2021

Evaluation of the Effect of the Clinical-Decision-Support Systems on Diabetes Management: A Multivariate Meta-Analysis Comparison with Univariate Meta-Analysis

Abdelfattah Elbarsha

Evaluation of the Effect of the Clinical-Decision-Support Systems on Diabetes

Management: A Multivariate Meta-Analysis Comparison with Univariate Meta-Analysis

_____

A Dissertation

Presented to

the Faculty of the Morgridge College of Education

University of Denver

_____

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

_____

by

Abdelfattah Elbarsha

June 2021

Advisor: Dr. Kathy Green

Author: Abdelfattah Elbarsha
Title: Evaluation of the Effect of the Clinical-Decision-Support Systems on Diabetes
Management: A Multivariate Meta-Analysis Comparison with Univariate Meta-Analysis
Advisor: Dr. Kathy Green
Degree Date: June 2021

Abstract

The advantage of using meta-analysis lies in its ability in providing a quantitative summary of the findings from multiple studies. The aim of this dissertation was first to conduct a simulation study in order to understand what factors (sample size, between-study correlation, and percent of missing data) have a significant effect on meta-analysis estimates and whether using univariate or multivariate meta-analysis would produce different estimates.

The second goal of this study was to evaluate the effect of clinical decision support systems CDSS on diabetes care management by conducting three separate univariate meta-analyses and one multivariate meta-analysis. CDSS are health information technology systems that analyze data within electronic health records (EHR) to help make decisions about a patient's care. Several studies reported inconsistent conclusions about how effective CDSSs are on diabetes care management based on three indicators. Low-density lipoproteins (LDL), glycated hemoglobin (HbA1c), and blood pressure (PB) have been used as indicators of diabetes care management according to the National Institute for Health and Care Excellence (NICE) guidelines. To combine the results from studies that evaluate the effect of CDSSs on diabetes care management, meta-analysis was used. The results of the two univariate and multivariate meta-analyses were compared.

The simulation study indicated that MVMA was less affected by missing values compared to UVMA. However, both methods performed equally when no missing data were present. The standard errors of the estimates in both methods were reduced by increasing the sample size with more reduction in standard errors found in MVMA. The results of UVMA and MVMAs of CDSSs' effect concluded that CDSSs had a significant effect on reducing levels of HbA1c. CDSSs was only significant on LDL when UVMA was applied while pulse pressure (PP) was only affected by CDSSs in the case of MVMA with deleted missing values. CDSSs in general could have a potential effect on diabetes care management.

The results of the simulation and the meta-analyses of the CDSSs indicated that MVMA performed slightly better at different sample sizes and percent of missingness levels than did UVMA.

Acknowledgments

First and foremost, for my brother, whom I love and miss so much, I dedicate this dissertation to your soul. And I know that you would be proud of me. You are always in my heart. I would also like to express my appreciation to my parents, my wife and my children. It would be impossible for me to complete my study without their marvelous understanding, support, and encouragement in the past few years. Finally, I am extremely thankful to Dr. Kathy Green, Dr. Antonio Olmos, Dr. Shimelis Assefa, and Dr. Frédérique Chevillot, for their priceless advice, continuous support, and patience throughout my PhD study. Their enormous knowledge and plentiful experience have encouraged me in my academic research and daily life.

Table of Contents

List of Tables

Page

List of Figures

Chapter One

Introduction and Literature Review

Findings from a single study generally need to be replicated by independent studies for the research and practitioner communities to treat the findings as accurate. It would be unusual for important decisions to be made about, for example, treatment efficacy, based on results of one single study. An analytical approach that integrates the results of independent studies and pools their results into a single typical result is needed. In 1976, Gene Glass referred to the statistical analysis of an extensive collection of results from independent studies for the purpose of integrating the findings as "meta-analysis" (O'Rourke, 2007). While systematic literature reviews provide a framework for the combination of studies, meta-analysis provides a quantitative summary of the findings from multiple studies. In this study, I compared estimates produced by conducting multiple univariate meta-analyses (UVMA) with estimates from multivariate meta-analyses (MVMA) which take into account multiple outcomes simultaneously. This comparison was done using simulation and an empirical comparison using real data from an applied context. In the simulation study, I show under what conditions the utilization of MVMA is optimal and then an empirical comparison between MVMA and UVMA using real data of clinical decision support systems (CDSS) and their effects on diabetes care management follows. There are multiple issues that might affect differences between

MVMA and UVMA. They include between and within study correlation, number of studies included (sample size), and the percent of missing data. Several studies that compared estimates of the UVMA and MVMA reported almost identical results. For instance, Schwarzer et al. (2015) examined the results of univariate and MVMA using a non-random sample of five studies of a systematic review and meta-analysis study done by Lloyd et al. (2010). The result of this comparison yielded almost identical estimates when the within-study correlation was zero and differed by within only one standard error (and so differences were nonsignificant) as the correlation moved from 0.9 to -0.9. Another study by Trikalinos et al. (2014) found that conclusions based on the main effects of each outcome were similar to either univariate or multivariate meta-analysis. So, while the within-study correlation would seem important, no support for its effects on outcomes has been found.

Another problem that might affect meta-analysis results is when there are missing outcome data. Missing data present a threat to the validity of any meta-analysis of research studies. In any analysis, it is assumed that the data are missing completely at random (MCAR) or missing at random (MAR). However, if the data are analyzed as if they were MCAR or MAR but in fact are not MAR, then bias typically occurs (Ellington et al., 2015). Missing data is a common issue in meta-analysis. For instance, in the case of UVMA, studies that do not have all outcomes of interest will be excluded from the study and that might be costly in terms of losing information. However, that is not the case in MVMA. Having only one outcome is enough for a study to be included.

Determining how missingness would affect the quality of UVMA and MVMA estimates is vital and this study was designed to assess effects of the amount of missing data.

In this study, the researcher explored under what conditions the estimates of a MVMA are similar to estimates from UVMA. The researcher explored effects of between-study correlation, percentage of missing data, and sample size through a simulation study followed by an empirical comparison of UVMA and MVMA using studies of CDSSs.

The evaluation of the impact of CDSSs in improving the quality of diabetes care was another interest in this study and was the context for an empirical comparison of UVMA and MVMA. A CDSS is "any electronic system designed to aid directly in clinical decision making, in which characteristics of individual patients are used to generate patient-specific assessments or recommendations that are then presented to clinicians for consideration" (Kawamoto et al., 2005). Several indicators were used as a guide in assessing the quality of care. Those indicators were: low-density lipid cholesterol (LDL-C), glycated hemoglobin (HbA1c), and blood pressure (BP). It has been shown that the simultaneous control of HbA1c, BP, and LDL-C reduces the risk of diabetes complications and death (Hu et al., 2016a). The indicators that the researcher used as a guide follow the National Institute for Health and Care Excellence (NICE) guidelines.

Research Questions and Hypotheses

This study was designed to answer six research questions and seven hypotheses. One question was answered using a simulation. The other questions and their associated

hypotheses were answered by conducting UVMA and MVMA. Hypotheses are stated in alternative form.

Simulation study research questions and hypotheses

Q1. Are there differences in parameter estimates when method (UVMA/MVMA), sample size, between-study correlation, and the percentage of missingness are varied?

H_1: There are interpretable main effects (Partial eta squared $\eta_p^2 \geq 0.01$) of sample size, between-study correlation, and the percentage of missingness on the parameter estimates (effect size, standard error).

H_2: There are interpretable interactions (Partial eta squared $\eta_p^2 \geq 0.01$) among sample size, between-study correlation, and the percentage of missingness.

H_3: There is an interpretable main effect (Partial eta squared $\eta_p^2 \geq 0.01$) of method (UVMA/MVMA) on parameter estimates.

UVMA research questions and hypotheses

Q2. What are the effects of CDSSs in controlling HbA1c levels among diabetic patients?

H_4: CDSSs have a significant effect in controlling HbA1c levels for diabetic patients.

Q3. What are the effects of CDSSs in controlling BP levels for diabetic patients?

H_5: CDSSs have a significant effect in controlling BP levels among diabetic patients.

Q4. What are the effects of CDSSs in controlling LDL-C levels among diabetic patients?

H_6: CDSSs have a significant effect in controlling LDL-C levels for diabetic patients.

MVMA research question and hypotheses

Q5. What are the effects of CDSSs on simultaneous control of HbA1c, BP, and LDL-C among diabetic patients?

H_7: CDSSs have a significant effect on simultaneous control of HbA1c, BP, and LDL-C among diabetic patients.

UVMA and MVMA comparison research question

Q6. Is there a difference in the estimates between multivariate meta-analysis and univariate meta-analysis of CDSSs?

This study is organized as follows. In the following section, the MVMA model and its applications are reviewed, with emphasis on studies that have used both MVMA and UVMA. UVMA and MVMA computation and fitting techniques are then addressed and followed by a conceptual comparison. Then, issues related to and limitations of MVMA are discussed. Finally, a summary of studies that evaluated the effect of the CDSSs in improving diabetes care are presented.

Review of the Literature

Meta-analysis has been promoted for over 30 years, with applications in education, dentistry, clinical trials, survival, marketing, surrogate outcomes, prognostic markers, diagnostic tests, and genetics among others. According to Jackson et al. (2011), the most common areas where the application of meta-analysis has been particularly successful are: (a) diagnostic test meta-analysis; (b) multiple effects in randomized controlled trials or observational studies; (c) multiple parameter models for exposure in observational studies; and (d) network meta-analysis. The bivariate meta-analysis of studies of diagnostic test quality is probably the most common medical application of

meta-analysis. In the case of multiple effects and in any situation where clinical trials or observational studies have more than one outcome of interest, an MVMA could be applied. In the third case listed above (c), the aim is to pool information across studies for exposure parameters that characterize effects of specific interest. In network analysis, multiple treatments are compared across studies that provide results for multiple treatment groups. UVMA dominates meta-analysis studies to date, though many research projects deliberately include multiple outcome measures which would make MVMA seem the obvious analytic choice.

Multiple approaches are available to compute multivariate effect sizes. Averaging the multivariate effect sizes within each study is one of them. It is adequate to apply univariate meta-analysis since averaged effect sizes are independent across studies. However, averaging is usually not suitable when the multivariate effect sizes are measuring non-combinable different constructs such as academic achievement and student engagement. Another approach is to meta-analyze each effect size separately. This approach is relatively easy to implement. However, the dependence among the effect sizes is ignored as is any differences in between-study dependencies. A third approach is modeling the multivariate effect sizes simultaneously where the dependency among the effect sizes is taken into account. This is usually more appropriate than conducting separate univariate meta-analyses since MVMA employs the correlation among the multivariate effect sizes (Cheung, 2013).

Univariate Meta-Analysis Effect Size Computation

In meta-analysis, an important step is computation of the effect size. Effect sizes should be carefully computed since they represent essential information that will be

extracted from the studies included. Pearson correlation coefficients (r), standardized mean differences (g), and odds ratios (OR) are common indices that represent effect sizes. It should be noted that significance tests are not effect sizes and *vice versa* (Card, 2015). In general, there are two models in meta-analysis, fixed-effects models and random-effects models, with mixed-effects models reflecting a mixture of fixed- and random-effects models. In the case of the fixed-effects models, it is assumed that the population effect sizes are the same across studies. However, in random-effects models, each study assumed to have its own effect sizes. Fixed-effects models are suitable if the researcher has included in the meta-analysis all (or mostly all) studies that are available for population of interest. In this case, the interest is to draw conclusions from the included studies. If the researcher intends to generalize findings, random- or mixed-effects models are more suitable. Next, the three indices of effect sizes are described, and their formulas presented before the advantages and disadvantages of applying MVMA are reviewed.

Pearson Correlation (r)

The Pearson correlation coefficient represents the association between two continuous variables and the formula used for the computation of r is:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{(N-1)s_x s_y} = \frac{\sum Z_X Z_Y}{N} \qquad \text{Eq. 1}$$

where

x_i and y_i are the values of individual i on the two variables

$\bar{x}$ and $\bar{y}$ are the sample means of the two variables.

N is the sample size

s_x and s_y represent the population estimated standard deviations of the two variables.

$Z_X = \frac{x_i - \bar{x}}{s_x}$ and $Z_Y = \frac{y_i - \bar{y}}{s_y}$ are the standardized scores.

A Correlation ρ itself is directly taken as an effect size. Even though Pearson's r is considered as an interpretable index of effect size for the association between two continuous variables, r is transformed prior comparing effect sizes across studies. One of the most common transformation is Fisher's transformation (Z_r) as shown below in Eq. 2.

$$Z_r = \frac{1}{2} ln \left(\frac{1+r)}{1-r} \right) \qquad \text{Eq. 2}$$

And the standard error of Z_r is

$$SE_{Z_r} = \frac{1}{\sqrt{N-3}} \qquad \text{Eq. 3}$$

where

Z_r represents Fisher's transformation of r.

r is Pearson's correlation coefficient.

N is the sample size of the study.

The reason behind transforming r to Z_r is that the distribution of r around a given nonzero population ρ is skewed, especially when the sample size is not large enough, whereas the distribution of Z_r around a nonzero population ρ is symmetric (Card, 2015).

