175	0.028	0.122	0.243
	45	Mean coverage for normal distribution	60	0.944	0.003	0.938	0.949
		Mean coverage for t distribution	60	0.950	0.003	0.944	0.954
		Mean 95% CI length for normal distribution	60	0.149	0.024	0.104	0.207
		Mean 95% CI length for t distribution	60	0.153	0.025	0.107	0.213
	55	Mean coverage for normal distribution	60	0.946	0.002	0.941	0.950
		Mean coverage for t distribution	60	0.951	0.002	0.946	0.955
		Mean 95% CI length for normal distribution	60	0.135	0.022	0.094	0.188
		Mean 95% CI length for t distribution	60	0.138	0.022	0.096	0.192

Table 4-6. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.22	5	Mean coverage for normal distribution	30	0.881	0.003	0.874	0.886
		Mean coverage for t distribution	30	0.953	0.002	0.948	0.956
		Mean 95% CI length for normal distribution	30	0.398	0.065	0.281	0.545
		Mean 95% CI length for t distribution	30	0.564	0.092	0.397	0.772
	15	Mean coverage for normal distribution	30	0.931	0.002	0.928	0.934
		Mean coverage for t distribution	30	0.951	0.002	0.948	0.954
		Mean 95% CI length for normal distribution	30	0.241	0.039	0.170	0.330
		Mean 95% CI length for t distribution	30	0.264	0.043	0.186	0.361
	25	Mean coverage for normal distribution	30	0.938	0.002	0.933	0.942
		Mean coverage for t distribution	30	0.950	0.002	0.946	0.953
		Mean 95% CI length for normal distribution	30	0.188	0.031	0.133	0.257
		Mean 95% CI length for t distribution	30	0.198	0.032	0.140	0.271
	35	Mean coverage for normal distribution	30	0.942	0.001	0.940	0.945
		Mean coverage for t distribution	30	0.950	0.001	0.948	0.954
		Mean 95% CI length for normal distribution	30	0.160	0.026	0.113	0.218
		Mean 95% CI length for t distribution	30	0.165	0.027	0.117	0.226
	45	Mean coverage for normal distribution	30	0.944	0.003	0.940	0.948
		Mean coverage for t distribution	30	0.950	0.002	0.947	0.954
		Mean 95% CI length for normal distribution	30	0.141	0.023	0.099	0.193
		Mean 95% CI length for t distribution	30	0.145	0.024	0.102	0.198
	55	Mean coverage for normal distribution	30	0.946	0.002	0.942	0.950
		Mean coverage for t distribution	30	0.951	0.002	0.948	0.955
		Mean 95% CI length for normal distribution	30	0.128	0.021	0.090	0.175
		Mean 95% CI length for t distribution	30	0.131	0.021	0.092	0.179

Table 4-7. Aim 3: Comparisons of UW_RR vs. W_RR

Outcomes	Mean	STD	Minimum	Maximum
Mean coverage for UW_RR	0.951	0.002	0.944	0.959
Mean coverage for W_RR	0.950	0.002	0.940	0.959
Mean 95% CI length for UW_RR	0.299	0.261	0.058	2.014
Mean 95% CI length for W_RR	0.291	0.253	0.058	1.976

Table 4-8. Aim 3: Comparisons of UW_RR vs. W_RR by event rate in treatment arm and number of studies

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.02	5	Mean coverage for UW_RR	150	0.955	0.002	0.951	0.959
		Mean coverage for W_RR	150	0.952	0.003	0.944	0.959
		Mean 95% CI length for UW_RR	150	1.312	0.252	0.862	2.014
		Mean 95% CI length for W_RR	150	1.273	0.238	0.861	1.976
	15	Mean coverage for UW_RR	150	0.953	0.002	0.947	0.957
		Mean coverage for W_RR	150	0.951	0.002	0.945	0.958
		Mean 95% CI length for UW_RR	150	0.603	0.116	0.397	0.918
		Mean 95% CI length for W_RR	150	0.585	0.109	0.396	0.901
	25	Mean coverage for UW_RR	150	0.951	0.002	0.946	0.955
		Mean coverage for W_RR	150	0.950	0.002	0.944	0.955
		Mean 95% CI length for UW_RR	150	0.452	0.087	0.297	0.687
		Mean 95% CI length for W_RR	150	0.438	0.081	0.297	0.674
	35	Mean coverage for UW_RR	150	0.950	0.002	0.946	0.956
		Mean coverage for W_RR	150	0.950	0.002	0.945	0.957
		Mean 95% CI length for UW_RR	150	0.376	0.072	0.248	0.571
		Mean 95% CI length for W_RR	150	0.365	0.068	0.247	0.560
	45	Mean coverage for UW_RR	150	0.950	0.003	0.945	0.955
		Mean coverage for W_RR	150	0.950	0.002	0.945	0.956
		Mean 95% CI length for UW_RR	150	0.330	0.063	0.217	0.500
		Mean 95% CI length for W_RR	150	0.320	0.059	0.217	0.491
	55	Mean coverage for UW_RR	150	0.951	0.002	0.945	0.957
		Mean coverage for W_RR	150	0.951	0.002	0.945	0.956
		Mean 95% CI length for UW_RR	150	0.297	0.057	0.195	0.450
		Mean 95% CI length for W_RR	150	0.288	0.053	0.195	0.442

Table 4-8. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.02	5	Mean coverage for UW_RR	150	0.955	0.002	0.951	0.959
		Mean coverage for W_RR	150	0.952	0.003	0.944	0.959
		Mean 95% CI length for UW_RR	150	1.312	0.252	0.862	2.014
		Mean 95% CI length for W_RR	150	1.273	0.238	0.861	1.976
	15	Mean coverage for UW_RR	150	0.953	0.002	0.947	0.957
		Mean coverage for W_RR	150	0.951	0.002	0.945	0.958
		Mean 95% CI length for UW_RR	150	0.603	0.116	0.397	0.918
		Mean 95% CI length for W_RR	150	0.585	0.109	0.396	0.901
	25	Mean coverage for UW_RR	150	0.951	0.002	0.946	0.955
		Mean coverage for W_RR	150	0.950	0.002	0.944	0.955
		Mean 95% CI length for UW_RR	150	0.452	0.087	0.297	0.687
		Mean 95% CI length for W_RR	150	0.438	0.081	0.297	0.674
	35	Mean coverage for UW_RR	150	0.950	0.002	0.946	0.956
		Mean coverage for W_RR	150	0.950	0.002	0.945	0.957
		Mean 95% CI length for UW_RR	150	0.376	0.072	0.248	0.571
		Mean 95% CI length for W_RR	150	0.365	0.068	0.247	0.560
	45	Mean coverage for UW_RR	150	0.950	0.003	0.945	0.955
		Mean coverage for W_RR	150	0.950	0.002	0.945	0.956
		Mean 95% CI length for UW_RR	150	0.330	0.063	0.217	0.500
		Mean 95% CI length for W_RR	150	0.320	0.059	0.217	0.491
	55	Mean coverage for UW_RR	150	0.951	0.002	0.945	0.957
		Mean coverage for W_RR	150	0.951	0.002	0.945	0.956
		Mean 95% CI length for UW_RR	150	0.297	0.057	0.195	0.450
		Mean 95% CI length for W_RR	150	0.288	0.053	0.195	0.442

Table 4-8. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.12	5	Mean coverage for UW_RR	150	0.952	0.003	0.945	0.958
		Mean coverage for W_RR	150	0.949	0.003	0.942	0.954
		Mean 95% CI length for UW_RR	150	0.526	0.096	0.326	0.802
		Mean 95% CI length for W_RR	150	0.512	0.091	0.325	0.789
	15	Mean coverage for UW_RR	150	0.951	0.002	0.947	0.956
		Mean coverage for W_RR	150	0.950	0.002	0.945	0.955
		Mean 95% CI length for UW_RR	150	0.245	0.045	0.152	0.373
		Mean 95% CI length for W_RR	150	0.239	0.043	0.152	0.367
	25	Mean coverage for UW_RR	150	0.950	0.001	0.946	0.954
		Mean coverage for W_RR	150	0.949	0.002	0.944	0.953
		Mean 95% CI length for UW_RR	150	0.184	0.033	0.115	0.280
		Mean 95% CI length for W_RR	150	0.180	0.032	0.114	0.276
	35	Mean coverage for UW_RR	150	0.950	0.001	0.947	0.954
		Mean coverage for W_RR	150	0.949	0.002	0.943	0.954
		Mean 95% CI length for UW_RR	150	0.154	0.028	0.096	0.233
		Mean 95% CI length for W_RR	150	0.150	0.026	0.096	0.230
	45	Mean coverage for UW_RR	150	0.950	0.002	0.945	0.954
		Mean coverage for W_RR	150	0.950	0.003	0.944	0.955
		Mean 95% CI length for UW_RR	150	0.135	0.024	0.084	0.204
		Mean 95% CI length for W_RR	150	0.131	0.023	0.084	0.201
	55	Mean coverage for UW_RR	150	0.951	0.002	0.945	0.955
		Mean coverage for W_RR	150	0.951	0.002	0.945	0.955
		Mean 95% CI length for UW_RR	150	0.121	0.022	0.075	0.184
		Mean 95% CI length for W_RR	150	0.118	0.021	0.075	0.181

Table 4-8. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.12	5	Mean coverage for UW_RR	150	0.952	0.003	0.945	0.958
		Mean coverage for W_RR	150	0.949	0.003	0.942	0.954
		Mean 95% CI length for UW_RR	150	0.526	0.096	0.326	0.802
		Mean 95% CI length for W_RR	150	0.512	0.091	0.325	0.789
	15	Mean coverage for UW_RR	150	0.951	0.002	0.947	0.956
		Mean coverage for W_RR	150	0.950	0.002	0.945	0.955
		Mean 95% CI length for UW_RR	150	0.245	0.045	0.152	0.373
		Mean 95% CI length for W_RR	150	0.239	0.043	0.152	0.367
	25	Mean coverage for UW_RR	150	0.950	0.001	0.946	0.954
		Mean coverage for W_RR	150	0.949	0.002	0.944	0.953
		Mean 95% CI length for UW_RR	150	0.184	0.033	0.115	0.280
		Mean 95% CI length for W_RR	150	0.180	0.032	0.114	0.276
	35	Mean coverage for UW_RR	150	0.950	0.001	0.947	0.954
		Mean coverage for W_RR	150	0.949	0.002	0.943	0.954
		Mean 95% CI length for UW_RR	150	0.154	0.028	0.096	0.233
		Mean 95% CI length for W_RR	150	0.150	0.026	0.096	0.230
	45	Mean coverage for UW_RR	150	0.950	0.002	0.945	0.954
		Mean coverage for W_RR	150	0.950	0.003	0.944	0.955
		Mean 95% CI length for UW_RR	150	0.135	0.024	0.084	0.204
		Mean 95% CI length for W_RR	150	0.131	0.023	0.084	0.201
	55	Mean coverage for UW_RR	150	0.951	0.002	0.945	0.955
		Mean coverage for W_RR	150	0.951	0.002	0.945	0.955
		Mean 95% CI length for UW_RR	150	0.121	0.022	0.075	0.184
		Mean 95% CI length for W_RR	150	0.118	0.021	0.075	0.181

Table 4-8. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.17	5	Mean coverage for UW_RR	130	0.952	0.003	0.945	0.957
		Mean coverage for W_RR	130	0.949	0.003	0.942	0.954
		Mean 95% CI length for UW_RR	130	0.447	0.083	0.265	0.673
		Mean 95% CI length for W_RR	130	0.436	0.079	0.265	0.663
	15	Mean coverage for UW_RR	130	0.951	0.002	0.947	0.956
		Mean coverage for W_RR	130	0.950	0.002	0.945	0.956
		Mean 95% CI length for UW_RR	130	0.209	0.039	0.124	0.314
		Mean 95% CI length for W_RR	130	0.204	0.037	0.124	0.310
	25	Mean coverage for UW_RR	130	0.950	0.001	0.946	0.954
		Mean coverage for W_RR	130	0.949	0.002	0.944	0.953
		Mean 95% CI length for UW_RR	130	0.157	0.029	0.093	0.235
		Mean 95% CI length for W_RR	130	0.153	0.028	0.093	0.232
	35	Mean coverage for UW_RR	130	0.950	0.002	0.947	0.954
		Mean coverage for W_RR	130	0.949	0.002	0.944	0.954
		Mean 95% CI length for UW_RR	130	0.131	0.024	0.078	0.196
		Mean 95% CI length for W_RR	130	0.128	0.023	0.078	0.194
	45	Mean coverage for UW_RR	130	0.950	0.002	0.946	0.955
		Mean coverage for W_RR	130	0.950	0.002	0.945	0.954
		Mean 95% CI length for UW_RR	130	0.114	0.021	0.068	0.172
		Mean 95% CI length for W_RR	130	0.112	0.020	0.068	0.170
	55	Mean coverage for UW_RR	130	0.951	0.002	0.946	0.954
		Mean coverage for W_RR	130	0.951	0.002	0.946	0.954
		Mean 95% CI length for UW_RR	130	0.103	0.019	0.061	0.155
		Mean 95% CI length for W_RR	130	0.101	0.018	0.061	0.153

Table 4-8. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.22	5	Mean coverage for UW_RR	120	0.952	0.003	0.944	0.957
		Mean coverage for W_RR	120	0.949	0.003	0.940	0.955
		Mean 95% CI length for UW_RR	120	0.400	0.070	0.249	0.593
		Mean 95% CI length for W_RR	120	0.391	0.068	0.248	0.584
	15	Mean coverage for UW_RR	120	0.951	0.002	0.948	0.955
		Mean coverage for W_RR	120	0.950	0.002	0.944	0.954
		Mean 95% CI length for UW_RR	120	0.187	0.033	0.116	0.276
		Mean 95% CI length for W_RR	120	0.183	0.032	0.116	0.273
	25	Mean coverage for UW_RR	120	0.950	0.002	0.945	0.955
		Mean coverage for W_RR	120	0.949	0.002	0.944	0.952
		Mean 95% CI length for UW_RR	120	0.140	0.024	0.088	0.207
		Mean 95% CI length for W_RR	120	0.138	0.024	0.087	0.205
	35	Mean coverage for UW_RR	120	0.950	0.002	0.946	0.954
		Mean coverage for W_RR	120	0.949	0.002	0.944	0.955
		Mean 95% CI length for UW_RR	120	0.117	0.020	0.073	0.173
		Mean 95% CI length for W_RR	120	0.115	0.020	0.073	0.171
	45	Mean coverage for UW_RR	120	0.950	0.002	0.945	0.954
		Mean coverage for W_RR	120	0.949	0.002	0.945	0.954
		Mean 95% CI length for UW_RR	120	0.103	0.018	0.064	0.151
		Mean 95% CI length for W_RR	120	0.101	0.017	0.064	0.150
	55	Mean coverage for UW_RR	120	0.951	0.002	0.946	0.955
		Mean coverage for W_RR	120	0.951	0.002	0.945	0.956
		Mean 95% CI length for UW_RR	120	0.092	0.016	0.058	0.136
		Mean 95% CI length for W_RR	120	0.091	0.016	0.058	0.135

Table 4-9. Aim 4-1: Comparisons of t distribution vs. normal distribution for UW_RR

Outcomes	Mean	STD	Minimum	Maximum
Mean coverage for normal distribution	0.931	0.022	0.870	0.952
Mean coverage for t distribution	0.951	0.002	0.944	0.959
Mean 95% CI length for normal distribution	0.255	0.186	0.056	1.422
Mean 95% CI length for t distribution	0.299	0.261	0.058	2.014

Table 4-10. Aim 4-1: Comparisons of t distribution vs. normal distribution for UW_RR by event rate in treatment arm and number of studies

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.02	5	Mean coverage for normal distribution	150	0.886	0.003	0.878	0.894
		Mean coverage for t distribution	150	0.955	0.002	0.951	0.959
		Mean 95% CI length for normal distribution	150	0.926	0.178	0.609	1.422
		Mean 95% CI length for t distribution	150	1.312	0.252	0.862	2.014
	15	Mean coverage for normal distribution	150	0.932	0.002	0.926	0.937
		Mean coverage for t distribution	150	0.953	0.002	0.947	0.957
		Mean 95% CI length for normal distribution	150	0.551	0.106	0.362	0.839
		Mean 95% CI length for t distribution	150	0.603	0.116	0.397	0.918
	25	Mean coverage for normal distribution	150	0.939	0.002	0.933	0.944
		Mean coverage for t distribution	150	0.951	0.002	0.946	0.955
		Mean 95% CI length for normal distribution	150	0.429	0.082	0.282	0.652
		Mean 95% CI length for t distribution	150	0.452	0.087	0.297	0.687
	35	Mean coverage for normal distribution	150	0.942	0.002	0.937	0.949
		Mean coverage for t distribution	150	0.950	0.002	0.946	0.956
		Mean 95% CI length for normal distribution	150	0.363	0.069	0.239	0.551
		Mean 95% CI length for t distribution	150	0.376	0.072	0.248	0.571
	45	Mean coverage for normal distribution	150	0.944	0.003	0.939	0.949
		Mean coverage for t distribution	150	0.950	0.003	0.945	0.955
		Mean 95% CI length for normal distribution	150	0.321	0.061	0.211	0.486
		Mean 95% CI length for t distribution	150	0.330	0.063	0.217	0.500
	55	Mean coverage for normal distribution	150	0.946	0.002	0.939	0.952
		Mean coverage for t distribution	150	0.951	0.002	0.945	0.957
		Mean 95% CI length for normal distribution	150	0.290	0.056	0.191	0.440
		Mean 95% CI length for t distribution	150	0.297	0.057	0.195	0.450

Table 4-10. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.07	5	Mean coverage for normal distribution	150	0.882	0.004	0.872	0.892
		Mean coverage for t distribution	150	0.953	0.003	0.945	0.958
		Mean 95% CI length for normal distribution	150	0.488	0.090	0.311	0.742
		Mean 95% CI length for t distribution	150	0.691	0.127	0.441	1.051
	15	Mean coverage for normal distribution	150	0.931	0.002	0.927	0.935
		Mean coverage for t distribution	150	0.952	0.002	0.947	0.956
		Mean 95% CI length for normal distribution	150	0.294	0.054	0.188	0.446
		Mean 95% CI length for t distribution	150	0.322	0.059	0.206	0.488
	25	Mean coverage for normal distribution	150	0.939	0.002	0.935	0.942
		Mean coverage for t distribution	150	0.951	0.002	0.947	0.955
		Mean 95% CI length for normal distribution	150	0.229	0.042	0.147	0.348
		Mean 95% CI length for t distribution	150	0.241	0.044	0.155	0.366
	35	Mean coverage for normal distribution	150	0.942	0.002	0.939	0.946
		Mean coverage for t distribution	150	0.950	0.002	0.947	0.955
		Mean 95% CI length for normal distribution	150	0.194	0.036	0.125	0.294
		Mean 95% CI length for t distribution	150	0.201	0.037	0.129	0.305
	45	Mean coverage for normal distribution	150	0.944	0.003	0.939	0.949
		Mean coverage for t distribution	150	0.950	0.003	0.945	0.955
		Mean 95% CI length for normal distribution	150	0.172	0.032	0.110	0.260
		Mean 95% CI length for t distribution	150	0.176	0.032	0.113	0.267
	55	Mean coverage for normal distribution	150	0.946	0.002	0.940	0.949
		Mean coverage for t distribution	150	0.951	0.002	0.946	0.954
		Mean 95% CI length for normal distribution	150	0.155	0.029	0.099	0.235
		Mean 95% CI length for t distribution	150	0.159	0.029	0.102	0.241

