

*Canadian Agency for
Drugs and Technologies
in Health*

*Agence canadienne
des médicaments et des
technologies de la santé*

HTA

**Caveats in the Meta-Analyses
of Continuous Data:
A Simulation Study**

MARCH 2008

Supporting Informed Decisions

Until April 2006, the Canadian Agency for Drugs and Technologies in Health (CADTH) was known as the Canadian Coordinating Office for Health Technology Assessment (CCOHTA).

Publications can be requested from:

CADTH
600-865 Carling Avenue
Ottawa ON Canada K1S 5S8
Tel: 613-226-2553
Fax: 613-226-5392
E-mail: pubs@cadth.ca

or download from CADTH's web site:
<http://www.cadth.ca>

Cite as: Banerjee S, Wells G.A, Chen L. *Caveats in the Meta-Analyses of Continuous Data: A Simulation Study*. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2008.

Production of this report is made possible by financial contributions from Health Canada and the governments of Alberta, British Columbia, Manitoba, New Brunswick, Newfoundland and Labrador, Northwest Territories, Nova Scotia, Nunavut, Ontario, Prince Edward Island, Saskatchewan, and Yukon. The Canadian Agency for Drugs and Technologies in Health takes sole responsibility for the final form and content of this report. The views expressed herein do not necessarily represent the views of Health Canada or any provincial or territorial government.

Reproduction of this document for non-commercial purposes is permitted provided appropriate credit is given to CADTH.

CADTH is funded by Canadian federal, provincial, and territorial governments.

Legal Deposit – 2008
National Library of Canada
ISBN: 978-1-897465-56-1 (print)
ISBN: 978-1-897465-57-8 (online)
H0356 – March 2008

PUBLICATIONS MAIL AGREEMENT NO. 40026386
RETURN UNDELIVERABLE CANADIAN ADDRESSES TO
CANADIAN AGENCY FOR DRUGS AND TECHNOLOGIES IN HEALTH
600-865 CARLING AVENUE
OTTAWA ON K1S 5S8

Canadian Agency for Drugs and Technologies in Health

**Caveats in the Meta-Analyses of Continuous Data:
A Simulation Study**

Srabani Banerjee, MSc, PhD¹
George A. Wells, MSc, PhD²
Li Chen, MSc²

March 2008

¹ Canadian Agency for Drugs and Technologies in Health, Ottawa ON

² Cardiovascular Research Methods Centre, University of Ottawa Heart Institute, Ottawa ON

Reviewers

These individuals kindly provided comments on this report.

External Reviewers

Larry W. Stitt, MSc
Assistant Director, Biostatistical Support Unit
Department of Epidemiology and Biostatistics
The University of Western Ontario
London ON

Keith O'Rourke, DPhil
Visiting Assistant Professor
Duke University
Durham, NC, USA

Brian Hutton, BSc, MSc, PhD
Biostatistician
Ottawa Health Research Institute
Ottawa ON

This report is a review of existing literature, studies, materials, and other information and documentation (collectively the “source documentation”) that are available to CADTH. The accuracy of the contents of the source documentation on which this report is based is not warranted, assured, or represented in any way by CADTH, and CADTH does not assume responsibility for the quality, propriety, inaccuracies, or reasonableness of any statements, information, or conclusions contained in the source documentation.

CADTH takes sole responsibility for the final form and content of this report. The statements and conclusions in this report are those of CADTH and not of its reviewers or Scientific Advisory Panel members.

Authorship

Srabani Banerjee, overall research lead for the project, contributed to the study concept and design, analysis and interpretation of data and prepared the draft.

George A. Wells contributed to the study concept and design, interpretation of data and provided statistical expertise.

Li Chen performed the simulations on SAS and contributed to data analysis.

All authors contributed to the revision of the report.

TABLE OF CONTENTS

1	INTRODUCTION.....	1
2	OBJECTIVE.....	1
3	METHOD.....	1
	3.1 Generation of baseline and final data.....	1
	3.2 Calculation of summary estimates (pooled estimates).....	2
4	RESULTS	3
	4.1 Trends in 95% CI.....	3
	4.1.1 Low correlation coefficient	4
	4.1.2 Moderate correlation coefficient.....	4
	4.1.3 High correlation coefficient.....	4
	4.2 Difference in Significance.....	4
	4.2.1 Low correlation coefficient	4
	4.2.2 Moderate correlation coefficient.....	5
	4.2.3 High correlation coefficient.....	5
5	DISCUSSION.....	18
6	CONCLUSION	19
7	REFERENCES.....	20
	APPENDIX 1	21
	APPENDIX 2A.....	22
	APPENDIX 2B.....	23
	APPENDIX 2C.....	24
	APPENDIX 3.....	25

1 INTRODUCTION

Certain outcomes (e.g., blood pressure, blood glucose levels, hemoglobin, height) are expressed as continuous data. These outcomes are reported by study investigators either as final (end of treatment) values or as change (change from baseline) values. When performing meta-analysis to derive summary estimates for such outcomes, there is some debate as to the method to use: whether one should pool final and change values in the same analysis, or convert the values to one type if there is sufficient information to enable such conversion or perform separate analyses.

2 OBJECTIVE

To determine whether results of meta-analyses with outcomes expressed as continuous data are different if final values (treatment end value) only are used as compared with the use of change values (treatment end value – baseline value) only, or a combination of both final and change values. A difference could potentially affect the significance of the results.

3 METHOD

To address the issue, a simulation study was undertaken. Randomized controlled trials (RCTs) with one experimental group and one control group were simulated by generating baseline and final values for a continuous outcome. These values were generated taking into consideration a number of different parameters [e.g., coefficient of variation (CV); effect size (ES); correlation coefficient (ρ) between the final and baseline values; and number of participants (N) in each group, in each RCT]. The parameter values were chosen to depict different relevant RCT scenarios. Meta-analyses were then undertaken with the simulated RCT data.

3.1 Generation of baseline and final data

Baseline and final values were simulated by generating bivariate normal random variables (x_1, x_2) from normal random variables (z_1, z_2) by using the equations below.^{1,2}

$$x_1 = m_1 + s_1 z_1 \quad (1)$$

$$x_2 = m_2 + s_2(z_1\rho + z_2\sqrt{1-\rho^2}) \quad (2)$$

$m_1 \pm s_1$ and $m_2 \pm s_2$ are mean \pm standard deviation of x_1 and x_2 respectively
 ρ = correlation coefficient between the baseline and final values.

RCT data were simulated by generating baseline and final values for outcomes for each participant in both the experimental and control groups. Mean baseline values were assumed to be 100 for both groups. No effect was assumed for the control group; and a small, medium, or large effect was assumed for the experimental group. Several different scenarios were simulated by varying the

values of CV, ES, ρ and N. The values of the parameters were chosen to cover a relevant spectrum to reflect RCT scenarios. Three values of the CV were used: 0.1, 0.3, and 0.5. Three effect sizes were used: 0.2, 0.5, and 0.9, which were considered to represent small, medium, and large respectively.³ Three values for ρ were considered: 0.1, 0.5, and 0.9; they represent low, moderate, and high correlation. The values used for N in each group (i.e. experimental and control groups) were 10, 50, and 100, giving a total number of participants in each RCT as 2N (i.e., 20, 100, and 200 respectively). Thus with different combinations of these parameters, 81 different scenarios were obtained. Appendix 1 shows the parameter values for the 27 scenarios where N=10. The parameter values listed in Appendix 1 were also used for the remaining 54 scenarios where N=50 (27 scenarios), and N=100 (27 scenarios).

A Monte Carlo simulation analysis was performed using SAS version 9.1.3 (SAS Institute Inc., Cary, NC, USA). The RANNOR function in SAS was used to generate random numbers from a normal distribution with mean 0 and variance 1. Different seeds were used for the experimental and control groups, and the same seeds were used for the different sets of generated data.

