Chapter 9
Meta-Analysis of Weighted Mean Difference

Mean difference can be used as an effect size measure if the outcome variable has the same unit of measurement for both the treatment/intervention and placebo/control groups. The raw mean difference can be scaled by the inverse variance weight to define weighted mean difference (WMD). Unlike the SMD, the WMD retains the same unit of measurement as the outcome variable.

Meta-analysis of weighted mean difference (WMD) is covered in this chapter. It provides meta-analysis of WMD under different statistical models along with subgroup analysis with illustrative examples.

9.1 Weighted Mean Difference

For two arms experiments/studies the difference of the two means of an outcome variable is a good starting point to measure the effect size. The raw mean difference is simply the difference of the means of the two arms. It is not essential to standardize the raw mean difference as an effect size measure unless the outcome variable is measured in different units. In many cases the raw mean difference is used as an effect size measure, but it is weighted by the inverse variance. The process produces the weighted mean difference (WMD) as a measure of effect size. This WMD measure retains the same unit of measurement as the outcome variable and is used in many meta-analyses.

9.2 Estimation of Effect Size

Consider an experiment or study with patients randomly divided into two arms, the treatment group with mean of the outcome variable (say Y) to be μ_1 (or μ_T) and control/placebo group having mean μ_2 (or μ_P). Based on a random sample of size

n_1 from the treatment group, let the sample mean be $\hat{\mu}_1 = \bar{Y}_1$ and sample variance be $\hat{\sigma}_1^2 = S_1^2$. Similarly, of another random sample of size n_2 for the placebo group, let the sample mean be $\hat{\mu}_2 = \bar{Y}_2$ and sample variance be $\hat{\sigma}_2^2 = S_2^2$. The sample means and variances are used as estimates of the respective population mean and variance. Assume that both samples are independent the populations are normally distributed.

Then the population raw mean difference $\delta = \mu_1 - \mu_2$ is estimated by the sample mean difference $\hat{\delta} = \bar{Y}_1 - \bar{Y}_2$. As discussed in the previous chapter, the standard error (SE) of the estimator of δ depends on whether the two population standard deviations are equal or not. If the equality of population standard deviations are unknown, we will assume that they are not equal and use appropriate formula to calculate the variance and SE.

For any individual study, let us define population weighted mean difference (WMD) as $\theta = \omega\delta$, where $\omega = 1/\sigma^2$, in which σ^2 is the population variance of δ, is the inverse variance weight of the population mean difference δ. The population WMD, θ is an unknown parameter. An estimator of θ is given by its sample counterpart $\hat{\theta} = w\hat{\delta}$, where $w = 1/v$ is the sample weight, in which $v = \hat{\sigma}^2$ is the sample variance, and $\hat{\delta} = \bar{Y}_1 - \bar{Y}_2$ is the estimate of unknown population mean difference, δ.

In a meta-analysis with $i = 1, 2, \ldots, k$ independent studies, the sample WMD of the ith study is defined as $\hat{\theta}_i = w_i\hat{\delta}_i$ with standard error $SE_i = \sqrt{v_i}$.

So the estimate of the common effect size of all studies is given by $\hat{\theta} = \sum_1^k w_i\hat{\delta}_i \Big/ \sum_1^k w_i$ and the standard error of the estimator of θ becomes $SE(\hat{\theta}) = \sqrt{1 \Big/ \sum_1^k w_i}$.

Then the $(1 - \alpha) \times 100\%$ confidence interval for population WMD, θ is given by the lower limit (LL) and upper limit (UL) as follows:

$$LL = \hat{\theta} - z_{\frac{\alpha}{2}} \times SE(\hat{\theta}) \text{ and}$$

$$LL = \hat{\theta} + z_{\frac{\alpha}{2}} \times SE(\hat{\theta}),$$

where $z_{\frac{\alpha}{2}}$ is the critical value of standard normal distribution leaving $\frac{\alpha}{2}$ area on the upper (or lower) tail of the normal curve.

Example 9.1 Consider the summary data on blood loss for the Laparoscopic-assisted Rectal Resection (LARR) versus Open Rectal Resection (ORR) for Carcinoma from eleven independent studies from Memon et al. 2018.

Table 9.1 Summary statistics of blood loss of eleven studies on sample size, mean and standard deviation of LARR and ORR groups

Study	LARR			ORR		
	n	Mean	Stdev	n	Mean	Stdev
Bonjer et al. (2015)	699	200	222	345	400	370
Fleshman et al. (2015)	240	256.1	305.8	222	318.4	331.7
Stevenson et al. (2015)	238	100	111	235	150	181.5
Jeong et al. (2014)	170	200	148	170	217.5	185
Ng et al. (2014)	40	141.8	500	40	361.1	623
Liu et al. (2011)	98	310	96	88	380	85
Lujan et al. (2009)	101	127.8	113.3	103	234.2	174.3
Ng et al. (2009)	76	280	500	77	337	423
Ng et al. (2008)	51	321.7	750	48	555.6	1188
Braga et al. (2007)	83	213	258	85	396	258
Zhou et al. (2004)	82	20	19	89	92	25

For the above data find the (i) raw mean difference (mean LARR–mean ORR) and standard error of the estimator of population mean difference, (ii) calculate 95% confidence interval for the population mean difference of each of the studies, and (iii) the weight for the first (Bonjer et al. 2015) study.

Solution:
The calculated values of the mean difference (MD), variance of mean difference (Var), standard error of mean difference (SE), weight as inverse variance (W), lower limit (LL) and upper limit (UL) of 95% confidence interval, and sum of W and sum of product of W and MD are shown in Table 9.2.

Explanations of calculations in Table 9.2
To answer the questions in Example 9.1, consider the calculations for the first study (Bonjier et al. 2015):

(i) The raw mean difference is $\hat{\delta}_i = $ MD $=$ difference of mean of LARR and ORR groups $= 200$–$400 = -200$.

The variance of the mean difference (assuming population variances are unequal) is

$$\text{Var} = \frac{S_L^2}{n_L} + \frac{S_O^2}{n_O} = \frac{222^2}{699} + \frac{370^2}{345} = 467.318.$$

Then the standard error becomes
SE $= \sqrt{467.318} = 21.6175$.

Table 9.2 Calculated values of the summary statistics for the mean difference of blood loss

Study	MD	Var	SE	LL	UL	W	WxMD
Bonjer et al. (2015)	−200	467.32	21.618	−242.4	−157.63	0.0021	−0.428
Fleshman et al. (2015)	−62.3	885.25	29.753	−120.6	−3.9839	0.0011	−0.0704
Stevenson et al. (2015)	−50	191.95	13.855	−77.15	−22.845	0.0052	−0.2605
Jeong et al. (2014)	−17.5	330.17	18.171	−53.11	18.114	0.003	−0.053
Ng et al. (2014)	−219	15953	126.31	−466.9	28.26	6E−05	−0.0137
Liu et al. (2011)	−70	176.14	13.272	−96.01	−43.987	0.0057	−0.3974
Lujan et al. (2009)	−106	422.05	20.544	−146.7	−66.134	0.0024	−0.2521
Ng et al. (2009)	−57	5613.2	74.921	−203.8	89.846	0.0002	−0.0102
Ng et al. (2008)	−234	40432	201.08	−628	160.21	2E−05	−0.0058
Braga et al. (2007)	−183	1585.1	39.813	−261	−104.97	0.0006	−0.1155
Zhou et al. (2004)	−72	11.425	3.3801	−78.62	−65.375	0.0875	−6.302
						0.10798	**−7.9085**

(ii)   The lower limit of the 95% confidence interval is

LL $= -200 - 1.96 \times 21.6175 = -242.37$ and upper limit is
UL $= -200 + 1.96 \times 21.6175 = -157.63$.