Table 4-10. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.12	5	Mean coverage for normal distribution	150	0.881	0.003	0.872	0.890
		Mean coverage for t distribution	150	0.952	0.003	0.945	0.958
		Mean 95% CI length for normal distribution	150	0.372	0.067	0.230	0.566
		Mean 95% CI length for t distribution	150	0.526	0.096	0.326	0.802
	15	Mean coverage for normal distribution	150	0.931	0.002	0.926	0.935
		Mean coverage for t distribution	150	0.951	0.002	0.947	0.956
		Mean 95% CI length for normal distribution	150	0.224	0.041	0.139	0.341
		Mean 95% CI length for t distribution	150	0.245	0.045	0.152	0.373
	25	Mean coverage for normal distribution	150	0.939	0.002	0.934	0.943
		Mean coverage for t distribution	150	0.950	0.001	0.946	0.954
		Mean 95% CI length for normal distribution	150	0.175	0.032	0.109	0.266
		Mean 95% CI length for t distribution	150	0.184	0.033	0.115	0.280
	35	Mean coverage for normal distribution	150	0.942	0.002	0.937	0.946
		Mean coverage for t distribution	150	0.950	0.001	0.947	0.954
		Mean 95% CI length for normal distribution	150	0.148	0.027	0.092	0.225
		Mean 95% CI length for t distribution	150	0.154	0.028	0.096	0.233
	45	Mean coverage for normal distribution	150	0.944	0.003	0.938	0.948
		Mean coverage for t distribution	150	0.950	0.002	0.945	0.954
		Mean 95% CI length for normal distribution	150	0.131	0.024	0.081	0.199
		Mean 95% CI length for t distribution	150	0.135	0.024	0.084	0.204
	55	Mean coverage for normal distribution	150	0.946	0.002	0.939	0.949
		Mean coverage for t distribution	150	0.951	0.002	0.945	0.955
		Mean 95% CI length for normal distribution	150	0.119	0.021	0.074	0.180
		Mean 95% CI length for t distribution	150	0.121	0.022	0.075	0.184

Table 4-10. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.17	5	Mean coverage for normal distribution	130	0.881	0.003	0.870	0.889
		Mean coverage for t distribution	130	0.952	0.003	0.945	0.957
		Mean 95% CI length for normal distribution	130	0.316	0.058	0.187	0.475
		Mean 95% CI length for t distribution	130	0.447	0.083	0.265	0.673
	15	Mean coverage for normal distribution	130	0.931	0.002	0.925	0.934
		Mean coverage for t distribution	130	0.951	0.002	0.947	0.956
		Mean 95% CI length for normal distribution	130	0.191	0.035	0.113	0.287
		Mean 95% CI length for t distribution	130	0.209	0.039	0.124	0.314
	25	Mean coverage for normal distribution	130	0.938	0.002	0.935	0.942
		Mean coverage for t distribution	130	0.950	0.001	0.946	0.954
		Mean 95% CI length for normal distribution	130	0.149	0.027	0.089	0.224
		Mean 95% CI length for t distribution	130	0.157	0.029	0.093	0.235
	35	Mean coverage for normal distribution	130	0.942	0.002	0.938	0.946
		Mean coverage for t distribution	130	0.950	0.002	0.947	0.954
		Mean 95% CI length for normal distribution	130	0.126	0.023	0.075	0.189
		Mean 95% CI length for t distribution	130	0.131	0.024	0.078	0.196
	45	Mean coverage for normal distribution	130	0.944	0.003	0.939	0.949
		Mean coverage for t distribution	130	0.950	0.002	0.946	0.955
		Mean 95% CI length for normal distribution	130	0.111	0.021	0.066	0.167
		Mean 95% CI length for t distribution	130	0.114	0.021	0.068	0.172
	55	Mean coverage for normal distribution	130	0.946	0.002	0.941	0.949
		Mean coverage for t distribution	130	0.951	0.002	0.946	0.954
		Mean 95% CI length for normal distribution	130	0.101	0.019	0.060	0.151
		Mean 95% CI length for t distribution	130	0.103	0.019	0.061	0.155

Table 4-10. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.22	5	Mean coverage for normal distribution	120	0.880	0.003	0.873	0.888
		Mean coverage for t distribution	120	0.952	0.003	0.944	0.957
		Mean 95% CI length for normal distribution	120	0.283	0.049	0.176	0.418
		Mean 95% CI length for t distribution	120	0.400	0.070	0.249	0.593
	15	Mean coverage for normal distribution	120	0.931	0.002	0.927	0.935
		Mean coverage for t distribution	120	0.951	0.002	0.948	0.955
		Mean 95% CI length for normal distribution	120	0.171	0.030	0.106	0.253
		Mean 95% CI length for t distribution	120	0.187	0.033	0.116	0.276
	25	Mean coverage for normal distribution	120	0.938	0.002	0.934	0.943
		Mean coverage for t distribution	120	0.950	0.002	0.945	0.955
		Mean 95% CI length for normal distribution	120	0.133	0.023	0.083	0.197
		Mean 95% CI length for t distribution	120	0.140	0.024	0.088	0.207
	35	Mean coverage for normal distribution	120	0.942	0.002	0.939	0.946
		Mean coverage for t distribution	120	0.950	0.002	0.946	0.954
		Mean 95% CI length for normal distribution	120	0.113	0.020	0.071	0.167
		Mean 95% CI length for t distribution	120	0.117	0.020	0.073	0.173
	45	Mean coverage for normal distribution	120	0.944	0.002	0.939	0.948
		Mean coverage for t distribution	120	0.950	0.002	0.945	0.954
		Mean 95% CI length for normal distribution	120	0.100	0.017	0.062	0.147
		Mean 95% CI length for t distribution	120	0.103	0.018	0.064	0.151
	55	Mean coverage for normal distribution	120	0.946	0.002	0.941	0.950
		Mean coverage for t distribution	120	0.951	0.002	0.946	0.955
		Mean 95% CI length for normal distribution	120	0.090	0.016	0.056	0.133
		Mean 95% CI length for t distribution	120	0.092	0.016	0.058	0.136

Table 4-11. Aim 4-2: Comparisons of t distribution vs. normal distribution for W_RR

Outcomes	Mean	STD	Minimum	Maximum
Mean coverage for normal distribution	0.929	0.024	0.860	0.952
Mean coverage for t distribution	0.950	0.002	0.940	0.959
Mean 95% CI length for normal distribution	0.249	0.180	0.056	1.395
Mean 95% CI length for t distribution	0.291	0.253	0.058	1.976

Table 4-12. Aim 4-2: Comparisons of t distribution vs. normal distribution for W_RR by event rate in treatment arm and number of studies

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.02	5	Mean coverage for normal distribution	150	0.882	0.005	0.868	0.892
		Mean coverage for t distribution	150	0.952	0.003	0.944	0.959
		Mean 95% CI length for normal distribution	150	0.899	0.168	0.608	1.395
		Mean 95% CI length for t distribution	150	1.273	0.238	0.861	1.976
	15	Mean coverage for normal distribution	150	0.931	0.003	0.923	0.937
		Mean coverage for t distribution	150	0.951	0.002	0.945	0.958
		Mean 95% CI length for normal distribution	150	0.535	0.100	0.362	0.824
		Mean 95% CI length for t distribution	150	0.585	0.109	0.396	0.901
	25	Mean coverage for normal distribution	150	0.938	0.003	0.932	0.943
		Mean coverage for t distribution	150	0.950	0.002	0.944	0.955
		Mean 95% CI length for normal distribution	150	0.416	0.077	0.282	0.640
		Mean 95% CI length for t distribution	150	0.438	0.081	0.297	0.674
	35	Mean coverage for normal distribution	150	0.942	0.002	0.937	0.949
		Mean coverage for t distribution	150	0.950	0.002	0.945	0.957
		Mean 95% CI length for normal distribution	150	0.352	0.065	0.238	0.541
		Mean 95% CI length for t distribution	150	0.365	0.068	0.247	0.560
	45	Mean coverage for normal distribution	150	0.943	0.002	0.939	0.949
		Mean coverage for t distribution	150	0.950	0.002	0.945	0.956
		Mean 95% CI length for normal distribution	150	0.311	0.058	0.211	0.477
		Mean 95% CI length for t distribution	150	0.320	0.059	0.217	0.491
	55	Mean coverage for normal distribution	150	0.946	0.002	0.940	0.952
		Mean coverage for t distribution	150	0.951	0.002	0.945	0.956
		Mean 95% CI length for normal distribution	150	0.281	0.052	0.191	0.432
		Mean 95% CI length for t distribution	150	0.288	0.053	0.195	0.442

Table 4-12. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.07	5	Mean coverage for normal distribution	150	0.878	0.005	0.866	0.887
		Mean coverage for t distribution	150	0.950	0.003	0.944	0.955
		Mean 95% CI length for normal distribution	150	0.474	0.085	0.311	0.729
		Mean 95% CI length for t distribution	150	0.672	0.120	0.440	1.033
	15	Mean coverage for normal distribution	150	0.930	0.003	0.922	0.935
		Mean coverage for t distribution	150	0.950	0.002	0.945	0.956
		Mean 95% CI length for normal distribution	150	0.286	0.051	0.188	0.439
		Mean 95% CI length for t distribution	150	0.313	0.056	0.205	0.480
	25	Mean coverage for normal distribution	150	0.938	0.002	0.931	0.943
		Mean coverage for t distribution	150	0.950	0.002	0.944	0.955
		Mean 95% CI length for normal distribution	150	0.223	0.040	0.147	0.342
		Mean 95% CI length for t distribution	150	0.235	0.042	0.154	0.360
	35	Mean coverage for normal distribution	150	0.941	0.002	0.935	0.946
		Mean coverage for t distribution	150	0.949	0.002	0.945	0.954
		Mean 95% CI length for normal distribution	150	0.189	0.034	0.124	0.289
		Mean 95% CI length for t distribution	150	0.196	0.035	0.129	0.300
	45	Mean coverage for normal distribution	150	0.944	0.003	0.938	0.950
		Mean coverage for t distribution	150	0.950	0.003	0.945	0.955
		Mean 95% CI length for normal distribution	150	0.167	0.030	0.110	0.255
		Mean 95% CI length for t distribution	150	0.172	0.031	0.113	0.263
	55	Mean coverage for normal distribution	150	0.946	0.002	0.941	0.949
		Mean coverage for t distribution	150	0.951	0.002	0.946	0.954
		Mean 95% CI length for normal distribution	150	0.151	0.027	0.099	0.231
		Mean 95% CI length for t distribution	150	0.155	0.028	0.101	0.237

Table 4-12. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.12	5	Mean coverage for normal distribution	150	0.876	0.005	0.861	0.886
		Mean coverage for t distribution	150	0.949	0.003	0.942	0.954
		Mean 95% CI length for normal distribution	150	0.362	0.064	0.230	0.557
		Mean 95% CI length for t distribution	150	0.512	0.091	0.325	0.789
	15	Mean coverage for normal distribution	150	0.929	0.003	0.923	0.935
		Mean coverage for t distribution	150	0.950	0.002	0.945	0.955
		Mean 95% CI length for normal distribution	150	0.219	0.039	0.139	0.336
		Mean 95% CI length for t distribution	150	0.239	0.043	0.152	0.367
	25	Mean coverage for normal distribution	150	0.937	0.002	0.932	0.942
		Mean coverage for t distribution	150	0.949	0.002	0.944	0.953
		Mean 95% CI length for normal distribution	150	0.171	0.030	0.109	0.262
		Mean 95% CI length for t distribution	150	0.180	0.032	0.114	0.276
	35	Mean coverage for normal distribution	150	0.941	0.002	0.935	0.946
		Mean coverage for t distribution	150	0.949	0.002	0.943	0.954
		Mean 95% CI length for normal distribution	150	0.145	0.026	0.092	0.222
		Mean 95% CI length for t distribution	150	0.150	0.026	0.096	0.230
	45	Mean coverage for normal distribution	150	0.943	0.002	0.938	0.949
		Mean coverage for t distribution	150	0.950	0.003	0.944	0.955
		Mean 95% CI length for normal distribution	150	0.128	0.023	0.081	0.196
		Mean 95% CI length for t distribution	150	0.131	0.023	0.084	0.201
	55	Mean coverage for normal distribution	150	0.946	0.002	0.939	0.950
		Mean coverage for t distribution	150	0.951	0.002	0.945	0.955
		Mean 95% CI length for normal distribution	150	0.116	0.020	0.074	0.177
		Mean 95% CI length for t distribution	150	0.118	0.021	0.075	0.181

Table 4-12. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.17	5	Mean coverage for normal distribution	130	0.876	0.005	0.861	0.886
		Mean coverage for t distribution	130	0.949	0.003	0.942	0.954
		Mean 95% CI length for normal distribution	130	0.308	0.056	0.187	0.468
		Mean 95% CI length for t distribution	130	0.436	0.079	0.265	0.663
	15	Mean coverage for normal distribution	130	0.929	0.003	0.923	0.933
		Mean coverage for t distribution	130	0.950	0.002	0.945	0.956
		Mean 95% CI length for normal distribution	130	0.186	0.034	0.113	0.283
		Mean 95% CI length for t distribution	130	0.204	0.037	0.124	0.310
	25	Mean coverage for normal distribution	130	0.937	0.002	0.932	0.941
		Mean coverage for t distribution	130	0.949	0.002	0.944	0.953
		Mean 95% CI length for normal distribution	130	0.145	0.027	0.088	0.221
		Mean 95% CI length for t distribution	130	0.153	0.028	0.093	0.232
	35	Mean coverage for normal distribution	130	0.941	0.002	0.935	0.946
		Mean coverage for t distribution	130	0.949	0.002	0.944	0.954
		Mean 95% CI length for normal distribution	130	0.123	0.022	0.075	0.187
		Mean 95% CI length for t distribution	130	0.128	0.023	0.078	0.194
	45	Mean coverage for normal distribution	130	0.943	0.002	0.938	0.949
		Mean coverage for t distribution	130	0.950	0.002	0.945	0.954
		Mean 95% CI length for normal distribution	130	0.109	0.020	0.066	0.165
		Mean 95% CI length for t distribution	130	0.112	0.020	0.068	0.170
	55	Mean coverage for normal distribution	130	0.946	0.002	0.941	0.949
		Mean coverage for t distribution	130	0.951	0.002	0.946	0.954
		Mean 95% CI length for normal distribution	130	0.099	0.018	0.060	0.149
		Mean 95% CI length for t distribution	130	0.101	0.018	0.061	0.153

Table 4-12. Continued

Event rate in treatment	No. of studies	Outcomes	No. of scenarios	Mean	STD	Minimum	Maximum
0.22	5	Mean coverage for normal distribution	120	0.875	0.005	0.860	0.884
		Mean coverage for t distribution	120	0.949	0.003	0.940	0.955
		Mean 95% CI length for normal distribution	120	0.276	0.048	0.175	0.412
		Mean 95% CI length for t distribution	120	0.391	0.068	0.248	0.584
	15	Mean coverage for normal distribution	120	0.929	0.003	0.923	0.934
		Mean coverage for t distribution	120	0.950	0.002	0.944	0.954
		Mean 95% CI length for normal distribution	120	0.167	0.029	0.106	0.249
		Mean 95% CI length for t distribution	120	0.183	0.032	0.116	0.273
	25	Mean coverage for normal distribution	120	0.937	0.002	0.930	0.943
		Mean coverage for t distribution	120	0.949	0.002	0.944	0.952
		Mean 95% CI length for normal distribution	120	0.131	0.023	0.083	0.195
		Mean 95% CI length for t distribution	120	0.138	0.024	0.087	0.205
	35	Mean coverage for normal distribution	120	0.941	0.002	0.935	0.946
		Mean coverage for t distribution	120	0.949	0.002	0.944	0.955
		Mean 95% CI length for normal distribution	120	0.111	0.019	0.070	0.165
		Mean 95% CI length for t distribution	120	0.115	0.020	0.073	0.171
	45	Mean coverage for normal distribution	120	0.943	0.002	0.938	0.947
		Mean coverage for t distribution	120	0.949	0.002	0.945	0.954
		Mean 95% CI length for normal distribution	120	0.098	0.017	0.062	0.146
		Mean 95% CI length for t distribution	120	0.101	0.017	0.064	0.150
	55	Mean coverage for normal distribution	120	0.946	0.002	0.940	0.951
		Mean coverage for t distribution	120	0.951	0.002	0.945	0.956
		Mean 95% CI length for normal distribution	120	0.089	0.015	0.056	0.132
		Mean 95% CI length for t distribution	120	0.091	0.016	0.058	0.135

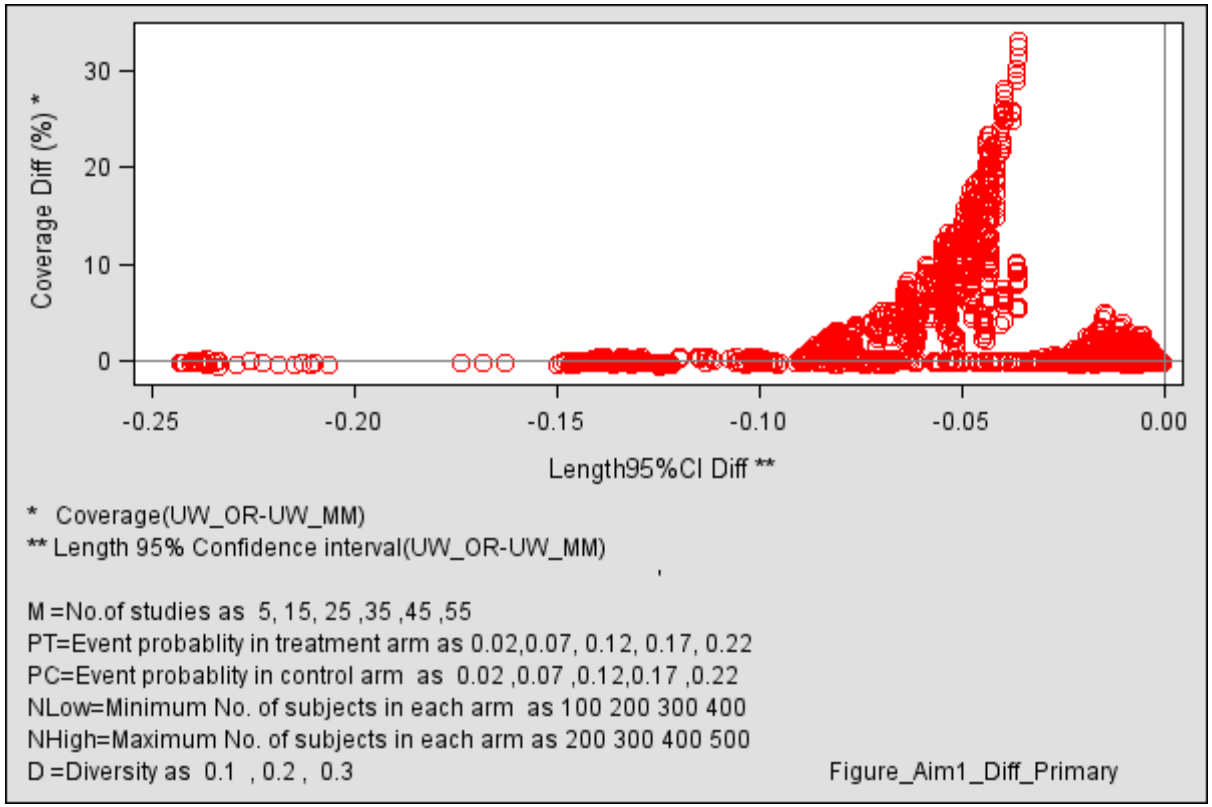

Figure 4-1. Aim 1: Comparisons of UW_OR vs. UW_MM in terms of difference for both mean coverage and mean length 95% CI