Standardized Mean Difference (g)

The standardized mean difference represents "the magnitude of difference between the means of two groups as a function of the group's standard deviation" (Card, 2015). According to Card 2015, there are three common indices of standardized mean

difference. These indices are Hodges's (g), Cohen's (d), and Glass's index (g_{Glass}). The equations below represent the three indices respectively.

$$g = \frac{M_1 - M_2}{S_{pooled}} \quad \text{Eq. 4,} \qquad d = \frac{M_1 - M_2}{Sd_{pooled}} \quad \text{Eq. 5,} \qquad g_{Glass} = \frac{M_1 - M_2}{S_1} \quad \text{Eq. 6}$$

where

M_1 and M_2 are the means of the two groups.

$S_{pooled} = \sqrt{\frac{(x_i - \bar{x})^2}{(n-1)}}$ represents the pooled estimates of the population standard deviation.

$Sd_{pooled} = \sqrt{\frac{(x_i - \bar{x})^2}{n}}$ is the pooled sample standard deviation.

S_1 is the estimate of the population standard deviation from the control group.

The standard error of g is given below

$$SE_g = \sqrt{\frac{n_1 + n_2}{n_1 n_2} + \frac{g^2}{2(n_1 + n_2)}} \approx \frac{4g^2}{N_{total}} \qquad \text{Eq. 7}$$

where

n_1 and n_2 represent the sample sizes of group 1 and group 2.

N_{total} is the total sample size of the study assuming equal sample size per group.

Correction is needed when the primary study sample size is small. The standardized mean difference has been shown to be a biased estimate when the sample sizes are small (less than 20). Therefore, the following adjustment should be applied for small sample size (French et al., n.d.):

$$g_{adjusted} = g - \frac{3g}{4(n_1 + n_2) - 9} \qquad \text{Eq. 8}$$

Odds Ratio (OR)

Another useful index of effect size of the association between two dichotomous variables is the odds ratio (OR). OR is defined as the probability of the event divided by the probability of the alternative (Card, 2015). The0 OR can be calculated from a 2×2 table using the following formula:

$$ES_{OR} = \frac{ad}{bc} = \frac{P_a P_d}{P_b P_c} = \frac{P_a/P_b}{P_c/P_d} = \frac{P_a(1-P_c)}{P_c(1-P_a)}$$

Eq. 9

where a, b, c, and d represent the cell frequencies and P_a, P_b P_c, and P_d are the proportion of each group in each status.

In meta-analysis, the natural log of the OR is used, which has an approximately normal distribution with a mean of 0 and a standard deviation of 1.83 (Lipsey & Wilson, 2001). By using the logged odds, the interpretation of the effect size becomes clearer and makes the calculation of the standard error easier. The logged OR, standard error, and inverse variance weight can be calculated as the following:

$$ES_{LOR} = Log(ES_{OR})$$

Eq. 10

$$SE_{LOR} = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

Eq. 11

$$W_{LOR} = \frac{1}{SE_{LOR}} = \frac{abcd}{ab(c+d)+cd(a+b)}$$

Eq. 12

Multivariate Meta-Analysis Effect Size Computation

I start with the model of bivariate meta-analysis and then extend it to MVMA model for simplicity. Therefore assume $j = 1,2$ so that each study i yields two estimated treatment effects: $y_i = (y_{i1}, y_{i2})'$. In the fixed effects case, the study estimates will follow a bivariate normal distribution (Mavridis & Salanti, 2013).

$$\begin{pmatrix} y_{i1} \\ y_{i2} \end{pmatrix} \sim BND \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_{i1}^2 & \rho_i \sigma_{i1} \sigma_{i2} \\ \rho_i \sigma_{i1} \sigma_{i2} & \sigma_{i2}^2 \end{pmatrix} \right) \qquad \text{Eq. 13}$$

where ρ_i represents the within-study correlation between outcomes $j = 1,2$ for study i.

The vector $(\mu_1, \mu_2)'$ is the vector of means for each outcome. The matrix $S_i =$

$\begin{pmatrix} \sigma_{i1}^2 & \rho_i \sigma_{i1} \sigma_{i2} \\ \rho_i \sigma_{i1} \sigma_{i2} & \sigma_{i2}^2 \end{pmatrix}$ represents the within-study covariance matrix. The overall

correlation in the two-dimensional case is split into two components, within-study

correlation (ρ_i) and between-study correlation (ρ_τ) where:

$$\begin{pmatrix} y_{i1} \\ y_{i2} \end{pmatrix} \Big| \begin{pmatrix} \theta_{i1} \\ \theta_{i2} \end{pmatrix} \sim BND \left(\begin{pmatrix} \theta_{i1} \\ \theta_{i2} \end{pmatrix}, \begin{pmatrix} \sigma_{i1}^2 & \rho_i \sigma_{i1} \sigma_{i2} \\ \rho_i \sigma_{i1} \sigma_{i2} & \sigma_{i2}^2 \end{pmatrix} \right) \qquad \text{Eq. 14}$$

The vector $\theta = (\theta_{i1} \quad \theta_{i2})'$ is the study-specific effects for each outcome. θ is also

normally distributed.

$$\begin{pmatrix} \theta_{i1} \\ \theta_{i2} \end{pmatrix} \sim BND \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \tau_1^2 & \rho_\tau \tau_1 \tau_2 \\ \rho_\tau \tau_1 \tau_2 & \tau_2^2 \end{pmatrix} \right) \qquad \text{Eq. 15}$$

where τ_j is defined as the between-study variation (heterogeneity) for effect size j.

When jointly meta-analyzed the outcomes, in addition to the matrix S_i, there is also a

between-study covariance matrix, $\Delta = \begin{pmatrix} \tau_1^2 & \rho_\tau \tau_1 \tau_2 \\ \rho_\tau \tau_1 \tau_2 & \tau_2^2 \end{pmatrix}$. By combining Eq. 14 and Eq.

15 we get

$$\begin{pmatrix} y_{i1} \\ y_{i2} \end{pmatrix} \sim BND \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_{i1}^2 + \tau_1^2 & \rho_i \sigma_{i1} \sigma_{i2} + \rho_\tau \tau_1 \tau_2 \\ \rho_i \sigma_{i1} \sigma_{i2} + \rho_\tau \tau_1 \tau_2 & \sigma_{i1}^2 + \tau_2^2 \end{pmatrix} \right) \qquad \text{Eq. 16}$$

Eq. 16 can be extended for the case of p outcomes as follows

$$\begin{pmatrix} y_{i1} \\ \vdots \\ y_{ip} \end{pmatrix} \sim MND \left(\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix}, \begin{pmatrix} \sigma_{i1}^2 + \tau_1^2 & \cdots & \rho_i \sigma_{i1} \sigma_{ip} + \rho_{\tau(1,p)} \tau_1 \tau_p \\ \vdots & \ddots & \vdots \\ \rho_i \sigma_{i1} \sigma_{ip} + \rho_{\tau(1,p)} \tau_1 \tau_p & \cdots & \sigma_{ip}^2 + \tau_p^2 \end{pmatrix} \right) \qquad \text{Eq. 17}$$

11

The following equation is the matrix representation of the random-effects model

$$y_i = \mu + \delta_i + \varepsilon_i \qquad\qquad \text{Eq. 18}$$

where δ_i is a vector of random effects of the study i where $\delta_i \sim MVN(0, \Delta)$ and ε_i

represents a vector of random sampling error of the study i. ε_i is independent of δ_i and it

is normally distributed $\varepsilon_i \sim MVN(0, S_i)$. The matrix Δ is the between-study covariance as

defined previously but in this case, it is involving two unknown parameters τ_j and $\rho_{\tau(jj')}$

$$\Delta = \begin{pmatrix} \tau_1^2 & \cdots & \rho_{\tau(1,p)}\tau_1\tau_p \\ \vdots & \ddots & \vdots \\ \rho_{\tau(1,p)}\tau_1\tau_p & \cdots & \tau_p^2 \end{pmatrix} \qquad\qquad \text{Eq. 19}$$

the variance–covariance matrix of y_i is $\Delta + S_i$ and by including l covariates in the model

in Eq. 17 so that: $\qquad\qquad \mu = X_i\beta y_i = X_i\beta + \delta_i + \varepsilon_i$

Eq. 20

The previous equation is the general random-effects multiple outcomes meta-

regression model which is also known as a mixed effects model. X_i represents the

$p \times (l + 1)$ the observed covariate values matrix for each study and β is the vector of l

coefficients and the constant term.

Fitting multivariate meta-analysis models

There are several approaches in order to estimate the MVMA model parameters.

In the MVMA model, the parameters of interest are $\boldsymbol{\mu}$, the vector of the effect estimates

for the $\boldsymbol{p}$ outcomes, $\boldsymbol{p} \times \boldsymbol{p}$ variance–covariance matrix $\boldsymbol{C}$, and the heterogeneities τ_j^2 and

between-studies correlations $\boldsymbol{\rho_\tau}$ represented by the matrix Δ. In this section, I summarize

several estimation methods in the case of the random-effects model since the fixed effects

estimates computations are simpler and are considered as a special case.

Likelihood methods

To estimate model parameters, likelihood methods can be used with the assumption of the independence of the studies. The likelihood is defined as:

$$L \approx -\frac{1}{2}\sum_{i=1}^{n} log|\Delta + S_i| - -\frac{1}{2}\sum_{i=1}^{n} e_i'(\Delta + S_i)^{-1} e_i \qquad \text{Eq.21}$$

The only disadvantage of the likelihood methods is that when the number of studies is large, they become computationally intensive and time consuming.

Maximum likelihood (ML) estimates

With the assumption that all studies have the same outcomes and there are no missing values, the effect estimates could be estimated by maximizing the likelihood as follows:

$$\hat{\mu} = \left(\sum_{i=1}^{n}(\hat{\Delta} + S_i)^{-1} \sum_{i=1}^{n}(\hat{\Delta} + S_i)\right)^{-1} y$$

Eq.22

The estimates produced by Eq.22 are approximately normally distributed with variance covariance matrix:

$$\hat{C} = \left(\sum_{i=1}^{n}(\hat{\Delta} + S_i)^{-1}\right)^{-1}$$

Eq.23

Restricted ML (REML)

This estimation approach is very common in the literature since it produces unbiased estimates of variance and covariance parameters. The REML is defined as:

$$RL \approx -\frac{1}{2}\left|\sum_{i=1}^{n}(\hat{\Delta} + S_i)^{-1}\right| \qquad \text{Eq.24}$$

By maximizing the likelihood functions, positive definiteness will be ensured in Δ.

Estimates of μ and C could be obtained by using the estimated $\hat{\Delta}$ in equations (22) and

(23).

Generalized least squares (GLS)

In this method, it is assumed that each outcome could be modelled by a regression

line and regression models are not independent and therefore, correlations are considered.

The matrix X when there are no covariates in the model is defined as:

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$$

where X_i is the identity $p \times p$ matrix. GLS could be maximized by

$$(y - X\mu)'S^{-1}(y - X\mu) \qquad\qquad \text{Eq.25}$$

In the case of the fixed effects, the GLS estimates could be obtained as the follows:

$$\hat{\mu} = (X'S^{-1}X)^{-1}X'S^{-1}y \qquad\qquad \text{Eq.26}$$

The variance-covariance matrix is $\hat{C} = (X'S^{-1}X)^{-1}$. The random effects estimator

estimates $\hat{\mu}$ iteratively until a successful convergent is gained, and $\hat{\Delta} = \dfrac{1}{(n-2)e'e - \frac{1}{n}\sum_{i=1}^{n} S_i}$

where $e = y - \hat{\mu}$.

In reviewing the effect size computation formulas for UVMA and MVMA,

differences appear in how both approaches are constructed. MVMA uses matrix notation

since more than one outcome variable is included in a simultaneous analysis while only

one outcome is handled in the UVMAs. A primary difference between UVMA and

MVMA is that in UVMA, outcomes are analyzed separately assuming that they are

independent. In MVMA, all outcomes are analyzed simultaneously. The importance of

MVMA comes when two or more related outcomes within the same study are of interest. The relationship between these outcomes is known as the "within-study correlation" while the relationship between the outcomes across studies is known as "between-studies correlation." Both within and between study correlations are assumed to be "zero" in UVMA, neglecting the relationships between the multiple outcome measures. Looking closely at the MVMA fitting models discussed earlier, it can be noticed that the within-study covariance (correlation) ρ_i appears on the total variance matrix $(\hat{\Delta} + S_i)$ and more specifically the inverse of this matrix $(\hat{\Delta} + S_i)^{-1}$. ρ_i is involved through the total variance matrix which in turn is involved in treatment effect and the heterogeneity covariance matrix computations. Therefore, ρ_i should have some impact on the estimation of the MVMA such as producing estimates that have smaller standard errors as well as improving the estimation of the between-study variances. As a result, we should generally expect more precise estimates (Jackson et al., 2011). This effect, however, has not been found in the literature.

As in MANOVA, MVMA allows us to measure several dependent variables simultaneously which will increase the probability of discovering which outcome is truly important. Additionally, using MVMA could decrease the chance of committing Type I errors that might occur if multiple UVMAs were conducted independently. Lastly, differences that are not discovered by UVMA could be revealed by using MVMA (French et al., 2008). MVMA models depend on iterative procedures rather than closed-form analytical solutions in the estimation of the parameters. Based on that, knowing when MVMA is ideal to use and when it would give different estimates compared to UVMA is

difficult to reach analytically (Ishak et al., 2008). In this case, conducting simulations would be more efficient and would help in finding cases when MVMA is ideally utilized.