3.2 Calculation of summary estimates (pooled estimates)

Two summary statistics [weighted mean difference (WMD) and standardized mean difference (SMD)] are used for meta-analysis of continuous data. When the outcome measurements in all the trials are on the same scale, the difference in means is used. In this case, the meta-analysis computes a weighted average of these differences in means; this is the WMD. When the outcome measurements in all the trials are not on the same scale, the standardized difference in means is used and the meta-analysis is conducted using the SMD method. For the analyses here, it was assumed that all the outcome measurements were on the same scale. The WMD and 95% confidence interval (CI) were computed using the equations given below.⁴

In general, for the i th trial, the mean difference (MD_i) between the experimental and control groups (Table 1), and the standard error (SE_i) were computed using the following equations:

$$MD_i = m_{ei} - m_{ci} \quad (3)$$

$$SE_i = \sqrt{\frac{S_{ei}^2}{n_{ei}} + \frac{S_{ci}^2}{n_{ci}}} \quad (4)$$

The pooled standard deviation (SD) for the control and experimental groups was computed using equation (5):

$$SD_i = s_i = \sqrt{\frac{(n_{ei} - 1)s_{ei}^2 + (n_{ci} - 1)s_{ci}^2}{(n_{ei} + n_{ci} - 2)}} \quad (5)$$

Table 1: Summary information for <i>i</i>th trial			
Study Arm	Mean	Standard Deviation	Number of Participants
Experimental	m_{ei}	S_{ei}	n_{ei}
Control	m_{ci}	S_{ci}	n_{ci}

For each case, changes between baseline and final follow-up were calculated. For each scenario, meta-analyses were performed using final values (f), change values (ch) or a combination of these. The summary estimates were derived by calculating a weighted average of the treatment effects from individual trials according to the following:

$$WMD = \frac{\sum w_i (m_{ei} - m_{ci})}{\sum w_i} \quad (6)$$

where w_i is the weight of the *i*th trial and is given by the inverse of the squared standard error (SE)

$$w_i = \frac{1}{(SE_i)^2} \quad (7)$$

The standard error for the pooled estimate (i.e. WMD) is given by

$$SE_{WMD} = \frac{1}{\sqrt{\sum w_i}} \quad (8)$$

Ten RCTs were pooled in each meta-analysis and the summary estimate was presented as WMD with 95% CI. Meta-analyses were conducted using either all 10 final values (10f) or all 10 change values (10ch), or their combination. When meta-analyses were conducted using a combination of change and final values, three types of combinations were used in the analyses: seven final values and three change values (7f3ch); five final values and five change values (5f5ch); and three final values and seven change values (3f7ch). Each meta-analysis was simulated 10,000 times. The accuracy of pooled WMD and 95% CI using different methods were compared based on 10,000 simulations.

The results of the meta-analyses were presented graphically using GraphPad Prizm 3 software (GraphPad Software Inc. San Diego, USA).

4 RESULTS

Eighty-one different scenarios were simulated and meta-analyses were performed with the simulated data.

4.1 Trends in 95% CI

In Figures 1 to 3, each of the plots (a to f) shows the mean WMD (95% CI) at $\rho=0.1$, 0.5, and 0.9. At each value of ρ , there are five sets of meta-analyses starting from the left with the meta-analysis using 10f values, followed in order by meta-analyses using 7f3ch, 5f5ch, 3f7ch, and 10ch values.

Each plot (a to f) in Figures 1 to 3 shows results obtained with a different combination of values for ES (0.2, 0.5, and 0.9), CV (0.1 and 0.3), and N (10, 50, and 100). A similar trend is seen in all three figures. A similar trend is also seen when CV=0.5 (Appendices 2A to 2C). The bias in the WMD value was calculated for each scenario and was found to be less than 0.07.

4.1.1 Low correlation coefficient

Figures 1 to 3 show that when ρ was low (i.e. 0.1), the 95% CI of the WMD was narrower (indicating less variation) when the meta-analysis was conducted using all final values (10f), as compared to that when all change values (10ch) were used in the meta-analysis. The 95% CI of the WMD gradually increased as the proportion of final values used in the meta-analyses decreased and the number of change values used in the meta-analyses increased (i.e., 7f3ch followed sequentially by 5f5ch, 3f7ch, and 10ch). This trend was observed for all the combinations of ES, CV, and N used for this set of analyses.

4.1.2 Moderate correlation coefficient

Figures 1 to 3 show that when the correlation coefficient (ρ) was moderate (i.e. 0.5), the 95% CI of the WMD was relatively unchanged, irrespective of whether all final values (10f) or a combination of final and change values (7f3ch, 5f5ch, and 3f7ch) or all change values (10ch) were used in the meta-analysis for scenarios where CV=0.1. For scenarios where CV=0.3 or 0.5, the 95% CIs decreased by a small extent as the number of final values used in the meta-analyses decreased.

4.1.3 High correlation coefficient

Figures 1 to 3 show that when the correlation coefficient (ρ) was high (i.e. 0.9), the 95% CI of the WMD was wider when the meta-analysis was conducted using all final values (10f), as compared with that when all change values (10ch) were used in the meta-analysis. The 95% CI of the WMD gradually decreased as the proportion of final values used in the meta-analyses decreased and the number of change values used in the meta-analyses increased (i.e. 7f3ch followed sequentially by 5f5ch, 3f7ch, and 10ch). This trend was observed for all the combinations of ES, CV, and N used for this set of analyses.

4.2 Difference in Significance

Tables 2 to 4 show the proportion (expressed as percentage and 95% CI) of meta-analyses with significant WMD (s-WMD) when using only final values (10f), but non-significant WMD (ns-WMD) when using combinations of final and change values (7f3ch, 5f5ch, and 3f7ch) or only change values (10ch). The three tables also show the proportion of meta-analyses with ns-WMD when using only final values (10f), but s-WMD when using combinations of final and change values (7f3ch, 5f5ch, and 3f7ch) or only change values (10ch). In each case the proportion was calculated from 10,000 meta-analyses and the corresponding 95% CI of the proportion was calculated (using equations⁵ in appendix 3). Scenarios of concern are indicated by shading (darker shading indicates greater concern).

4.2.1 Low correlation coefficient

Table 2 shows the percentage of 10,000 simulated meta-analyses where the significance of WMD using final values (10f) differed from that obtained using other data types (i.e., 7f3ch, 5f5ch, 3f7ch, or 10ch)

when the correlation was low (i.e. $\rho=0.1$). When the sample size was small ($N=10$) and the effect size was small (i.e. $ES=0.2$), in 13.9% to 14.8% of the meta-analyses, the WMDs obtained using only final values were significant when the corresponding WMDs obtained using only change values were non-significant, irrespective of the CV value used. Also when $N=10$ and $ES=0.2$, in 10.9% to 11.6%, 9.0% to 9.4% and 6.3% to 6.8% of the meta-analyses, the WMDs obtained using only final values (10f) were significant when the corresponding WMDs obtained using combinations of final and change values (3f7ch, 5f5ch, and 7f3ch values respectively) were non-significant, irrespective of the CV value used. When $N=10$ and $ES=0.5$, in 23.4% to 24.6% of the meta-analyses the WMDs obtained using only final values were significant when the corresponding WMDs obtained using only change values were non-significant, irrespective of the CV value used. When $N=50$ and $ES=0.2$, in 29.6% to 30.6% of the meta-analyses, the WMDs obtained using only final values were significant when the corresponding WMDs obtained, using only change values, were non-significant.

4.2.2 Moderate correlation coefficient

Table 3 shows the percentage of 10,000 simulated meta-analyses where the significance of WMD using final values (10f) differed from that obtained using other data types (i.e., 7f3ch, 5f5ch, 3f7ch, or 10ch) when the correlation was moderate (i.e. $\rho=0.5$). When $N=10$ and $ES=0.2$, in 12.7% to 13.7% of the meta-analyses, the WMDs obtained using only final values were significant when the corresponding WMDs obtained using only change values were non-significant, irrespective of the CV value used. Also when $N=10$ and $ES=0.2$, in 10.5% to 11.2%, 8.9% to 9.6% and 6.8% to 7.0% of the meta-analyses, the WMDs obtained using only final values (10f) were significant when the corresponding WMDs obtained using combinations of final and change values (3f7ch, 5f5ch, and 7f3ch values respectively) were non-significant, irrespective of the CV value used.

4.2.3 High correlation coefficient

Table 4 shows the percentage of 10,000 simulated meta-analyses where the significance of WMD using final values (10f) differed from that obtained using other data types (i.e., 7f3ch, 5f5ch, 3f7ch, or 10ch) when the correlation was high (i.e. $\rho=0.9$). When $N=10$ and $ES=0.2$, in 59.1% to 60.1% of the meta-analyses, the WMDs obtained using only final values (10f) were non-significant when the corresponding WMDs obtained using only change values (10ch) were significant, irrespective of the CV value used. Also when $N=10$ and $ES=0.2$, in 48.9% to 50.1%, 38.9% to 40.9% and 26.1% to 27.4% of the meta-analyses, the WMDs obtained using only final values (10f) were non-significant when the corresponding WMDs obtained using combinations of final and change values (3f7ch, 5f5ch, and 7f3ch values respectively) were significant, irrespective of the CV value used. When $N=50$ and $ES=0.2$, in 18.5% to 22.3% of the meta-analyses, the WMDs obtained using only final values were non-significant when the corresponding WMDs obtained using only change values (10ch), or combination of change and final values (3f7ch, 5f5ch, or 7f3ch) were significant. When $N=10$ and $ES=0.5$ in 11.1% to 18.2% of the meta-analyses, the WMDs obtained using only final values were non-significant when the corresponding WMDs obtained using only change values or combinations of final and change values (3f7ch, 5f5ch, and 7f3ch) were significant.