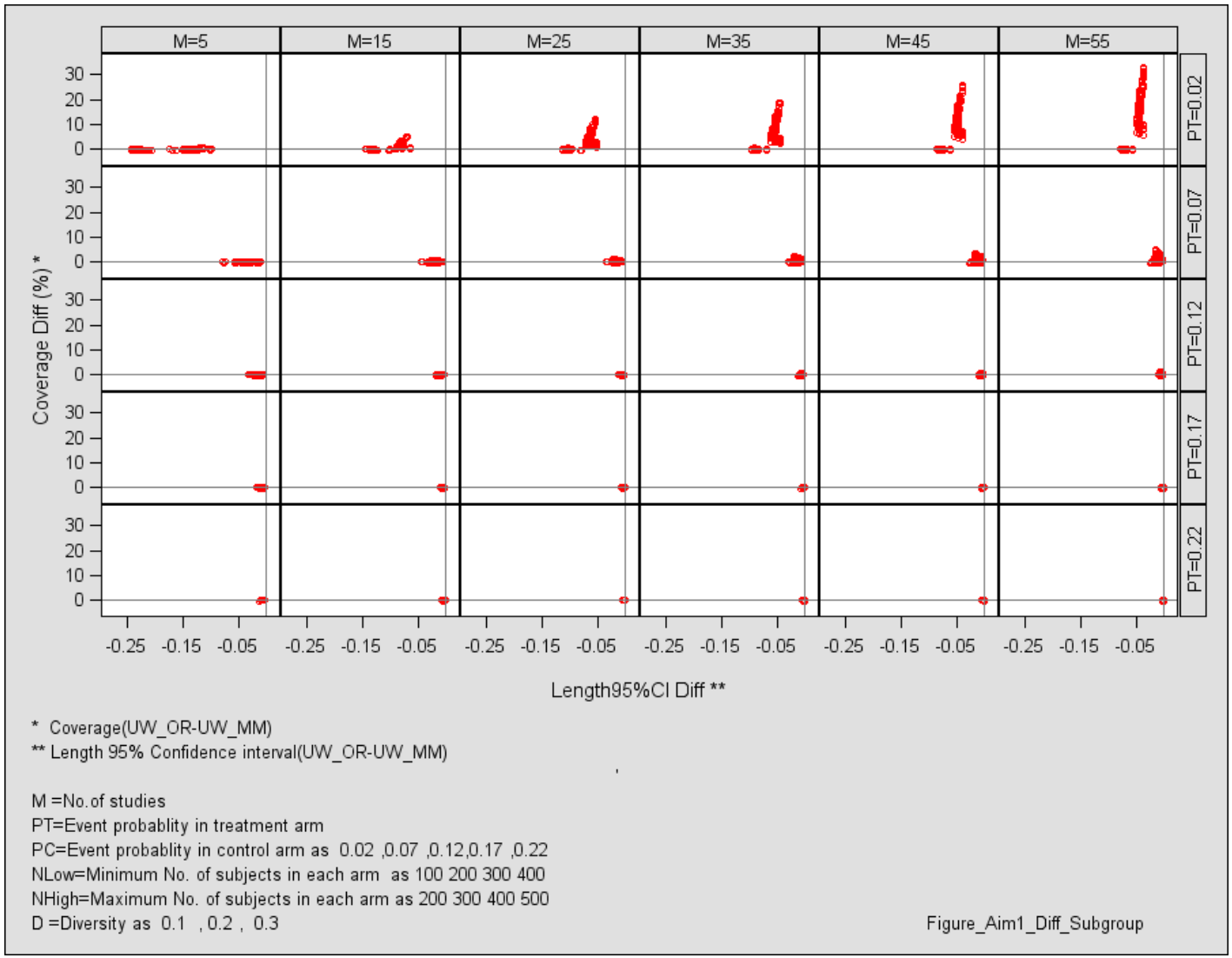

Figure 4-2. Aim 1: Comparisons of UW_OR vs. UW_MM in terms of difference for both mean coverage and mean length 95% CI by event rate in treatment arm and number of studies

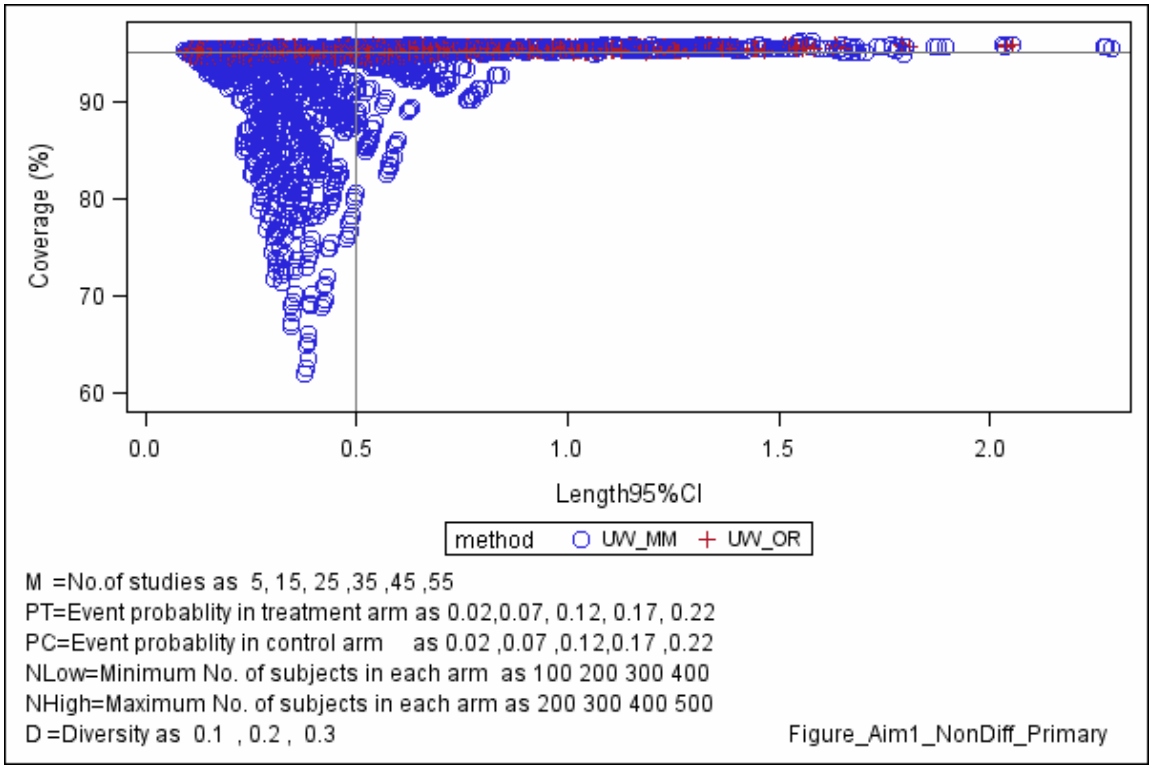

Figure 4-3. Aim 1: Comparisons of UW_OR vs. UW_MM in terms of both mean coverage and mean length 95% CI

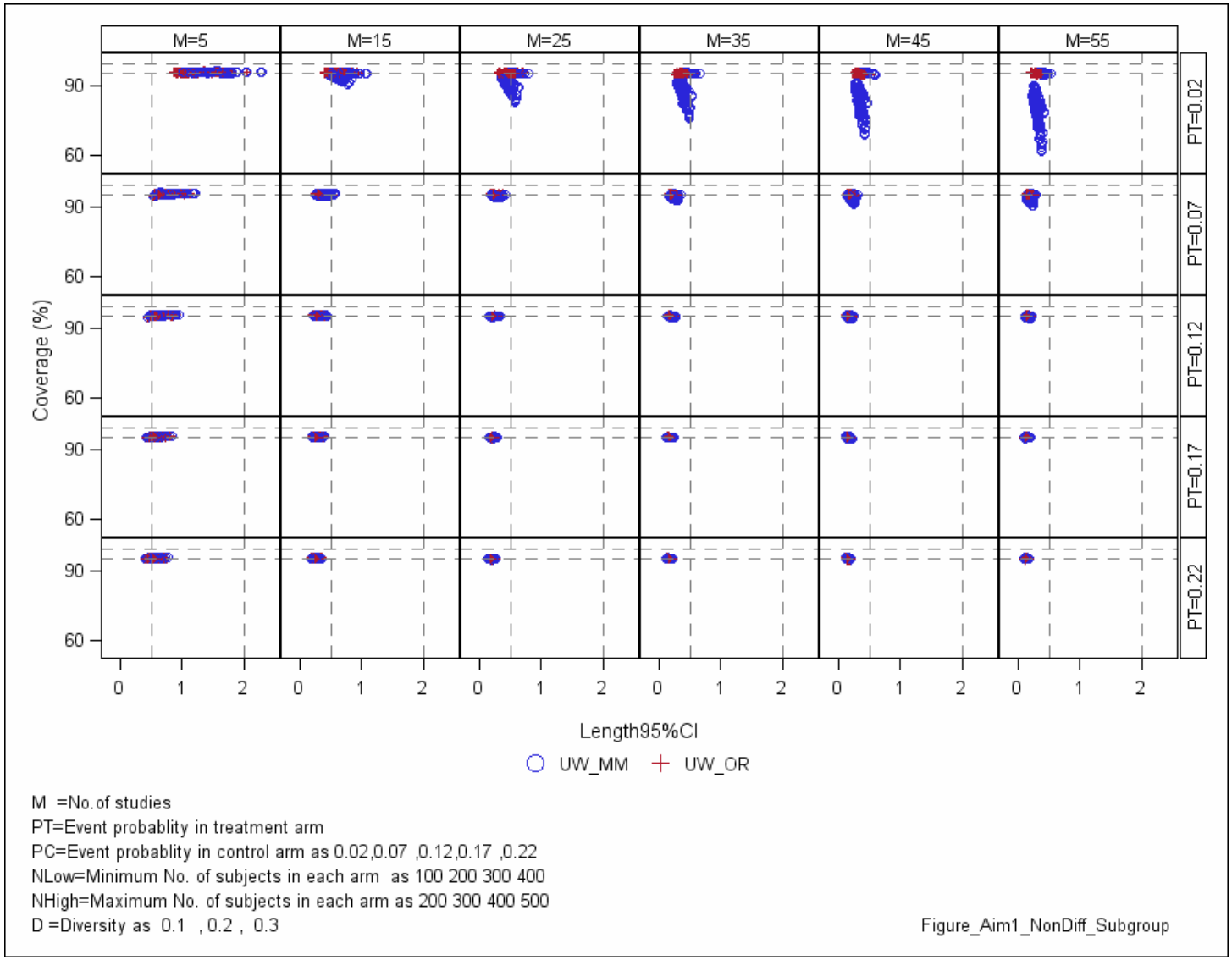

Figure 4-4. Aim 1: Comparisons of UW_OR vs. UW_MM in terms of both mean coverage and mean length 95% CI by event rate in treatment arm and number of studies

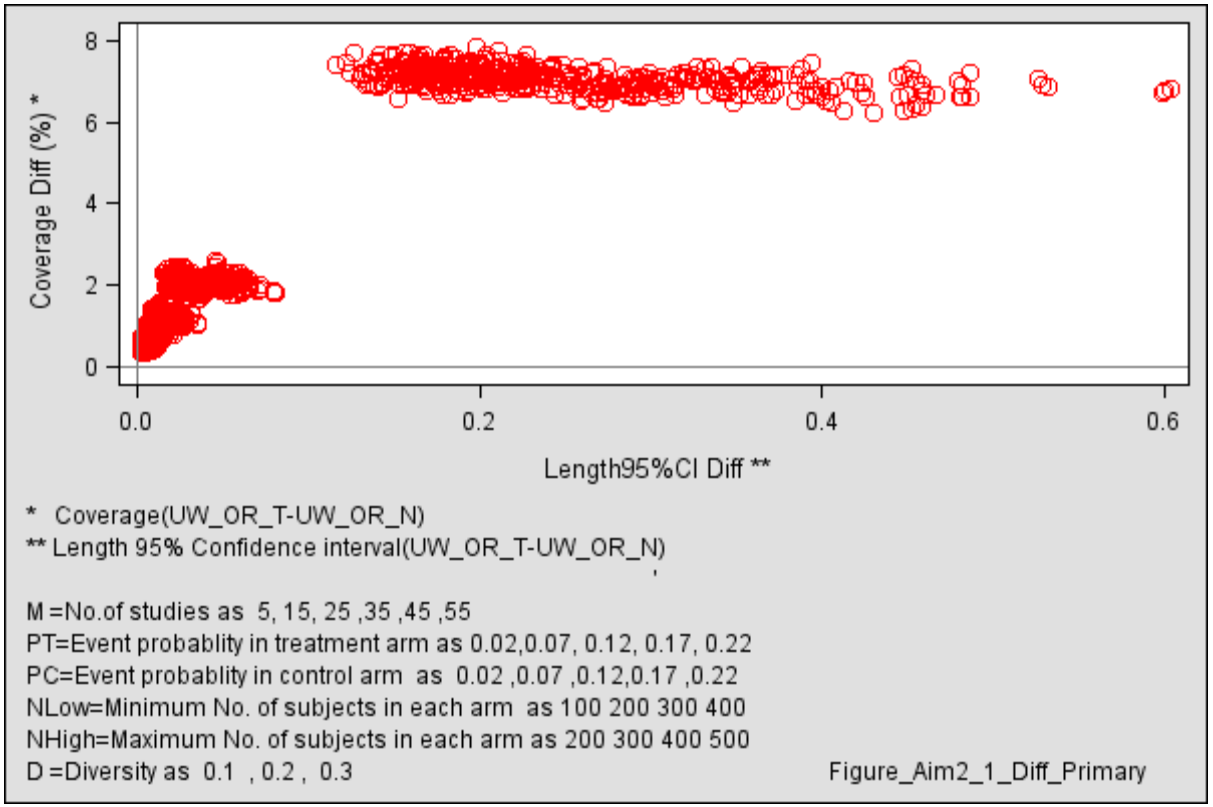

Figure 4-5. Aim 2-1: Comparisons of t distribution vs. normal distribution for UW_OR in terms of difference for both mean coverage and mean length 95% CI

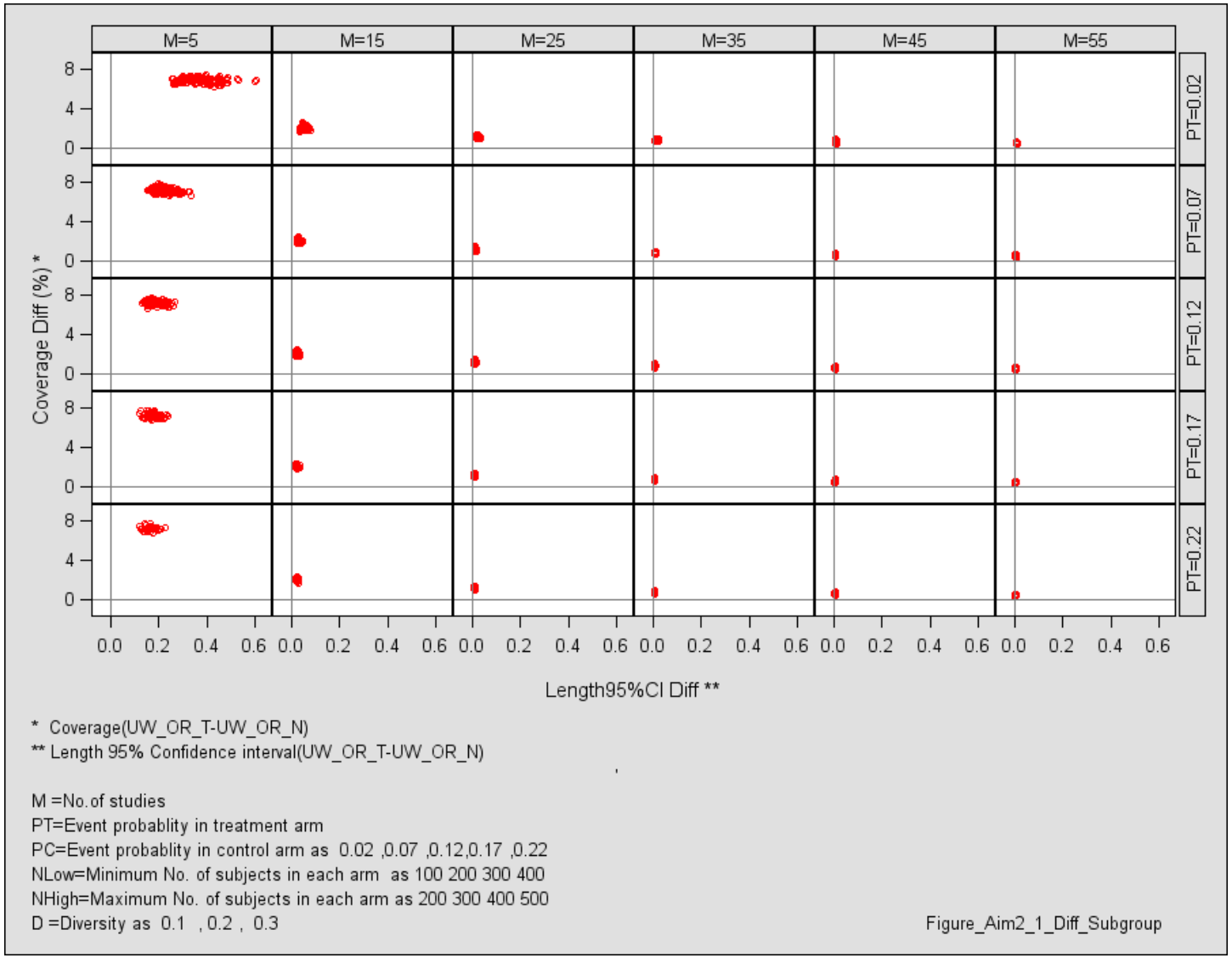

Figure 4-6. Aim 2-1: Comparisons of t distribution vs. normal distribution for UW_OR in terms of difference for both mean coverage and mean length 95% CI by event rate in treatment arm and number of studies

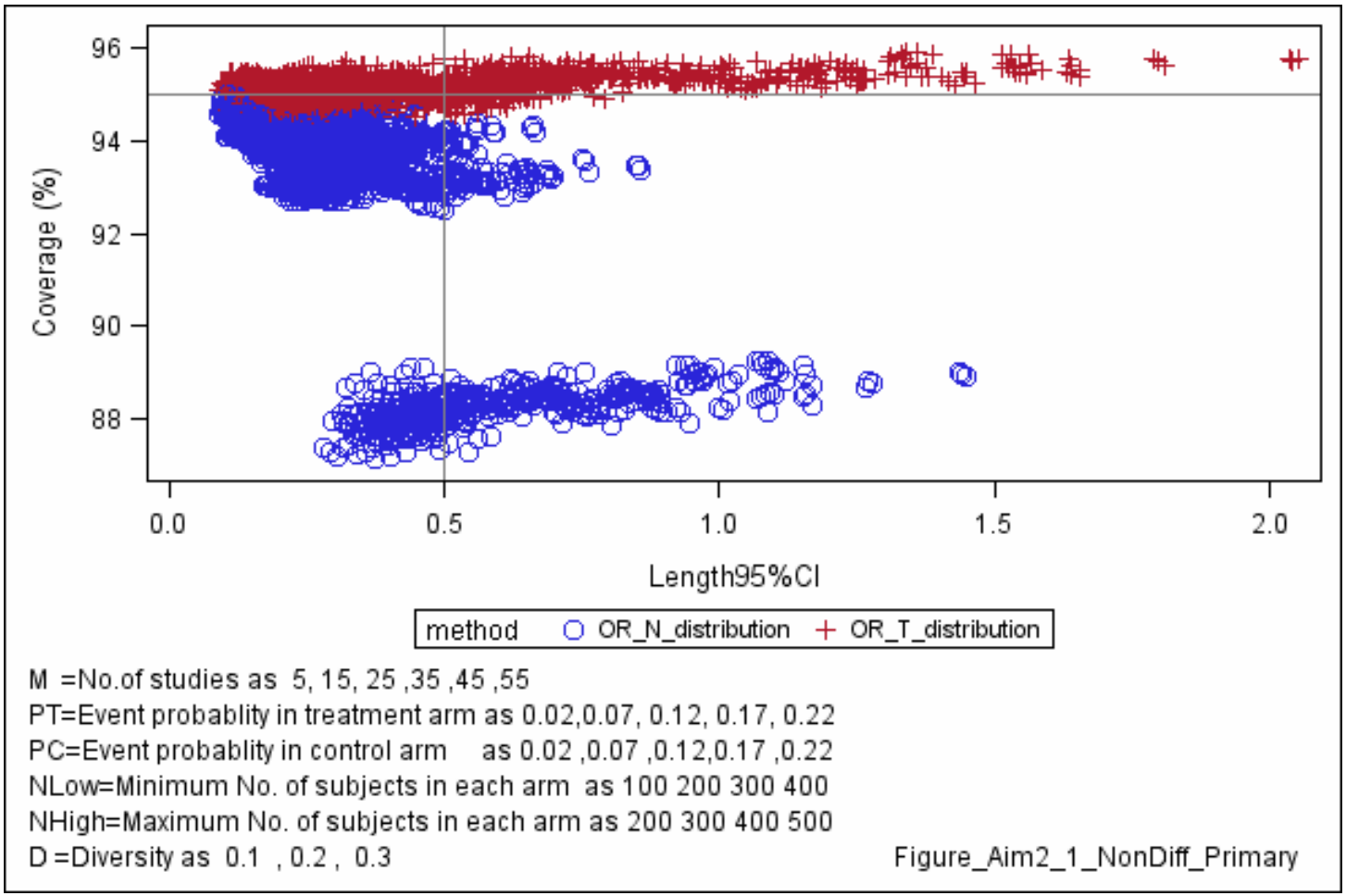

Figure 4-7. Aim 2-2: Comparisons of t distribution vs. normal distribution for UW_OR in terms of both mean coverage and mean length 95% CI