Advantages and Disadvantages of MVMA

Several advantages are offered by MVMA that make it superior to separate UVMA of each outcome, which are conducted under the assumption of independent outcomes (Riley, 2009). According to Jackson et al. (2011), the utilization of MVMA methods can be beneficial and can provide estimates with better statistical properties than UVMA. However, these benefits in return require meeting more assumptions which may not result in better inference in every case. Typical assumptions made with MVMA are (a) the multivariate normality assumption, (b) a multivariate linear relationship between outcomes, and (c) a constant between-studies covariance matrix. However, for instance, the multivariate normality assumption is usually hard to uphold. Furthermore, a linear relationship between studies' effects is needed since it is hard to estimate nonlinear relationships with a limited number of studies.

Researchers may obtain different conclusions when using MVMA compared to UVMA. Jackson et al. (2011) stated that conclusions drawn from a MVMA might sometimes vary from those from a UVMA. The authors supported their claim using an example where the aim was to describe the relationship between fasting glucose levels and cardiovascular disease. The univariate meta-analysis yielded a significant log hazard ratio; however, the multivariate meta-analysis gave a lesser, non-significant log hazard ratio. Carrying out a single MVMA would be more efficient than doing many univariate ones. All MVMA parameter estimates are simultaneously provided in a single analysis. Therefore, it would be easier to compare the results from different outcome variables

which may lead to a different conclusion than from separate UVMAs. Additionally, the researchers reported that the utilization of a MVMA method could reduce bias due to partial reporting since in UVMA, studies that do not have all the variables of interest are excluded.

The quality of the estimates of MVMA and UVMA has been compared in situations with two parameters. In the fixed- and random-effects meta-analysis, as the number of parameters increases, MVMA benefits can increase substantially. However, in the case of the random-effects meta-analysis and when high between-study variability is present, the possible improvement would be small. When all studies have a common between-study covariance matrix, the covariance matrices become even more similar as between-study variance increases, reducing the benefit of MVMA (Boca et al., 2017). Additionally, the actual improvement with MVMA is further shrunken by the need to estimate an increasingly large between-study covariance matrix. Also, when the between-study variability is minimal, or zero, the loss of the effectiveness by choosing random-effects meta-analysis over fixed-effects meta-analysis increases as the number of parameters increases (Boca et al., 2017).

The importance of MVMA is highlighted when there is more than a single outcome of interest, which might present a challenge since the within-study correlations must be present. Knowledge of the within-study correlations is usually *unavailable* in practice since the correlations are not typically reported. In order to address this limitation, several analysis methods for dealing with unknown within-study correlations have been proposed. Wei & Higgins (2013) proposed an approach for the approximation of the within-study covariances based on data about possible correlations between

outcomes under study. The authors argued that when heterogeneity of effects across studies is present and when there is a high correlation within studies, the proposed approach to approximation of covariances performs better than others. Another method was proposed by Y. Chen et al. (2016) that introduced a simple non-iterative method. The authors claimed that the method could be helpful for MVMA since the within-study correlations are not required for the analysis. The proposed method is based on the use of standard univariate methods for the marginal effects as well as producing a joint inference for multiple parameters. In addition, the researchers stated that, based on simulation studies, the proposed method provides unbiased estimates, good estimated standard errors, and good confidence intervals. This method is claimed to have high relative effectiveness when compared with classic MVMA where the within-study correlations are known. In a recent study by Lin & Chu (2018), a new approach called "multivariate meta-analysis of multiple factors" has been introduced to synthesize data from all available factors simultaneously. The authors claim that MVMA of multiple factors can improve statistical efficiency and reduce biases compared with separate analyses by carrying information across factors. A Bayesian hybrid model is used to conduct MVMA of multiple factors in order to account for both within- and between-study correlations (that are usually unavailable from published articles). The performance of MVMA of multiple factors and the traditional methods were compared by the researchers using simulations. The hybrid model was found to be effective in reducing complexity by specifying a joint marginal correlation matrix for all studies. However, the researchers stated that if the collected studies' marginal correlation matrices vary extremely, a poor fit might be produced by the hybrid model.

Despite these challenges, MVMA has been successfully applied in many different settings. It has been used increasingly in the educational and medical sciences. In education, for instance, MVMA has been applied in numerous studies in order to assess the effects of an intervention. For example, evidence for Classroom Management Self-Efficacy (CMSE) in relation to three dimensions of burnout: emotional exhaustion, depersonalization, and (lowered) personal accomplishment was examined using a MVMA. The authors stated that their study was the only meta-analysis that examined classroom management self-efficacy and teacher burnout; additionally, they stated that it was the first MVMA conducted within the educational psychology field. The use of MVMA was adopted in this study since they had three outcomes (dimensions of burnout using a measure of association) and MVMA allowed the inclusion of the correlations among the three dimensions of burnout. The results suggested that there was a significant relationship between classroom management self-efficacy and the three dimensions of burnout, meaning that teachers with lower levels of CMSE were more likely to experience the feelings of burnout and *vice versa* (Aloe, Amo et al., 2014). In a similar study, (Aloe, Shisler, et al., 2014) explored the relationship between student misbehavior and the same three dimensions of teacher burnout (i.e., emotional exhaustion, depersonalization, and personal accomplishment). They included a total of 21 independent studies and the results suggested that students' misbehavior was significantly associated with the three dimensions of teacher burnout.

The use of MVMA has also increased in clinical research where multiple outcomes are likely to be the case. MVMA can play an essential role in determining which treatment should be recommended for a specific condition. For instance, a study,

conducted by Del Re et al. (2013), reported that an earlier review found that the effect sizes of oral naltrexone, which is an FDA-approved medication for treating alcohol use disorders on relapse to heavy drinking and, to a lesser extent, percent days drinking, were smaller in more recent trials and in multicenter trials than in single-site studies. They examined whether these results apply when considering studies from 2004 to 2009 and whether single-site versus multicenter trials, the use of placebo run-in periods, and placebo group improvement accounted for variation in naltrexone effects and reducing effects over time. In another example of a study that used MVMA in a clinical setting, the power of a MVMA was evaluated by applying the method to existing two-dimensional gel electrophoresis data from human prostate and colon tumors to extract valuable information since numerous cancer two-dimensional gel electrophoresis studies have stated partially redundant lists of differently expressed proteins. Fourteen proteins were identified with a common trend between the tumor types (prostate and colon). By utilizing multivariate meta-analysis, a common protein profile for two malign tumor types was successfully determined, which would not be the case if data sets were analyzed separately (Rosenberg et al., 2010). In a more recent study that used MVMA for a clinical goal, Y. Q. Zhang et al. (2017) evaluated the severe effects of daily mean temperature, cold spells, and heat waves on stroke mortality in 12 counties in China. Researchers gathered data associated with daily mortality from stroke and meteorology in the 12 counties through 2009-2012. In this study, a MVMA was utilized in order to understand the community-specific associations between temperature and stroke mortality. In addition, they were also interested in understanding the effect of cold- and-heat-associated risks on mortality at different lag days. There are numerous additional

multivariate studies that have been applied in different settings, but for the present study purpose only a few examples have been discussed.

Comparisons of MVMA and UVMA

In meta-analysis, it is very common to have multiple outcomes that can be analyzed separately by conducting independent meta-analyses (UVMA) or by analyzing them jointly in a single model (MVMA). In this section, I review articles that aimed to compare separate (UVMA) with joint (MVMA) meta-analysis.

Trikalinos and Olkin (2012) showed an example of a comparison of the multivariate model with the UVMA at multiple time-points. They found that the results of UVMA and MVMA analyses were almost identical, with a slight difference in the values of the effect sizes and the relative standard errors. They reported that when the within-study covariances are zero or are all equal, estimates were the same in MVMA and UVMA. However, they noted that the standard errors of the estimates were generally slightly different between UVMA and MVMA.

Simulations have been used to compare UVMA with MVMA. Trikalinos et al. (2013) conducted a comparison using real data and a simulation study. The Cochrane Library of Systematic Reviews was screened to identify UVMA studies of categorical outcomes that could be jointly analyzed (MVMA). The data were then analyzed with UVMA and MVMA. The summary estimates and the relative standard errors of the UVMA and MVMA were compared in an accompanying simulation study. The difference in summary effects and their confidence intervals between UVMA and MVMA was almost always small in both the empirical sample and the simulation study. The author

suggested using MVMA in estimating differences between outcome-specific summary treatment effects.

The performance of the MVMA approach was compared with the common UVMA inverse-variance weighted approach in an extensive simulation. The study explored different meta-analytic scenarios of genetic association studies of correlated end points. In this simulation, the findings suggested that the performance of the MVMA approach produced similar or better estimates than the UVMA method when the within- or between-studies correlations are at least moderate. The study showed that the MVMA approach yields smaller bias and root mean square error (RMSE) estimates (Neupane & Beyene, 2015).

Another comparison study was conducted by Lin and Chu (2018) utilized Bayesian multivariate meta-analysis. The Bayesian MVMA made it feasible to import informative prior distributions, specifically on correlations in the MVMA model. In this study, the Bayesian MVMA was used to synthesize data on correlated outcomes in rheumatoid arthritis and to embody informative prior data in the model. A Bayesian hybrid model was used to perform MVMA since it accommodates both within- and between-study covariances which are commonly unavailable from published articles. The five-dimensional health-related quality of life measure (EuroQol) was used to map the estimates of a health assessment questionnaire, and then the effect was compared with mapping the health assessment questionnaire obtained from the UVMA. UVMA yielded larger bias and root mean square errors. The hybrid model that was used effectively minimized model complications by assigning a common marginal correlation matrix for all studies.

By reviewing the literature, it appears that results are consistent in that the use of MVMA would increase the quality of the estimates and allow the inclusion of extra studies. Also, results agreed that even though the within and between study correlations were taken into account, the summary estimates of UVMA and MVMA were almost identical, with no explanation found regarding this. The literature suggested using MVMA when the within- or between-studies correlations are at least moderate, might produce similar or sometimes better estimates compared to UVMA. The question that one might ask here is why estimates do not vary much when comparing UVMA and MVMA. If UVMA gives similar summary estimates compared to MVMA, should MVMA still be used given that it entails more assumptions which can cause estimation difficulties? Finally, by conducting separate UVMA, several studies will be excluded and if the correlation between studies is assumed to be zero, how significant would that exclusion (resulting in a reduced sample size) be that on the estimates of statistical quality? The impact of between-study correlation and data missingness has yet to be clearly determined.

Clinical Decision Support Systems (CDSS) and Diabetes Care Management

Diabetes care management is critical since unwatched or untreated diabetes could result in severe complications that may damage many vital organs and most likely lead to premature death (Centers for Disease Control and Prevention (CDC), 2000). One of the approaches to diabetes care management explored in this study is use of clinical decision support systems (CDSS). Conclusions of systematic reviews studies that reviewed CDSSs had inconsistent conclusions (Jia et al., 2019a). Therefore, there is a need to evaluate the CDSSs empirically which could improve the integrity and accuracy of the

research. As an assistant lecturer at the University of Benghazi, department of health informatics, I had been working closely with a public health specialist. CDSSs implementation in the Libyan health system was always in debate since the effectiveness of CDSSs is still uncertain. A study that confirms whether the use CDSSs is effective or not will help in making a decision regarding their use. It has been shown that clinical decision support systems have improved in their design, utilization, and effectiveness to manage and improve quality of diabetes care. Current CDSSs have high use rates and high clinician/user satisfaction rates. Also, the use of CDSSs has significantly improved blood pressure control, glucose control, and cardiovascular risk trajectories in diabetic patients. Based on that, CDSSs will likely become essential technologies that help to guide clinician and patient decision-making (Patrick J. O'Connor & Sperl-Hillen, 2019).

A considerable amount of literature has been published where MVMA was the primary analysis method. However, the goal of this study is to compare the effect size estimates of MVMA and UVMA in the context of CDSS use in improving the quality of diabetes care. Therefore, the researcher next sheds light on studies that examined the impact of different CDSSs in order to get a clear picture of the use of the CDSSs and their benefits regarding the quality of diabetes care prior to conducting MVMA and UVMA on CDSS studies. The sequence of the review starts with a summary of studies that examined a specific CDSS in improving diabetes care in order to get an initial insight into whether the use of CDSSs was beneficial or not. Later, systematic reviews that discuss the benefits of using CDSSs in enhancing diabetes care are reviewed in order to reach a more general conclusion from narrative reviews and UVMAs.

The effect of CDSSs in improving different conditions has been explored in numerous studies. However, limited studies have discussed the effectiveness of CDSSs in enhancing the quality of diabetes care. A recent study examined the effect of an electronic health record-based clinical decision support tool on diabetes management in primary care practices participating in Delaware's patient-centered medical home project (Gill et al., 2019). In the quantitative analysis phase, bivariate analyses were conducted to describe the data and to compare outcome measures for patients in the groups. Glycemic and lipid control were analyzed using multivariate regression analyses to control for possible confounding factors. Qualitatively, the staff at each primary care office were interviewed and a research assistant summarized the interview transcripts. One of the authors reviewed and interpreted all results and put together the concluding summary. In sum, the researchers found that the use of clinical-decision-support systems was linked with superior improvements from baseline in hemoglobin A1c and low-density lipoprotein cholesterol. Based on the interviews, physicians and staff stated that the clinical decision support toolkit allowed them to be more engaged in clinical decision-making and thus helping to improve diabetes care (Gill et al., 2019).