(iii)   The weight (for the first study) is W $= \frac{1}{Var} = \frac{1}{467.32} = 0.00214$ and WxMD $= 0.00214 \times (-200) = -0.428$.

9.3   Tests on Effect Size

To test the significance of the unknown common effect size, θ, test the null hypothesis
$H_0 : \theta = 0$ against $H_A : \theta \neq 0$ using the test statistic
$Z = \frac{\hat{\theta}}{SE(\hat{\theta})}$ which follows a standard normal distribution.
For a two-tailed test, reject H_0 at the α level of significance (in favour of the alternative hypothesis) if the observed (or calculated) value of Z statistic satisfies $|z_0| \geq z_{\alpha/2}$; otherwise don't reject the null hypothesis (of a two-sided test).

Example 9.2 Consider the blood loss data from eleven independent studies in Table 9.1

Test the significance of the common effect size, θ.

Solution:
To test the significance of the unknown common effect size θ, test the null hypothesis
$H_0 : \theta = 0$ against $H_A : \theta \neq 0$ use the test statistic Z as

$$z_0 = \frac{\hat{\theta}}{SE(\hat{\theta})} = \frac{-73.2411}{3.043199} = -24.0671. \text{[see Example 9.3 for details]}$$

The P-value is $P(|Z| > 24.07) = 2 \times P(Z > 24.07) = 0$.

Since the P-value is 0 the test is highly significant. Thus there is strong sample evidence that the mean difference is significantly different from 0. In other words, the mean blood loss in LARR group is significantly different from that of the ORR group.

9.4 Fixed Effect (FE) Model

The fixed effect (FE) model is used if there is no significant heterogeneity of effect size among the independent studies. In this section, the FE model is presented in a general framework for the meta-analysis of RR with example. An introduction to the FE model is found in (Borenstein et al. 2010).

In meta-analysis, results from all the k independent studies are combined by pooling the summary statistics of primary studies to a single point estimate and confidence interval for the common population effect size θ. Under the fixed effect model, the common effect size estimator, WMD ($=\theta$) is given by

$\hat{\theta}_{FE} = \sum_{i=1}^{k} w_i \hat{\delta}_i \Big/ \sum_{i=1}^{k} w_i$ and the variance of the estimator of the common effect

size is $Var(\hat{\theta}_{FE}) = 1 \Big/ \sum_{i=1}^{k} w_i$. Hence the standard error of the estimator of the

common effect size is $SE(\hat{\theta}_{FE}) = \sqrt{1 \Big/ \sum_{i=1}^{k} w_i}$.

The confidence interval
The $(1 - \alpha) \times 100\%$ confidence interval for the common population WMD θ based on the sample estimates is given by the lower limit (LL) and upper limit (UL) as follows:

$$LL = \hat{\theta}_{FE} - z_{\alpha/2} \times SE(\hat{\theta}_{FE}) \text{ and}$$
$$UL = \hat{\theta}_{FE} + z_{\alpha/2} \times SE(\hat{\theta}_{FE}).$$

Here $z_{\alpha/2}$ is the $\frac{\alpha}{2}$ th cut-off point of standard normal distribution and $SE(\hat{\theta}_{FE}) = \sqrt{Var(\hat{\theta}_{FE})}$.

To compute the confidence interval and perform test on the population effect size θ, under the FE model, we need to compute the point estimate and standard error of the estimator for all studies.

Example 9.3 Consider the summary data on blood loss for the Laparoscopic-assisted Rectal Resection (LARR) versus Open Rectal Resection (ORR) for Carcinoma from Table 9.1.

Using the summary statistics at the bottom of Table 9.2 calculate the (i) point esti-mate of the population WMD, (ii) standard error of estimator and (iii) 95% confidence interval for the population WMD of blood loss under the fixed effect model.

Solution:

To answer the above questions, we use the summary statistics in the last row of Table 9.2 as follows:

(i) The estimate of the population WMD is obtained as the sample WMD as

$$\hat{\theta}_{FE} = \frac{\sum_1^{11} w_i \hat{\delta}_i}{\sum_1^{11} w_i} = \frac{\sum_1^{11} WMD}{\sum_1^{11} W} = \frac{-7.9085}{0.10798} = -73.2411 \approx -73.24.$$

(ii) The standard error is $SE(\hat{\theta}_{FE}) = \sqrt{\frac{1}{\sum_1^{11} w_i}} = \sqrt{\frac{1}{\sum_1^{11} W}} = \sqrt{\frac{1}{0.10798}} = 3.043199.$

(iii) The 95% confidence interval for the common population WMD, θ is given by

LL $= \hat{\theta}_{FE} - 1.96 \times SE(\hat{\theta}_{FE}) = -73.2411 - 1.96 \times 3.043199 = -79.2058 \approx$ -79.21 and

UL $= \hat{\theta}_{FE} + 1.96 \times SE(\hat{\theta}_{FE}) = -73.2411 + 1.96 \times 3.043199 = -67.2764 \approx$ $-67.28.$

Comment The above point estimate (-73.24) and confidence limits $(-79.21, -67.28)$ are displayed in the bottom row of forest plot and represented by the diamond as in Fig. 9.1.

Measuring Heterogeneity

Here we consider two popular methods to identify and measure the extent of heterogeneity among the effect sizes of independent studies.
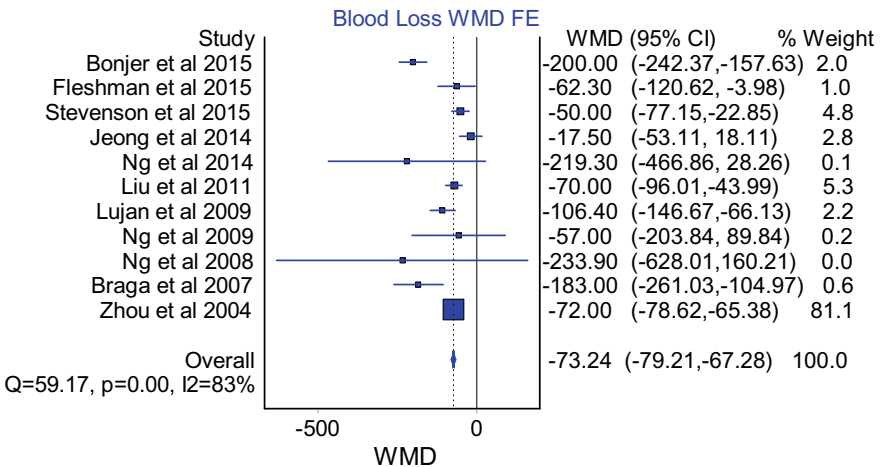

Fig. 9.1 Forest plot of meta-analysis on blood loss for the LARR and ORR groups under FE model

Cochran's Q Statistic (Cochran, 1973)
The Cochran's Q is defined as

$$Q = \sum_{i=1}^{k} w_i \hat{\theta}_i^2 - \frac{\left(\sum_{i=1}^{k} w_i \hat{\theta}_i\right)^2}{\sum_{i=1}^{k} w_i},$$

where w_i is the weight and $\hat{\theta}_i$ is the effect size estimate of the ith study.