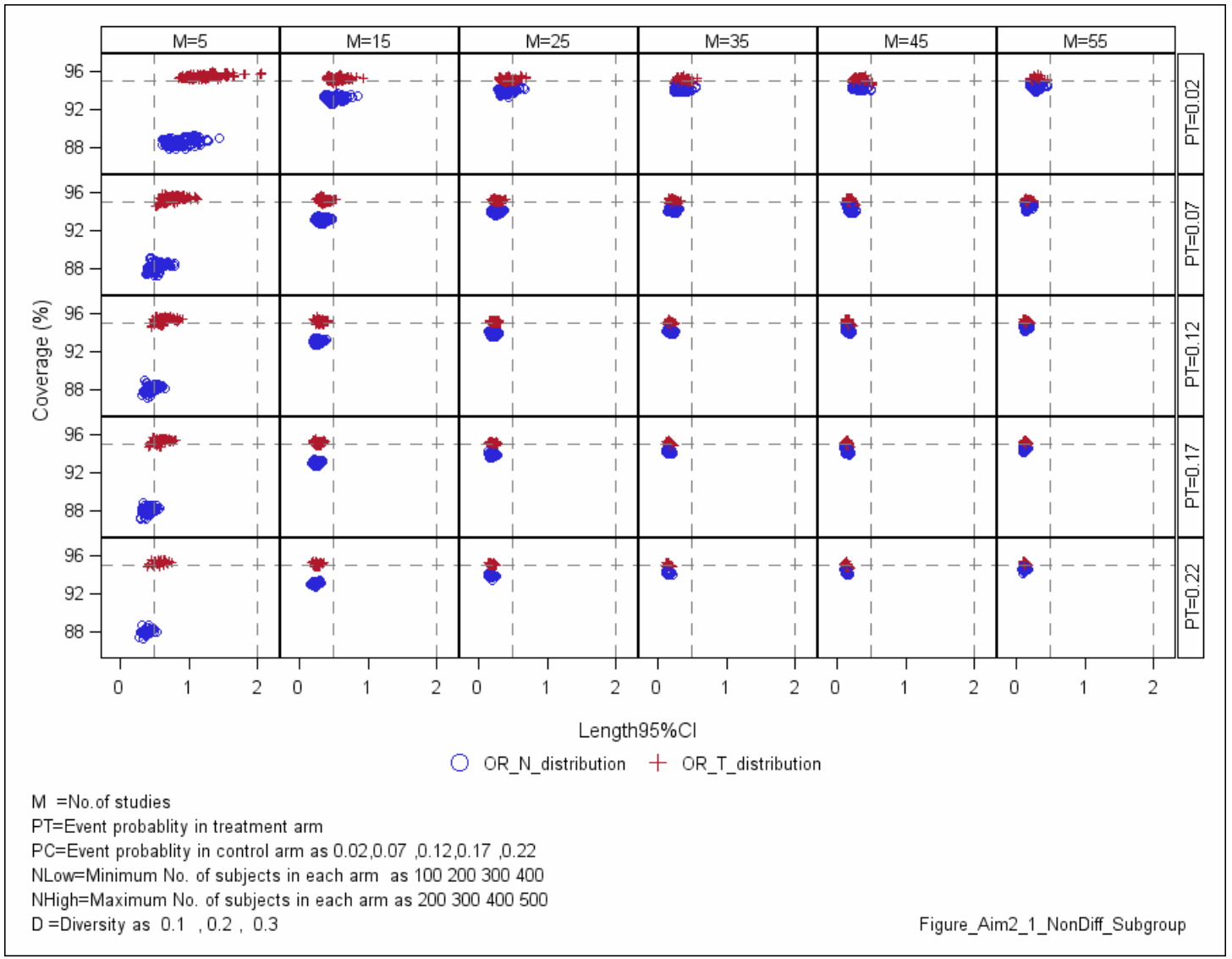

Figure 4-8. Aim 2-2: Comparisons of t distribution vs. normal distribution for UW_OR in terms of both mean coverage and mean length 95% CI by event rate in treatment arm and number of studies

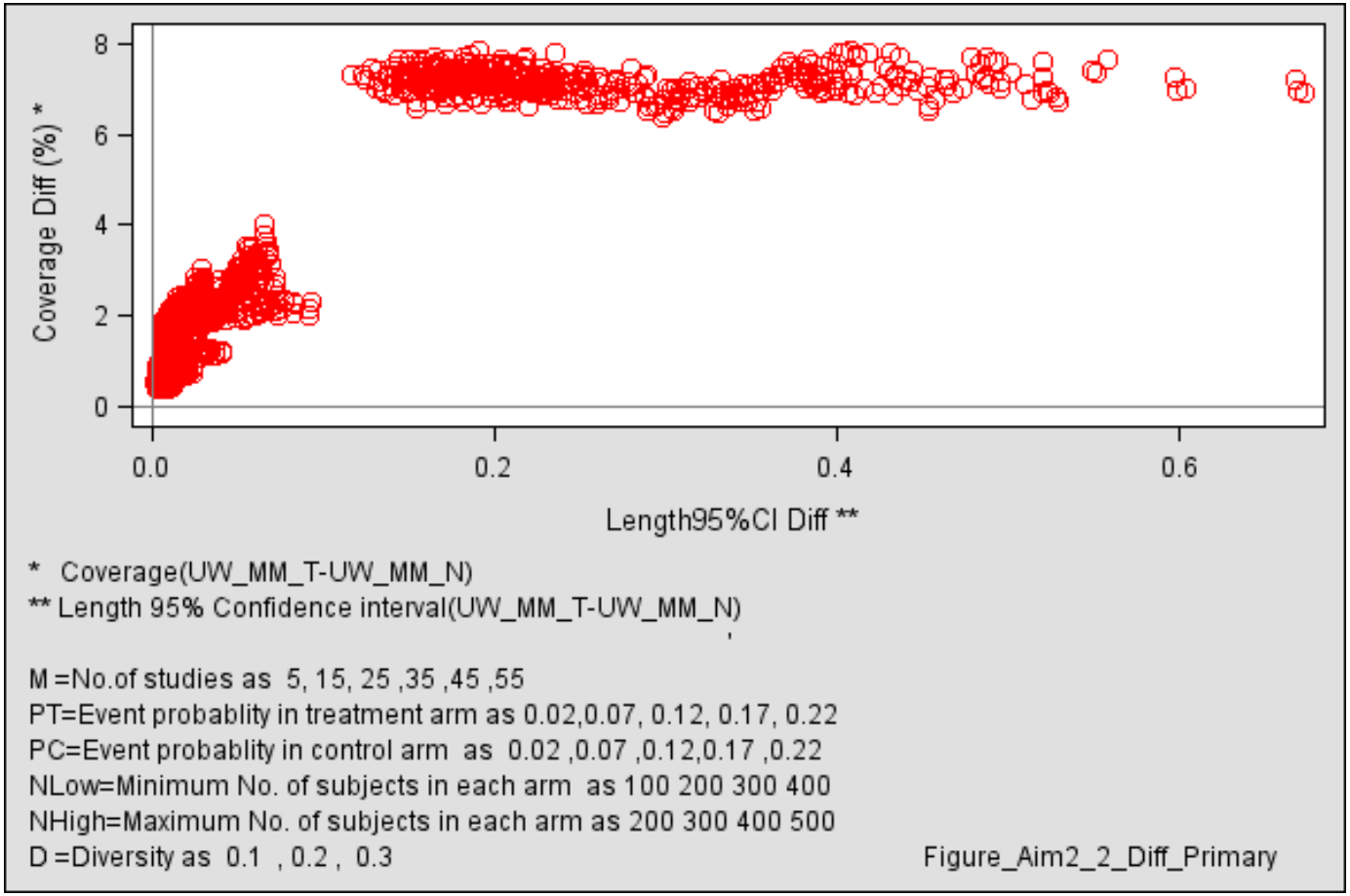

Figure 4-9. Aim 2-2: Comparisons of t distribution vs. normal distribution for UW_MM in terms of difference for both mean coverage and mean length 95% CI

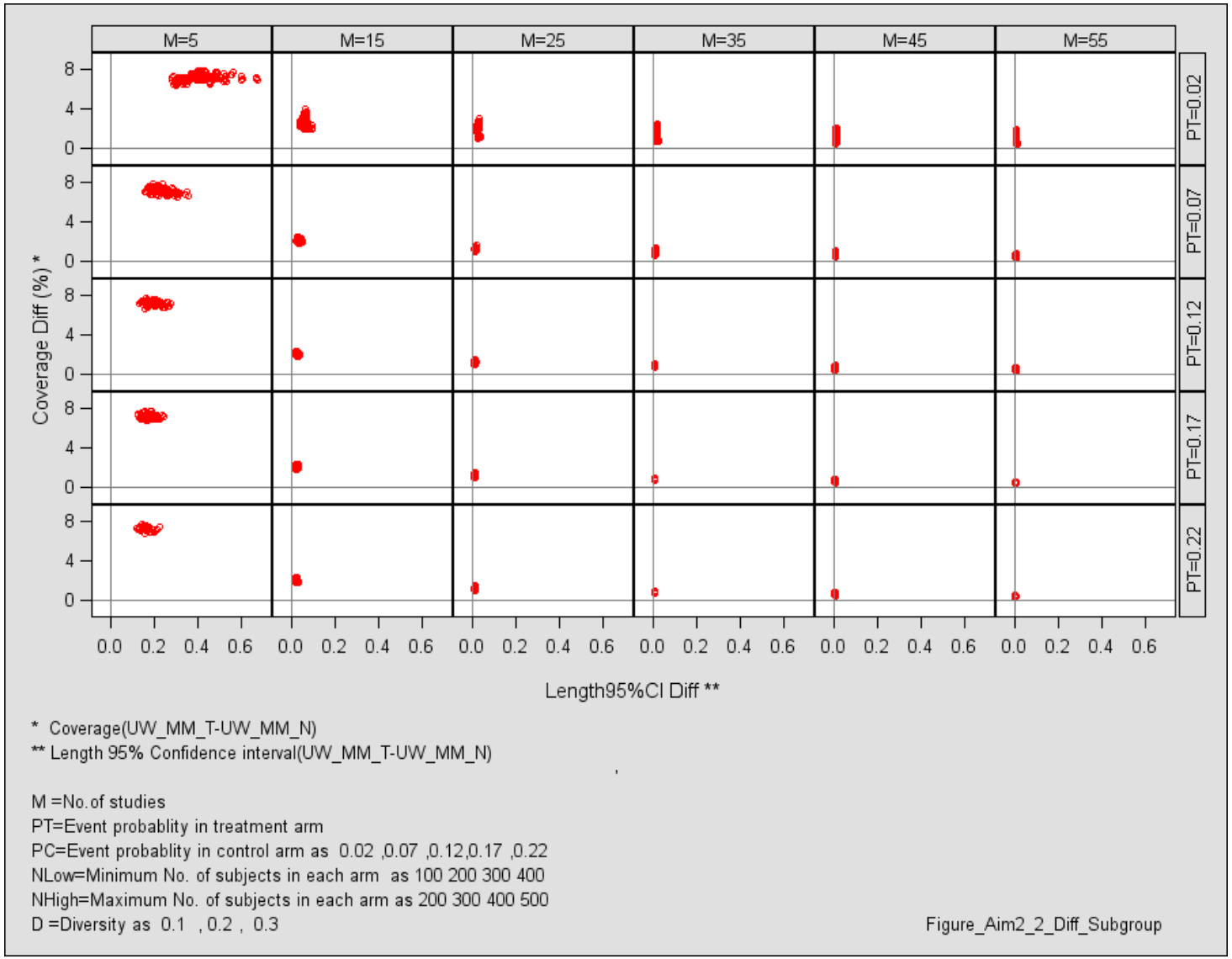

Figure 4-10. Aim 2-2: Comparisons of t distribution vs. normal distribution for UW_MM in terms of difference for both mean coverage and mean length 95% CI by event rate in treatment arm and number of studies

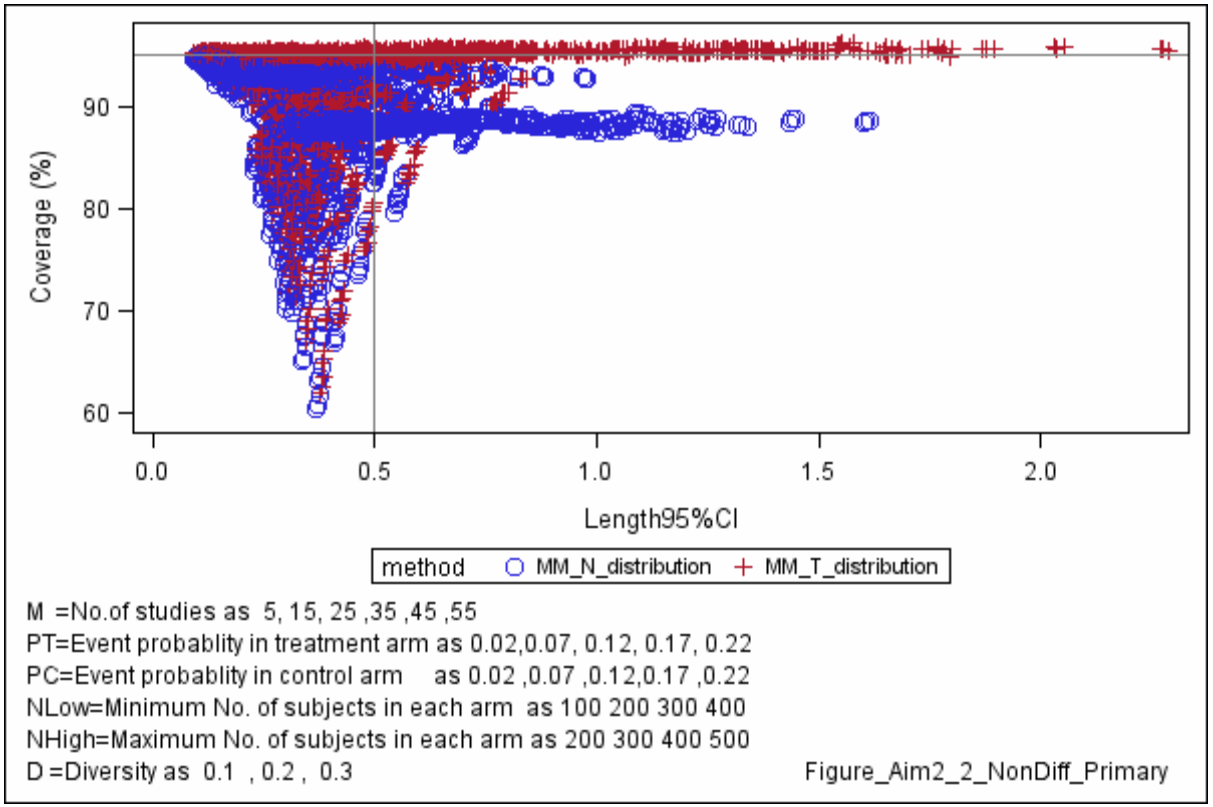

Figure 4-11. Aim 2-2: Comparisons of t distribution vs. normal distribution for UW_MM in terms of both mean coverage and mean length 95% CI

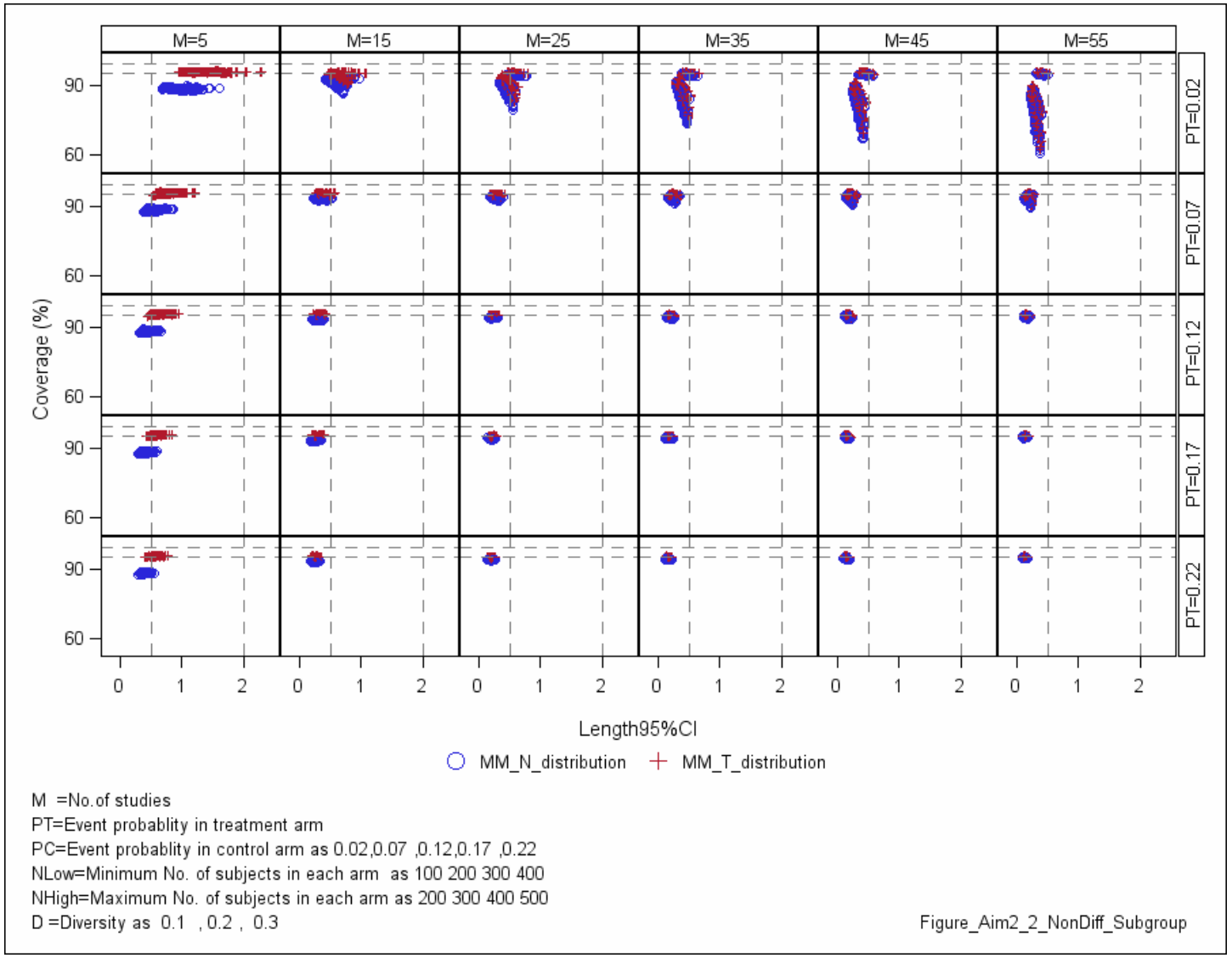

Figure 4-12. Aim 2-2: Comparisons of t distribution vs. normal distribution for UW_MM in terms of both mean coverage and mean length 95% CI by event rate in treatment arm and number of studies