Management of diabetes, which is considered a complex chronic disease, may require the integration and the understanding of multiple laboratory test results. Traditional electronic health records tend to visualize lab results in a disorganized and separated way which makes the interpretation of results associated with diabetes care challenging. Sim et al. (2017) developed a diabetes-specific CDSS interface that displays glycemic, lipid, and renal function results. The CDSS graphically summarized all related laboratory results and presented them in a color-coded system which allowed easy and

quick interpretation of the metabolic control of the patients. It also has an alert module

that notifies users of any tests that had to be rerun. An interactive graph module was

added to the CDSS for better graphical visualization of the trends of the lab results. In a

pilot study, the developed CDSS significantly improved the understanding of abnormal

laboratory results when compared to the existing laboratory reporting interface. However,

no significant improvement was found in the identification of patients needing treatment

modification. The researchers reported that the diabetes-specific CDSS interface they

developed could improve the management of diabetes and they expected that this CDSS

would be helpful when applied in an outpatient setting.

The implementation of electronic health records (EHR) is assumed to improve the

quality of ambulatory care, especially for chronic clinical conditions such as diabetes.

However, there had been no comparative studies of longer-term observation of the

quality of care in practices using electronic health records with those using paper records.

To address this gap, Crosson et al. (2012) examined data collected over three years to

conduct this comparison. Some practices had utilized electronic health records previous

to initial data collection and kept using the system during the observation period, while

the other practices used the typical paper records. They analyzed data from 16 practices

that utilized electronic health records and 26 that did not. Measures of care were

evaluated for 798 patients with diabetes. They also noted that they used hierarchical

linear models to examine the relationship between electronic health records use and

obligation to evidence-based diabetes care guidelines. Hierarchical logistic models were

also used in order to compare rates of improvement over three years. Electronic health

records use was not significantly related to better adherence to care guidelines. Patients in

practices that did not use an EHR were more likely to meet all three intermediate outcomes goals for hemoglobin A1c, low-density lipoprotein cholesterol, and blood pressure at the 2-year follow-up. However, the quality of diabetes care improved among all practices. As a conclusion of their study, the authors stated that consistently using EHR over three years might *not* successfully improve the quality of diabetes care.

In some cases, CDSSs provide information that can be shared and discussed by both patient and physician which might improve the management of diabetes. However, according to Holbrook et al. (2009), this has rarely been examined in community-based primary care. Based on that, Holbrook et al. conducted a study in order to assess the effectiveness of a Web-based diabetes color-coded tracker shared between patient and primary care providers in improving the quality of diabetes management in community-based primary care. The researchers randomly assigned adult primary care patients with type 2 diabetes to obtain either the intervention or traditional care. Forty-six primary care providers were sequentially recruited and 511 of their patients. A significantly better process composite score was found for patients in the intervention group compared to control patients. They reported that 61.7% of patients in the intervention group showed improvement compared to an improvement of 42.6% of the control group patients. Additionally, a significant improvement was found in more variables in the intervention group. A significantly higher decline was also found in blood pressure and glycated hemoglobin in the intervention group patients. Greater satisfaction with their diabetes care in the intervention group patients was reported. As a conclusion, the researchers stated that the shared electronic decision-support system improved the process of care and some clinical indicators of the quality of diabetes care.

Continuing with CDSS evaluation, in a study by P. J. O'Connor et al. (2011), an electronic health record-based diabetes clinical decision support system was evaluated. The goal was to explore its impact on the control of hemoglobin A1c (glycated hemoglobin), blood pressure, and low-density lipoprotein (LDL) cholesterol levels in adults with diabetes. In this study, a clinic-randomized trial was conducted where 11 clinics with 41 primary care physicians and 2,556 patients with diabetes were included. Patients were randomly assigned to either intervention group (receive) or control group (not to receive) an electronic health record-based clinical decision support system. General and generalized linear mixed models with repeated time measurements were utilized in order to comply with the nested data structure. Patients in the intervention group had significantly better hemoglobin A1c, systolic blood pressure control, and slightly better maintenance of diastolic blood pressure control. However, no improvement was detected in low-density lipoprotein cholesterol levels compared to patients in the control group. The researchers concluded that electronic health record-based diabetes clinical decision support has a significant effect in improving glucose control and partially improved blood pressure control in patients with type 2 diabetes.

Coronary artery disease (CAD) is a significant factor in the long-term prediction of disease progression in patients with diabetes mellitus (DM) and examining the impact of a CDSS in improving the care of these patients is important (Aronson & Edelman, 2014). Several studies have examined the possible effect of using CDSSs for this condition. In an evaluation of the impact of two different CDSSs in improving and addressing deficiencies in the care of patients with coronary artery disease and diabetes mellitus, Sequist et al. (2005) aimed to examine the effect of an integrated patient-

specific electronic clinical reminder system on diabetes care. They were also interested in evaluating the impact of the introduced CDSS on coronary artery disease care. A total of 194 primary care physicians, 4549 patients with diabetes, and 2199 patients with coronary artery disease at 20 ambulatory clinics were included in the study. Clinics were randomly assigned to the intervention so that physicians received either evidence-based electronic reminders or typical care. The researchers found that electronic reminders increased the chances of recommended diabetes and coronary artery disease care. However, the effect of individual reminders was inconsistent. In general, the researchers argued that an integrated electronic reminder system led to an improvement in care for both diabetes and coronary artery disease.

Following a complex medication routine might cause a struggle with self-care and in managing blood glucose levels for patients with diabetes. In a study by Schnipper et al. (2010), the goal was to evaluate a new documentation-based CDSS (Smart Form) effectiveness in addressing deficiencies in the care of patients with coronary artery disease and diabetes mellitus. In their controlled randomized trial, they randomly assigned primary care physicians in 10 ambulatory practices to usual care or the coronary artery disease /diabetes mellitus Smart Form. Patients of intervention primary care physicians had a better proportion of deficiencies addressed compared with controls. The authors stated that the use of the Smart Form was limited, and a modest improvement in management was detected (Schnipper et al., 2010; Sequist et al., 2005).

Medication nonadherence is common among patients with diabetes mellitus and relates to critical adverse results. Therefore, interventions are needed in order to improve medication management so patients can achieve the promising benefit of prescribed

treatments (Ho et al., 2006). As a contribution to that, Morrow et al. (2012) conducted a project that aimed to improve self-management of medications and related health outcomes by introducing system support. They presented an Electronic Medical Record (EMR)-integrated system designed to enhance patient-physician collaboration required for medication management. The researchers said that the new EMR "helps providers and patients work together to create effective medication schedules that are easy to implement" (Morrow et al., 2012). The researchers stated that an evaluation study to examine the usefulness of the Medtable™ in improving care control condition among diabetic patients struggling to manage multiple medications was planned. However, the effectiveness of Medtable™ has not been evaluated yet.

CDSSs have been widely used in developed countries and their effects have been evaluated in many studies. However, only a few works in the literature demonstrate the effectiveness of CDSSs in developing countries. In Brazil, a study was conducted to analyze the possibility, usability, and clinical influence of a clinical decision support system in Brazilian primary care diabetes patients. A quasi-experimental design was performed and type-2 diabetes primary care patients older than 40 years of age were included. Patients were evaluated before and after the implementation of the CDSS. The CDSS application included clinical assessments and blood glucose measurements and produced detailed recommendations built on the data analyzed. The total number of patients included was 145 patients and 70.0% of them had been diagnosed with diabetes more than five years ago. There was no improvement found in median hemoglobin A1c. The subgroup analysis showed that a significant decrease in median hemoglobin A1c level was observed in patients with a hemoglobin A1c level of ≥ 9% at baseline.

However, this reduction happened before the implementation of CDSS. Healthcare practitioners stated that the CDSS was easy to use and claimed that it provided valuable information for patient care. However, it was concluded that the implementation of the CDSS did not improve the hemoglobin A1c level, and that might have happened because of the short follow-up and/or infrequent CDSS use by the healthcare practitioners (Xavier et al., 2016). Another study conducted in India by Prabhakaran et al. (2019) evaluated an integrated CDSS for multiple chronic condition management in primary care. The researchers were specifically interested in assessing the impact of a "mHealth system mWellcare" for the integrated management of hypertension, diabetes mellitus, current tobacco and alcohol use, and depression compared to the improved usual care among patients with hypertension and diabetes mellitus in India. Community health was randomly assigned to either receive the "mWellcare" or improved usual care. The result of this cluster-randomized controlled trial that involved 40 community health centers yielded a non-significant difference between the two groups for systolic blood pressure and glycated hemoglobin. Likewise, there were no differences between the two groups regarding tobacco and alcohol use or other secondary outcomes. As a conclusion of this study, the researchers concluded that the use of "mWellcare" was not beneficial in the management of the chronic conditions studied.

The evaluation of CDSSs remains limited. More research is needed in developing countries to reach an accurate conclusion about whether the use of CDSSs improves the quality of care in these countries.

Numerous randomized and non-randomized controlled trials have been conducted on the topic and several systematic reviews (not meta-analyses) discussed the findings. A

31

review conducted by Garg et al. (2005) aimed to review randomized and non-randomized controlled trials that evaluated the impacts of computerized CDSSs. Their data were based on searching the MEDLINE, EMBASE, Cochrane Library, Inspec, and ISI databases as well as checking reference lists through September 2004. For inclusion criteria, the authors included all randomized and non-randomized controlled trials that assessed the effect of a CDSS compared with care given without a CDSS on practitioner performance or patient outcomes. The researchers found one hundred studies that met their inclusion criteria. They found that CDSS improved practitioner performance in 62 of the 97 studies evaluating practitioner performance. Regarding patient outcomes, they found that 52 studies evaluated one or more patient outcomes with only seven trials stating that patient outcomes were improved. The researchers concluded that many types of CDSSs improved practitioner performance. However, they reported that the effect on patient outcomes was inconsistent. Finding systematic reviews and meta-analysis on this topic is extremely beneficial since high-quality evidence and extensive references to primary studies relevant to the research topic are provided. However, there is a very limited number of relevant systematic reviews and meta-analysis studies.

A systematic review and meta-analysis conducted by Jeffery et al. (2013) that reviewed randomized trials evaluated the effects of computerized CDSSs in ambulatory diabetes management compared with a non-computerized clinical decision support system control. They used a comprehensive computerized CDSS overview conducted in January 2010. They used EMBASE, MEDLINE, INSPEC/COMPENDEX and Evidence-Based Medicine Reviews (EBMR) from January 2010 to April 2012 in their search. In addition to the previous sources, reference lists of related reviews included articles and

Clinicaltrials.gov were also used in their search. Based on the researchers' inclusion criteria, randomized controlled trial studies of diabetes in ambulatory care settings that also compare a computerized CDSS intervention with a non-computerized CDSS control and measuring either a process of care or a patient outcomes were included in the review. Two reviewers were independently responsible for screening of studies, data extraction, and evaluating risk of bias and quality of evidence assessments. The systematic review included 15 trials. Hemoglobin A1C (HbA1c) and quality of life and hospitalization were all not statistically significant, but all favored the computerized CDSSs over the control. Computerized CDSSs were also superior compared to control in terms of triglycerides and practitioner performance. Even though outcomes seem to be leaning toward support of computerized CDSS interventions, the effects were small, and the quality of the evidence was low. Additionally, no improvements were detected in important patient outcomes. The researchers concluded that a marginal improvement in clinical outcomes might be gained by utilizing computerized CDSSs in diabetes management. However, they claimed that because of the risk of bias, inconsistency, and imprecision, confidence in the evidence is low.

Another review of randomized controlled trials of medical record powered CDSSs to improve the quality of diabetes care was performed by Ali et al. (2016). The goal was to evaluate the effectiveness of CDSSs in improving quality of type II diabetes care. Inconsistent and variable results for the quality of diabetes care measures were found in the review. The process of care for all three measures of quality (Glycated hemoglobin (HbA1c), low-density lipid cholesterol (LDL-C), and blood pressure (BP)) of diabetes care were significantly improved. However, weak to modest positive results were

observed for the clinical measures of the diabetes care indicators. In addition to this, the technology adoption of CDSS was found to be consistently low.

Some systematic reviews have shown that CDSSs have potentially improved diabetes care. However, it is not clear whether CDSSs are effective in improving diabetes management care since different methods of measuring and presenting outcomes were used with inconsistent conclusions. In order to address this issue, Jia et al. (2019) in their recent work conducted a comprehensive overview to evaluate the effects of CDSSs on diabetes care as well as examining methodological and reporting qualities. PubMed, EMBASE, and Cochrane Library were the primary search sources through February 2017. The researchers included systematic reviews that examined the effects of CDSS on diabetes care. The outcomes in the overview were defined and evaluated separately for the process of care and patient outcomes. Methodological quality was assessed by an instrument for critically evaluating systematic reviews of randomized controlled clinical trials ("AMSTAR") and reporting qualities were assessed by an evidence-based minimum set of items for reporting in systematic reviews and meta-analyses ("PRISMA"). The total number of systematic reviews included was 17 studies. These studies had 222 unique randomized controlled trials and 102 non-randomized controlled trials. In 32 of 102 unique studies, CDSSs were found to be significantly effective in improving patient outcomes. The process of care was found in 117 out of 143 unique studies to be substantially affected by CDSS. Overall scores of AMSTAR resulted in a mean score of 6.5 where AMSTAR ranged from 8 to 11 is considered high quality, 4 to 7 is of medium quality, and 0 to 3 is low quality (Sharif et al., 2013). As a conclusion of this comprehensive overview, the researchers claimed that CDSSs improved the quality

of diabetes care by improving the process of care or patient outcomes. There was also evidence that CDSSs that provide alerts, reminders, or feedback to participants were most likely to influence diabetes care. However, poor reporting of methodological domains and qualitative or narrative methods to combine findings was reported.