The above Q statistic follows a chi-squared distribution with $df = (k-1)$, where k is the number of studies included in the meta-analysis. Since the expected value of a chi-squared variable is its degrees of freedom, the expected value of Q is (k-1), that is, $E(Q) = (k-1) = df$.

Test of Heterogeneity
To test the null hypothesis of the equality of effect sizes (i.e. equality excluding random error) across all studies, test

$H_0 : \theta_1 = \theta_2 = \ldots = \theta_k = \theta$ against H_a : not all θ_i's are equal (at least one of them is different), using the Cochran's Q statistic as defined above.

Reject the null hypothesis at the α level of significance if the observed value of the Q statistic is larger than or equal to $\chi^2_{k-1,1-\alpha}$, the level α critical value of the chi-squared distribution with (k-1) df, such that $P\left(\chi^2_{k-1} \geq \chi^2_{k-1,1-\alpha}\right) = \alpha$; otherwise don't reject it.

The small P-value leads to the conclusion that there is true difference among the effect sizes. However the non-significant P-value may not mean that the effect sizes are not different as this could happen due to low power of the test. The test should not be used to measure the magnitude of the true dispersion.

The I^2 Statistic (Higgins et al. 2003)
The I^2 statistic is a ratio of excess variation to the total variation expressed in percentages as follows:

$$I^2 = \left(\frac{Q - df}{Q}\right) \times 100\%,$$

and is viewed as the proportion of between studies variation and total variation (within plus between studies variation).

Comment
The values of Q and I^2 statistics are calculated from the sample summary data and they are not dependent on any statistical models.

Example 9.4 Consider the summary data on blood loss for the Laparoscopic-assisted Rectal Resection (LARR) versus Open Rectal Resection (ORR) for Carcinoma from Table 9.1.

Find the value of (i) Q statistic and (ii) I^2 statistic for the blood loss data.

Solution:
The following Table 9.3 provides summary calculations for finding Q and I^2 statistics.
   From the summary statistics in Table 9.3.

(i)   The Q statistic is calculated as

$$Q = \sum_{i=1}^{k} w_i \hat{\delta}_i^2 - \frac{\left(\sum_{i=1}^{k} w_i \hat{\delta}_i\right)^2}{\sum_{i=1}^{k} w_i} = \sum_{i=1}^{k} WMD^2 - \frac{\left(\sum_{i=1}^{k} WMD\right)^2}{\sum_{i=1}^{k} W}$$

$$= 638.3923 - \frac{(-7.9085)^2}{0.10798} = 59.17.$$

To test the heterogeneity of effect sizes the P-value is found from the chi-squared
Table (with $df = (11 - 1) = 10$) as
   $P\left(\chi_{10}^2 \geq 59.17\right) = 0$. Note there is no area left under the chi-squired density curve
to the right of 59.17. Since the P-value is close to 0, we reject the null hypothesis (of
equal effect sizes).
   The I^2 statistic is found to be

$$I^2 = \left(\frac{Q - df}{Q}\right) \times 100\% = \left(\frac{59.17 - (11 - 1)}{59.17}\right) \times 100\%$$

$$= 0.83098 \times 100\% = 83\%,$$

where $df = (k - 1) = (11 - 1) = 10$.

Table 9.3 Calculated values of the summary statistics for the mean difference of blood loss data

Study	MD	Var	W	WxMD	W^2	WxDM^2
Bonjer et al. (2015)	−200	467.318	0.00214	−0.428	4.6E-06	85.5948
Fleshman et al. (2015)	−62.3	885.2478	0.00113	−0.0704	1.3E-06	4.38441
Stevenson et al. (2015)	−50	191.9487	0.00521	−0.2605	2.7E-05	13.0243
Jeong et al. (2014)	−17.5	330.1706	0.00303	−0.053	9.2E-06	0.92755
Ng et al. (2014)	−219	15953.23	6.3E-05	−0.0137	3.9E-09	3.01459
Liu et al. (2011)	−70	176.1431	0.00568	−0.3974	3.2E-05	27.8183
Lujan et al. (2009)	−106	422.0541	0.00237	−0.2521	5.6E-06	26.8235
Ng et al. (2009)	−57	5613.227	0.00018	−0.0102	3.2E-08	0.57881
Ng et al. (2008)	−234	40432.41	2.5E-05	−0.0058	6.1E-10	1.3531
Braga et al. (2007)	−183	1585.082	0.00063	−0.1155	4E-07	21.1276
Zhou et al. (2004)	−72	11.42491	0.08753	−6.302	0.00766	453.745
			0.10798	**−7.9085**	**0.00774**	**638.392**

The above value of Q = 59.17, its P-value = 0, and $I^2 = 83\%$ are presented in the forest plot produced by MetaXL on the left panel of Fig. 9.1.

Forest plot for WMD under FE model using MetaXL

The forest plot under the FE model (indicated by "IV" in the code) is constructed using MetaXL code

=MAInputTable("Blood Loss WMD FE","WMD","IV",B6:H16)

Remark: Explanations of MetaXL Code

For this type of meta-analyses in MetaXL the 'opening' code starts with MA Input Table ' = MAInputTable'. This is followed by an open parenthesis inside which the first quote contains the text that appears as the 'title of the output of the forest plot' e.g. "Blood Loss WMD FE" in the above code (user may choose any appropriate title here, but FE is chosen to indicate fixed effect model). Then in the second quote enter the type of effect measure, e.g. "WMD" in the above code which tells that the *weighted mean difference is the effect size.* Within the third quote enter the statistical model, e.g. "IV" in the above code stands for the *fixed effect* (abbreviated by FE) model. Each quotation is followed by a comma, and after the last comma enter the data area in Excel Worksheet, e.g. B6:H16 in the above code tells that the data on the independent studies are taken from the specified cells of the Excel Worksheet. The code ends with a closing parenthesis.

The forest plot of the meta-analysis using the above MetaXL code is found in Fig. **9.1**.

Interpretation

From the above forest plot of WMD under the FE model, the estimated common effect size is −73.24, and the 95% confidence interval is (−79.21, −67.28). The effect size is highly statistically significant (as 0 is not included in the 95% confidence interval).

Here Cochran's Q = 59.17 with P-value = 0 indicates highly significant heterogeneity among the mean difference of blood loss between the LARR and ORR groups of independent studies. The $I^2 = 83\%$ also reflects that there is high heterogeneity among the studies.

Remark

The sign of the mean difference (MD = $\hat{\delta}$) and subsequent estimates of the common effect size are all negative because of the way the mean difference is defined here, mean of LARR group minus mean of ORR group. If the order of difference is reversed, that is, if the mean of ORR group is subtracted from that of LARR group then the sign of the MD and other estimates will interchange (negative to positive and vice versa). The results of the meta-analysis on the reversed ordered MD = $\hat{\delta}$ is shown in Fig. 9.2.