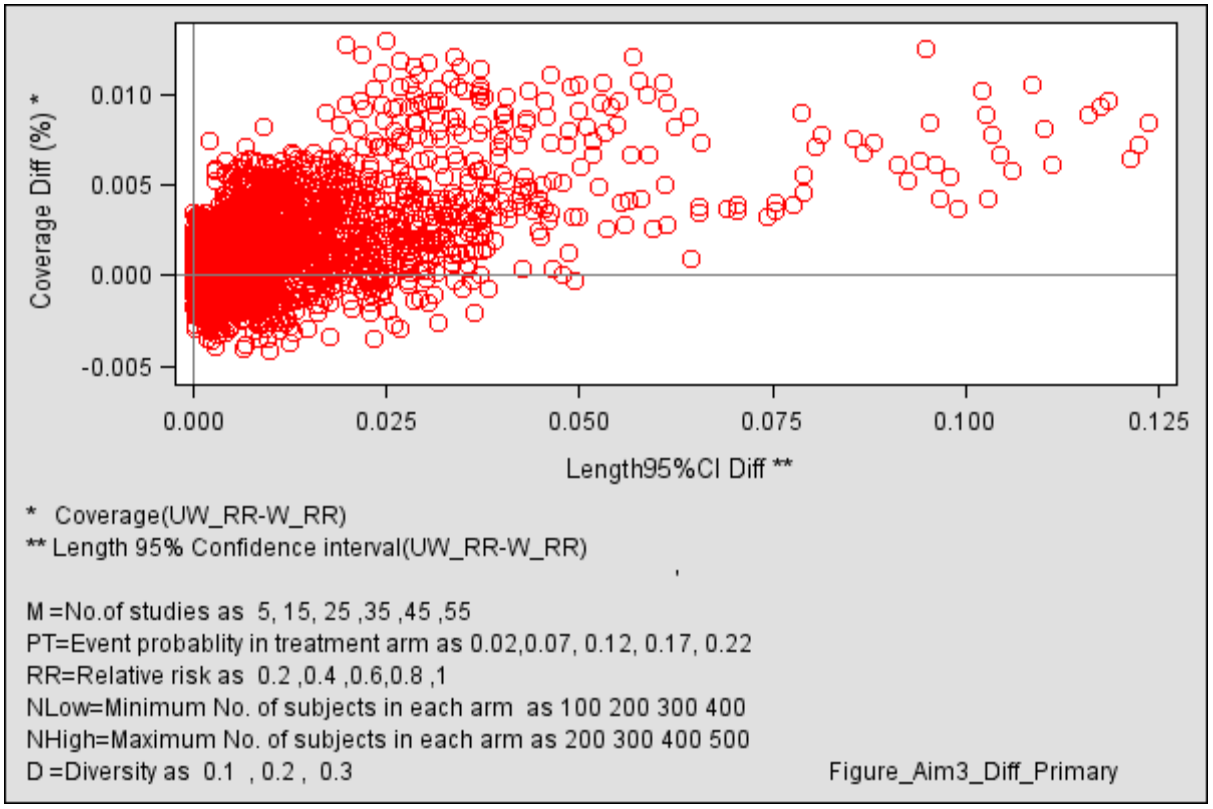

Figure 4-13. Aim 3: Comparisons of UW_RR vs. W_RR in terms of difference for both mean coverage and mean length 95% CI

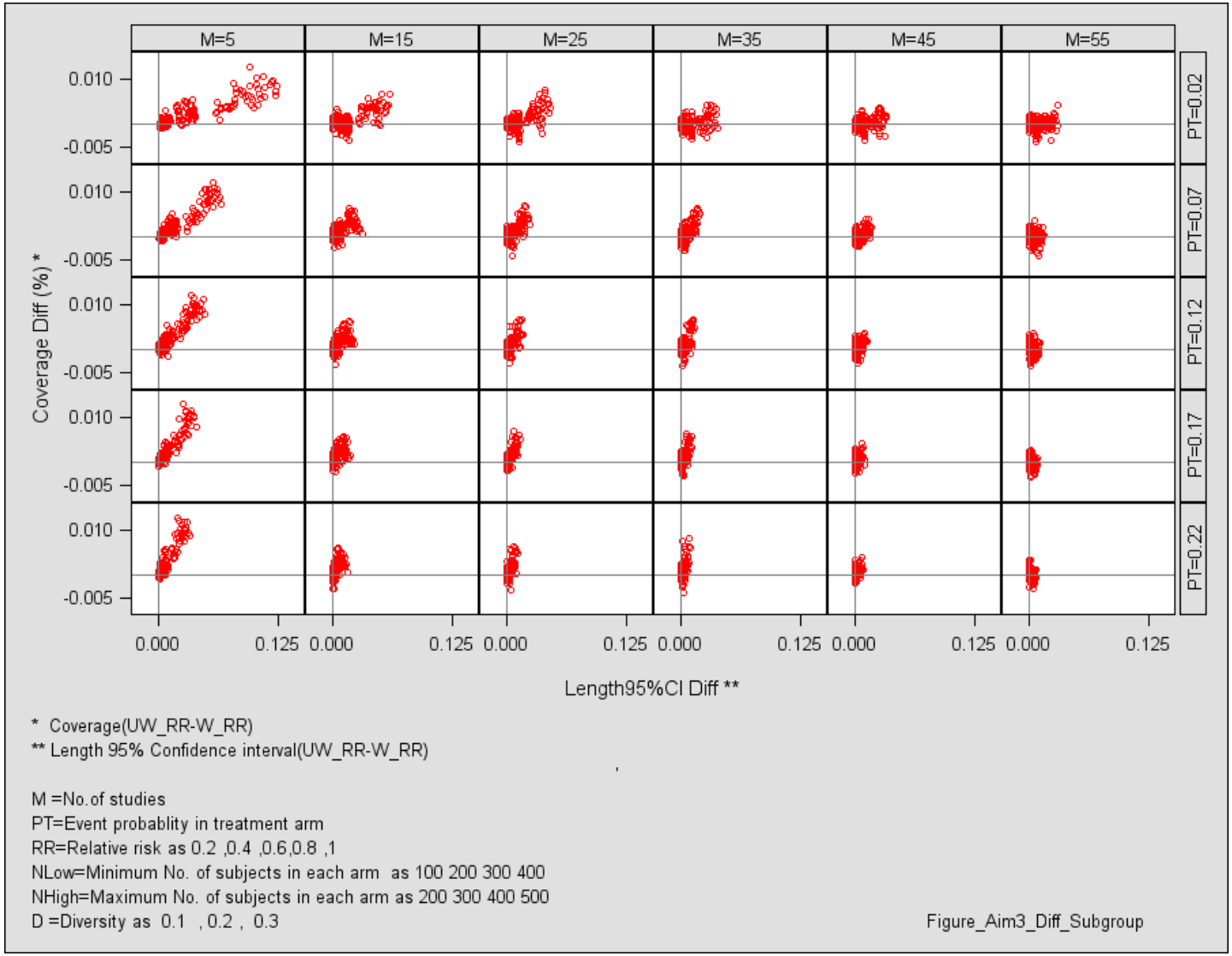

Figure 4-14. Aim 3: Comparisons of UW_RR vs. W_RR in terms of difference for both mean coverage and mean length 95% CI by event rate in treatment arm and number of studies

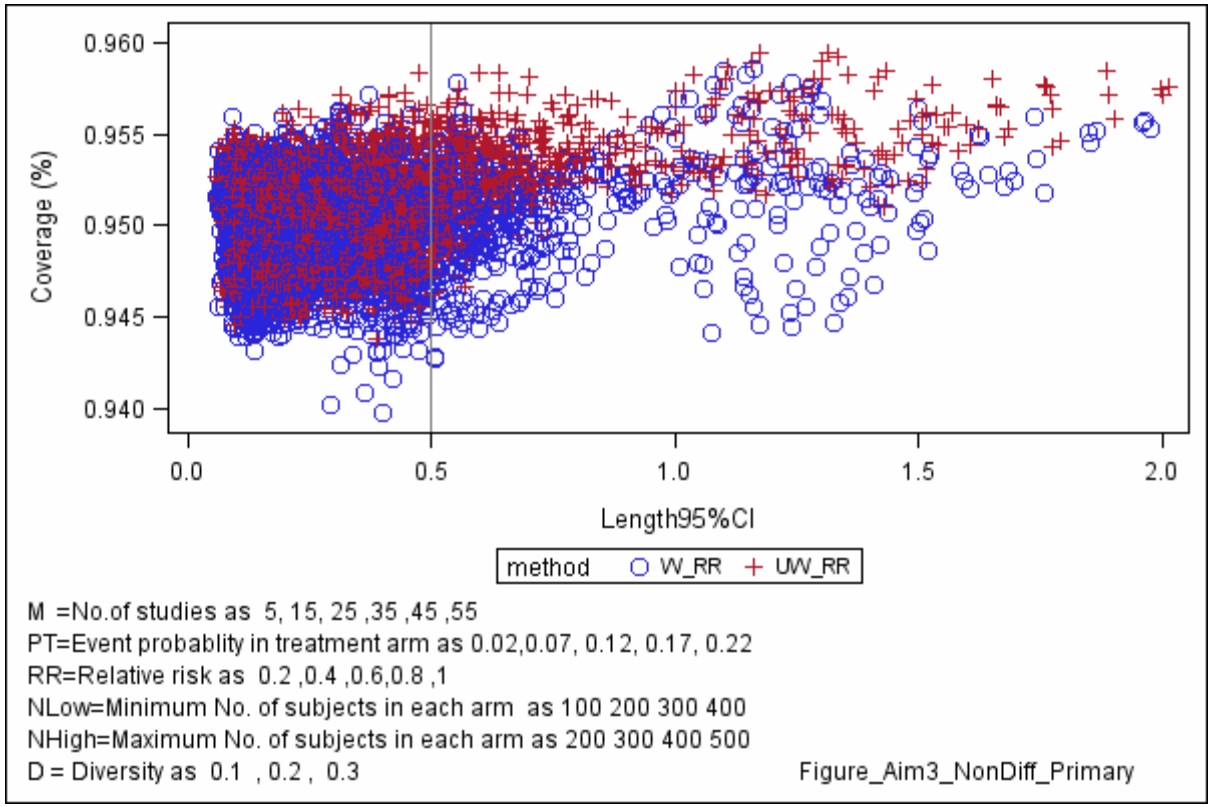

Figure 4-15. Aim 3: Comparisons of UW_RR vs. W_RR in terms of both mean coverage and mean length 95% CI

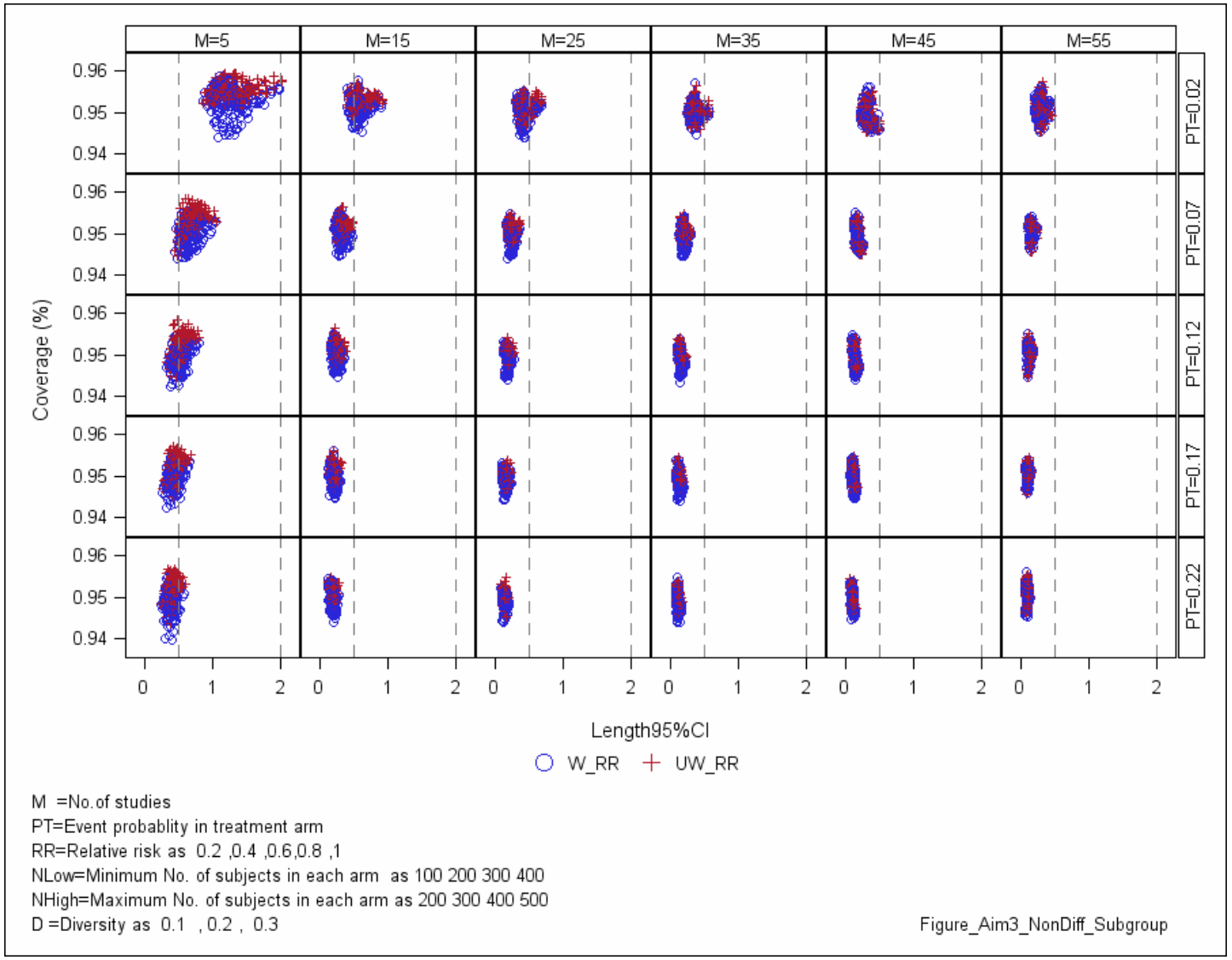

Figure 4-16. Aim 3: Comparisons of UW_RR vs. W_RR in terms of both mean coverage and mean length 95% CI by event rate in treatment arm and number of studies

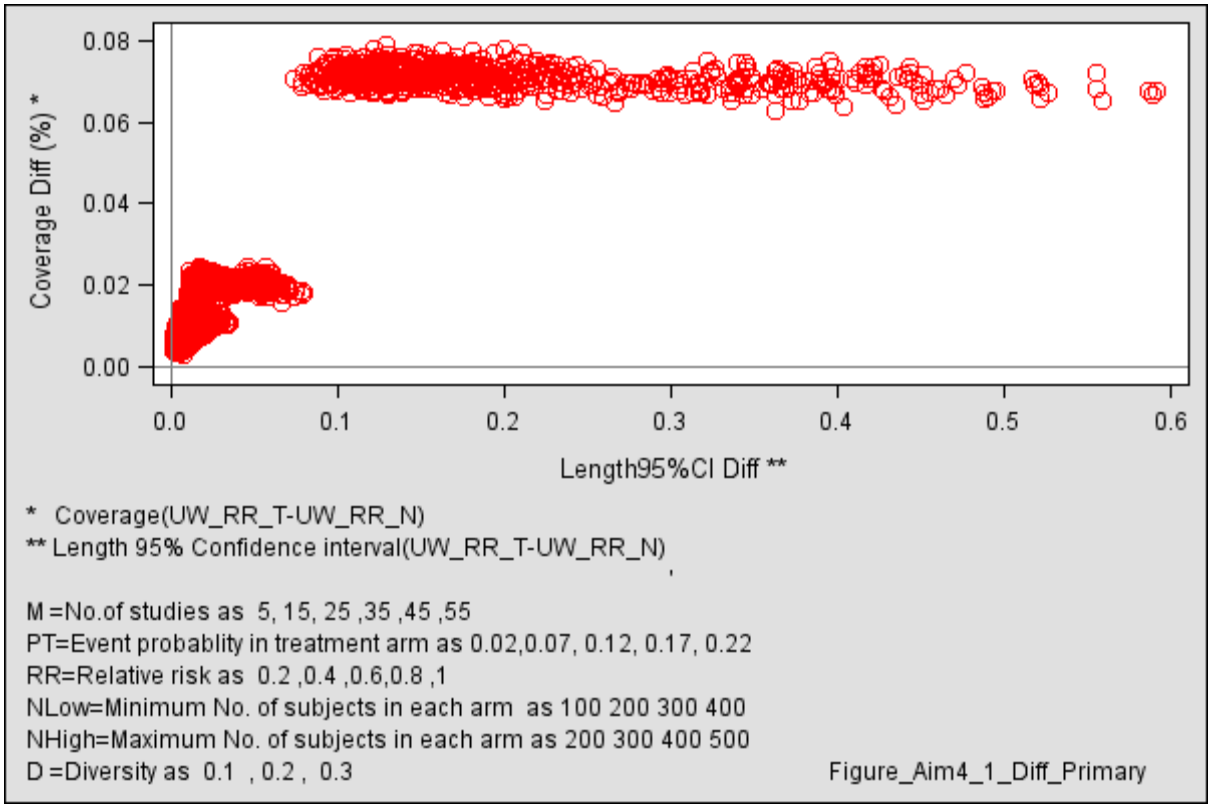

Figure 4-17. Aim 4-1: Comparisons of t distribution vs. normal distribution for UW_RR in terms of difference for both mean coverage and mean length 95% CI

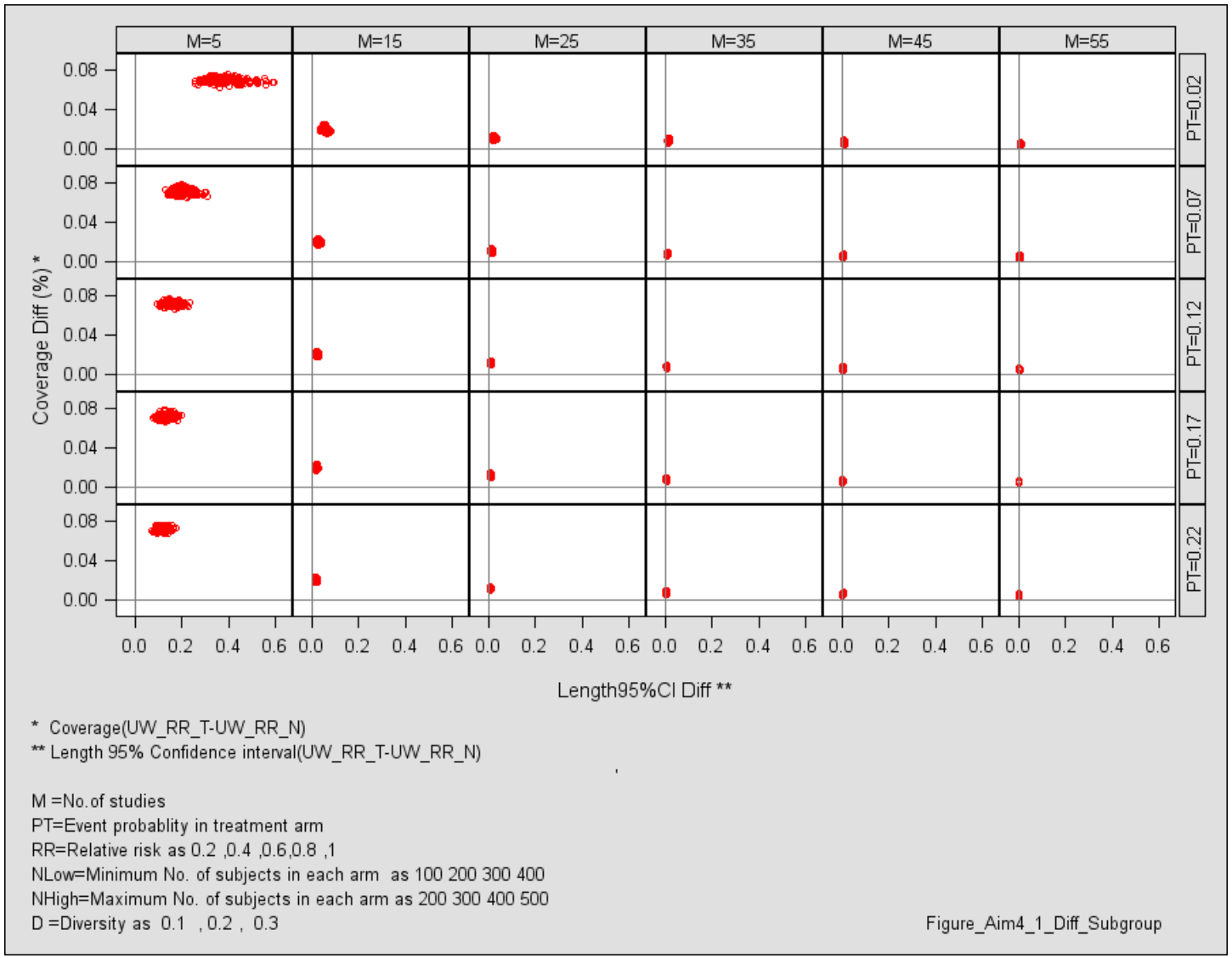

Figure 4-18. Aim 4-1: Comparisons of t distribution vs. normal distribution for UW_RR in terms of difference for both mean coverage and mean length 95% CI by event rate in treatment arm and number of studies