Table 5 summarizes the results for the different scenarios shown in Tables 2 to 4. When the effect size was large ($ES=0.9$), results of the meta-analyses with the different data types (10f, 7f3ch, 5f5ch, 3f7ch, and 10ch) were similar (significant or non-significant) irrespective of sample size, CV and ρ used in the analyses. When the effect size was small ($ES=0.2$) or the sample size was small, then there were instances when different significance levels were obtained for WMD depending on the method used for analyses (i.e. the data type used in the analyses).

Figure 1

Means of the WMD and 95% CI obtained from 10,000 meta-analyses are plotted at three values of the correlation coefficient ($\rho=0.1, 0.5, 0.9$). At each value of ρ there are a set of five meta-analyses results, starting from the left with results obtained using 10f (■), 7f3ch (▲), 5f5ch (▼), 3f7ch (◆), and 10ch (●). For each meta-analysis 10 RCTs were pooled and each RCT had 10 participants ($N=10$) in each group (i.e. a total of 20 participants). In each plot the combination of ES and CV values is different.

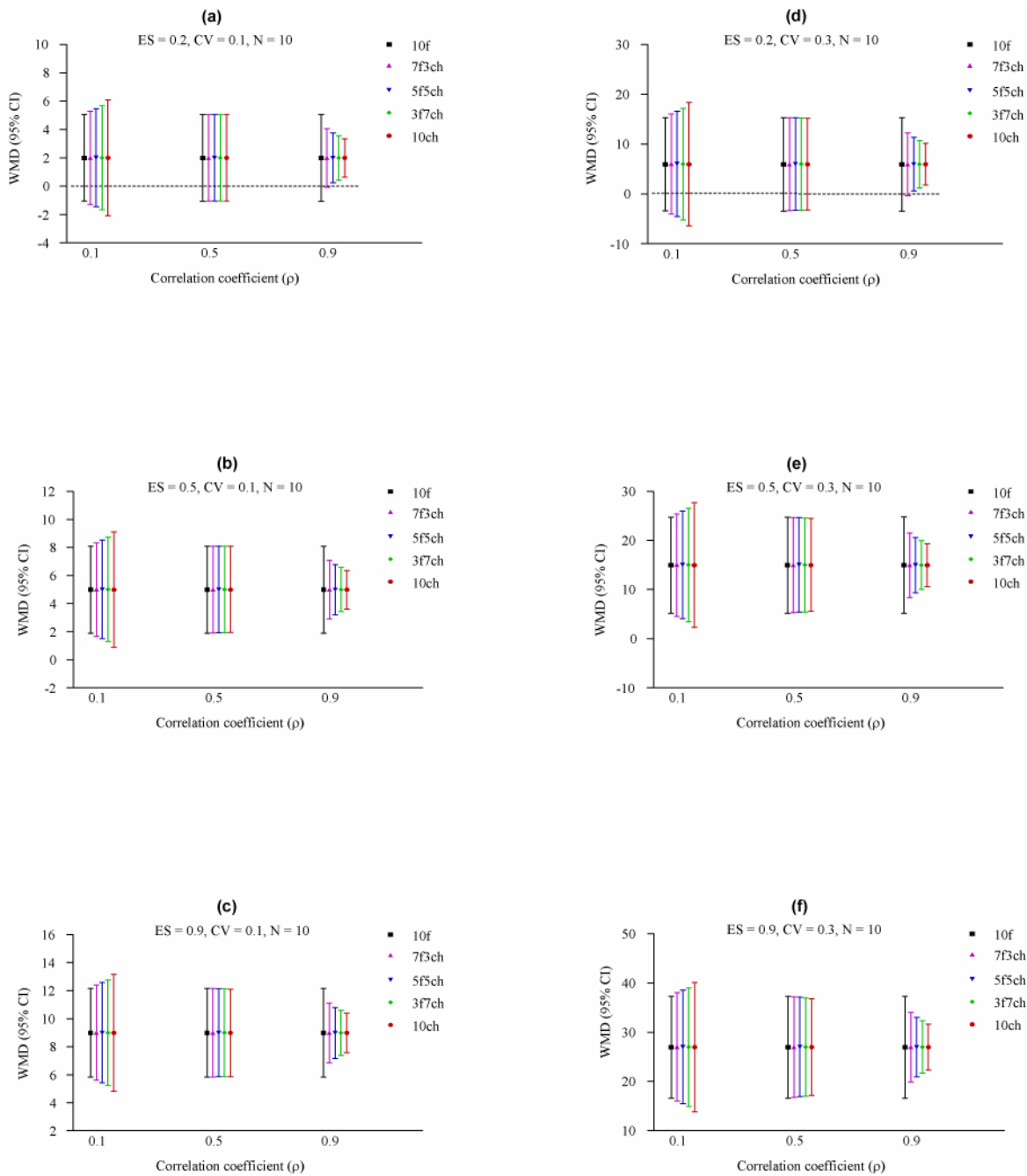


Figure 2

Means of the WMD and 95% CI obtained from 10,000 meta-analyses are plotted at three values of the correlation coefficient ($\rho=0.1, 0.5, 0.9$). At each value of ρ there are a set of five meta-analyses results, starting from the left with results obtained using 10f (■), 7f3ch (▲), 5f5ch (▼), 3f7ch (◆), and 10ch (●). For each meta-analysis 10 RCTs were pooled and each RCT had 50 participants ($N=50$) in each group (i.e. a total of 100 participants). In each plot the combination of ES and CV values is different.

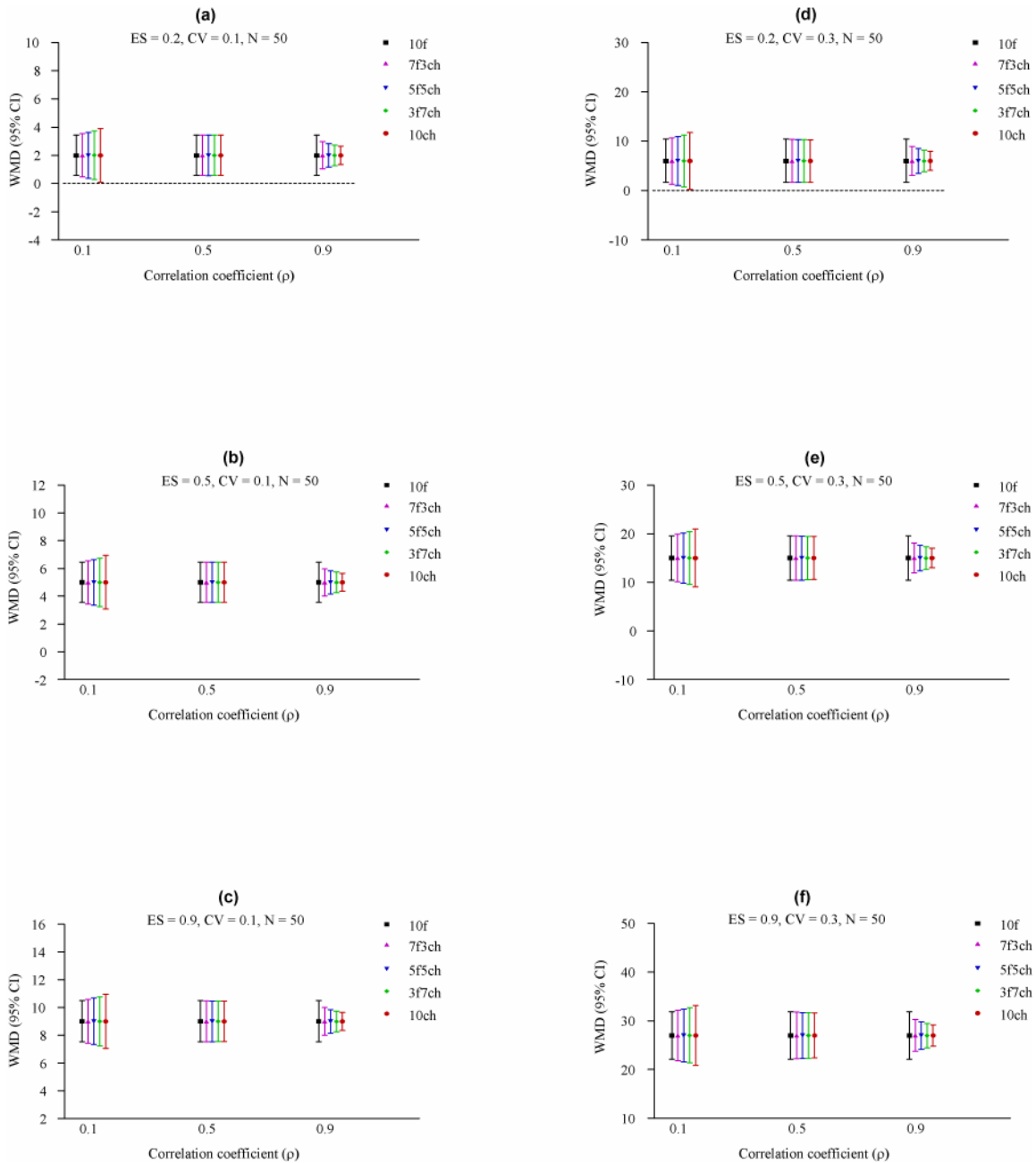


Figure 3

Means of the WMD and 95% CI obtained from 10,000 meta-analyses are plotted at three values of the correlation coefficient ($\rho=0.1, 0.5, 0.9$). At each value of ρ there are a set of five meta-analyses results, starting from the left with results obtained using 10f (■), 7f3ch (▲), 5f5ch (▼), 3f7ch (◆), and 10ch (●). For each meta-analysis 10 RCTs were pooled and each RCT had 100 participants ($N=100$) in each group (i.e. a total of 200 participants). In each plot the combination of ES and CV values is different.