Comment The interpretation of the forest plot in Fig. 9.2, significantly different WMD and significant heterogeneity, is the same as that in Fig. 9.1. All the estimates, confidence limits, value of Q statistic and P-value here are the same in magnitude
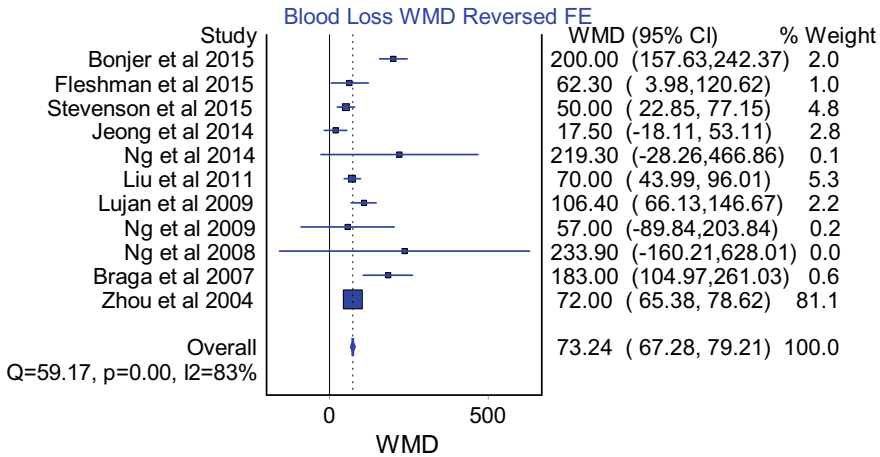
Fig. 9.2 Forest plot of meta-analysis on blood loss for the LARR and ORR groups in reversed order (mean ORR–mean LARR) under FE model

as the previous forest plot but the minus sign is replaced by the plus sign. Thus the order of the raw mean difference does not impact on the final conclusion of the meta-analysis.

9.5   Random Effects (REs) Model

Random effects (REs) model is used when the effect size across the independent studies is significantly heterogeneous. This model was introduced by DerSimonian and Laird, 1986. In spite of its frequent use, some valid criticisms of this model and its poor performance compared with inverse variance heterogeneity (IVhet) and quality effect (QF) models are provided in Doi et al. (2015c, b, c).

Under the random effects model, the population variance of the effect size is the sum of the variance of $\hat{\theta}_i$ about θ (σ^2), the within-study variance, and between-study variances, τ^2. So, for the ith study, the unknown modified variance becomes $\sigma_i^{*2} = \sigma_i^2 + \tau^2$ is estimated by its sample counterpart $v_i^* = v_i + \hat{\tau}^2$, where v_i is the estimate of σ_i^2 and $\hat{\tau}^2$ is the estimate of τ^2. Therefore, the weight assigned to the ith study is defined as

$$w_i^* = \frac{1}{v_i + \hat{\tau}^2} \text{ for i} = 1, 2, \dots, k.$$

The common effect size under the REs model is estimated by

$$\hat{\theta}_{RE} = \sum_{i=1}^{k} w_i^* \hat{\delta}_i \Bigg/ \sum_{i=1}^{k} w_i^*.$$

The standard error of the estimator of the common effect size is given by

$$SE(\hat{\theta}_{RE}) = \sqrt{1 \Bigg/ \sum_{i=1}^{k} w_i^*}.$$

The $(1 - \alpha) \times 100\%$ confidence interval for the effect size θ under the REs model is given by the lower limit (LL) and upper limit (UL) as follows:

$$LL = \hat{\theta}_{RE} - z_{\alpha/2} \times SE(\hat{\theta}_{RE}) \text{ and}$$
$$UL = \hat{\theta}_{RE} + z_{\alpha/2} \times SE(\hat{\theta}_{RE}),$$

where $z_{\alpha/2}$ is the $\frac{\alpha}{2}$ th cut-off point of standard normal distribution.

Estimation of τ^2

The between studies variance is estimated as a scaled excess variation as follows

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

where

$$Q = \sum_{i=1}^{k} w_i \hat{\delta}_i^2 - \frac{\left(\sum_{i=1}^{k} w_i \hat{\delta}_i\right)^2}{\sum_{i=1}^{k} w_i}, \quad C = \sum_{i=1}^{k} w_i - \frac{\sum_{i=1}^{k} w_i^2}{\sum_{i=1}^{k} w_i}$$

and $df = (k - 1)$ in which k is the number of studies.

Example 9.5 Consider the summary data on blood loss for the Laparoscopic-assisted Rectal Resection (LARR) versus Open Rectal Resection (ORR) for Carcinoma from Table 9.1.
Find the estimated value of between studies variance, $\hat{\tau}^2$, for the blood loss data.

Solution:
Using the summary statistics in Table 9.3 of the previous example we calculate the values of Q and C statistics which are required to find the value of $\hat{\tau}^2$. Here

$$Q = \sum_{i=1}^{k} w_i \hat{\delta}_i^2 - \frac{\left(\sum_{i=1}^{k} w_i \hat{\delta}_i \right)^2}{\sum_{i=1}^{k} w_i} = 638.3923 - \frac{(-7.9085)^2}{0.10798} = 59.17,$$

$$C = \sum_{i=1}^{11} w_i - \frac{\sum_{i=1}^{11} w_i^2}{\sum_{i=1}^{11} w_i} = 0.10798 - \frac{0.007742}{0.10798} = 0.0363 \text{ and } df = k-1 = 11-1 = 10.$$

Then, the estimate of the between studies variance becomes
$\hat{\tau}^2 = \frac{Q-df}{C} = \frac{59.17-10}{0.0363} = 1355.029.$

Illustration of REs Model for WMD

Example 9.6 Consider the summary data on blood loss for the Laparoscopic-assisted Rectal Resection (LARR) versus Open Rectal Resection (ORR) for Carcinoma from Table 9.1.

Find the (i) point estimate of the combined population WMD, (ii) standard error of the estimator, and (iii) 95% confidence interval of the population effect size, WMD under the random effects model.

In Table 9.4 the combined variance (Var*) is the sum of Var and Tau^2, and modified weight (W*) is the weight under the REs model calculated as the reciprocal of Var*.

As an illustration, for **Study 1** (Bonjer et al. 2015), the modified variance for the REs model is found to be

$v_1^* = v_1 + \hat{\tau}^2 = 467.318 + 1355.03 = 1822.3$ and the modified weight becomes $w_1^* = 1 / Var_1^* = \frac{1}{1822.3} = 0.00055.$

Now using the summary statistics from Table 9.4 we get

Table 9.4 Calculated values of the summary statistics for the REs model

Study	MD	Var	Tau^2	Var*	W*	W*xMD
Bonjer et al. (2015)	−200	467.318	1355.03	1822.3	0.00055	−0.10975
Fleshman et al. (2015)	−62.3	885.2478	1355.03	2240.3	0.00045	−0.02781
Stevenson et al. (2015)	−50	191.9487	1355.03	1547	0.00065	−0.03232
Jeong et al. (2014)	−17.5	330.1706	1355.03	1685.2	0.00059	−0.01038
Ng et al. (2014)	−219	15953.23	1355.03	17308	5.8E-05	−0.01267
Liu et al. (2011)	−70	176.1431	1355.03	1531.2	0.00065	−0.04572
Lujan et al. (2009)	−106	422.0541	1355.03	1777.1	0.00056	−0.05987
Ng et al. (2009)	−57	5613.227	1355.03	6968.3	0.00014	−0.00818
Ng et al. (2008)	−234	40432.41	1355.03	41787	2.4E-05	−0.0056
Braga et al. (2007)	−183	1585.082	1355.03	2940.1	0.00034	−0.06224
Zhou et al. (2004)	−72	11.42491	1355.03	1366.5	0.00073	−0.05269
					0.00475	**−0.42723**