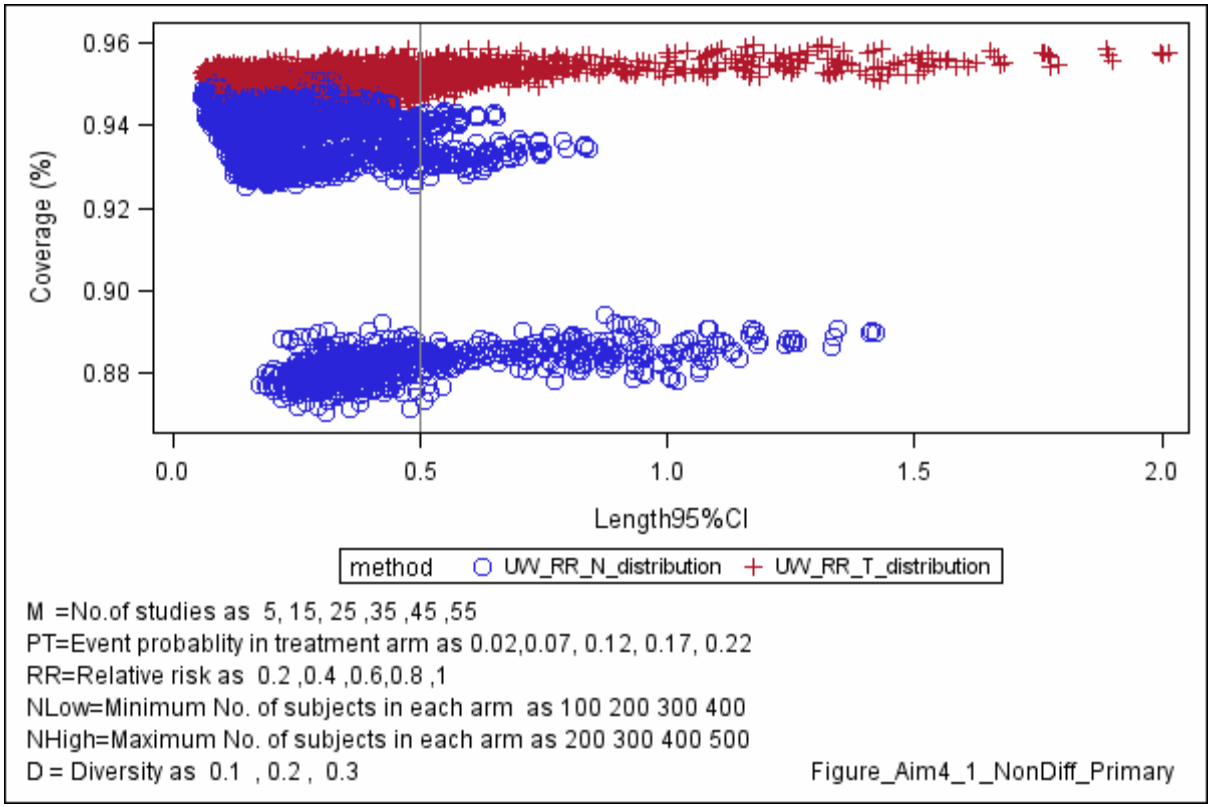

Figure 4-19. Aim 4-1: Comparisons of t distribution vs. normal distribution for UW_RR in terms of both mean coverage and mean length 95% CI

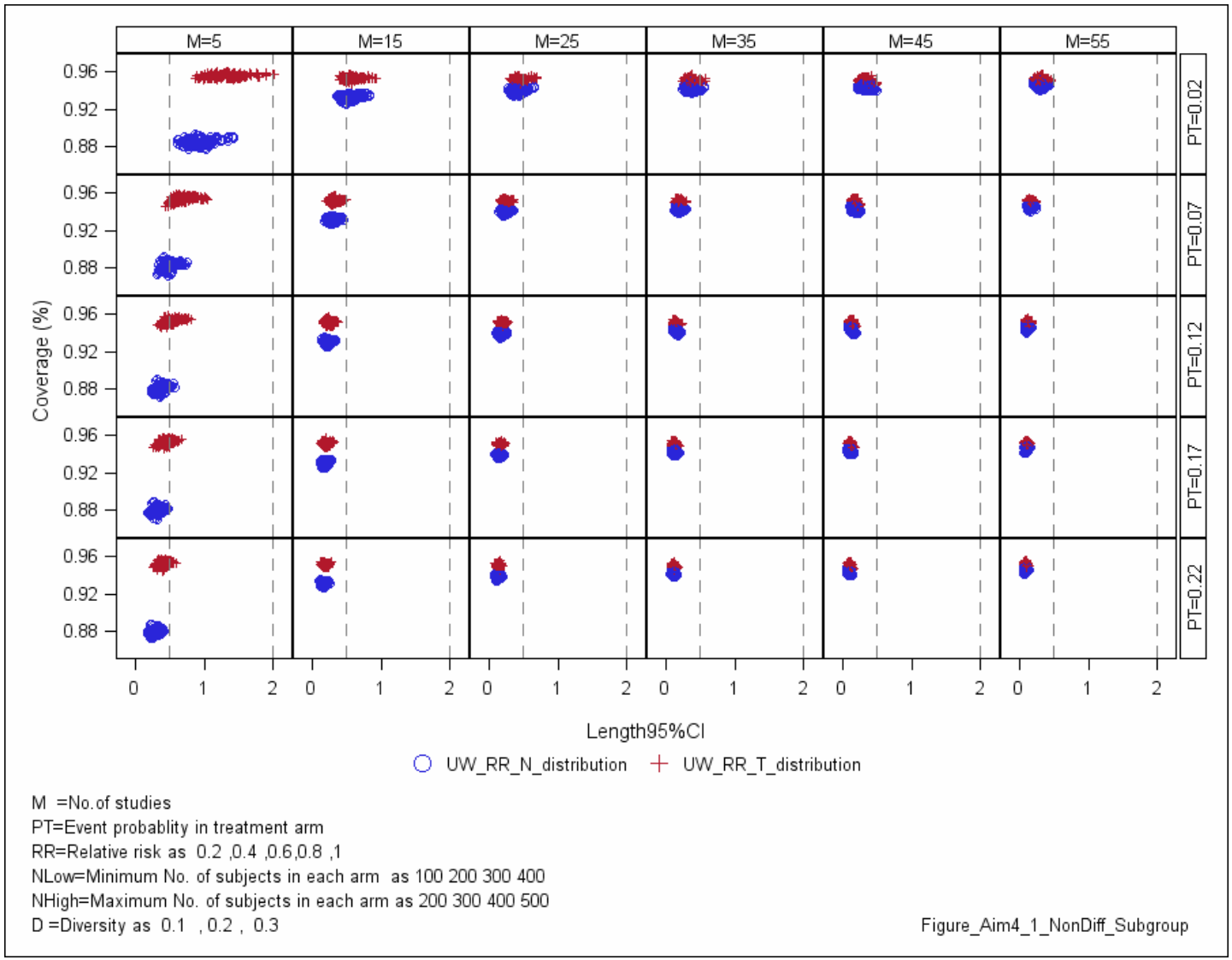

Figure 4-20. Aim 4-1: Comparisons of t distribution vs. normal distribution for UW_RR in terms of both mean coverage and mean length 95% CI by event rate in treatment arm and number of studies

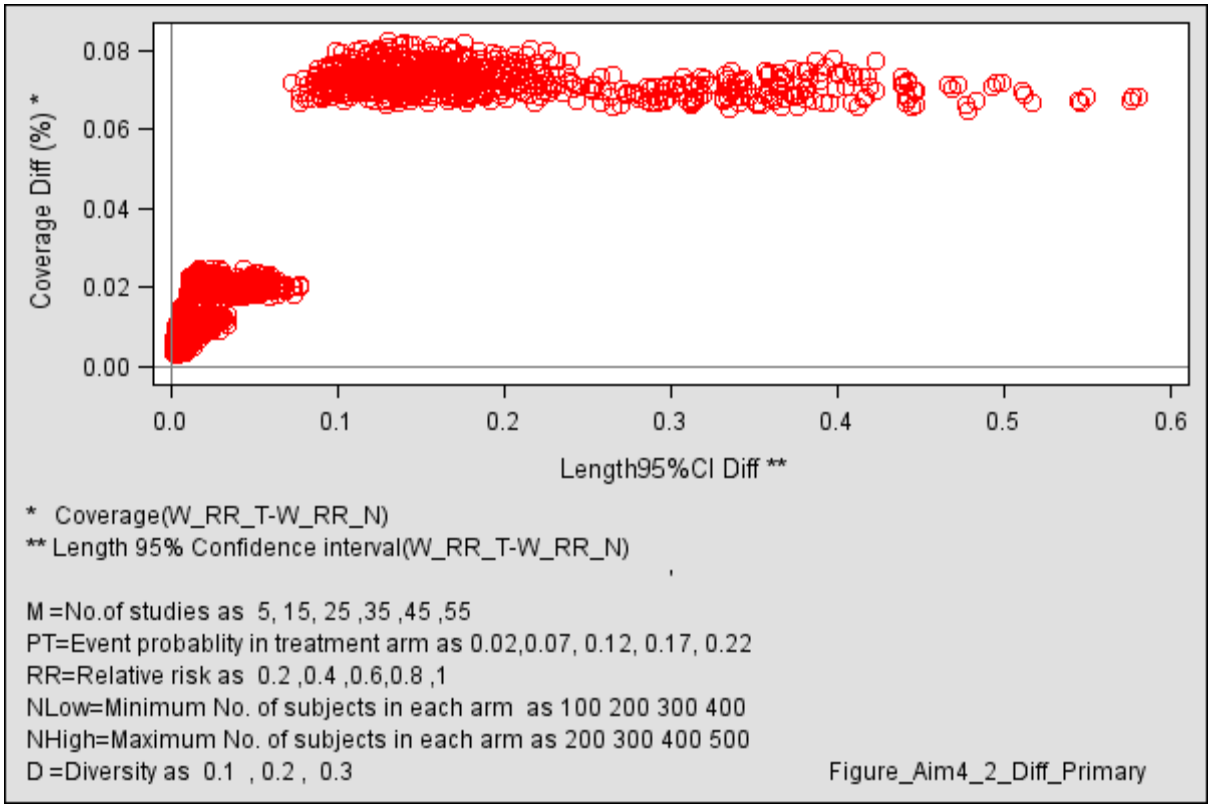

Figure 4-21. Aim 4-2: Comparisons of t distribution vs. normal distribution for W_RR in terms of difference for both mean coverage and mean length 95% CI

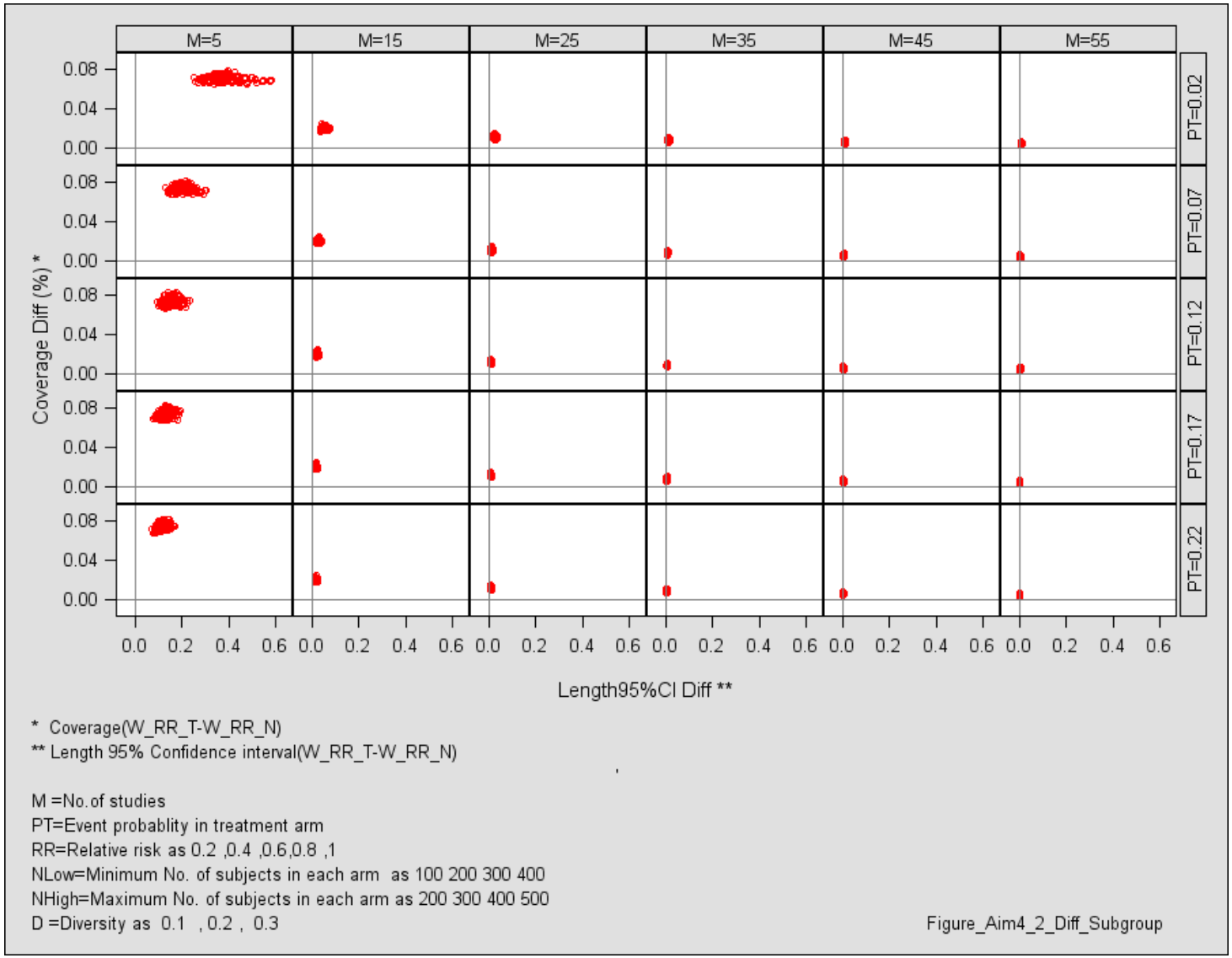

Figure 4-22. Aim 4-2: Comparisons of t distribution vs. normal distribution for W_RR in terms of difference for both mean coverage and mean length 95% CI by event rate in treatment arm and number of studies

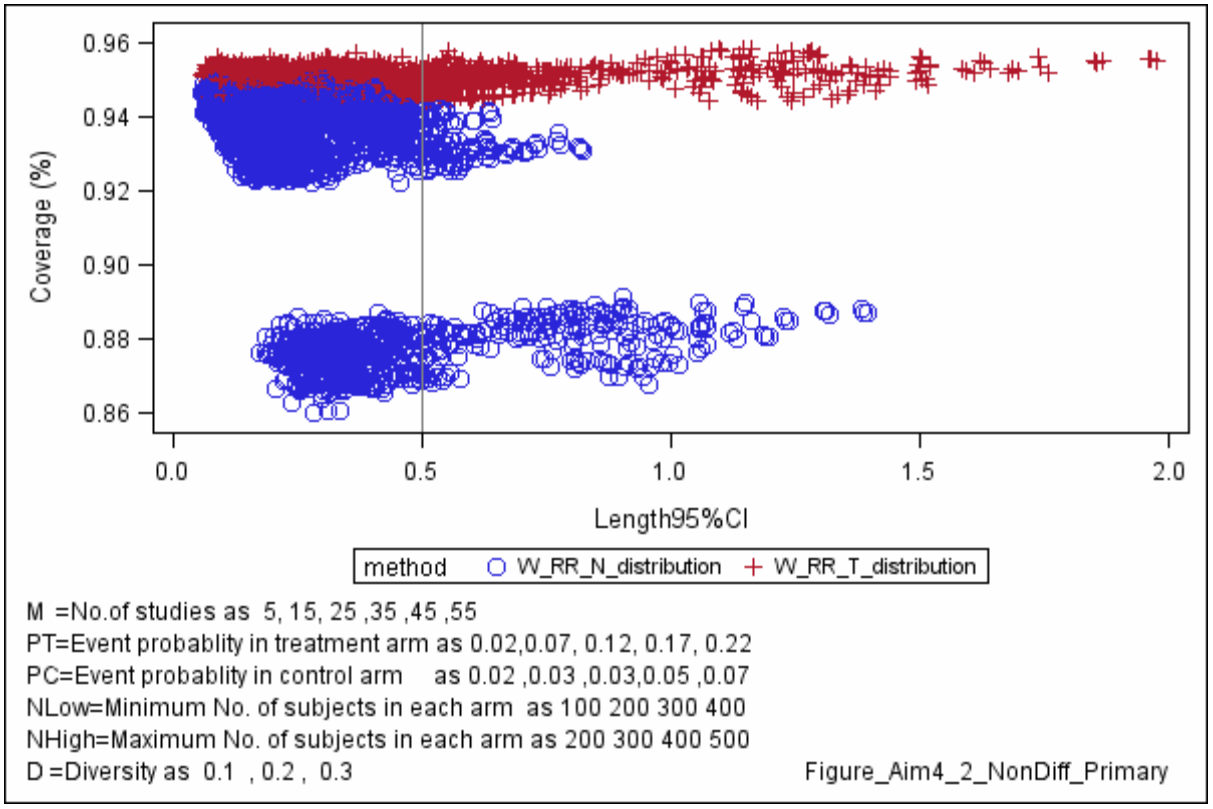

Figure 4-23. Aim 4-2: Comparisons of t distribution vs. normal distribution for W_RR in terms of both mean coverage and mean length 95% CI