The effect of the CDSSs has been evaluated in numerous studies and the results are not consistent. Additionally, several systematic reviews have studied the effect of the CDSSs. However, a statistical procedure to combine the numerical data from multiple separate studies is needed. Therefore, conducting a MVMA will help to get a clearer picture and more precise and powerful results.

Purpose of the Study

MVMA is becoming more commonly used, especially in clinical research where there is no single, "gold standard" outcome measure. In sum, it has been suggested that MVMA obtains estimates for all effects under study simultaneously, defines the relationship between the effects, and provide estimates with better statistical properties (i.e., lower standard errors) than univariate meta-analysis. Even though MVMA can be useful and provide better statistical estimates, these benefits could depend on making additional assumptions regarding both what is reported in source studies and also about the nature of the data. MVMA requires within and between study correlation estimates which are not always available. Despite the additional complications and issues the MVMA brings, it can make a real contribution to the field of meta-analysis (Jackson et al., 2011)

Reviewing the literature leads to the gap previously discussed. **Although some authors have conducted comparative studies, the question of why the results**

produced by MVMA with the consideration of the correlation and univariate meta-analysis are almost identical is insufficiently explored. In order to address this issue and after reviewing the formulas of UVMA and MVMA to get a better understanding of how both methods are constructed, a simulation study was utilized to see how and when estimates produced by MVMA would be different and more statistically precise compared to estimates produced by separate UVMAs. In this simulation, three factors were allowed to vary. Sample size had three levels, the percentage of missingness had three levels, and finally the between-study correlation had two levels. The effect sizes were generated using standardized mean differences to simplify the comparison procedure. In addition, no one to the best of the researcher's knowledge has applied MVMA to study the effect of CDSSs. Therefore, to address these gaps in the literature, the researcher applied MVMA and UVMA to examine the effects of CDSSs on the quality of diabetes care management and explore reasons behind the similarity and differences in the effect estimates of the two methods.

Definition of terms

Coding

Coding is a procedure that extracting information necessary to perform a meta-analysis

from the primary studies (Card, 2015).

Effect size (EF)

EF is a standardized scale-free estimate of the relationship between an exposure and an outcome. Any difference in the outcome between the study groups such as mean

differences, relative risk, odds ratio, and risk difference could be defined as an effect sizes (Delgado-Rodríguez, 2001).

Exclusion criteria:

Exclusion criteria is a criterion that is used by the researcher in order to specify which studies should be excluded from a meta-analysis**.**

Fixed-effects model (RE)

In FEM, effects are assumed to be homogeneous across the studies in which effect sizes have a common true value for all studies (Delgado-Rodríguez, 2001).

Funnel plot

A funnel plot is a graphical method to display any possibility of publication bias. It simply displays the relation between the effect size of the study and its size. When publication bias is not exists, the funnel shape should be symmetric (Delgado-Rodríguez, 2001).

Heterogeneity

Heterogeneity means that there is between study variation. When there is heterogeneity, it means that there could be more than one true effect sizes in the combined studies (Delgado-Rodríguez, 2001).

Inclusion criteria

Inclusion criteria is a criterion that is used by the researcher in order to specify which studies should be included in a meta-analysis.

Gray literature

Gray literature is rarely included in meta-analyses. These kinds of literature are not controlled by commercial publishers. Gray literature typically has a limited dissemination and are difficult to be retrieved (Card, 2015).

Meta-analysis (MA)

Meta-analysis can be defined as a statistical method to calculate an overall effect of single, independent studies by systematically synthesize their findings (Shorten & Shorten, 2013).

Univariate meta-analysis (UVMA)

In UVMA treatment effects for multiple outcomes are meta-analyzed separately ignoring the possible correlation between the effects (Trikalinos et al., 2014).

Multivariate meta-analysis (MVMA)

MVMA is an extension of the standard UVMA. In MVMA, effect estimates are jointly synthesized which allows for the accounting of within-study and between-study correlations of the outcomes (Jackson et al., 2011).

Publication bias

Publication bias could happen when the published studies carried out on a specific topic do not represent all the relative studies (Delgado-Rodríguez, 2001).

Random-effects model (REM)

Unlike FEMs, the REM does not assume that the effects across studies being pooled are homogeneous. That means that each sample of studies has it is own a true effect size (Delgado-Rodríguez, 2001).

Chapter Two

Method

In this chapter, I first describe how data were simulated to examine the effects of sample size, amount of missing data, and between-study correlation on the estimates of UVMA and MVMA when they are allowed to vary. I then describe the process used in terms of steps taken to perform the meta-analysis.

Data Simulation

Before I empirically compared UVMA and MVMA using real data, I conducted a simulation which allowed me to control some variables and hold others constant in an attempt to understand when the utilization of MVMA (where two or more outcomes are simultaneously analyzed), gives more precise estimates. For data simulation, a Monte Carlo simulation technique was applied. The Monte Carlo approach is a very common technique to test theoretical hypotheses by generating datasets that meet specified conditions (Paxton et al., 2001). In this simulation, two outcomes were considered for simplicity. Outcomes one and two were set to have a true effect size of 0.7 and 0.5, respectively. The sample estimates for the studies were generated from the normal distribution using the true effect sizes while sample within study standard deviations were generated from a gamma distribution. Variables that were manipulated were sample size, between-study correlation, and the percentage of missing data. For sample size, and

according to Jackson and Turner (2017), at least five studies are needed to consistently gain power from random-effects meta-analysis. Therefore, sample size had three levels--small = 5, medium = 20, large = 50 studies. The between-study correlation had two levels (weak = 0.1 and strong = 0.9). Finally, the degree of missing data had three levels (0%, 30%, and 70%). Missing data were generated to be MCAR.

Table 1 below shows a summary of the simulation conditions across all manipulated factors, with 3x2x3=18 conditions.

Table 1
Summary of Conditions across the Varying Factors

Sample size	Between-study correlation	Degree of missingness
Small (5)	Weak (0.1) Strong (0.9)	0%
	Weak (0.1) Strong (0.9)	30%
	Weak (0.1) Strong (0.9)	70%
Medium (20)	Weak (0.1) Strong (0.9)	0%
	Weak (0.1) Strong (0.9)	30%
	Weak (0.1) Strong (0.9)	70%
Large (50)	Weak (0.1) Strong (0.9)	0%
	Weak (0.1) Strong (0.9)	30%
	Weak (0.1) Strong (0.9)	70%

In each scenario, the generated data were used to conduct UVMA and MVMA. The result in each scenario then was compared in terms of the statistical properties (i.e., how different are the outcome estimates? are there any standard error and confidence interval improvements? is there any change in conclusions?). To analyze results of this simulation study, a 3x2x3 three-way analysis of variance (ANOVA) was used in order to examine main effects and interactions among the three factors for each of the two dependent variables, ES = .7 and ES = .5. Subsequently, a further 2x3x3 (method x missingness x sample size) MANOVA was conducted to compare results of methods UVMA and MVMA directly. A MANOVA was used to simultaneously analyze ES = 7 and ES = .5 as the two dependent variables. In this analysis, the between-study correlation factor was eliminated as it did not apply to the UVMA and further, it showed little effect in the ANOVAs.

The simulations for all scenarios were conducted using the R statistical software program (R Core Team (2020). For each scenario 1,000 simulation runs (replications) were carried out in order to reach sufficient and stable estimates (Belias et al., 2019; Carter et al., 2019; Huang et al., 2015; Mittlböck & Heinzl, 2006). Partial eta squared (η_p^2) rules of thumb for small =.01, medium =.06, and large =.14 were used to identify interpretable effects of the controlled factors rather than statistical significance due to the large sample size. The missing data were only applied on the second ES = 0.5 in order to see if the presence of missing data in one outcome could affect the other.

Meta-analysis Procedure

Literature Search

In this study, multiple sources were used in the searching procedure, including The Cochrane Controlled Trials Register (CCTR), PubMed which uses MEDLINE as a primary database, checking of reference lists, and hand-searching of key journals. In order to minimize publication bias, references in published studies, computerized databases searching of unpublished material, conference proceedings, and graduate dissertations were considered in the searching procedure (Ab, 2010).

Study Selection

In meta-analysis, in order to produce reliable results, randomized controlled trials (RCTs) are usually used since they provide the highest level of evidence with least bias (Ahn & Kang, 2018). Therefore, only RCTs of the effects of quality of diabetes care with the use of clinical decision support systems were considered for inclusion. The keywords that were used for the search include electronic medical record; electronic health record; computerized clinical decision support system; quality of diabetes care; diabetes patient outcomes; health information technology; cholesterol management; hyperlipidemia; low-density lipoprotein cholesterol. Boolean statements were utilized to either expand or reduce the search recall and return a precise result. Published and unpublished RCTs between 2010 and the present were included.

Table 2 below displays the search string that the researcher developed in order to conduct the initial search.

Table 2

Search Strings and Databases Used

DATABASE	SEARCH STRING	FILTER
PUBMED (MEDLINE) The Cochrane Controlled Trials Register (CCTR)	(Health Information Technology OR HIT OR Computerized decision support OR Electronic health record) AND (diabetes care management OR diabetes management) Filters: Randomized Controlled Trial (clinical decision support) AND (diabetes care management OR diabetes management) Filters: Randomized Controlled Trial	Date published: 2010-present

Inclusion and Exclusion Criteria

There were two screening steps once the initial search of the prospective studies

was done. The purpose of these two screening steps is to enhance the efficiency of the

selection and avoid the risk of leaving out any pertinent studies. In the first step, the

abstracts of the articles were examined for their relevance. For instance, abstracts

containing the keywords "*clinical-decision-support systems*" and "*diabetes care*

management" were retained for the next step. Abstracts that did not contain such words

were still acceptable if they contained any other relevant indirect keywords.

The articles that passed step one then went through the second screening step. In

this step, articles were fully screened. Articles qualified to be included in the meta-

analysis if they passed the exclusion criteria. The first criterion is that articles would be

excluded if they do not have empirical data, such as qualitative studies which do not hold

any experimental or empirical data. Secondly, articles would be excluded if they do not

provide enough statistics and data to estimate the required effect size and its associated

variance. The third exclusion criterion is that studies would be excluded if they do not

state at least one of the following indicators: LDL-C, HbA1c, or BP. Fourth, the study

43

also would be excluded from the analysis if it does not mention the use of at least one

CDSS and does not include patients with diabetes mellitus.

Regarding grey literature and unpublished studies, ProQuest Dissertations was the

primary database used for searching the grey literature or unpublished studies. The search

string that was used in order to retrieve studies from ProQuest Dissertations was:

ab(clinical decision support systems) AND ab(diabetes)

In addition to ProQuest Dissertations, the National Institute of Health (NIH,

http://www.nimh.nih.gov/index) and OpenGrey (http://www.opengrey.eu/) were used in the

search for grey literature and unpublished studies. Figure 1 shows the flow diagram of the

search strategy and selection of articles.