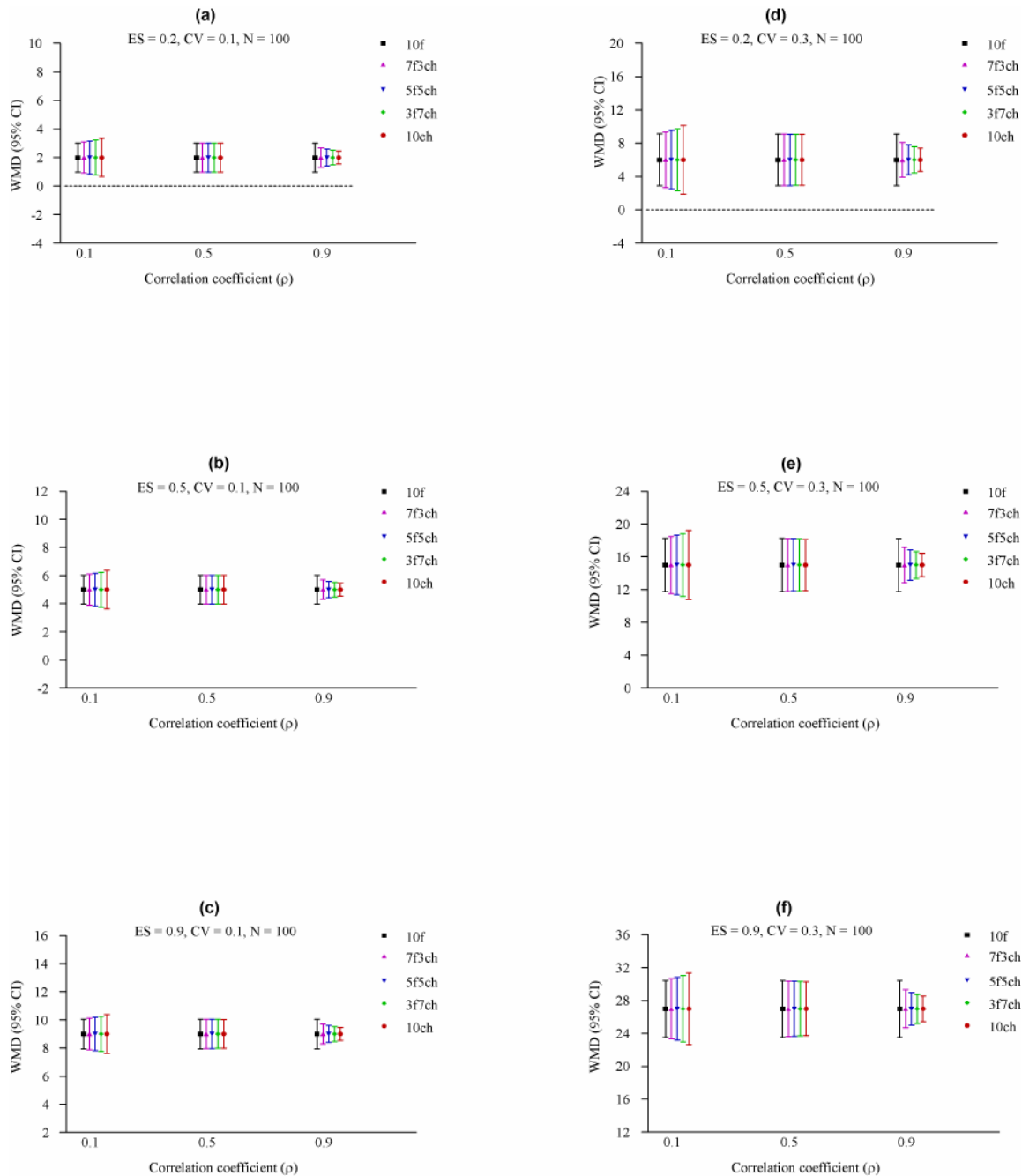


Table 2: Percentage of meta-analyses* where the significance of WMD obtained using final values (10f) differed from that obtained using other data types (i.e., 7f3ch, 5f5ch, 3f7ch, or 10ch). Correlation coefficient (ρ)=0.1

Scenario			Percentage [†] of meta-analyses showing pooled estimates as							
Coefficient of Variation (CV)	Number of Participants in Each Group (N)	Effect Size (ES)	s-WMD Using 10f and ns-WMD Using:				ns-WMD Using 10f and s-WMD Using:			
			7f3ch	5f5ch	3f7ch	10ch	7f3ch	5f5ch	3f7ch	10ch
0.1	10	0.2 (small)	6.78 (6.30, 7.29)	9.39 (8.83, 9.98)	11.56 (10.94, 12.21)	14.81 (14.12, 15.52)	4.49 (4.10, 4.92)	5.17 (4.75, 5.63)	5.33 (4.90, 5.79)	5.88 (5.43, 6.36)
0.1	10	0.5 (medium)	6.35 (5.88, 6.85)	10.31 (9.72, 10.93)	15.37 (14.67, 16.10)	24.57 (23.73, 25.43)	2.06 (1.79, 2.36)	2.20 (1.93, 2.51)	2.41 (2.12, 2.73)	2.43 (2.14, 2.76)
0.1	10	0.9 (large)	0.05 (0.02, 0.12)	0.10 (0.05, 0.19)	0.22 (0.14, 0.34)	0.90 (0.73, 1.11)	0.02 (0.00, 0.08)	0.02 (0.00, 0.08)	0.02 (0.00, 0.08)	0.02 (0.00, 0.08)
0.1	50	0.2 (small)	8.62 (8.08, 9.19)	14.10 (13.43, 14.80)	20.12 (19.34, 20.92)	30.56 (29.66, 31.48)	2.49 (2.20, 2.82)	2.68 (2.38, 3.02)	2.98 (2.66, 3.34)	2.97 (2.65, 3.33)
0.1	50	0.5 (medium)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.01 (0.00, 0.06)	0.03 (0.01, 0.10)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.1	50	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.1	100	0.2 (small)	1.77 (1.52, 2.05)	3.47 (3.12, 3.85)	5.99 (5.54, 6.48)	13.23 (12.58, 13.91)	0.27 (0.18, 0.40)	0.35 (0.25, 0.49)	0.31 (0.21, 0.45)	0.23 (0.15, 0.35)
0.1	100	0.5 (medium)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.1	100	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.3	10	0.2 (small)	6.58 (6.11, 7.09)	9.21 (8.65, 9.80)	11.29 (10.68, 11.93)	14.37 (13.69, 15.08)	4.34 (3.95, 4.76)	5.04 (4.62, 5.49)	5.22 (4.80, 5.68)	5.81 (5.36, 6.29)

Table 2: Percentage of meta-analyses^{*} where the significance of WMD obtained using final values (10f) differed from that obtained using other data types (i.e., 7f3ch, 5f5ch, 3f7ch, or 10ch). Correlation coefficient (ρ)=0.1

Scenario			Percentage [†] of meta-analyses showing pooled estimates as							
Coefficient of Variation (CV)	Number of Participants in Each Group (N)	Effect Size (ES)	s-WMD Using 10f and ns-WMD Using:				ns-WMD Using 10f and s-WMD Using:			
			7f3ch	5f5ch	3f7ch	10ch	7f3ch	5f5ch	3f7ch	10ch
0.3	10	0.5 (medium)	6.92 (6.43, 7.44)	11.14 (10.53, 11.78)	15.84 (15.13, 16.57)	24.03 (23.20, 24.88)	2.53 (2.24, 2.86)	2.75 (2.44, 3.09)	3.02 (2.70, 3.38)	2.90 (2.58, 3.25)
0.3	10	0.9 (large)	0.12 (0.07, 0.22)	0.19 (0.12, 0.30)	0.48 (0.36, 0.64)	1.53 (1.30, 1.80)	0.02 (0.00, 0.08)	0.03 (0.01, 0.10)	0.02 (0.00, 0.08)	0.03 (0.01, 0.10)
0.3	50	0.2 (small)	8.88 (8.33, 9.46)	14.29 (13.61, 14.99)	20.10 (19.32, 20.90)	30.00 (29.10, 30.91)	2.57 (2.27, 2.90)	3.02 (2.70, 3.38)	3.14 (2.81, 3.51)	3.16 (2.83, 3.53)
0.3	50	0.5 (medium)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.01 (0.00, 0.06)	0.06 (0.02, 0.14)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.3	50	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.3	100	0.2 (small)	2.07 (1.80, 2.37)	3.78 (3.42, 4.18)	6.39 (5.92, 6.89)	13.65 (12.99, 14.34)	0.30 (0.21, 0.43)	0.37 (0.26, 0.52)	0.38 (0.27, 0.53)	0.31 (0.21, 0.45)
0.3	100	0.5 (medium)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.3	100	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.5	10	0.2 (small)	6.30 (5.84, 6.80)	9.00 (8.45, 9.58)	10.92 (10.32, 11.55)	13.90 (13.23, 14.60)	3.95 (3.58, 4.36)	4.89 (4.48, 5.34)	5.04 (4.62, 5.49)	5.70 (5.26, 6.18)
0.5	10	0.5 (medium)	7.26 (6.76, 7.79)	11.52 (10.90, 12.17)	16.23 (15.52, 16.97)	23.41 (22.59, 24.26)	2.77 (2.46, 3.12)	3.21 (2.88, 3.58)	3.63 (3.28, 4.02)	3.47 (3.12, 3.85)