(i) the point estimate of the common effect size, WMD under the REs model to be

$$\hat{\theta}_{RE} = \sum_{i=1}^{11} w_i^* \hat{\delta}_i \Big/ \sum_{i=1}^{11} w_i^* = \sum_{i=1}^{11} W^* MD \Big/ \sum_{i=1}^{11} W^* = \frac{-0.42723}{0.00475} = -89.98.$$

(ii) The standard error of the estimator of the common effect size is

$$SE(\hat{\theta}_{RE}) = \sqrt{1 \Big/ \sum_{i=1}^{11} w_i^*} = \sqrt{1 \Big/ \sum_{i=1}^{11} W^*} = \sqrt{\frac{1}{0.00475}} = 14.513.$$

(iii) The 95% confidence interval for the population effect size θ under the REs model is given by the lower limit (LL) and upper limit (UL) as follows:

$$LL = \hat{\theta}_{RE} - 1.96 \times SE(\hat{\theta}_{RE}) = -89.98 - 1.96 \times 14.513 = -118.43$$
$$UL = \hat{\theta}_{RE} + 1.96 \times SE(\hat{\theta}_{RE}) = -89.98 + 1.96 \times 14.513 = -61.54.$$

Comment The above point estimate (-89.98) and the 95% confidence interval ($-0118.43, -61.54$) are presented at the bottom row of the forest plot and represented by a diamond as shown in Fig. 9.3.

Forest plot for WMD under REs model using MetaXL
The forest plot under the RE model (indicated by "RE" in the code) is constructed using MetaXL code

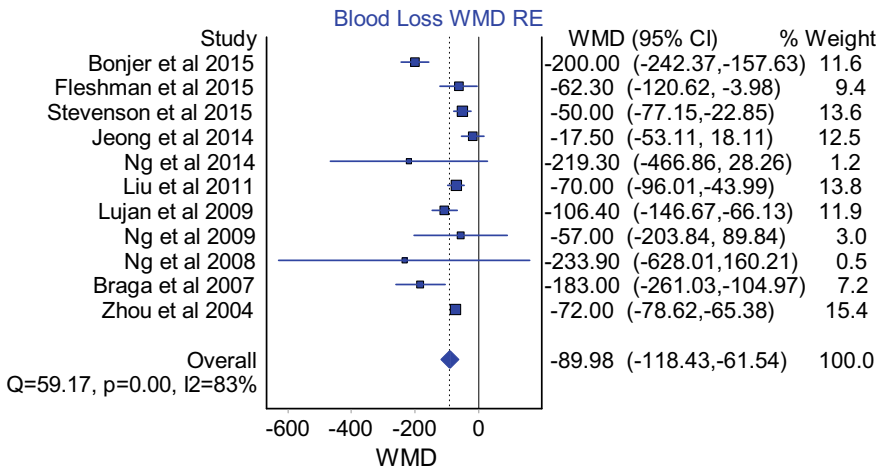| Study	WMD (95% CI)	% Weight
Bonjer et al 2015	-200.00 (-242.37,-157.63)	11.6
Fleshman et al 2015	-62.30 (-120.62, -3.98)	9.4
Stevenson et al 2015	-50.00 (-77.15,-22.85)	13.6
Jeong et al 2014	-17.50 (-53.11, 18.11)	12.5
Ng et al 2014	-219.30 (-466.86, 28.26)	1.2
Liu et al 2011	-70.00 (-96.01,-43.99)	13.8
Lujan et al 2009	-106.40 (-146.67,-66.13)	11.9
Ng et al 2009	-57.00 (-203.84, 89.84)	3.0
Ng et al 2008	-233.90 (-628.01,160.21)	0.5
Braga et al 2007	-183.00 (-261.03,-104.97)	7.2
Zhou et al 2004	-72.00 (-78.62,-65.38)	15.4
Overall	-89.98 (-118.43,-61.54)	100.0

Blood Loss WMD RE

Q=59.17, p=0.00, I2=83%

Fig. 9.3 Forest plot of meta-analysis on blood loss for the LARR and ORR groups under REs model

=MAInputTable("Blood Loss WMD RE","WMD","RE",B6:H16)

Remark: Explanations of MetaXL Code
For this type of meta-analyses in MetaXL the 'opening' code starts with MA Input Table ' =MAInputTable'. This is followed by an open parenthesis inside which the first quote contains the text that appears as the 'title of the output of the forest plot' e.g. "Blood Loss WMD RE" in the above code (user may choose any appropriate title here, but RE is chosen to indicate random effects model). Then in the second quote enter the type of effect measure, e.g. "WMD" in the above code which tells that the *weighted mean difference is the effect size.* Within the third quote enter the statistical model, e.g. "RE" in the above code stands for the *random effects* (abbreviated by REs) model. Each quotation is followed by a comma, and after the last comma enter the data area in Excel Worksheet, e.g. B6:D16 in the above code tells that the data on the independent studies are taken from the specified cells of the Excel Worksheet. The code ends with a closing parenthesis.

The forest plot of the meta-analysis using the above MetaXL code is found in Fig. 9.3.

Interpretation
From the above forest plot of WMD of blood loss under the REs model, the estimated common effect size is -89.98, and the 95% confidence interval is $(-118.43, -61.54)$. The effect size is highly statistically significant (as 0 is not included in the confidence interval).

Here Cochran's $Q = 59.17$ with P-value $= 0$ indicates highly significant heterogeneity among the mean difference of blood loss between the LARR and ORR groups of independent studies. The $I^2 = 83\%$ also reflects that there is high heterogeneity among the studies.

9.6   Inverse Variance Heterogeneity (IVhet) Model

The IVhet model is used when there is significant heterogeneity in the effect size across all the independent studies. Details on this model is found in (Doi et al., 2015a).

The estimator of the common effect size WMD $(=\theta)$ under the inverse variance heterogeneity (IVhet) model is given by

$$\hat{\theta}_{IVhet} = \frac{\sum_{i=1}^{k} w_i \hat{\delta}_i}{\sum_{i=1}^{k} w_i}.$$

Then the variance of the estimator under the IVhet model is given by

$$Var(\hat{\theta}_{IVhet}) = \sum_{i=1}^{k} \left[\left(\frac{1}{v_i} \bigg/ \sum_{i=1}^{k} \frac{1}{v_i} \right)^2 (v_i + \hat{\tau}^2) \right]$$

$$= \sum_{i=1}^{k} \left[\left(w_i \bigg/ \sum_{i=1}^{k} w_i \right)^2 \times v_i^* \right].$$

For the computation of the confidence interval of the common effect size based on the IVhet model use the following estimated standard error

$$SE(\hat{\theta}_{IVhet}) = \sqrt{Var(\hat{\theta}_{IVhet})}.$$

Then, the $(1 - \alpha) \times 100\%$ confidence interval for the common effect size θ under the IVhet model is given by the lower limit (LL) and upper limit (UL) as follows:

$$LL = \hat{\theta}_{IVhet} - z_{\alpha/2} \times SE(\hat{\theta}_{IVhet})$$
$$UL = \hat{\theta}_{IVhet} + z_{\alpha/2} \times SE(\hat{\theta}_{IVhet}),$$

where $z_{\alpha/2}$ is the $\frac{\alpha}{2}$ th cut-off point of standard normal distribution.