Figure 1

Flow Diagram of Search Strategy and Selection of Articles.

```
┌─────────────────────────────────┐          ┌─────────────────────────────────┐
│ Studies Identified through       │          │ Additional studies identified    │
│ database search (n = 257)        │─────────→│ through other sources (n = 65)   │
└─────────────────────────────────┘          └─────────────────────────────────┘
                │                                              │
                ▼                                              │
┌─────────────────────────────────┐          ┌─────────────────────────────────┐
│ Studies screened                 │          │ Studies excluded                 │
│ (n = 322)                        │─────────→│ (n = 175)                        │
└─────────────────────────────────┘          └─────────────────────────────────┘
                │
                ▼
┌─────────────────────────────────┐          ┌─────────────────────────────────┐
│ Full-text articles assessed for  │          │ Full-text articles excluded      │
│ eligibility (n = 147)            │─────────→│ (n = 106)                        │
└─────────────────────────────────┘          └─────────────────────────────────┘
                │
                ▼
┌─────────────────────────────────┐
│ Studies included in meta-analyses│
│ (n = 41)                         │
└─────────────────────────────────┘
                │
    ┌───────────┼───────────┬───────────┐
    ▼           ▼           ▼           ▼
┌────────┐ ┌──────────┐ ┌──────────┐ ┌────────┐
│ MVMA   │ │ UVMA of  │ │ UVMA of  │ │ UVMA of│
│        │ │ LDL      │ │ HbA1c    │ │ PP     │
│(n = 41)│ │ (n = 16) │ │ (n =38)  │ │(n = 20)│
└────────┘ └──────────┘ └──────────┘ └────────┘
```

45
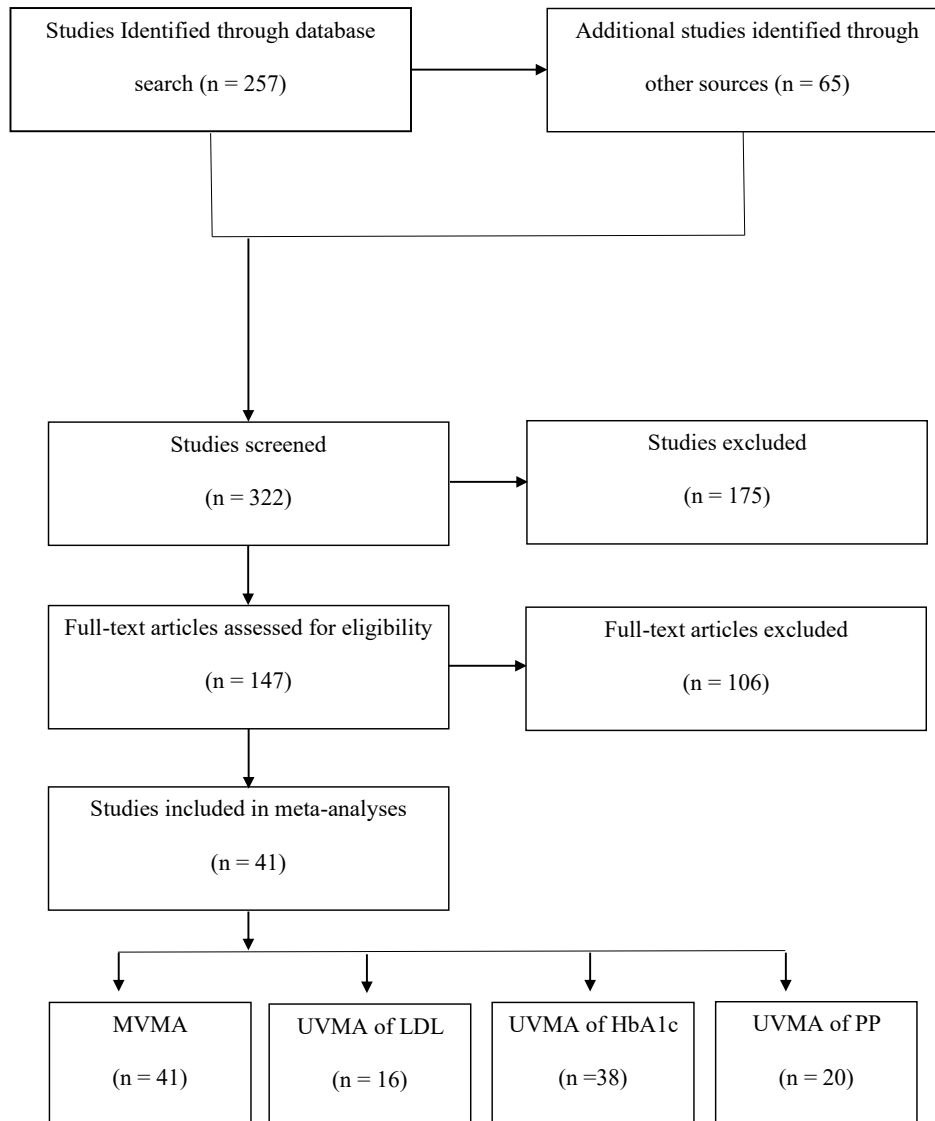
Publication bias

To the best of my knowledge, no literature was identified that addresses a specific method on how publication bias might be captured when a MVMA is used. Therefore, to deal with any possible biases, this study considered the unpublished studies found by using ProQuest Dissertations, HIB, and OpenGrey as search sources. Additionally, some publication bias evaluations, such as funnel plots, Orwin's fail-safe N, and p-curve analysis, were used to evaluate publication bias (Card, 2015)

Sources of bias.

Typically, it is more likely that studies with statistically significant findings are published compared to studies that report non-significant results. Those published studies are found to have a larger effect size. The possibility of producing bias in the significance of the effect sizes might be very large, especially if studies have relatively small sample sizes. Therefore, published studies are more likely to be included in a meta-analysis (Dickersin, 2005). If a researcher conducting a systematic review was able to find studies that in the grey literature, then publication bias would not be a problem for meta-analysis. However, this is not usually the case, for instance, looking in the first 1000 Cochrane systematic reviews, it has been found that almost half of them contained no data from grey or unpublished sources (Mallet et al., 2002).

Publication status is not the only source of bias, there are other factors that can lead to a bias in effect size. First, language bias, where English-language databases and journals are more likely to be searched. Secondly, availability bias, where the selective inclusion of studies is more likely for those studies easily accessible to the researcher. Third, cost bias where the selective inclusion of studies is more likely designed to include

available free or low-cost studies. Similarity bias means the selective inclusion of studies

based on one's own opinion. Another source of bias is duplication bias where studies with

statistically significant results are more likely to be published in more than one journal.

Finally, citation bias can exist where studies with statistically significant results are easier

to identify since they are more likely to be cited in other studies (Egger et al., 1997;

Gøtzsche & Johansen, 1997; Jüni et al., 2002; Ravnskov, 1992; Tramèr et al., 1997).

Bias diagnostics methods

Overestimation of the actual effect size is a problem when the studies in a meta-

analysis are based on a biased sample of studies, and that should be accommodated.

Several methods have been designed to assess the potential impact of bias on a given

meta-analysis. Each one of these methods aims to answer specific inquires. The first

method is the funnel plot which detects evidence of bias (Light & Pillemer, 1984). The

second method is Orwin's Fail-safe N. This method checks if the entire effect is an

artifact of bias (Orwin, 1983). Finally, Duval and Tweedie's Trim and Fill method that

allows knowing the impact that the bias has (Duval & Tweedie, 2000). These three

methods are explained in more detail below.

Funnel plot

Funnel plots have been commonly used to check bias in meta-analyses. A funnel

plot is a scatter plot of the effect estimates from individual studies versus some measure

of each study's size, which is usually the standard error of the effect estimate. If no bias is

present and there is between-study heterogeneity, the plot will follow a symmetric

inverted funnel. However, when the publication bias is present, the plot is symmetric at

the top, with a few missing studies in the middle, and more missing studies close to the bottom. (Sterne et al., 2011).

Orwin's Fail-safe N

This method enables the researcher to discover the number of missing studies that would lead the overall effect to a non-zero level. The researcher selects a value that would represent the smallest effect assumed to be of substantive utility and determine how many missing studies it would take to make the summary effect fall below this point. Additionally, Orwin's method allows the researcher to determine the mean effect in the missing studies as a non-zero value (Becker, 2005; Begg & Mazumdar, 1994).

Duval and Tweedie's Trim and Fill

Duval and Tweedie's Trim and Fill approach allows estimation of the unbiased effect. The idea of the approach is that it uses an iterative method to eliminate the most extreme small studies from the funnel plot and, more specifically, the positive side and computes the effect size at each iteration until symmetry is gained. As a result, theoretically, this will produce an unbiased estimate of the effect size. It also shrinks the variance of the effects, generating a narrow confidence interval. The original studies then will be added back into the analysis and a mirror image for each will be imputed (Duval & Tweedie, 2000).

In meta-analysis, it is essential to include an evaluation of publication bias in order for the result to be robust and to guarantee the integrity of the individual meta-analysis. Ignoring the potential for bias might lead to the conclusion that the current meta-analyses cannot be trusted.

Variables and Coding

As mentioned in the introduction, three indicators (outcomes) were used in UNMAs and MVMAs as a guide in assessing the quality of diabetes care. Those indicators were low-density lipid cholesterol (LDL-C), glycated hemoglobin (HbA1c), and blood pressure (BP). Reduction in these indicators reduces the risk of diabetes complications and death (Hu et al., 2016b). In the following paragraphs, I briefly summarize them and their recommended levels.

Low-density lipid cholesterol (LDL-C*)*

LDL-C, also called the bad cholesterol, is most of the body's cholesterol. The risk of heart disease and stroke increases as LDL cholesterol increases (CDC, 2020). A goal of <100 mg/dl (2.60 mmol/l) for LDL cholesterol is recommended for patients with diabetes without preexisting CVD according to The American Diabetes Association (ADA).

Glycated hemoglobin (HbA1c)

Glycated Hemoglobin (HbA1c) is used routinely to evaluate glycemic control in diabetics to achieve treatment goals and limit long term complications. The ADA has recently recommended the use of glycated hemoglobin (HbA1c) as an indicator to diagnose diabetes mellitus (Tay et al., 2011). HbA1c levels ranged from 5.7% to 6.4% indicates a higher risk of getting diabetes. Levels of 6.5% or higher mean you have diabetes (Jagannathan et al., 2016).

Blood pressure (BP)

Blood pressure (BP) is the pressure of circulating blood on the walls of blood vessels (Sa et al., 2014). It is recommended that a blood pressure goal is less than 130/85 mm Hg in patients with hypertension and diabetes mellitus (Bakris, 2001).

49

Coding process

A typical code sheet and codebook was developed for the MVMA and UVMA. The code sheets in this meta-analysis contained three primary categories of variables: (a) methodological and substantive features; (b) study quality; and (c) outcome data. Methodological and substantive features are significant variables to code in every meta-analysis. Information such as year of publication, type of research design, and inclusion criteria is important to describe the literature and hence can relate these characteristics to study findings (Sa et al., 2014). The strategy that was used in this study in order to develop and adopt the coding sheet was a review of a random subset of studies to be synthesized and adopting all related coding variables during the review. After including the adopted variables, the coding sheet was pilot tested on a different subset of studies. As soon as the development of the code sheet was completed, a codebook was developed to lead the coding process (Brown et al., 2003).

Evaluation coding decisions

According to Card (2015), there are two essential qualities of the coding system that are linked to aspects of transparency and replicability. Additionally, it is significant to account for the reliability of the coding. In any meta-analysis study, enough details of the coding process should be presented in order for the audience to know how the coding decisions were made (transparency). Replicability is the ability approaching the same coding decisions as the researchers did if an audience member were to use the coding strategy the researcher developed to the studies included in the current meta-analysis (Card, 2015).

Coding reliability

Card (2015), has extensively explained the mechanism of evaluation of the reliability of the coding. To empirically assess the replicability of the coding, the reliability of independent coding procedure of the same studies was evaluated. There are two ways to evaluate the reliability, either by using intercoder reliability or intracoder reliability. Intercoder reliability is when reliability is evaluated between two independent coders. Intercoder reliability is evaluated by having two independent coders assigned to code a subset of overlapping studies. The number of the studies that coder should independently code should be large to ensure an adequate estimate of reliability. A sample of 20 to 50 studies is recommended and the researcher's decision to choose a sample size within this range should depend on the researcher understating of the coding interface level. For example, a lower interface level suggests that a lower number of studies is needed in order to confirm intercoder agreement and vice versa. In contrast, intracoder reliability is when the evaluation occurs within the same coder. This approach is not ideal in assessing the reliability of the coding system as Card (2015) noted that " intracoder agreement is not a perfect substitute for intercoder agreement because one coder might hold potential biases or consistently make the same coding errors during both coding sessions" (p.75).

Therefore, in this study, intercoder agreement was adopted as a reliability measure of the coding and one Ph.D. student and the researcher coded a subset of overlapping studies. The reliability was quantified using the Agreement Rate (AR). Even though AR does not account for base rates of coding, according to Card (2015), it is the simplest and

51

commonly used index of coding reliability. The intercoder agreement of this study was 88% which is considered very good agreement (Hartmann, 1977; Stemler, 2004).

Coding study characteristics

At least four characteristics should be coded as Card (2015) recommended. These four study characteristics are: characteristics of the sample, measurement, design, and source characteristics.

Regarding sample characteristics, aspects of sampling procedure and demographic features of the sample were coded. Some characteristics that might be coded are the setting, sampling technique, gender, socioeconomic status, etc. It is not necessary to code all sample characteristics. The researcher only coded relevant characteristics.

For measurement characteristics, according to Card (2015), knowing the strengths and the weaknesses of the measurement processes would be extremely helpful in guiding descriptions about what measurement characteristics should be coded. For instance, some potential variables that might be coded are the source of information and specific features of the measurement process.

Study-design characteristics is another set of characteristics that Card (2015) recommended to code. Since only RCT studies were included, the type of RCT used and aspects of the control groups were coded to have an idea about the design features. Finally, coding whether the study is published or not is essential to evaluate evidence of publication bias. As Card mentioned, coding the year of publication might be useful, especially in the evaluation of year as a moderator to clarify any historic trends in the effect size across time. More characteristics might arise when the actual coding procedure

begins. Since studies that were included in this research are RCTs, the outcomes were more likely to be a continuous variable. When there are treatment and control groups, the mean difference is typically used to quantify the treatment effect (Cheung, 2013).

<div align="center">**Chapter Three**</div>

<div align="center">**Results**</div>

Results from the data analysis are reported in this Chapter. First, the results of the simulation are presented. After presenting the simulation results, the empirical study comparison findings are presented. Finally, the evaluation of the effect of the CDSSs is summarized.