Table 2: Percentage of meta-analyses^{*} where the significance of WMD obtained using final values (10f) differed from that obtained using other data types (i.e., 7f3ch, 5f5ch, 3f7ch, or 10ch). Correlation coefficient (ρ)=0.1

Scenario			Percentage [†] of meta-analyses showing pooled estimates as							
Coefficient of Variation (CV)	Number of Participants in Each Group (N)	Effect Size (ES)	s-WMD Using 10f and ns-WMD Using:				ns-WMD Using 10f and s-WMD Using:			
			7f3ch	5f5ch	3f7ch	10ch	7f3ch	5f5ch	3f7ch	10ch
0.5	10	0.9 (large)	0.42 (0.31, 0.57)	0.58 (0.44, 0.75)	1.08 (0.89, 1.31)	2.42 (2.13, 2.75)	0.07 (0.03, 0.15)	0.12 (0.07, 0.22)	0.10 (0.05, 0.19)	0.11 (0.06, 0.20)
0.5	50	0.2 (small)	9.03 (8.48, 9.61)	14.30 (13.62, 15.00)	20.17 (19.39, 20.97)	29.59 (28.70, 30.50)	2.78 (2.47, 3.13)	3.19 (2.86, 3.56)	3.45 (3.11, 3.83)	3.35 (3.01, 3.73)
0.5	50	0.5 (medium)	0.00 (0.00, 0.05)	0.01 (0.00, 0.06)	0.01 (0.00, 0.06)	0.06 (0.00, 0.14)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.5	50	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.5	100	0.2 (small)	2.21 (1.94, 2.52)	4.10 (3.72, 4.51)	6.86 (6.38, 7.38)	14.01 (13.34, 14.71)	0.37 (0.26, 0.52)	0.43 (0.32, 0.58)	0.46 (0.34, 0.62)	0.41 (0.30, 0.56)
0.5	100	0.5 (medium)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.5	100	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)

^{*}10,000 meta-analyses were simulated and for each meta-analysis 10 trials were pooled; [†]percentage and the corresponding 95% CI; ch=change value (i.e. final value – baseline value); f=final value; ns=non-significant; ns-WMD=non-significant WMD; s=significant; s-WMD=significant WMD; WMD=weighted mean difference.

Scenario of concern (i.e. scenario where WMD is significant or non-significant depending on whether f, ch, or a combination of f and ch were used in the meta-analysis) is indicated by shading. When the percentage value in at least one cell is >10, the shading is dark (indicating greater concern) and if it is ≤10 but >3, the shading is light (indicating lesser concern).

Table 3: Percentage of meta-analyses* where the significance of WMD obtained using final values (10f) differed from that obtained using other data types (i.e. 7f3ch, 5f5ch, 3f7ch, or 10ch). Correlation coefficient (ρ)=0.5

Scenario			Percentage [†] of Meta-Analyses Showing Pooled Estimates As:							
Coefficient of Variation (CV)	Number of Participants in Each Group (N)	Effect Size (ES)	s-WMD Using 10f and ns-WMD Using:				ns-WMD Using 10f and s-WMD Using:			
			7f3ch	5f5ch	3f7ch	10ch	7f3ch	5f5ch	3f7ch	10ch
0.1	10	0.2 (small)	7.00 (6.51, 7.52)	9.58 (9.01, 10.18)	11.19 (10.58, 11.83)	13.71 (13.04, 14.40)	7.24 (6.74, 7.77)	9.39 (8.83, 9.98)	11.02 (10.42, 11.65)	13.56 (12.90, 14.25)
0.1	10	0.5 (medium)	3.89 (3.52, 4.29)	5.12 (4.70, 5.57)	6.13 (5.67, 6.62)	7.07 (6.58, 7.59)	4.43 (4.04, 4.86)	5.75 (5.31, 6.23)	6.85 (6.37, 7.37)	7.91 (7.39, 8.46)
0.1	10	0.9 (large)	0.01 (0.00, 0.06)	0.01 (0.00, 0.06)	0.01 (0.00, 0.06)	0.00 (0.00, 0.05)	0.02 (0.00, 0.08)	0.02 (0.00, 0.08)	0.03 (0.00, 0.10)	0.03 (0.00, 0.10)
0.1	50	0.2 (small)	5.80 (5.35, 6.28)	7.72 (7.21, 8.26)	8.99 (8.44, 9.57)	10.56 (9.97, 11.18)	6.02 (5.57, 6.51)	7.74 (7.23, 8.29)	9.27 (8.71, 9.86)	11.10 (10.49, 11.74)
0.1	50	0.5 (medium)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.1	50	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.1	100	0.2 (small)	0.73 (0.58, 0.92)	0.93 (0.76, 1.14)	1.04 (0.85, 1.26)	1.11 (0.92, 1.34)	0.82 (0.66, 1.02)	1.03 (0.85, 1.25)	1.17 (0.97, 1.41)	1.30 (1.09, 1.55)
0.1	100	0.5 (medium)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.1	100	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.3	10	0.2 (small)	6.95 (6.46, 7.47)	9.21 (8.65, 9.80)	10.81 (10.21, 11.44)	13.13 (12.48, 13.81)	7.13 (6.64, 7.66)	9.15 (8.60, 9.74)	10.82 (10.22, 11.45)	13.19 (12.54, 13.87)

Table 3: Percentage of meta-analyses* where the significance of WMD obtained using final values (10f) differed from that obtained using other data types (i.e. 7f3ch, 5f5ch, 3f7ch, or 10ch). Correlation coefficient (ρ)=0.5

Scenario			Percentage [†] of Meta-Analyses Showing Pooled Estimates As:							
Coefficient of Variation (CV)	Number of Participants in Each Group (N)	Effect Size (ES)	s-WMD Using 10f and ns-WMD Using:				ns-WMD Using 10f and s-WMD Using:			
			7f3ch	5f5ch	3f7ch	10ch	7f3ch	5f5ch	3f7ch	10ch
0.3	10	0.5 (medium)	4.18 (3.80, 4.60)	5.51 (5.07, 5.98)	6.60 (6.13, 7.11)	7.43 (6.93, 7.97)	5.37 (4.94, 5.83)	6.92 (6.43, 7.44)	8.44 (7.91, 9.01)	9.81 (9.24, 10.41)
0.3	10	0.9 (large)	0.01 (0.00, 0.06)	0.02 (0.00, 0.08)	0.02 (0.00, 0.08)	0.00 (0.00, 0.05)	0.06 (0.02, 0.14)	0.08 (0.04, 0.16)	0.08 (0.04, 0.16)	0.09 (0.04, 0.18)
0.3	50	0.2 (small)	5.80 (5.35, 6.28)	7.72 (7.21, 8.26)	8.92 (8.37, 9.50)	10.49 (9.90, 11.11)	6.59 (6.12, 7.10)	8.52 (7.89, 9.09)	10.14 (9.56, 10.75)	12.33 (11.70, 12.99)
0.3	50	0.5 (medium)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.3	50	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.3	100	0.2 (small)	0.86 (0.69, 1.07)	1.13 (0.94, 1.36)	1.16 (0.96, 1.39)	1.16 (0.96, 1.39)	1.04 (0.85, 1.26)	1.21 (1.01, 1.45)	1.38 (1.16, 1.63)	1.55 (1.32, 1.82)
0.3	100	0.5 (medium)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.3	100	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.5	10	0.2 (small)	6.76 (6.28, 7.27)	8.91 (8.36, 9.49)	10.45 (9.86, 11.07)	12.56 (11.92, 13.23)	6.99 (6.50, 7.51)	8.98 (8.43, 9.56)	10.49 (9.90, 11.11)	12.85 (12.20, 13.53)
0.5	10	0.5 (medium)	4.59 (4.19, 5.02)	5.77 (5.32, 6.25)	6.89 (6.41, 7.41)	7.74 (7.23, 8.29)	6.34 (5.87, 6.84)	8.27 (7.74, 8.83)	9.84 (9.27, 10.44)	11.68 (11.06, 12.33)
0.5	10	0.9 (large)	0.10	0.04	0.05	0.07	0.17	0.18	0.21	0.22