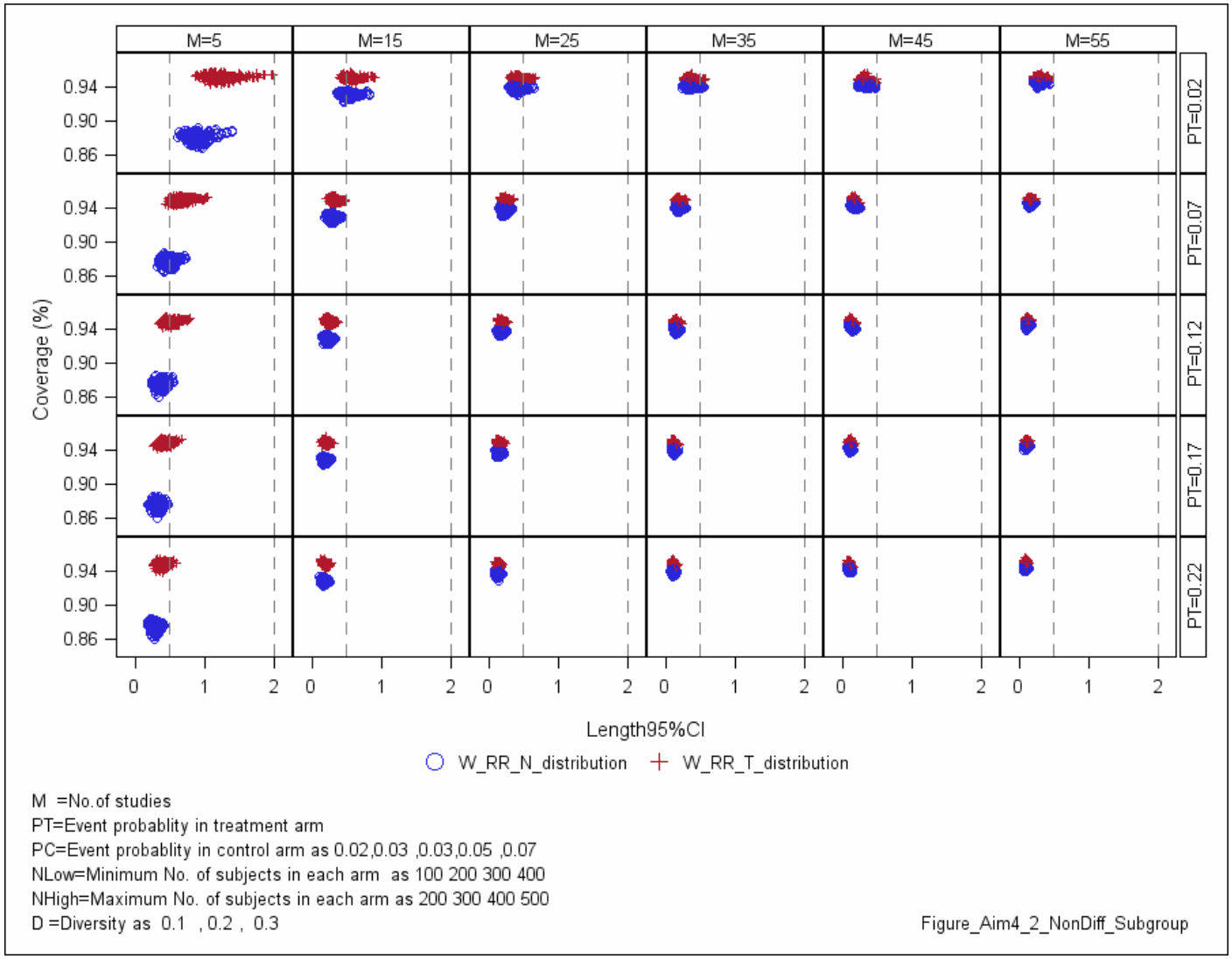

Figure 4-24. Aim 4-2: Comparisons of t distribution vs. normal distribution for W_RR in terms of both mean coverage and mean length 95% CI by event rate in treatment arm and number of studies

CHAPTER 5
DISCUSSION

To investigate post-marketing safety signals, researchers usually adopt three approaches, meta-analyses, large clinical safety trials, and observational studies. Due to the practical limitations of the clinical trials and observational studies [33], the evidence from a meta-analysis on existing clinical trials often is viewed as the highest authority, especially in making consensus on a controversial issue. The FDA first issued regulatory guidelines to industry on conducting a meta-analysis on clinical trials, which were especially for cardiovascular risks of antidiabetics in 2008 [34]. In March, 2011, the FDA co-sponsored the Drug Information Association (DIA) conference on the DIA/FDA best practices for regulatory information synthesis of randomized controlled trials for product safety evaluation [35]. Considering the highest authority given to evidence arrived from a meta-analysis, selecting the appropriate meta-analytical method is imperative and important.

This chapter is organized as follows. First, we discuss the specific findings. Second, we elaborate in depth regarding the statistical meta-analysis methods. Third, we provide an empirical example using the UW_RR and the W_RR as compared with the empirical weighted method (W_DL). Fourth, we highlight the strengths of this study. Fifth, we address the limitations of this study. Sixth, we identify the need for further studies. At the end, we draw our conclusions based upon this study.

Specific Findings

Aim 1

Findings for hypothesis 1. When a small number of studies, such as five, are combined for a meta-analysis, the results of both approaches of the UW_OR and the

UW_MM are similar regardless of event rate. However, at low event rates in the treatment arm (PT), for example, PT=0.02, as the number of studies (M) increases, the mean coverage of the UW_MM drops, whereas at high event rates in the treatment arm (PT), such as PT=0.22, the UW_MM does not drop apparently. The reason is that the individual study estimates of the odds ratio are biased at low event rates. This bias remains the same regardless of the number of studies combined for a meta-analysis study. As the number of studies (M) goes up, the standard errors get smaller. Therefore, the confidence intervals are tightened so that the 95% CIs are more likely to exclude the global true OR value. At M=5, the 95 % CIs are wider, and are more likely to cover the global true value of OR than at M=55.

The UW_OR provides a robust estimate of the summary effect size in that the UW_OR estimates the summary effect size using first the summation of each proportion for each arm in each study, i.e. risk for each arm in each study, then doing ratio for the relative risk or the odds ratio on the summary estimates of proportions. This provides evidence of the robustness of the UW_OR when applied to a large number of studies contributing to a meta-analysis. Since each study is weighted equally, no single trial dominates this unweighted analysis. This also lessens, but does not eliminate, concerns about publication bias involving a small fraction of missed studies. In responses to two letters to *Statistics in Medicine* by Carpenter et al. [36] and Rucker et al. [29], Shuster et al. [37] demonstrated, at least empirically, that the UW_OR is robust to both (a) exclusion of a small number of trials [37] and (b) to early termination of a large trial [29].

As for a meta-analysis, in order to have more accurate estimates of summary effect sizes, the large sample theory comes into play in two ways. One is a large sample

size for each individual study. The other is a large number of studies combined for conducting a meta-analysis. The theory explains that the efficiency of both the UW_OR and the UW_MM is getting closer as the number of studies increases or event rates increase.

Implication for findings 1. There is no precision difference between the UW_OR and the UW_MM when rare events are unlikely regardless of the number of studies.

Aim 2

Findings for hypothesis 2-1. The UW_OR is developed with the assumption that the summary effect size is a t distribution instead of the widely used normal distribution. Therefore, the 95% CI is calculated using the t cutoff point with the number of studies (M) -1 degrees of freedom. When M is small, the t score is large (appendix A). Therefore, the 95% CI is wider, leading to the more accurate coverage of 95%. On the other hand, the UW_OR with a normal distribution has a narrower 95% CI which leads the underestimation of the coverage. However, as the number of studies increases, the t score decreases towards the normal z score (appendix A). This is the reason that the UW_OR with a t distribution is close to the UW_OR with a normal distribution as the number of studies combined for a meta-analysis increases.

Implication for findings 2-1. The UW_OR with a t approximation is more accurate, which provides the mean coverage closer to the expected 95% coverage than the UW_OR with a normal approximation when a small number of studies are combined for a meta-analysis regardless whether or not there are low event rates.

Findings for hypothesis 2-2. The UW_MM is developed with the assumption that the summary effect size is with a t distribution instead of the widely used normal distribution. Therefore, the 95% CI is calculated using the t cutoff point with the number

of studies (M) -1 degrees of freedom. When M is small, the t score is large. Therefore, the 95% CI is wider, leading to the more accurate coverage of 95%. On the other hand, the UW_MM with a normal distribution has a narrower 95% CI which leads to the underestimation of the coverage. However, as the number of studies increases, the t score decreases towards the normal z score (appendix A). This is the reason that the UW_MM with a t distribution is close to the UW_MM with a normal distribution as the number of studies increases.

As we discussed in aim 1, the 95% coverage for the UW_MM with a t distribution drops dramatically when the event rate is low in that the estimates of the odds ratios from individual studies are biased at low event rates. This bias remains the same regardless of the number of studies combined in a meta-analysis study. Moreover, as M increases, the t score and standard errors become smaller. Therefore, the 95% CIs are tightened so that they are more likely to exclude the true value of OR.

Implication for findings 2-2. The UW_MM with a t distribution is more accurate, providing the mean coverage closer to the expected 95% coverage than the one with a normal distribution when a small number of studies are combined for a meta-analysis and rare rates are unlikely.

Aim 3

Findings for Hypothesis 3. The study shows that the method with the sample size weighting on the proportion (W_RR) has similar summary effect size as compared with the UW_RR but the W_RR shows shorter mean length of the 95% CI for certain number of scenarios. We know that these two approaches do not estimate the same summary effects, which the UW_RR estimates the unweighted summary effect while the W_RR estimates the weighted summary effect. In this study, we forced two

methods to estimate the same summary effect size by assuming that there is no association between the effect size and the sample size.

One important observation centers on the rate of convergence of the asymptotic distributions for the weighted (W_RR) vs. the unweighted method (UW_RR). While we cannot mathematically prove superior convergence, the central limit theorem works more rapidly when the risk of outliers is relatively low. The UW_RR uses the bivariate summary proportions, which are relatively low, and unlikely to be prone to large outliers. On the other hand, the W_RR uses the adjusted number of events in its ratio estimate of relative risk. As such, these will have much higher variability, since large scale trials tend to have many more adverse events than smaller ones.

Implication for Finding 3. As a special case, the UW_RR and the W_RR estimate the same summary effect size when there is no association between the effect size and the sample size. But the W_RR is more precise in terms of the shorter 95% CI. As the number of studies combined increases, the precision of both approaches is increased simultaneously.

Aim 4

Findings for hypothesis 4-1. The UW_RR, like the UW_OR, is developed with the assumption that the summary effect size is a t distribution instead of the widely used normal distribution. Therefore, the 95% CI is calculated using the t cutoff point with the number of studies (M) -1 degrees of freedom. When M is small, the t score is large (Appendix A). Therefore, the 95% CI is wider, leading to the more accurate coverage of 95%. On the other hand, the UW_RR with a normal distribution has a narrower 95% CI which leads to underestimate the coverage. However, as the number of studies increases, the t score decreases towards the normal z score (appendix A). This is the

reason that the UW_RR with a t distribution is close to the UW_RR with a normal distribution as the number of studies combined for a meta-analysis increases.

Implication for findings 4-1. The UW_RR with a t approximation is more accurate, providing the mean coverage closer to the expected 95% coverage than the UW_RR with a normal approximation when a small number of studies are combined for a meta-analysis regardless whether or not there are low event rates.

Findings for Hypothesis 4-2. The W_RR is developed with the assumption that the summary effect size is a t distribution instead of the widely used normal distribution. Therefore, the 95% CI is calculated using the t cutoff point with the number of studies (M) -1 degrees of freedom. When M is small, the t score is large. Therefore, the 95% CI is wider leading to the more accurate coverage of 95%. On the other hand, the W_RR with a normal distribution has a narrower 95% CI which leads to the underestimation of the coverage. However, as the number of studies increases, the t score decreases towards the normal z score (appendix A). This is the reason that the W_RR with a t distribution is close to the W_RR with a normal distribution as the number of studies combined for a meta-analysis increases.

Implication for Findings 4-2. The W_RR with a t approximation is more accurate, and provides the mean coverage closer to the 95% coverage than the one with a normal approximation when a small number of studies are combined for a meta-analysis regardless whether or not there are low event rates.

Meta-Analysis in General

As mentioned in Chapter II (Review of the Literature), the most widely used meta-analysis method is the random-effects model.

Under the umbrella of the random-effects model, two categories of weighted and unweighted approaches are frequently used, both of which are non-Bayesian. In regards to the weighted random-effects approaches, there are the empirical DerSimonian-Laird approach with weights derived from that study's variance [17], and the new approach developed by Shuster et al. [6] with the total individual study'sample size weighting on the proportions. There are also two unweighted random-effects approaches, one with the summarized proportions in the form of odds ratio or relative risk [7, 9] and one with the mean of means in the form of odds ratio [8].

In this subsection, we focus on issues related to the random-effects model in depth. We start with the weighting issue and the estimated summary effect size. Next, we discuss the low event rates and continuity adjustment when zero event cells occur. Thirdly, we address asymptotic properties when a small number of studies are combined for a meta-analysis. Finally, we discuss the three typical metrics for binomial outcomes, odds ratio, relative risk and risk difference.

Weighting Issues and Summary Effect Size Estimates

The weights for the empirical random-effects model (W_DL) are volatile random variables, because the weights are derived from both the within-study variance and the between-study variance [5]. The W_DL needs to be avoided in low event rate applications for following two reasons. First, it is unclear as to what the W_DL is estimating, given that the weights are random variables but treated as non-random. Second, in low event rate applications, the effect sizes and weights probably are correlated [8, 38]. For example, a) In drug development, early smaller studies may be pure (drug only vs. placebo), while later larger studies may use the drug vs. placebo in an adjuvant setting. Therefore, larger studies expect to have a smaller difference in

efficacy. b) For side effects, however, an adjuvant therapy interaction with the experimental drug may trigger a larger differential, yielding the opposite correlation. c) Better designed studies may lower the sampling error, thereby increasing the weight, while the greater skills of these investigators may lead to larger advantages for the experimental therapy over the controls. This is especially problematic in surgery device trials. d) Unknown to a meta-analyst, some studies may have been terminated early for efficacy, yielding smaller weights than those that run to completion. These arguments suggest that the correlations of weights and effect sizes can be expected, and it is statistically risky to assume they do not exist, especially at low event case. A positive correlation may produce a positive bias and a negative correlation may produce a negative bias.

The W_RR is one solution to overcome the issues from the empirical approach (W_DL), which the weights are derived from the sample size, a random variable. The weights are applied to the proportion before the summary effect size is calculated and they are treated as random variables. In contrast, some sample size based methods such as Emerson et al. [39, 40], use weights proportional to the product of the sample size in each arm divided by the sum of the sample sizes of both arms. This implies that more balanced randomizations carry more weight. However, it is hard to interpret its results. The advantage of the weighted approach (W_RR) is that it tends to have narrower confidence limits at the cost of slower convergence to normality as the number of studies being pooled becomes large.

As alternative solutions, the methods of UW_OR/UW_RR and UW_MM estimate the unweighted summary effect size. These methods can be viewed unweighted or

weighted equally, i.e. each study carries 1/M weight. The only method that has a similar spirit was developed by Follmann et al. [41] This method weights studies equally, but they apply their test to the mean of the individual study effect size, rather than a ratio estimate approach. This approach is complementary to the UW_OR/UW_RR approach in that its application is to a meta-analysis of a small number of larger studies [7]. However, when rare event occurs, the method would be problematic for relative risk estimates when any study has zero event in either arm, which yields an undefined individual effect size estimate.

Low Event Issue and Continuity Adjustment

Despite warnings in DerSimonian-Laird, Bradburn et al.[42], and Shuster et al. [7], the W_DL remains widely used for meta-analyses with low event rate binomial trials [43-56]. When zero events occur on one or both arms, a continuity adjustment is needed to compute the effect size for the individual study. There are several ways to adjust for continuity [57]. As for the W_DL, the continuity correction [57] is used by adding 0.5 to each of the four cells of the 2-by-2 outcome table that has zero event cells, and studies with no event in both arms are excluded. This way of continuity adjustment is the default option for the two most popular software packages (Comprehensive Meta-analysis 2.0 and RevMan 5.0) [5].

In one of the proposed solutions, the UW_MM uses the sample size adjusted continuity method when zero event happens on one or both arms [57]. Therefore, no studies are excluded. In this simulation study, since the number of subjects in both arms is the same, the sample size adjusted continuity is the same as adding 0.5 in the studies with zero in one or both arms.

By contrast, other proposed solutions of the unweighted method (UW_OR/UW_RR) and the weighted method (W_RR) do not need continuity adjustment. The reason is that the summary effect size is not computed until the proportion (risk) in each arm from each study is added together.

Asymptotic Properties and Large Sample Theory

As we mentioned before, for a meta-analysis, the large sample theory should apply at two levels of the individual study and the number of individual studies being combined for a meta-analysis. However, the most random-effects meta-analysis studies do not have a large number of studies to combine but presumes a normal distribution on the summary effect size on a log scale. Methods along these lines can be found in DerSimonian and Laird [17], Smith et al. [23], Hartung and Knapp [58], Brockwell and Gordon [21]], Warn et al.[24], and Burr and Doss [25].

Considering that we do not often have a large number of studies available for a meta-analysis, we shall use t-approximations, rather than normal approximations. The t approximations are equivalent to normal cutoffs in the limit, as the number of studies being combined tends to infinity. The advantage of using a t-distribution is that the mean coverage of the 95% CI is closer to the 95% coverage, while a method with a normal distribution often tends to underestimate (narrow confidence limits).

Selection of Measurement Metrics for Binomial Outcomes

As for a meta-analysis with binomial outcomes, a researcher often selects the metrics risk ratio or odds ratio over risk difference. The reason is that the risk ratio and odds ratio are relative measures, and therefore tend to be relatively insensitive to differences in baseline events [5]. By contrast, the risk difference is an absolute measure and it is very sensitive to the baseline risk. However, the risk difference is

sometimes more clinically meaningful [5]. In any event, the risk difference provides an unbiased metric for a meta-analysis due to no issues with low event and no assumption for the summary effect size. The risk difference can be predicted afterwards for any given baseline risk [5].

Elaborating Example on Mortality for Erythropoiesis Stimulating Agents

In this subsection, we illustrate how the empirically weighted model (W_DL), the sample size weighted model (W_RR), and the unweighted (UW_RR) work on the study of the mortality risk of erythropoiesis-stimulating agents in cancer population.

Bohlius et al. [59] conducted a patient level meta-analysis with respect to potentially increased mortality risk of erythropoiesis-stimulating agents in cancer population. This study combined 53 randomized cancer trials and found a highly significant increase in mortality with these agents as compared to controls. The FDA, in part on the basis of these data, issued a recommendation that these agents should be used under a risk management program known as a risk evaluation and mitigation strategy (REMS) [60].

Table 5-1 provides the summary effect size through four meta-analysis approaches: the published hazard ratio (instantaneous relative risk) analysis, the empirical weighted DL approach (W_DL), the weighted analysis of relative risk (W_RR), and the unweighted analysis of relative risk (UW_RR). We note the close agreement between the published results [59] and the other two weighted (W_DL, W_RR) results that use study level data.

We discuss the findings as follows. First, although the analysis of Bohlius et al. [59] was based upon actual patient level data, we have to utilize study level data due to the lack of access to the patient level data. However, as long as the follow-up time is

equally rigorous on both treatment arms and event rates are low, in such situations, the ratio of person-years at risk within a study (Treatment 2: Treatment 1) is closely approximated by the ratio of the sample sizes, which explains that the results from the actual patient level data are close to the result from the empirically weighted approach (W_DL). Second, the reason that the results from the W_RR are different from the results from the UW_RR is because they do not estimate the same summary effect size (weighted vs. unweighted) (Figure 5-1). Third, the situations that the results from the W_DL are close to the results from the UW_RR are either when (a) the diversity between studies goes to infinity, (b) sample size goes to infinity, or (c) there is no relationship between effect sizes and weights. In this case, the results from the W_DL are different from the results from the UW_RR, indicating that the W_DL does not estimate the unweighted summary effect size as claimed.

In summary, we learn two points from this example as follows. First, the W_DL does not estimate the unweighted summary effect size as claimed. Instead the W_DL may estimate the weighted summary effect size in this case (though not in general), i.e. the results from the W_DL are close to the results from the W_RR. Second, the results from the W_RR (1.14, 95% CI: 1.04-1.24) are different from the results from the UW_RR (1.05, 95% CI: 0.95-1.17). The contradicting results are fine since they estimate different summary effect size (unweighted for the UW_RR vs. weighted for the W_RR). The consensus of the W_RR and the UW_RR provides more confidence to make the inference, but no consensus was obtained here.