Results of the simulation

<div align="center">**Two-way ANOVAs of the UVMA Simulation**</div>

A two-way ANOVA was conducted to determine the effects of sample size and percent of missing data on effect size coefficient estimation and the corresponding standard errors. Due to the large sample size used in the simulations, partial eta squared values were used to evaluate the importance of main effects and interactions instead of statistical significance. Partial eta squared (η_p^2) rules of thumb for small =.01, medium =.06, and large =.14 were used to interpret the main effects and interactions (Kittler et al., 2007) with values of ≥.01 used to define interpretable effects.

Effect Size

Effect Size (.7). The results of the two-way ANOVA suggested that no effects of sample size and percent of missing data that were at least small in magnitude were detected on the effect size estimate of ES = .7 coefficient estimates ($p = .41$, $\eta_p^2 = .000$), ($p =.26$, η_p^2

= .000), respectively (Tables 3 and 4). Moreover, the interaction between sample size and percent of missing data was also negligible (p = .31, η_p^2 = .001).

Table 3
Anova Summary Table for ES = .7

	Sum of Squares	df	Mean Square	F	p	Partial Eta Squared (η_p^2)
SZ	.02	2	.01	.90	.407	.000
Miss	.02	2	.01	1.29	.275	.000
SZ * Miss	.04	4	.01	1.19	.313	.001
Error	79.33	8991	.01			
Total	79.41	8999				

Note: SZ = Sample size (5,20,50); Miss = The percent of missing data (0%,30%,70%).

Table 4
Effect Size Estimate Means of ES = .7 by Levels of Sample Size

Sample Size	Mean
Small = 5	.699
Medium = 20	.697
Large = 50	.700
Total	.699

Effect Size (.5). Regarding the effect size estimate of ES = .5, the two-way ANOVA showed that sample size and percent of missing data had no significant effects (p = .30, η_p^2 = .000), (p = .22, η_p^2 = .000), respectively (Tables 5 and 6). Additionally, the interaction term between sample size and percent of missing data was also nonsignificant (p = .42, η_p^2 = .000).

Table 5
Anova Summary Table for ES = .5

Source	Sum of Squares	df	Mean Square	F	p	Partial Eta Squared (η_p^2)
SZ	.02	2	.01	1.21	.300	.000
Miss	.03	2	.01	1.51	.222	.000
SZ * Miss	.04	4	.01	.97	.421	.000
Error	79.91	8991	.01			
Total	80.00	8999				

Table 6
Effect Size Estimate Means of ES = .5 by Level of Sample Size

Sample Size	Mean
Small = 5	.497
Medium = 20	.500
Large = 50	.500
Total	.499

Standard Errors

Effect Size = .7. Regarding the first outcome standard error estimates, the main effects of

sample size and percent of missing data were both statistically significant and had a

meaningful partial eta squared value ($p < .001$, $\eta_p^2 = .351$) and ($p < .001$, $\eta_p^2 = .082$),

respectively, for the standard error of the coefficients for ES = .7 (Table 7). The two-way

interaction between sample size and percent of missing was also significant ($p < .001$, η_p^2

= .022) meaning that different levels of sample size and percent of missing data had at

least a small effect on the standard errors of the UVMA coefficient estimates (Figures 2

and 3). A simple main effects analysis was conducted since interpreting the main effect in

the presence of a significant interaction might be misleading.

Table 7

Anova Summary Table for Standard Error Estimates for ES = .7

Source	Sum of Squares	df	Mean Square	F	p	Partial Eta Squared (η_p^2)
SZ	4.95	2	2.47	2434.70	<.001	.351
Miss	.82	2	.41	402.30	<.001	.082
SZ * Miss	.21	4	.05	51.24	<.001	.022
Error	9.13	8991		<.001		
Total	15.10	8999				

Note: SZ = Sample size (5,20,50); Miss = The percent of missing data (0.0%,30%,70%).

Figure 2

Mean Standard Error of Level of Missingness by Sample Size
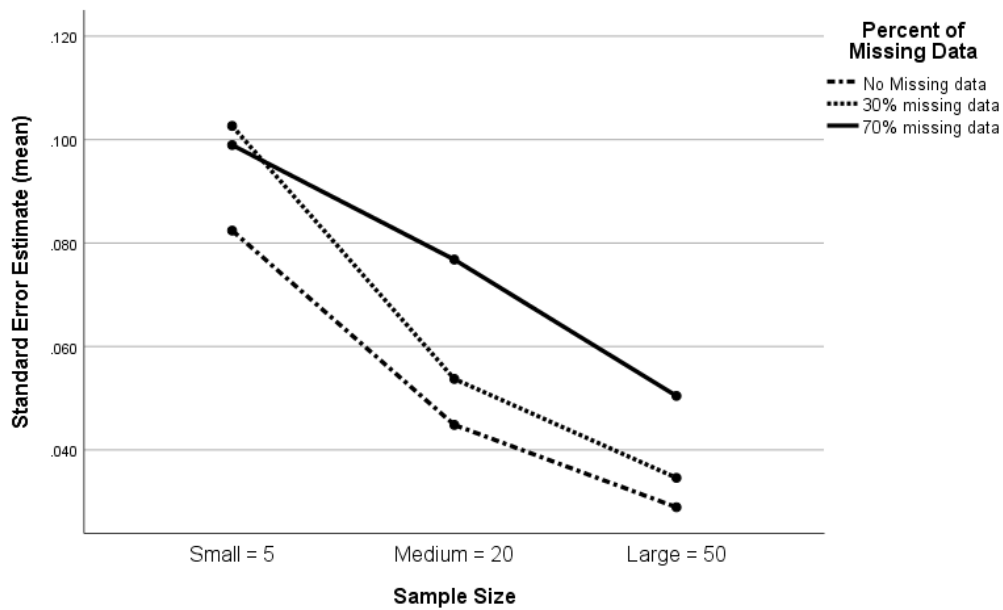
Figure 3

Mean Standard Error of Sample Size by Percent of Missing Data

Simple main effects for percent of missing data by sample size

Effect size = .7. There was a small effect of the main effect of percent of missing data on standard error for ES = .7 when the sample size was small or large (p < .001, partial η_p^2 = .025), (p < .001, partial η_p^2 = .027) (Table 8). However, the main effect of percent of missing data had a moderate effect when the sample size was medium. Since this particular simple main effect was statistically significant, the difference in mean standard errors between percent of missing data for the three levels of sample size was considered next. Table 9 below presents the means by cell and Table 10 presents the results of the pairwise comparisons.

Table 8

Simple Main Effects of Percent Missing by Sample Size for ES = .7

Sample Size		Sum of Squares	df	Mean Square	F	p	Partial Eta Squared (η_p^2)
Small = 5	Contrast	.232	2	.12	114.29	<.001	.025
	Error	9.130	8991	.00			
Medium = 20	Contrast	.544	2	.27	268.03	<.001	.056
	Error	9.130	8991	.00			
Large = 50	Contrast	.249	2	.12	122.46	<.001	.027
	Error	9.130	8991	.00			

In the situation when 30% and 70% of data were missing and the sample size was small, the mean standard error was .020 and 0.017, higher when 30% and 70% of data were missing than when there were no missing data. In addition, when the sample size was small, the mean standard error was .004, higher for 30% of data missing than for 70% of missing data.

In the case of medium sample size, when there were no missing data, the mean standard error was .009, higher when 30% of data were missing than when there were no missing data. The mean standard error was .032 higher when 70% of data were missing than when there were no missing data. Additionally, when the sample size was medium, the mean standard error was .023, higher when 70% of data were missing than for 30% of missing data. For large sample sizes and when there were no missing data, the mean standard error was .006, higher when 30% of data were missing than when there were no missing data. When there was 70% of missingness, the mean standard error was .022, higher when 70% of data were missing than for 30% of missingness. The mean standard error was .016, higher when 70% of data were missing than when 30% of data were missing.

Table 9

Descriptive statistics for sample size as a function of percent of missing data

Sample Size	Percent of Missing Data	Mean	SD	N
Small = 5	No Missing data	.082	.034	1000
	30% missing data	.103	.056	1000
	70% missing data	.099	.063	1000
	Total	.095	.053	3000
Medium = 20	No Missing data	.045	.007	1000
	30% missing data	.054	.011	1000
	70% missing data	.077	.026	1000
	Total	.058	.022	3000
Large = 50	No Missing data	.029	.003	1000
	30% missing data	.035	.004	1000
	70% missing data	.050	.006	1000
	Total	.038	.010	3000
Total	No Missing data	.052	.030	3000
	30% missing data	.064	.044	3000
	70% missing data	.075	.044	3000
	Total	.064	.041	9000

Note: N represents the sample size of the particular level of missingness

Table 10

Pairwise Comparison Results

Sample Size	(I) Percent of Missing Data	(J) Percent of Missing Data	Mean Difference (I-J)	SD	*p*
Small = 5	No Missing data	30% missing data	-.020	.001	<.001
		70% missing data	-.017	.001	<.001
	30% missing data	70% missing data	.004	.001	.028
Medium = 20	No Missing data	30% missing data	-.009	.001	<.001
		70% missing data	-.032	.001	<.001
	30% missing data	70% missing data	-.023	.001	<.001
Large = 50	No Missing data	30% missing data	-.006	.001	<.001
		70% missing data	-.022	.001	<.001
	30% missing data	70% missing data	-.016	.001	<.001

Effect size = .5. For the second outcome, the main effects of sample size on standard error were also significant and had a large partial eta squared value ($p < .001$, $\eta_p^2 = .305$). The percent of missing data was significant and had a small to moderate effect ($p < .001$, $\eta_p^2 = .046$) on the standard error (Table 11). The two-way interaction between sample size and percent of missing data was also significant and had a larger effect on the standard error estimate when compared to the first outcome ($p < .001$, $\eta_p^2 = .078$), meaning that different levels of sample size and percent of missing data had at least a small effect on the value of the UVMA standard error estimates for outcome two. Since the interaction of sample size and percent of missing data had an interpretable effect size, simple main effects analysis was conducted (Figures 4 and 5).

Table 11
Anova Summary Table for Standard Error Estimates for ES = .5

Source	Sum of Squares	df	Mean Square	F	p	Partial Eta Squared (η_p^2)
SZ	3.48	2	1.74	1972.10	<.001	.305
Miss	.38	2	.19	216.97	<.001	.046
SZ * Miss	.67	4	.17	190.87	<.001	.078
Error	7.94	8991	.00			
Total	12.48	8999				

Note. SZ = Sample size (5,20,50); Miss = The percent of missing data (0.0%,30%,70%).

Figure 4
Mean standard error of Level of Missingness by Sample Size
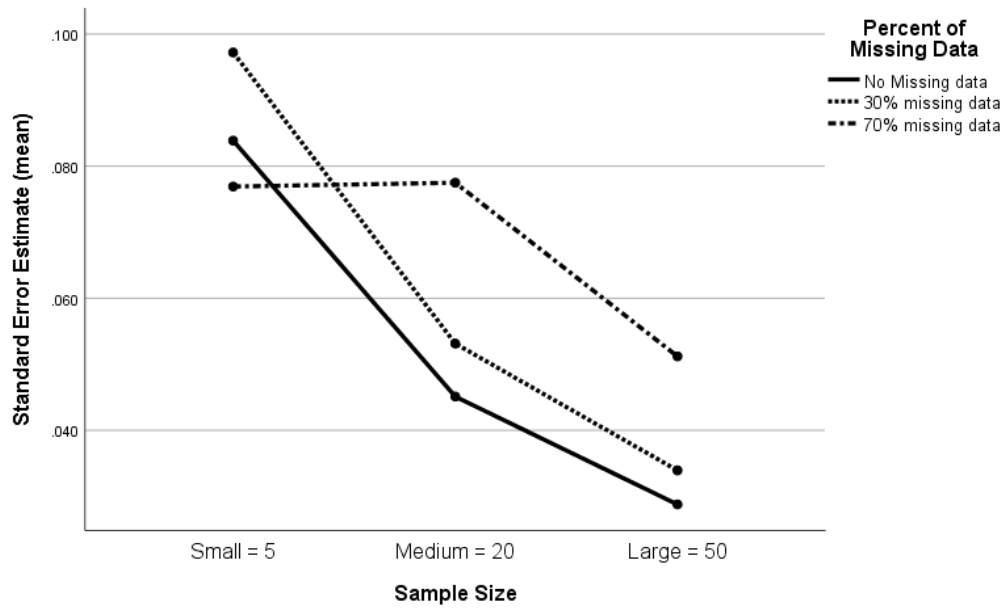
Figure 7
Mean standard error of Sample Size by Level of Missingness

Simple main effects for percent of missing data by sample size

The simple main effect of percent of missing data in ES = .5 standard error was small when the sample size was either small or large (p < .001, partial η_p^2 = .026; p < .001, partial η_p^2 = .034). However, a moderate effect of the main effect of percent of missing data was detected in coefficient two standard error when the sample size was medium (Table 12).

Table 12
Simple Main Effects of Percent Missing by Sample Size for ES = .5

Sample Size		Sum of Squares	df	Mean Square	F	p	Partial Eta Squared (η_p^2)
Small = 5	Contrast	.21	2	.11	120.46	<.001	.026
	Error	7.94	8991	.00			
Medium = 20	Contrast	.57	2	.28	322.05	<.001	.067
	Error	7.94	8991	.00			
Large = 50	Contrast	.28	2	.14	156.20	<.001	.034
	Error	7.94	8991	.00			

Pairwise comparisons were carried out since the simple main effect of percent of missing data was statistically significant. The pairwise comparisons table is presented in (Table 13). From Table 13, when there were no missing data, the mean standard error was .013, higher when 30% of data were missing than when there were no missing data. In the situation when 70% of data were missing, the mean standard error was .007, higher when no missing data were present than when there were 70% of data missing. In addition, when the sample size was small, the mean standard error was .020, higher for 30% of data missing than for 70% of missingness.