Table 3: Percentage of meta-analyses* where the significance of WMD obtained using final values (10f) differed from that obtained using other data types (i.e. 7f3ch, 5f5ch, 3f7ch, or 10ch). Correlation coefficient (ρ)=0.5

Scenario			Percentage [†] of Meta-Analyses Showing Pooled Estimates As:							
Coefficient of Variation (CV)	Number of Participants in Each Group (N)	Effect Size (ES)	s-WMD Using 10f and ns-WMD Using:				ns-WMD Using 10f and s-WMD Using:			
			7f3ch	5f5ch	3f7ch	10ch	7f3ch	5f5ch	3f7ch	10ch
			(0.05, 0.19)	(0.01, 0.11)	(0.02, 0.12)	(0.03, 0.15)	(0.10, 0.28)	(0.11, 0.29)	(0.13, 0.33)	(0.14, 0.34)
0.5	50	0.2 (small)	5.91 (5.46, 6.39)	7.65 (7.14, 8.19)	8.90 (8.35, 9.48)	10.55 (9.96, 11.17)	6.97 (6.48, 7.49)	8.91 (8.36, 9.49)	10.83 (10.23, 11.46)	13.09 (12.44, 13.77)
0.5	50	0.5 (medium)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.5	50	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.5	100	0.2 (small)	0.91 (0.74, 1.12)	1.25 (1.05, 1.49)	1.34 (1.13, 1.59)	1.34 (1.13, 1.59)	1.23 (1.03, 1.47)	1.54 (1.31, 1.81)	1.70 (1.46, 1.98)	1.96 (1.70, 2.26)
0.5	100	0.5 (medium)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.5	100	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)

*10,000 meta-analyses were simulated and for each meta-analysis 10 trials were pooled; [†]percentage and the corresponding 95% CI; ch=change value (i.e., final value – baseline value); f=final value; ns=non-significant; ns-WMD=non-significant WMD; s=significant; s-WMD=significant WMD; WMD=weighted mean difference.

Scenario of concern (i.e. scenario where WMD is significant or non-significant depending on whether f, ch, or a combination of f and ch were used in the meta-analysis) is indicated by shading. When the percentage value in at least one cell is >10, the shading is dark (indicating greater concern) and if it is ≤10 but >3, the shading is light (indicating lesser concern).

Table 4: Percentage of meta-analyses* where the significance of WMD obtained using final values (10f) differed from that obtained using other data types (i.e., 7f3ch, 5f5ch, 3f7ch, or 10ch). Correlation coefficient (ρ)=0.9

Scenario			Percentage [†] of Meta-Analyses Showing Pooled Estimates As:							
Coefficient of Variation (CV)	Number of Participants in Each Group (N)	Effect Size (ES)	s-WMD Using 10f and ns-WMD Using:				ns-WMD Using 10f and s-WMD Using:			
			7f3ch	5f5ch	3f7ch	10ch	7f3ch	5f5ch	3f7ch	10ch
0.1	10	0.2 (small)	5.31 (4.88, 5.77)	4.76 (4.35, 5.20)	4.09 (3.71, 4.50)	2.72 (2.41, 3.06)	27.38 (26.51, 28.27)	40.89 (39.93, 41.86)	50.06 (49.08, 51.04)	60.12 (59.15, 61.08)
0.1	10	0.5 (medium)	0.05 (0.02, 0.12)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	11.14 (10.53, 11.78)	11.33 (10.72, 11.97)	11.33 (10.72, 11.97)	11.33 (10.72, 11.97)
0.1	10	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.03 (0.00, 0.10)	0.03 (0.00, 0.10)	0.03 (0.00, 0.10)	0.03 (0.00, 0.10)
0.1	50	0.2 (small)	0.13 (0.07, 0.23)	0.01 (0.00, 0.06)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	18.46 (17.71, 19.24)	18.95 (18.19, 19.74)	18.99 (18.23, 19.78)	18.99 (18.23, 19.78)
0.1	50	0.5 (medium)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.1	50	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.1	100	0.2 (small)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	1.38 (1.16, 1.63)	1.38 (1.16, 1.63)	1.38 (1.16, 1.63)	1.38 (1.16, 1.63)
0.1	100	0.5 (medium)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.1	100	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.3	10	0.2 (small)	5.20 (4.78, 5.66)	4.54 (4.14, 4.97)	3.81 (3.45, 4.21)	2.46 (2.17, 2.79)	26.75 (25.89, 27.63)	39.74 (38.78, 40.71)	49.67 (48.69, 50.65)	59.63 (58.66, 60.59)

Table 4: Percentage of meta-analyses* where the significance of WMD obtained using final values (10f) differed from that obtained using other data types (i.e., 7f3ch, 5f5ch, 3f7ch, or 10ch). Correlation coefficient (ρ)=0.9

Scenario			Percentage [†] of Meta-Analyses Showing Pooled Estimates As:							
Coefficient of Variation (CV)	Number of Participants in Each Group (N)	Effect Size (ES)	s-WMD Using 10f and ns-WMD Using:				ns-WMD Using 10f and s-WMD Using:			
			7f3ch	5f5ch	3f7ch	10ch	7f3ch	5f5ch	3f7ch	10ch
0.3	10	0.5 (medium)	0.05 (0.02, 0.12)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	14.33 (13.65, 15.04)	14.78 (14.09, 15.49)	14.80 (14.11, 15.51)	14.80 (14.11, 15.51)
0.3	10	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.06 (0.02, 0.14)	0.06 (0.02, 0.14)	0.06 (0.02, 0.14)	0.06 (0.02, 0.14)
0.3	50	0.2 (small)	0.19 (0.12, 0.30)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	19.61 (18.84, 20.40)	20.34 (19.56, 21.15)	20.39 (19.61, 21.20)	20.39 (19.61, 21.20)
0.3	50	0.5 (medium)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.3	50	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.3	100	0.2 (small)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	1.82 (1.57, 2.11)	1.82 (1.57, 2.11)	1.82 (1.57, 2.11)	1.82 (1.57, 2.11)
0.3	100	0.5 (medium)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.3	100	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.5	10	0.2 (small)	4.99 (4.58, 5.44)	4.34 (3.95, 4.76)	3.63 (3.28, 4.02)	2.32 (2.04, 2.64)	26.08 (25.22, 26.95)	38.87 (37.91, 39.83)	48.86 (47.88, 49.84)	59.11 (58.14, 60.07)
0.5	10	0.5 (medium)	0.10 (0.05, 0.19)	0.02 (0.00, 0.08)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	17.32 (16.59, 18.08)	18.11 (17.36, 18.88)	18.21 (17.46, 18.98)	18.22 (17.47, 18.99)

Table 4: Percentage of meta-analyses* where the significance of WMD obtained using final values (10f) differed from that obtained using other data types (i.e., 7f3ch, 5f5ch, 3f7ch, or 10ch). Correlation coefficient (ρ)=0.9

Scenario			Percentage [†] of Meta-Analyses Showing Pooled Estimates As:							
Coefficient of Variation (CV)	Number of Participants in Each Group (N)	Effect Size (ES)	s-WMD Using 10f and ns-WMD Using:				ns-WMD Using 10f and s-WMD Using:			
			7f3ch	5f5ch	3f7ch	10ch	7f3ch	5f5ch	3f7ch	10ch
0.5	10	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.27 (0.18, 0.40)	0.27 (0.18, 0.40)	0.27 (0.18, 0.40)	0.27 (0.18, 0.40)
0.5	50	0.2 (small)	0.21 (0.13, 0.33)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	21.29 (20.49, 22.11)	22.20 (21.39, 23.03)	22.28 (21.47, 23.11)	22.28 (21.47, 23.11)
0.5	50	0.5 (medium)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.5	50	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.5	100	0.2 (small)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	2.22 (1.94, 2.53)	2.22 (1.94, 2.53)	2.22 (1.94, 2.53)	2.22 (1.94, 2.53)
0.5	100	0.5 (medium)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)
0.5	100	0.9 (large)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)	0.00 (0.00, 0.05)

*10,000 meta-analyses were simulated and for each meta-analysis 10 trials were pooled; [†]percentage and the corresponding 95% CI; ch=change value (i.e., final value – baseline value); f=final value; ns=non-significant; ns-WMD=non-significant WMD; s=significant; s-WMD=significant WMD; WMD=weighted mean difference.