Illustration of IVhet Model for WMD

Example 9.7 Consider the summary data on blood loss for the Laparoscopic-assisted Rectal Resection (LARR) versus Open Rectal Resection (ORR) for Carcinoma from Table 9.1.

Find the (i) point estimate of the population WMD, (ii) standard error of the estimator, and (iii) 95% confidence interval of the population effect size, WMD under the inverse variance heterogeneity model.

Solution:
  To answer the questions we need to compute the values in Table 9.5.
  In Table 9.5, Var* is the combined variance

$\left(Var_i^* = v_i + \hat{\tau}^2 \right)$ and W* is the modified weight under the IVhet model calculated as $W_i^* = \left[\left(\frac{1}{v_i} \bigg/ \sum_{i=1}^{k} \frac{1}{v_i} \right)^2 (v_i + \hat{\tau}^2) \right] = \left(w_i \bigg/ \sum_{1}^{k} w_i \right)^2 \times Var_i^*$ for the ith study.
For example, for the first study (Bonjer et al. 2015).
  $Var_1^* = 467.318 + 1355.03 = 1822.3$, and

$W_1^* = \left({0.00214} \big/ {0.10798} \right)^2 \times 1822.3 = 0.71569.$
  Now using the summary statistics in Table 9.5, we answer the questions in Example 9.7.

Table 9.5 Calculated values of the summary statistics for the IVhet model

Study	MD	Var	Tau^2	Var*	W	W*	WxMD
Bonjer et al. (2015)	−200	467.318	1355.03	1822.3	0.00214	0.71569	−0.428
Fleshman et al. (2015)	−62.3	885.2478	1355.03	2240.3	0.00113	0.24518	−0.0704
Stevenson et al. (2015)	−50	191.9487	1355.03	1547	0.00521	3.6011	−0.2605
Jeong et al. (2014)	−17.5	330.1706	1355.03	1685.2	0.00303	1.32585	−0.053
Ng et al. (2014)	−219	15953.23	1355.03	17308	6.3E-05	0.00583	−0.0137
Liu et al. (2011)	−70	176.1431	1355.03	1531.2	0.00568	4.23267	−0.3974
Lujan et al. (2009)	−106	422.0541	1355.03	1777.1	0.00237	0.85564	−0.2521
Ng et al. (2009)	−57	5613.227	1355.03	6968.3	0.00018	0.01897	−0.0102
Ng et al. (2008)	−234	40432.41	1355.03	41787	2.5E-05	0.00219	−0.0058
Braga et al. (2007)	−183	1585.082	1355.03	2940.1	0.00063	0.10036	−0.1155
Zhou et al. (2004)	−72	11.42491	1355.03	1366.5	0.08753	897.864	−6.302
					0.10798	**908.968**	**−7.9085**

(i) The point estimate of population WMD under the IVhet model is

$$\hat{\theta}^*_{IV het} = \frac{\sum\limits_{i=1}^{11} w_i \hat{\delta}_i}{\sum\limits_{i=1}^{11} w_i} = \frac{\sum\limits_{i=1}^{11} W \times MD}{\sum\limits_{i=1}^{11} W} = \frac{-7.9085}{0.10798} = -73.24.$$

(ii) The standard error of the estimator is

$$SE(\hat{\theta}^*_{IV het}) = \sqrt{\sum\limits_{i=1}^{11} W_i^*} = \sqrt{908.96775} = 30.1491.$$

(iii) The 95% confidence interval of the effect size under the IVhet model is given by the lower limit (LL) and upper limit (UL) as follows:

$$LL = \hat{\theta}_{IV het} - 1.96 \times SE(\hat{\theta}_{IV het}) = -73.24 - 1.96 \times 30.1491 = -132.33 \text{ and}$$

$$UL = \hat{\theta}_{IV het} + z_{\alpha/2} \times SE(\hat{\theta}_{IV het}) = -73.24 + 1.96 \times 30.1491 = -14.15.$$

Comment The above point estimate (−73.24) and the 95% confidence interval (−132.33, −14.15) are presented at the bottom row of the forest plot and represented by a diamond as found in Fig. 9.4.

Forest plot for WMD under IVhet model using MetaXL
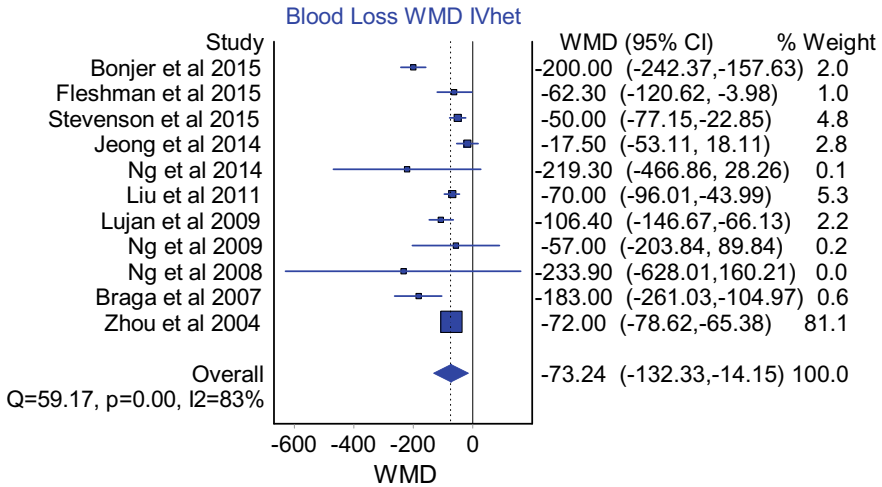The forest plot under the IVhet model (indicated by "IVhet" in the code) is constructed using MetaXL code

Fig. 9.4 Forest plot of meta-analysis on blood loss for the LARR and ORR groups under IVhet model

=MAInputTable("Blood Loss WMD IVhet","WMD","IVhet",B6:H16)

Remark: Explanations of MetaXL Code

For this type of meta-analyses in MetaXL the 'opening' code starts with MA Input Table ' =MAInputTable'. This is followed by an open parenthesis inside which the first quote contains the text that appears as the 'title of the output of the forest plot' e.g. "Blood Loss WMD IVhet" in the above code (user may choose any appropriate title here, but IVhet is chosen to indicate inverse variance heterogeneity model). Then in the second quote enter the type of effect measure, e.g. "WMD" in the above code which tells that the *weighted mean difference is the effect size.* Within the third quote enter the statistical model, e.g. "IVhet" in the above code stands for the *inverse variance heterogeneity* (abbreviated by IVhet) model. Each quotation is followed by a comma, and after the last comma enter the data area in Excel Worksheet, e.g. B6:D16 in the above code tells that the data on the independent studies are taken from the specified cells of the Excel Worksheet. The code ends with a closing parenthesis.

Interpretation

From the above forest plot, the estimated common effect size is -73.24, and the 95% confidence interval is $(-132.33, -14.15)$. The effect size is highly statistically significant (as 0 is not included in the confidence interval).