Strengths

Given the existing theoretical issues with the empirical random-effects meta-analysis, Shuster et al. proposed several ways to overcome their obstacles, the

UW_MM, the UW_OR /UW_RR and the W_RR. This is the first study to evaluate the efficiency of these methods by using data simulations. The findings of this study are built upon thousands of simulated scenarios, providing solid evidence to guide a researcher to select an appropriate meta-analysis statistical method in the future.

Limitations

This study also has some limitations. First, the results are generated based upon data simulations with somewhat limited scenarios. More scenarios with data simulations will be more helpful to make a general statement. Second, in all of these simulations, we keep the number of subjects the same for both treatment and control arms in each individual study. Therefore, the impact of variation of the number of subjects in the treatment arm and in the control arm is not reflected in existing scenarios. Finally, since scenarios are generated based upon predetermined parameters, each scenario is not randomly determined. Therefore, statistical tests are not helpful in determining the difference between compared methods of interest. We use graphics and tables to describe the efficiency differences across methods of interest.

Further studies

Since we draw conclusions based upon limited scenarios, we can extend this study by adding more scenarios. For example, we could add variation in the number of subjects between the treatment arm and the control arm.

As for the relative risk data simulation, although the UW_RR estimates the unweighted summary effect size, whereas the W_RR estimates the weighted summary effect size, our study forced the summary effect size from both the UW_RR and the W_RR to be mathematically equal. Furthermore, data simulation is needed to compare

the UW_RR vs. the W_RR when the association between the effect size and the sample size exists.

Also as a further study, the empirical validation on both the sample size weighted approach (W_RR) and the unweighted on proportion (UW_RR) may apply to some known associations of exposures and outcomes.

Conclusions

Aim 1: The UW_OR is more efficient and robust regardless of the number of studies combined and with or without low events, compared with the UW_MM. The UW_MM can be used when low events are unlikely.

Aim 2-1: The UW_OR with a t distribution provides a more accurate coverage regardless of the number of studies combined and with or without low events, compared with the UW_OR with a normal distribution. The UW_OR with a t distribution and the UW_OR with a normal distribution are very close when synthesizing a large number of studies.

Aim 2-2: The UW_MM with a t distribution has a more accurate coverage than the UW_MM with a normal distribution when a small number of studies are combined and no rare events occur. The UW_MM with a t distribution and the UW_MM with a normal distribution are very close when synthesizing a large number of studies.

Aim 3: Both the UW_RR and the W_RR are efficient when they are forced to estimate the same summary effect size. The UW_RR provides more accurate and more stable coverage than the W_RR, whereas the W_RR also provides an accurate coverage, and a relatively narrow confidence interval in some scenarios.

Aim 4-1: The UW_RR with a t distribution provides a more accurate coverage regardless of the number of studies combined and with or without low events, compared

with the UW_RR with a normal distribution. The UW_RR with a t distribution and the UW_RR with a normal distribution are very close when synthesizing a large number of studies.

Aim 4-2: The W_RR with a t distribution provides a more accurate coverage regardless of the number of studies combined and with or without low events, compared with the W_RR with a normal distribution. The W_RR with a t distribution and the W_RR with a normal distribution are very close when synthesizing a large number of studies.

In summary, when in doubt, it might be a good statistical practice to report both the sample size weighted (W_RR) and the unweighted analyses (UW_RR) (one as primary and one as secondary). If the two approaches agree qualitatively, this will add strength to make inference. But since they estimate completely different summary effect size (weighted vs. unweighted), it is not a contradiction when the qualitative conclusions are different.

Table 5-1. Mortality for erythropoiesis stimulating agents in cancer population using different random-effects approaches in meta-analyses

Approaches	Summary effect size	95% CL	P value (2-sided)
HR as Published [59]	1.17	1.06-1.30	0.001
W_DL (RR)	1.15	1.05-1.26	0.003
W_RR	1.14	1.04-1.24	0.005
UW_RR	1.05	0.95-1.17	0.330

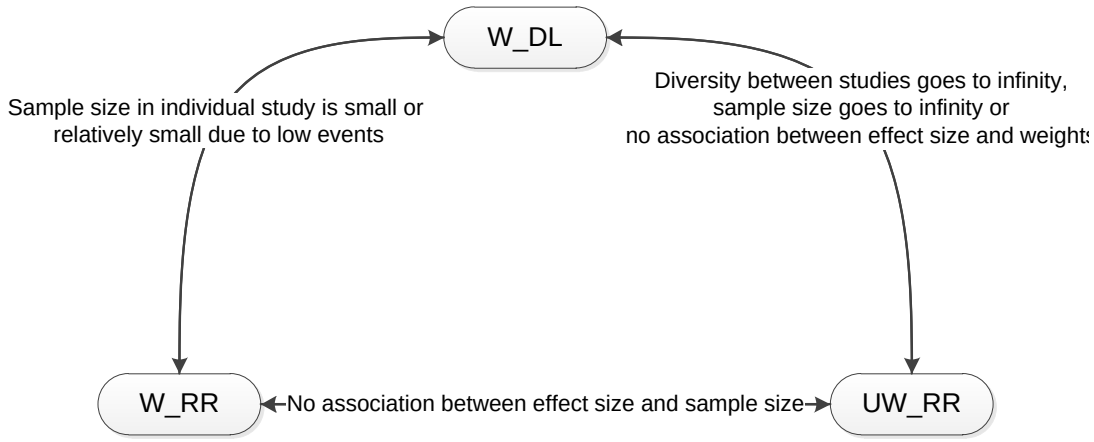

Figure 5-1. Illustration of associations among W_DL, W_RR and UW_RR

APPENDIX
T TABLE

Significance level = α

Degrees of Freedom	.005 (1-tail) .01 (2-tails)	.01 (1-tail) .02 (2-tails)	.025 (1-tail) .05 (2-tails)	.05 (1-tail) .10 (2-tails)	.10 (1-tail) .20 (2-tails)	.25 (1-tail) .50 (2-tails)
1	63.657	31.821	12.706	6.314	3.078	1.000
2	9.925	6.965	4.303	2.920	1.886	.816
3	5.841	4.541	3.182	2.353	1.638	.765
4	4.604	3.747	2.776	2.132	1.533	.741
5	4.032	3.365	2.571	2.015	1.476	.727
6	3.707	3.143	2.447	1.943	1.440	.718
7	3.500	2.998	2.365	1.895	1.415	.711
8	3.355	2.896	2.306	1.860	1.397	.706
9	3.250	2.821	2.262	1.833	1.383	.703
10	3.169	2.764	2.228	1.812	1.372	.700
11	3.106	2.718	2.201	1.796	1.363	.697
12	3.054	2.681	2.179	1.782	1.356	.696
13	3.012	2.650	2.160	1.771	1.350	.694
14	2.977	2.625	2.145	1.761	1.345	.692
15	2.947	2.602	2.132	1.753	1.341	.691
16	2.921	2.584	2.120	1.746	1.337	.690
17	2.898	2.567	2.110	1.740	1.333	.689
18	2.878	2.552	2.101	1.734	1.330	.688
19	2.861	2.540	2.093	1.729	1.328	.688
20	2.845	2.528	2.086	1.725	1.325	.687
21	2.831	2.518	2.080	1.721	1.323	.686
22	2.819	2.508	2.074	1.717	1.321	.686
23	2.807	2.500	2.069	1.714	1.320	.685
24	2.797	2.492	2.064	1.711	1.318	.685
25	2.878	2.485	2.060	1.708	1.316	.684
26	2.779	2.479	2.056	1.706	1.315	.684
27	2.771	2.473	2.052	1.703	1.314	.684
28	2.763	2.467	2.048	1.701	1.313	.683
29	2.756	2.462	2.045	1.699	1.311	.683
Large	2.575	2.327	1.960	1.645	1.282	.675

LIST OF REFERENCES

1. Hartzema AG, Tilson HH, Chan KA. *Pharmacoepidemiology and Therapeutic Risk Management.* (First edition edn). Harvey Whitney Books Company, 2008.

2. Temple R. Meta-analysis and epidemiologic studies in drug development and postmarketing surveillance. *JAMA* 1999; **281**: 841-844.

3. Temple RJ. The regulatory evolution of the integrated safety summary. *Drug Information Journal* 1991; **25**: 485-492.

4. O'Neill RT, Anello C. Does research synthesis have a place in drug regulatory policy? Synopsis of issues: assessment of efficacy and drug approval. *Clinical Research and Regulatory Affairs* 1996; **13**: 23-29.

5. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. *Introduction to meta-analysis.* Wiley, 2009.

6. Shuster JJ, Guo JD, Skyler JS. Meta-analysis of safety for low event-rate binomial trials. In Meta-analysis of safety for low event-rate binomial trials, Editor (ed)^(eds). City, 2011.

7. Shuster JJ, Jones LS, Salmon DA. Fixed vs random effects meta-analysis in rare event studies: the rosiglitazone link with myocardial infarction and cardiac death. *Stat Med* 2007; **26**: 4375-4385.

8. Shuster JJ. Empirical vs natural weighting in random effects meta-analysis. *Stat Med* 2010; **29**: 1259-1265.

9. Shuster JJ. Auxiliary Tools for Unweighted Meta-Analysis. In Auxiliary Tools for Unweighted Meta-Analysis, Editor (ed)^(eds). City, 2010.

10. Hyperstat. Efficiency. In Efficiency. 2010. http://davidmlane.com/hyperstat/A12977.html

11. Kleinbaum DG, Kupper LL, Muller KE, Nizam A. *Applied Regression Analysis and Other Multivarible Methods.* (3rd edn). Duxbury Press, 1998.

12. Chernoff H. Large-sample theory: Parametric case. *The Annals of Mathematical Statistics* 1956; **27**: 1-22.

13. Rice J. *Mathematical Statistics and Data Analysis (Second ed.), Duxbury Press, ISBN* (2 edn). Duxbury Press, 1995.

14. Oehlert GW. A note on the delta method. *The American Statistician* 1992; **46**: 27-29.

15. Stockburger DW. *Introductory Statistics: Concepts, Models, And Applications*, 1998.

16. jpkc.njmu.edu.cn. Normal distribution vs. t distribution. In Normal distribution vs. t distribution. 2011. http://jpkc.njmu.edu.cn/course/tongjixue/file/jxzy/ybzzSD/LiB0067.html

17. DerSimonian R, Laird N. Meta-analysis in clinical trials* *Controlled clinical trials* 1986; **7**: 177-188.

18. Leviton LC, Cook TD. What differentiates meta-analysis from other forms of review1. *Journal of Personality* 1981; **49**: 231-235.

19. Glass GV. Primary, secondary, and meta-analysis of research. *Educational researcher* 1976; **5**: 3-8.

20. Suny. The Evidence Pyramid. In The Evidence Pyramid. 2010. http://phpartners.org/tutorial/04-ebph/2-keyConcepts/4.2.7.html

21. Brockwell SE, Gordon IR. A comparison of statistical methods for meta analysis. *Statistics in medicine* 2001; **20**: 825-840.

22. Kuhnert R, Bohning D. A comparison of three different models for estimating relative risk in meta-analysis of clinical trials under unobserved heterogeneity. *Stat Med* 2007; **26**: 2277-2296.

23. Smith TC, Spiegelhalter DJ, Thomas A. Bayesian approaches to random-effects meta-analysis: a comparative study. *Statistics in medicine* 1995; **14**: 2685-2699.

24. Warn DE, Thompson SG, Spiegelhalter DJ. Bayesian random effects meta-analysis of trials with binary outcomes: methods for the absolute risk difference and relative risk scales. *Statistics in medicine* 2002; **21**: 1601-1623.

25. Burr D, Doss H. A Bayesian semiparametric model for random-effects meta-analysis. *Journal of the American Statistical Association* 2005; **100**: 242-251.

26. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies. *J Natl Cancer Inst* 1959; **22**: 719-748.

27. Cochran WG. The combination of estimates from different experiments. *Biometrics* 1954; **10**: 101-129.

28. Borenstein M, Hedges L, Rothstein H. Meta-Analysis Fixed effect vs. random effects In Meta-Analysis Fixed effect vs. random effects 2007. http://www.meta-analysis.com/downloads/Meta-analysis%20fixed%20effect%20vs%20random%20effects_orig.pdf

29. Shuster JJ, Hatton RC, Hendeles L, Winterstein AG. Reply to discussion of 'Empirical vs natural weighting in random effects meta analysis'. *Statistics in medicine* 2010; **29**: 1272-1281.

30. Friedrich JO, Adhikari NKJ, Beyene J. The ratio of means method as an alternative to mean differences for analyzing continuous outcome variables in meta-analysis: a simulation study. *BMC medical research methodology* 2008; **8**: 32.

31. Sterne JAC, Gavaghan D, Egger M. Publication and related bias in meta-analysis:: Power of statistical tests and prevalence in the literature. *Journal of Clinical Epidemiology* 2000; **53**: 1119-1129.

32. Eickhoff SB, Laird AR, Grefkes C, Wang LE, Zilles K, Fox PT. Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. *Hum Brain Mapp* 2009; **30**: 2907-2926.

33. Hudson M, Suissa S. Avoiding common pitfalls in the analysis of observational studies of new treatments for rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2010; **62**: 805-810.

34. FDA. Guidance for Industry: Diabetes mellitus-Evaluationg Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes. In Guidance for Industry: Diabetes mellitus-Evaluationg Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes. 2008. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071627.pdf

35. FDA/DIA. DIA/FDA Best Practices for Regulatory Information Synthesis of Randomized Trials for Product Safety Evaluation. In DIA/FDA Best Practices for Regulatory Information Synthesis of Randomized Trials for Product Safety Evaluation. 2011. http://www.diahome.org/DIAHome/Education/FindEducationalOffering.aspx?productID=25159&eventType=Meeting

36. Carpenter J, Rucker G, Schwarzer G. Comments on 'Fixed vs random effects meta-analysis in rare event studies: the rosiglitazone link with myocardial infarction and cardiac death'. *Stat Med* 2008; **27**: 3910-3912; author reply 3912-3914.

37. Shuster JJ, Jones LS, Salmon DA. Rebuttal to Carpenter et al. comments on 'Fixed vs. random effects meta-analysis in rare event studies: The rosiglitazone link with myocardial infarction and cardiac death'. *Statistics in medicine* 2008; **27**: 3912-3914.

38. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**: 629-634.

39. Emerson JD, Hoaglin DC, Mosteller F. Simple robust procedures for combining risk differences in sets of 2 x 2 tables. *Stat Med* 1996; **15**: 1465-1488.

40. Emerson JD, Hoaglin DC, Mosteller F. A modified random-effect procedure for combining risk difference in sets of 2?2 tables from clinical trials. *Statistical Methods and Applications* 1993; **2**: 269-290.

41. Follmann DA, Proschan MA. Valid inference in random effects meta-analysis. *Biometrics* 1999; **55**: 732-737.

42. Bradburn MJ, Deeks JJ, Berlin JA, Russell Localio A. Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. *Stat Med* 2007; **26**: 53-77.

43. Myung SK, McDonnell DD, Kazinets G, Seo HG, Moskowitz JM. Effects of Web- and computer-based smoking cessation programs: meta-analysis of randomized controlled trials. *Arch Intern Med* 2009; **169**: 929-937.

44. de Denus S, Sanoski CA, Carlsson J, Opolski G, Spinler SA. Rate vs rhythm control in patients with atrial fibrillation. A meta-analysis. *Archives of internal medicine* 2005; **165**: 258-262.

45. Bath PMW, Gray LJ. Association between hormone replacement therapy and subsequent stroke: a meta-analysis. *BMJ* 2005; **330**: 342.

46. Strippoli GFM, Navaneethan SD, Johnson DW, Perkovic V, Pellegrini F, Nicolucci A, Craig JC. Effects of statins in patients with chronic kidney disease: meta-analysis and meta-regression of randomised controlled trials. *BMJ* 2008; **336**: 645.

47. Piccini JP, Berger JS, O'onnor CM. Amiodarone for the prevention of sudden cardiac death: a meta-analysis of randomized controlled trials. *European heart journal* 2009; **30**: 1245.

48. Mottillo S, Filion KB, Beisle P, Joseph L, Gervais A, O'Loughlin J, Paradis G, Pihl R, Pilote L, Rinfret S. Behavioural interventions for smoking cessation: a meta-analysis of randomized controlled trials. *European heart journal* 2009; **30**: 718.

49. Verdecchia P, Angeli F, Cavallini C, Gattobigio R, Gentile G, Staessen JA, Reboldi G. Blood pressure reduction and renin杌ngiotensin system inhibition for prevention of congestive heart failure: a meta-analysis. *European heart journal* 2009; **30**: 679.

50. De Luca G, Ucci G, Cassetti E, Marino P. Benefits from small molecule administration as compared with abciximab among patients with ST-segment elevation myocardial infarction treated with primary angioplasty: a meta-analysis. *Journal of the American College of Cardiology* 2009; **53**: 1668.

51. Abbate A, Biondi-Zoccai GGL, Appleton DL, Erne P, Schoenenberger AW, Lipinski MJ, Agostoni P, Sheiban I, Vetrovec GW. Survival and Cardiac Remodeling Benefits in Patients Undergoing Late Percutaneous Coronary Intervention of the Infarct-Related Artery:: Evidence From a Meta-Analysis of Randomized Controlled Trials. *Journal of the American College of Cardiology* 2008; **51**: 956-964.

52. Bavry AA, Kumbhani DJ, Rassi AN, Bhatt DL, Askari AT. Benefit of Early Invasive Therapy in Acute Coronary Syndromes:: A Meta-Analysis of Contemporary Randomized Clinical Trials. *Journal of the American College of Cardiology* 2006; **48**: 1319-1325.

53. Brar SS, Leon MB, Stone GW, Mehran R, Moses JW, Brar SK, Dangas G. Use of Drug-Eluting Stents in Acute Myocardial Infarction:: A Systematic Review and Meta-Analysis. *Journal of the American College of Cardiology* 2009; **53**: 1677-1689.

54. Collet JP, Montalescot G, Le May M, Borentain M, Gershlick A. Percutaneous Coronary Intervention After Fibrinolysis:: A Multiple Meta-Analyses Approach According to the Type of Strategy. *Journal of the American College of Cardiology* 2006; **48**: 1326-1335.

55. Berger JS, Krantz MJ, Kittelson JM, Hiatt WR. Aspirin for the prevention of cardiovascular events in patients with peripheral artery disease. *JAMA: the journal of the American Medical Association* 2009; **301**: 1909.

56. Bellamy L, Casas JP, Hingorani AD, Williams D. Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis. *The Lancet* 2009; **373**: 1773-1779.

57. Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Statistics in medicine* 2004; **23**: 1351-1375.

58. Hartung J, Knapp G. A refined method for the meta-analysis of controlled clinical trials with binary outcome. *Statistics in medicine* 2001; **20**: 3875-3889.

59. Bohlius J, Schmidlin K, Brillant C, Schwarzer G, Trelle S, Seidenfeld J, Zwahlen M, Clarke M, Weingart O, Kluge S, Piper M, Rades D, Steensma DP, Djulbegovic B, Fey MF, Ray-Coquard I, Machtay M, Moebus V, Thomas G, Untch M, Schumacher M, Egger M, Engert A. Recombinant human erythropoiesis-stimulating agents and mortality in patients with cancer: a meta-analysis of randomised trials. *Lancet* 2009; **373**: 1532-1542.

60. FDA. FDA Drug Safety Communication: Erythropoiesis-Stimulating Agents (ESAs): Procrit, Epogen and Aranesp. In FDA Drug Safety Communication: Erythropoiesis-Stimulating Agents (ESAs): Procrit, Epogen and Aranesp. 2010. http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm200297.htm

BIOGRAPHICAL SKETCH

The author was born and grew up in Beijing, China. She holds a master's degree in pharmaceutical outcome and policy at the University of North Carolina at Chapel Hill with a prior pharmacology background. In 2008, she joined the Department of Pharmaceutical Outcomes and Policy for a doctorate and she majors in pharmacoepidemology. Her research interest is pharmacoepidemiology study designs and analyses including meta-analysis, pharmacovigilance, and risk management.