Scenario of concern (i.e. scenario where WMD is significant or non-significant depending on whether f, ch, or a combination of f and ch were used in the meta-analysis) is indicated by shading. When the percentage value in at least one cell is >10, the shading is dark (indicating greater concern) and if it is ≤10 but >3, the shading is light (indicating lesser concern).

Table 5: Summary of the different scenarios (i.e. summary of Tables 2, 3 and 4) with scenarios of concern indicated.

		Scenario									
		$\rho \rightarrow$	0.1	0.1	0.1	0.5	0.5	0.5	0.9	0.9	0.9
		CV \rightarrow	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5
N \downarrow	ES \downarrow										
10	0.2 (small)										
10	0.5 (medium)										
10	0.9 (large)										
50	0.2 (small)										
50	0.5 (medium)										
50	0.9 (large)										
100	0.2 (small)										
100	0.5 (medium)										
100	0.9 (large)										

CV=coefficient of variation; ES=effect size; N=number of patients in each group; ρ =correlation coefficient.

Scenario of concern (i.e. scenario where WMD is significant or non-significant depending on whether f, ch, or a combination of f and ch were used in the meta-analysis) is indicated by shading (darker shading indicates greater concern). When the percentage value in at least one cell of a row (in Tables 2, 3 or 4) is >10 , the shading is dark (indicating greater concern) and if it is ≤ 10 but >3 , the shading is light (indicating lesser concern).

5 DISCUSSION

This simulation study was undertaken to address the important issue of how summary estimates of continuous outcome data using the WMD approach were affected when different reporting structures of trial-level data (i.e. different data types) were used for meta-analysis. The results of this study provide valuable insights into the potential problems and limitations of the method of analysis chosen in different scenarios. When the sample size was small or the effect size was small, the significance of the WMD value obtained from a meta-analysis could be different depending on the method used (i.e., whether final values, change values, or a combination of final and change values were used). This was more of an issue when the ρ was either low or high, as compared to that when the ρ was moderate. However, when the sample size was large (i.e. $N \geq 100$) and the effect size was large, it did not matter which method was used.

When the sample size is small and the effect size is small, then the data could be transformed to one particular type. This depends on sufficient information being available or being imputed.⁶⁻⁸ Otherwise a sensitivity analysis of final and change values should be conducted and interpreted to determine the robustness of results based on the different summary statistics.

Our results show that when correlation was low, the 95% CI of the WMD was wider (indicating greater variation) when change values were used for the analysis, as compared to that when final values were used. When correlation was high, the reverse was observed. This reflects the fact that the variance of two highly correlated measures is generally less than that of a single measure with greater variation. Our results are in agreement with those of Vickers *et al.*⁹ They reported that if correlation is low, using the change score will add variation to the results and the follow-up score is more likely to show a significant result, and when correlation is high the reverse will occur. When the correlation

was moderate ($\rho=0.5$), the 95% CI was almost similar between the different methods used (i.e. the various data types used in the meta-analyses).

The results presented in this report were based on meta-analyses including 10 RCTs. The number of RCTs included in a meta-analysis will vary depending on the area (e.g., disease state, intervention, population) of the study because in some areas large numbers of trials have been conducted, but in some areas only a few trials have been conducted. To capture all such cases would be a computationally difficult task. For this report, it was determined that 10 trials represented a reasonable compromise of the number of small versus large trials that one may have available for a meta-analysis. It also allowed a number of different combinations of data types (i.e. final and change) to be investigated. In reality, there could be many different scenarios which could be described by a wide range of parameter values; however, it is not feasible to address all possible scenarios. Hence a number of parameter values were chosen to reflect relevant and interesting scenarios that are likely to occur.

Note that the CI limits shown in Figures 1 to 3 represent the mean limits obtained from 10,000 meta-analyses. Hence the proportion of meta-analyses with CI encompassing zero (i.e. meta-analyses showing ns-WMD values) would not be apparent from these figures unless the proportion was substantial. This explains why for some scenarios, Tables 2 to 4 indicate that there were ns-WMDs, which were not apparent in the figures. For example in Figure 1(b) for scenario with $\rho=0.1$, $ES=0.5$, $CV=0.1$, and $N=10$, the 95% CI using change values (10ch) does not encompass zero, although Table 2 shows that 25% of the meta-analyses using change values (10ch) had ns-WMD when the corresponding meta-analyses using final values (10f) had s-WMD.

6 CONCLUSION

When the ρ was high, the 95% CI for WMD was wider (indicating greater variation) when the meta-analysis was conducted using all final values as compared to that when all change values were used in the meta-analysis. When the ρ was low the reverse was true.

When the sample size was small or the effect size was small, the significance of the WMD value obtained from a meta-analysis could be different in some scenarios, depending on the method used. Such scenarios call for caution. But when the sample size was large and when the effect size was large, it did not matter which method was used.

The results of this study provide valuable insights into the potential problems and limitations of the method of analysis chosen in different scenarios.

7 REFERENCES

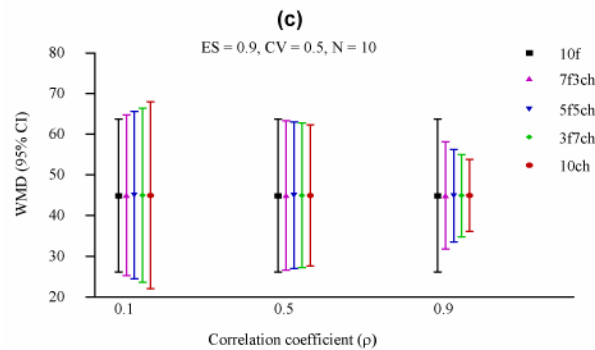
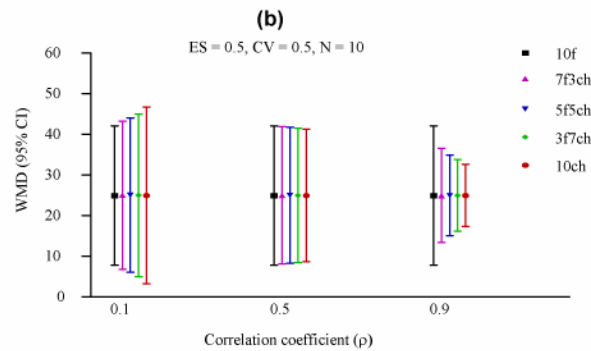
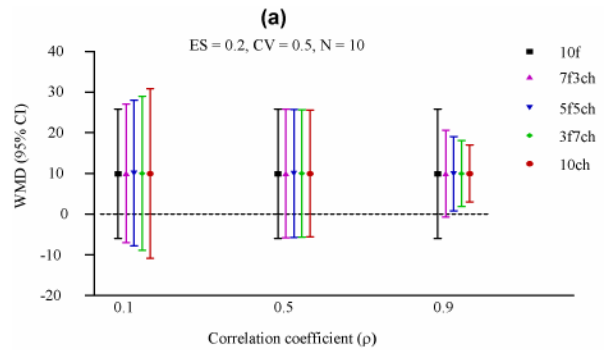
1. Naylor TH, Balintfy JL, Burdick DS, Chu K. *Computer simulation techniques*. New York: John Wiley & Sons Inc; 1966.
2. Takamitsu H. *Preparation* 2000. Available: <http://homepage2.nifty.com/hashimoto-t/try/prepar-e.html> (accessed 2007 Nov 16).
3. Cohen J. The t test for means. In: *Statistical power analysis for the behavioral sciences*. Hillsdale (NJ): Lawrence Erlbaum Associates, Inc.; 1988. p.19-75.
4. Egger M, Davey Smith G, Altman DG, editors. *Systematic reviews in health care: meta-analysis in context*. 2nd ed. London: BMJ Publishing Group; 2001.
5. Fleiss JL, Levin BA, Paik MC. *Statistical methods for rates and proportions*. 3rd ed. Hoboken (NJ): John Wiley and Sons; 2003.
6. Follmann D, Elliott P, Suh I, Cutler J. Variance imputation for overviews of clinical trials with continuous response. *J Clin Epidemiol* 1992;45(7):769-73.
7. Thiessen PH, Barrowman N, Garg AX. Imputing variance estimates do not alter the conclusions of a meta-analysis with continuous outcomes: a case study of changes in renal function after living kidney donation. *J Clin Epidemiol* 2007;60(3):228-40.
8. Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. *J Clin Epidemiol* 2006;59(1):7-10.
9. Vickers AJ, Altman DG. Statistics notes: analysing controlled trials with baseline and follow up measurements. *BMJ* 2001;323(7321):1123-4. Available: <http://bmj.com/cgi/reprint/323/7321/1123.pdf>

APPENDIX 1: Parameter Values for the Different Scenarios Simulated with Number of Participants (N) in Each Group Being 10