Here Cochrane's $Q = 59.17$ with P-value $= 0$ indicates highly significant heterogeneity among the mean difference of blood loss between the LARR and ORR groups of independent studies. The $I^2 = 83\%$ also reflects that there is high heterogeneity among the studies.

9.7   Subgroup Analysis

Consider the Blood loss data in Example 9.5. To illustrate subgroup analysis for the data, let's divide the studies into two groups: studied published before 2010 (Old Studies) and after 2010 (Recent Studies) to see if there is any difference in the effect size between the two subgroups. Using MetaXL we produce the forest plot of subgroup analysis as in Figs. 9.5, 9.6 and 9.7 representing the meta-analyses of the subgorups under FE, REs and IVhet models.

Interpretation (FE)
From the forest plot (FE) in Fig. 9.5, under the FE model, the estimated common effect size of Recent Studies is −70.91, and the 95% confidence interval is (−85.83, −55.98), and that for the Old Studies are −73.68 and (−80.19, −67.18) respectively. The effect size is highly statistically significant (as 0 is not included in the confidence interval) for both subgroups as well as the pooled results of all studies (−73,24, CI: −79.21, 67.28).

For the Recent Studies Cochran's Q = 48.05 with P-value = 0 indicating highly significant heterogeneity among the mean difference of blood loss between the LARR
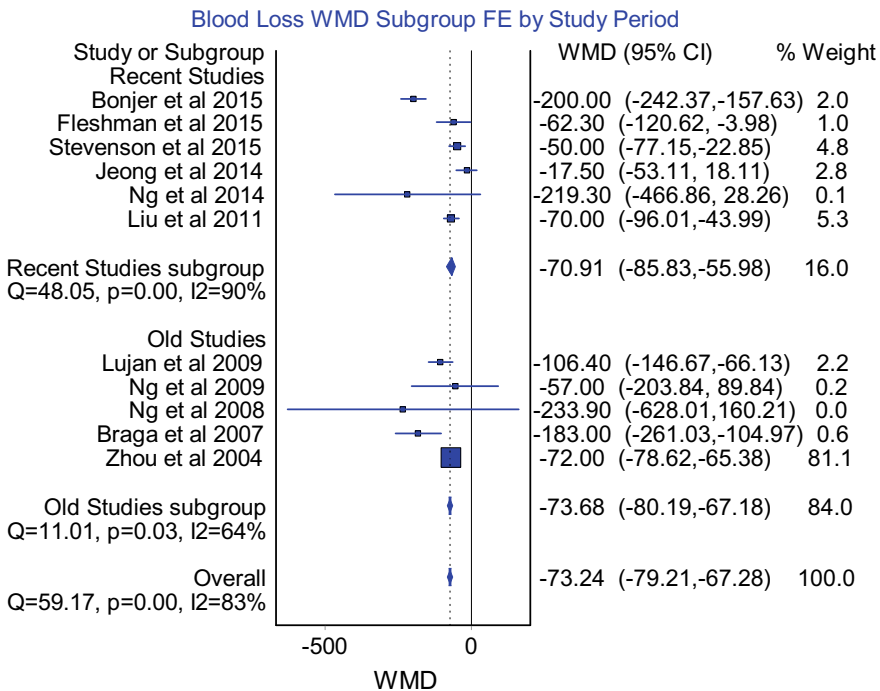

Fig. 9.5 Subgroup analysis by Older and Recent Studies of blood loss after LARR and ORR procedures under the *FE model*

Blood Loss WMD Subgroup RE by Study Period

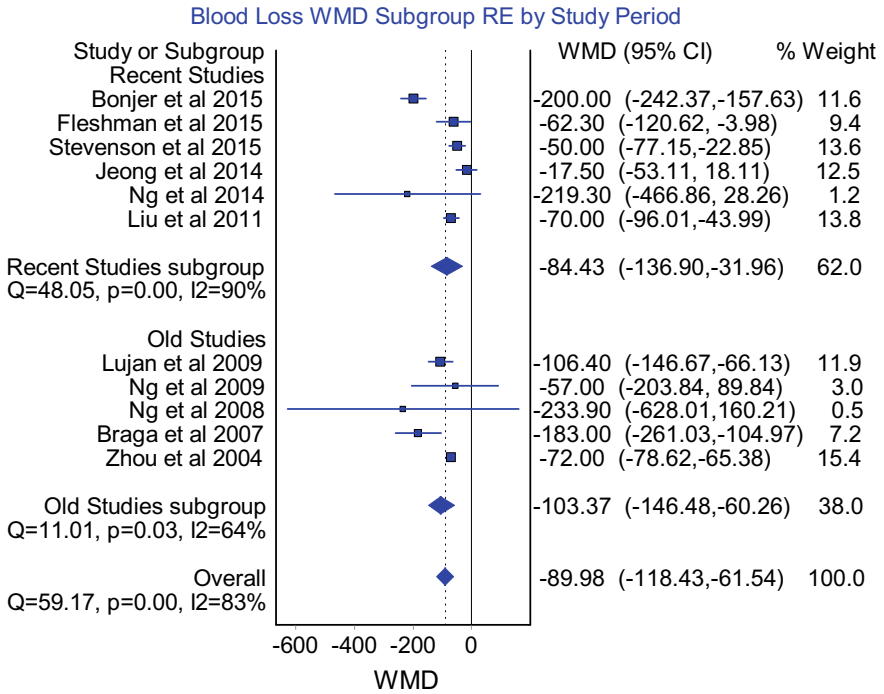| Study or Subgroup	WMD (95% CI)	% Weight
Recent Studies		
Bonjer et al 2015	-200.00 (-242.37,-157.63)	11.6
Fleshman et al 2015	-62.30 (-120.62, -3.98)	9.4
Stevenson et al 2015	-50.00 (-77.15,-22.85)	13.6
Jeong et al 2014	-17.50 (-53.11, 18.11)	12.5
Ng et al 2014	-219.30 (-466.86, 28.26)	1.2
Liu et al 2011	-70.00 (-96.01,-43.99)	13.8
Recent Studies subgroup Q=48.05, p=0.00, I2=90%	-84.43 (-136.90,-31.96)	62.0
Old Studies		
Lujan et al 2009	-106.40 (-146.67,-66.13)	11.9
Ng et al 2009	-57.00 (-203.84, 89.84)	3.0
Ng et al 2008	-233.90 (-628.01,160.21)	0.5
Braga et al 2007	-183.00 (-261.03,-104.97)	7.2
Zhou et al 2004	-72.00 (-78.62,-65.38)	15.4
Old Studies subgroup Q=11.01, p=0.03, I2=64%	-103.37 (-146.48,-60.26)	38.0
Overall Q=59.17, p=0.00, I2=83%	-89.98 (-118.43,-61.54)	100.0

-600 -400 -200 0

WMD

Fig. 9.6 Subgroup analysis by Older and Recent Studies of blood loss after LARR and ORR procedures under the *REs model*

and ORR groups of independent studies. The $I^2 = 90\%$ also reflects that there is high heterogeneity among the studies in this subgroup.