Scenario	Coefficient of Variation (CV)	Effect Size (ES)	Correlation Coefficient (ρ)	Experimental and Control Groups		Experimental Group		Control Group	
				Baseline Value		Final Value		Final Value	
				Mean	SD	Mean	SD	Mean	SD
1	0.1	0.2	0.1	100	10	102	10.2	100	10
2	0.1	0.2	0.5	100	10	102	10.2	100	10
3	0.1	0.2	0.9	100	10	102	10.2	100	10
4	0.1	0.5	0.1	100	10	105	10.5	100	10
5	0.1	0.5	0.5	100	10	105	10.5	100	10
6	0.1	0.5	0.9	100	10	105	10.5	100	10
7	0.1	0.9	0.1	100	10	109	10.9	100	10
8	0.1	0.9	0.5	100	10	109	10.9	100	10
9	0.1	0.9	0.9	100	10	109	10.9	100	10
10	0.3	0.2	0.1	100	30	106	31.8	100	30
11	0.3	0.2	0.5	100	30	106	31.8	100	30
12	0.3	0.2	0.9	100	30	106	31.8	100	30
13	0.3	0.5	0.1	100	30	115	34.5	100	30
14	0.3	0.5	0.5	100	30	115	34.5	100	30
15	0.3	0.5	0.9	100	30	115	34.5	100	30
16	0.3	0.9	0.1	100	30	127	38.1	100	30
17	0.3	0.9	0.5	100	30	127	38.1	100	30
18	0.3	0.9	0.9	100	30	127	38.1	100	30
19	0.5	0.2	0.1	100	50	110	55.0	100	50
20	0.5	0.2	0.5	100	50	110	55.0	100	50
21	0.5	0.2	0.9	100	50	110	55.0	100	50
22	0.5	0.5	0.1	100	50	125	62.5	100	50
23	0.5	0.5	0.5	100	50	125	62.5	100	50
24	0.5	0.5	0.9	100	50	125	62.5	100	50
25	0.5	0.9	0.1	100	50	145	72.5	100	50
26	0.5	0.9	0.5	100	50	145	72.5	100	50
27	0.5	0.9	0.9	100	50	145	72.5	100	50

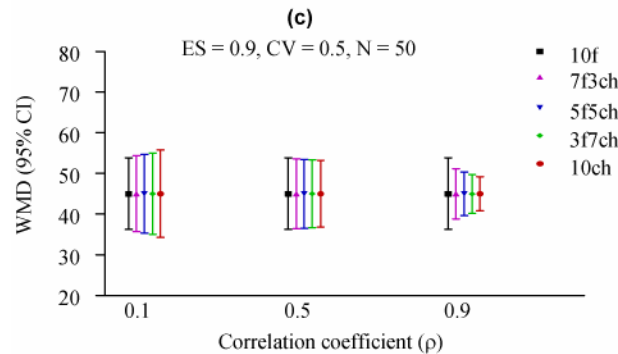
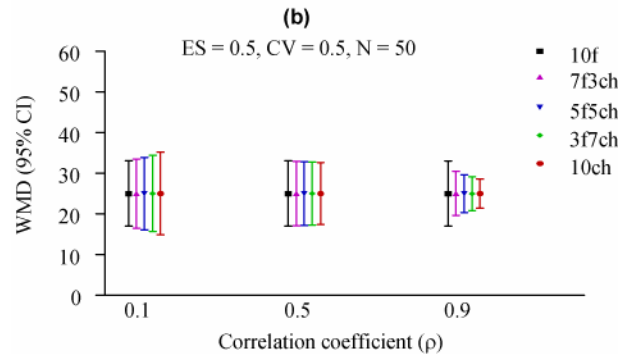
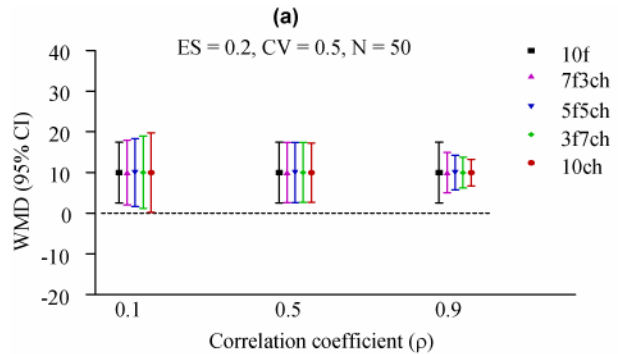
APPENDIX 2A: Weighted Mean Difference (N = 10)

Means of the weighted mean difference (WMD) and 95% confidence interval (CI) obtained from 10,000 meta-analyses are plotted at three values of the correlation coefficient ($\rho=0.1, 0.5, 0.9$). At each value of ρ there are a set of five meta-analyses results, starting from the left with results obtained using 10f (■), 7f3ch (▲), 5f5ch (▼), 3f7ch (◆), and 10ch (●). For each meta-analysis 10 RCTs were pooled and each RCT had 10 participants ($N=10$) in each group (i.e. a total of 20 participants). Graphs for coefficient of variation (CV)=0.5



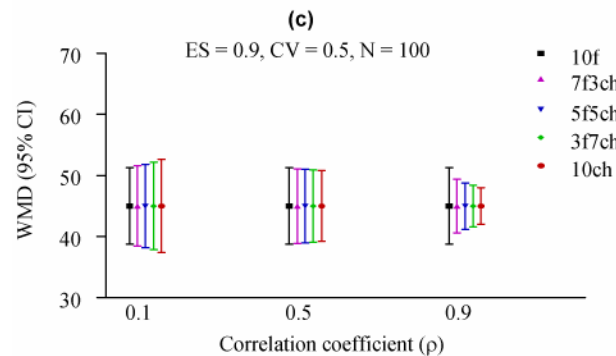
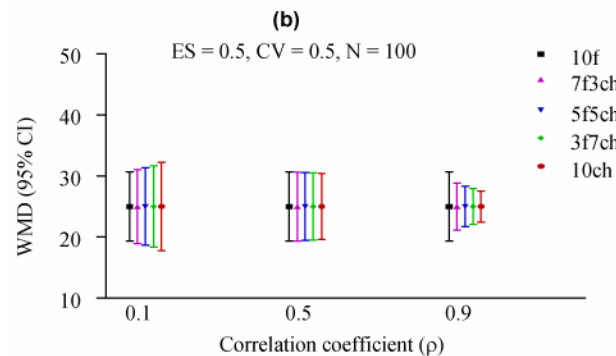
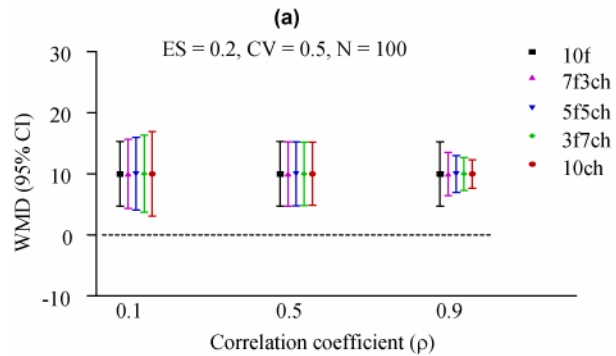
APPENDIX 2B: Weighted Mean Difference (N = 50)

Means of the WMD and 95% CI obtained from 10,000 meta-analyses are plotted at three values of ρ ($\rho=0.1, 0.5, 0.9$). At each value of ρ there are a set of five meta-analyses results; starting from the left with results obtained using 10f (■), 7f3ch (▲), 5f5ch (▼), 3f7ch (◆), and 10ch (●). For each meta-analysis 10 RCTs were pooled and each RCT had 50 participants ($N=50$) in each group (i.e. a total of 100 participants). Graphs for $CV=0.5$



APPENDIX 2C: Weighted Mean Difference (N = 100)

Means of the WMD and 95% CI obtained from 10,000 meta-analyses are plotted at three values of the ρ ($\rho=0.1, 0.5, 0.9$). At each value of ρ there are a set of five meta-analyses results, starting from the left with results obtained using 10f (■), 7f3ch (▲), 5f5ch (▼), 3f7ch (◆), and 10ch (●). For each meta-analysis 10 RCTs were pooled and each RCT had 100 participants ($N=100$) in each group (i.e. a total of 200 participants). Graphs for $CV=0.5$



APPENDIX 3: Equations for Calculations

a) Bias

Bias=Mean of WMD - True MD = Mean of WMD – Effect size×SD_{baseline}

b) 95% confidence interval (CI) for proportion (ρ):

$$\text{Lower 95\% CI} = \frac{(2np + z_{\alpha/2}^2 - 1) - z_{\alpha/2} \sqrt{z_{\alpha/2}^2 - [2 + (1/n)] + 4p(nq + 1)}}{2(n + z_{\alpha/2}^2)}$$

$$\text{Upper 95\% CI} = \frac{(2np + z_{\alpha/2}^2 + 1) + z_{\alpha/2} \sqrt{z_{\alpha/2}^2 + [2 - (1/n)] + 4p(nq - 1)}}{2(n + z_{\alpha/2}^2)}$$

where in the present case n=number of meta-analyses; q=1-ρ, and $z_{\alpha/2}=1.96$