For the Od Studies Cochran's $Q = 11.01$ with P-value $= 0.03$ indicating significant heterogeneity among the mean difference of blood loss between the LARR and ORR groups of independent studies. The $I^2 = 64\%$ also reflects that there is high heterogeneity among the studies in this subgroup. But there is more heterogeneity among the Recent Studies than the Old Studies.

Interpretation (REs)

From the forest plot of WMD on blood loss under the REs model, the estimated common effect size of Recent Studies is -84.43, and the 95% confidence interval is $(-136.90, -31.96)$, and that for the Old Studies are -103.37 and $(-146.48, -60.26)$ respectively. The effect size is highly statistically significant (as 0 is not included in the confidence interval) for both subgroups.
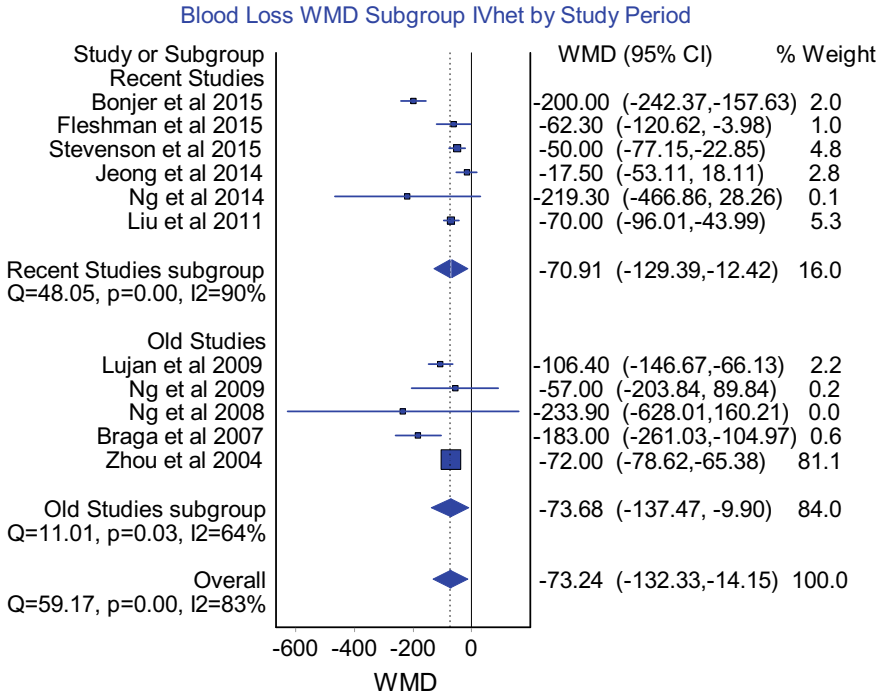
Fig. 9.7 Subgroup analysis by Older and Recent Studies of blood loss after LARR and ORR procedures under the *IVhet model*

The comments on the heterogeneity (Q statistic and P-value) remain the same as that for Fig. 9.5 as these are not dependent on any model.

Interpretation (IVhet)

From the forest plot of WMD on blood loss under the IVhet model, the estimated common effect size of Recent Studies is -70.91, and the 95% confidence interval is $(-129.39, -12.42)$, and that for the Old Studies are -73.68 and $(-137.47, -9.90)$ respectively. The effect size is highly statistically significant (as 0 is not included in the confidence interval) for both subgroups.

9.8   Publication Bias

The study of publication bias for WMD is very similar to that of SMD in Sect. 9.8.8 of the previous chapter. It is not necessary to re-produce thsem again here. Readers interested to produce funnel plot or Doi plot and their interpretation are referred to that Section.

9.9 Conclusions

The weighted mean difference (WMD) method of meta-analysis is covered in this chapter. It is applicable for continuous (numerical) outcome variables for two arms studies. In addition to introducing the WMD method of meta-analysis with step by step illustrations to apply the method on real-life data sets, forest plots are produced under different statistical models by using the MetaXL codes.

The comparison of results for different statistical models show variation in the point estimates and confidence intervals. The heterogeneity among the studies are also studied using Q and I^2 statistics. Subgroup analysis is also provided for the Recent Studies and Old Studies.

Appendix 9—Stata Codes for Meta-Analysis of WMD

A9.1 Blood loss dataset

Study	n_larr	mean_larr	sd_larr	n_orr	mean_orr	sd_orr	study_period
Bonjer et al. (2015)	699	200	222	345	400	370	recent
Fleshman et al. (2015)	240	256.1	305.8	222	318.4	331.7	recent
Stevenson et al. (2015)	238	100	111	235	150	181.5	recent
Jeong et al. (2014)	170	200	148	170	217.5	185	recent
Ng et al. (2014)	40	141.8	500	40	361.1	623	recent
Liu et al. (2011)	98	310	96	88	380	85	recent
Lujan et al. (2009)	101	127.8	113.3	103	234.2	174.3	old
Ng et al. (2009)	76	280	500	77	337	423	old
Ng et al. (2008)	51	321.7	750	48	555.6	1188	old
Braga et al. (2007)	83	213	258	85	396	258	old
Zhou et al. (2004)	82	20	19	89	92	25	old

A9.2 Stata Codes for the Blood loss dataset

. ssc install admetan

Codes for WMD meta-analysis of blood loss

. admetan n_larr mean_larr sd_larr n_orr mean_orr sd_orr, wmd
. admetan n_larr mean_larr sd_larr n_orr mean_orr sd_orr, wmd re
. admetan n_larr mean_larr sd_larr n_orr mean_orr sd_orr, wmd ivhet

A9.3 Codes for WMD Subgroup analysis of blood loss by study period

. admetan n_larr mean_larr sd_larr n_orr mean_orr sd_orr, wmd by(study_period)
. admetan n_larr mean_larr sd_larr n_orr mean_orr sd_orr, wmd re by(study_period)
. admetan n_larr mean_larr sd_larr n_orr mean_orr sd_orr, wmd ivhet by(study_period)

References

Borenstein M, Hedges LV, Higgins JP, Rothstein HR (2010) A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods 1(2):97–111

Cochran WG (1973) Experiments for nonlinear functions (R.A. Fisher Memorial Lecture). J Am Stat Assoc 68(344):771−781. https://doi.org/10.1080/01621459.1973.10481423

DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. Control Clin Trials 7(3):177–188

Doi SA, Barendregt JJ, Khan S, Thalib L, Williams GM (2015a) Advances in the meta-analysis of heterogeneous clinical trials I: the inverse variance heterogeneity model. Contemp Clin Trials 45(Pt A):130–138. https://doi.org/10.1016/j.cct.2015.05.009

Doi SA, Barendregt JJ, Khan S, Thalib L, Williams GM (2015b) Advances in the meta-analysis of heterogeneous clinical trials II: the quality effects model. Contemp Clin Trials 45(Pt A):123–129. https://doi.org/10.1016/j.cct.2015.05.010

Doi SA, Barendregt JJ, Khan S, Thalib L, Williams GM (2015c) Simulation comparison of the quality effects and random effects methods of meta-analysis. Epidemiology 26(4):e42–44

Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. BMJ 327(7414):557–560

Memon MA, Awaiz A, Yunus RM, Memon B, Khan S (2018) Meta-analysis of histopathological outcomes of laparoscopic assisted rectal resection (LARR) vs open rectal resection (ORR) for carcinoma. Am J Surg 216(5):1004–1015. https://doi.org/10.1016/j.amjsurg.2018.06.012