Table 13
Pairwise Comparisons Results

Sample Size	(I) Percent of Missing Data	(J) Percent of Missing Data	Mean Difference (I-J)	SD	p
Small = 5	No Missing data	30% missing data	-.013[*]	.001	<.001
		70% missing data	.007[*]	.001	<.001
	30% missing data	70% missing data	.020[*]	.001	<.001
Medium = 20	No Missing data	30% missing data	-.008[*]	.001	<.001
		70% missing data	-.032[*]	.001	<.001
	30% missing data	70% missing data	-.024[*]	.001	<.001
Large = 50	No Missing data	30% missing data	-.005[*]	.001	<.001
		70% missing data	-.022[*]	.001	<.001
	30% missing data	70% missing data	-.017[*]	.001	<.001

In the case of medium sample size, when there were no missing data, the mean standard error was .008, higher when 30% of data were missing than when there were no missing data. The mean standard error was .032 higher when 70% of data were missing than when there were no missing data. Additionally, when the sample size was medium, the mean standard error was .024, higher when 70% of data were missing than for 30% of missingness.

For large sample sizes and when there was no missing data, the mean standard error was .005, higher when 30% of data were missing than when there were no missing data. When there was 70% missingness, the mean standard error was .022, higher when 70% of data were missing than for 30% missingness. The mean standard error was .017, higher when 70% of data were missing than when 30% were missing. Table of means is provided as Table 14.

Table 14

Descriptive Statistics for Percent of Missing Data by Sample Size

Sample Size	Percent of Missing Data	Mean	SD	N
Small = 5	No Missing data	.084	.034	1000
	30% missing data	.097	.054	1000
	70% missing data	.077	.053	1000
	Total	.086	.049	3000
Medium = 20	No Missing data	.045	.008	1000
	30% missing data	.053	.011	1000
	70% missing data	.077	.028	1000
	Total	.059	.022	3000
Large = 50	No Missing data	.029	.003	1000
	30% missing data	.034	.004	1000
	70% missing data	.051	.010	1000
	Total	.038	.012	3000
Total	No Missing data	.053	.031	3000
	30% missing data	.061	.041	3000
	70% missing data	.069	.037	3000
	Total	.061	.037	9000

Note: N is the sample size of the particular level of missingness

Three-way ANOVA of the MVMA Simulation

Effect Size

Effect Size = .7. As was shown previously in the two-way ANOVA of the UVMA, the

result of the three-way ANOVA of the MVMA also suggested that no interpretable

effects of sample size and percent of missing data were detected on the effect size

estimate of ES = .7 ($p = .015$, $\eta_p^2 = .000$), ($p = .302$, $\eta_p^2 = .000$), respectively (Table 15).

In addition to sample size and percent of missing data, between-study correlation was

added as a third independent variable in this analysis. However, the results of the

ANOVA showed that between-study correlation had no interpretable effect on the effect

size estimate of ES = .7 (p = .96, η_p^2 = .000). Moreover, the interaction terms, sample size

with between-study correlation, sample size with percent of missing data, between-study

correlation with percent of missing data, and the three way interaction were too small to

interpret (p = 1.00, η_p^2 = .000), (p = .16, η_p^2 = .000), (p = .99, η_p^2 = .000), and (p = 1.00,

η_p^2 = .000), respectively.

Table 15
Anova Summary Table for Effect Size Coefficient Estimate for ES = .7

Source	Sum of Squares	df	Mean Square	F	p	Partial Eta Squared (η_p^2)
SZ	.03	2	.02	4.23	.015	.000
BSC	1.08E-5	1	.00	.00	.959	.000
Miss	.01	2	.00	1.20	.302	.000
SZ * BSC	5.77E-7	2	.00	.00	1.000	.000
SZ * Miss	.03	4	.01	1.63	.164	.000
BSC * Miss	6.28E-5	2	.00	.01	.992	.000
SZ * BSC * Miss	5.58E-5	4	.00	.00	1.000	.000
Error	72.88	17982	.00			
Total	72.95	17999				

Note: SZ = Sample size (5,20,50); BSC = Between-study correlation (0.1,0.9); Miss = The percent of missing data (0.0%,30%,70%).

Effect Size = .5. Regarding the effect size estimate of ES = .5, the three-way ANOVA

showed that sample size, between-study correlation, and percent of missing data had no

interpretable effects (p = .10, η_p^2 = .000), (p = .90, η_p^2 = .000), and (p = .13, η_p^2 = .000)

respectively (Table 16). Additionally, the interaction term between sample size and

between-study correlation, sample size and percent of missing, between-study correlation

and percent of missing data, were also too small to interpret (p = 0.99, η_p^2 = .000; p = .28,

η_p^2 = .000; p = .95, η_p^2 = .000), respectively. Finally, the three-way interaction of the three

independent variables had a negligible effect (p = .99, η_p^2 = .000).

Table 16

Anova Summary Table for Effect Size Coefficient Estimate for ES = .5

Source	Sum of Squares	df	Mean Square	F	p	Partial Eta Squared η_p^2
SZ	.05	2	.03	2.33	.098	.000
BSC	.00	1	.00	.01	.903	.000
Miss	.04	2	.02	2.04	.131	.000
SZ * BSC	.00	2	.00	.02	.985	.000
SZ * Miss	.05	4	.01	1.27	.280	.000
BSC * Miss	.00	2	.00	.05	.951	.000
SZ * BSC * Miss	.00	4	.00	.03	.998	.000
Error	193.95	17982	.01			
Total	194.10	17999				

Note: SZ = Sample size (5,20,50); BSC = Between-study correlation (0.1,0.9); Miss = The percent of missing data (0.0%,30%,70%).

Standard Errors

Effect size = .7. The three-way ANOVA of the standard error for ES = .7 suggested that different levels of sample size had a large effect on the standard error estimates ($p < .001$, $\eta_p^2 = .571$) (Table 17). However, between-study correlation, percent of missing data, and interaction terms were all nonsignificant and had no effect on the standard error estimates. Mean standard error estimates for ES = .7 by levels of sample size are presented in Table 18. Scheffé's post-hoc test was conducted to find out which pairs of means were significantly different. The results of the tests are presented in Table 19.

The Scheffé post hoc test for significance indicated that the average standard error was significantly lower when the sample size was large (M = .029, SD =.003) than when the sample size was either small or medium (M = .084, SD = .033), (M = .045, SD = .008) (p = <.001) and was significantly lower for medium than for small sample sizes.

Table 17

Anova Summary Table for Standard Error Estimates for ES = .7

Source	Sum of Squares	df	Mean Square	F	p	Partial Eta Squared η_p^2
SZ	9.47	2	4.74	11956.72	<.001	.571
BSC	.00	1	5.77E-5	.15	.703	.000
Miss	.00	2	.00	2.34	.097	.000
SZ * BSC	.00	2	2.14E-5	.05	.947	.000
SZ * Miss	.00	4	.00	1.59	.174	.000
BSC * Miss	.00	2	8.41E-6	.02	.979	.000
SZ * BSC * Miss	.00	4	7.79E-6	.02	.999	.000
Error	7.12	17982	.00			
Corrected Total	16.60	17999				

Note: SZ = Sample size (5,20,50); BSC = Between-study correlation (0.1,0.9); Miss = The percent of missing data (0.0%,30%,70%).

Table 18

Standard Error Mean Estimate for ES = .7 by Level of Sample Size

Sample Size	Mean	Std. Deviation
Small = 5	.084	.033
Medium = 20	.045	.008
Large = 50	.029	.003
Total	.052	.030

Table 19

Scheffé's Comparison for Standard Error Mean Estimate for ES = .7

(I) Sample Size	(J) Sample Size	Mean Difference (I-J)	SD	Sig.
Small = 5	Medium = 20	.039	.000	<.001
	Large = 50	.055	.000	<.001
Medium = 20	Large = 50	.016	.000	<.001

Effect size = .5. The three-way ANOVA of the standard error for ES = .5 showed

interpretable main effects for sample size and percent missing but not for between-study

correlation. Further, there was a statistically significant interaction between sample size and percent missing (Table 21). Figure 7 displays the interaction. No other interactions were significant. Simple main effects were then conducted due to the significant interaction.

Table 20
Descriptive statistics for Percent of Missing Data by Sample Size for ES = .5

Sample Size	Mean	SD	N
Small = 5	.088	.048	6000
Medium = 20	.057	.021	6000
Large = 50	.037	.011	6000
Total	.061	.037	18000

Table 21
Anova Summary Table for Standard Error Estimates for ES = .5

Source	Sum of Squares	df	Mean Square	F	p	Partial Eta Squared η_p^2
SZ	7.76	2	3.88	4368.59	<.001	.327
BSC	.00	1	.00	.12	.728	.000
Miss	.82	2	.41	461.81	<.001	.049
SZ * BSC	.00	2	.00	.03	.967	.000
SZ * Miss	.56	4	.14	157.12	<.001	.034
BSC * Miss	.00	2	.00	.02	.977	.000
SZ * BSC * Miss	.00	4	.00	.17	.953	.000
Error	15.98	17982	.00			
Total	25.13	17999				

Note: SZ = Sample size (5,20,50); BSC = Between-study correlation (0.1,0.9); Miss = The percent of missing data (0.0%,30%,70%).

Simple main effects for percent of missing data by sample size

In the situation where the sample size was either small or large, the effect of simple main effect of percent of missing data in coefficient two standard error was small (p < .001, partial η_p^2 = .007), (p < .001, partial η_p^2 = .029) (Table 22). However, a

moderate main effect of percent of missing data was detected in coefficient two standard

error when the sample size was medium (p < .001, partial η_p^2 = .046).

Table 22

Simple Main Effects of Percent Missing by Sample Size for ES = .5

Sample Size		Sum of Squares	df	Mean Square	F	p	Partial Eta Squared η_p^2
Small = 5	Contrast	.12	2	.06	67.41	<.001	.007
	Error	15.98	17982	.00			
Medium = 20	Contrast	.78	2	.39	437.14	<.001	.046
	Error	15.98	17982	.00			
Large = 50	Contrast	.48	2	.24	271.50	<.001	.029
	Error	15.98	17982	.00			

Since the simple main effects were all significant, pairwise comparisons were

carried out. Table 23 below presents the pairwise comparisons and Table 24 provides the

cell means and standard deviations. Results for small sample size were significant but

with an effect too small to be interpretable but are presented below for completeness.

According to Table 23, when the sample size was small, the mean standard error was

.010, higher when 30% of data were missing than when there were no missing data. In

the situation when 70% of data were missing, the mean coefficient two standard error

was .002, higher when no missing data were present of than when there was 70% of data

missing. In addition, when the sample size was small, the mean ES = .5 standard error

was .009, higher for 30% of data missing than for 70% of missingness.

Table 23
Pairwise Comparison Results for ES = .5

Sample Size	(I) Percentage of Missingness	(J) Percentage of Missingness	Mean Difference (I-J)	SD	*p*
Small = 5	No Missing data	30% missing data	-.010*	.00	<.001
		70% missing data	-.002	.00	.290
	30% missing data	70% missing data	.009*	.00	<.001
Medium = 20	No Missing data	30% missing data	-.007*	.00	<.001
		70% missing data	-.027*	.00	<.001
	30% missing data	70% missing data	-.020*	.00	<.001
Large = 50	No Missing data	30% missing data	-.005*	.00	<.001
		70% missing data	-.021*	.00	<.001
	30% missing data	70% missing data	-.016*	.00	<.001

Table 24
Descriptive Statistics for Percent of Missing Data by Sample Size

Sample Size	Percentage of Missingness	Mean	SD	N
Small = 5	No Missing data	.084	.033	2000
	30% missing data	.094	.054	2000
	70% missing data	.086	.054	2000
	Total	.088	.048	6000
Medium = 20	No Missing data	.045	.008	2000
	30% missing data	.052	.011	2000
	70% missing data	.072	.027	2000
	Total	.057	.021	6000
Large = 50	No Missing data	.029	.003	2000
	30% missing data	.034	.004	2000
	70% missing data	.050	.010	2000
	Total	.037	.011	6000
Total	No Missing data	.053	.030	6000
	30% missing data	.060	.040	6000
	70% missing data	.069	.039	6000
	Total	.061	.037	18000

Note: N is the sample size of the particular level of missingness

For medium sample size, when there were no missing data, the mean ES = .5 standard error was .007, higher when 30% of data were missing than when there was no missing data. The mean ES = .5 standard error was .027 higher when 70% of data were missing than when there were no missing data. Moreover, when the sample size was medium, the mean ES = .5 standard error was .020, higher when 70% of data were missing than for 30% of missingness.

For large sample sizes and when there was no missing data, the mean ES = .5 standard error was .005, higher when 30% of data were missing than when there was no missing data. In the case of 70% of data missing, the mean ES = .5 standard error was .021, higher when 70% of data were missing than for 30% of missingness. The mean ES = .5 standard error was .016, higher when 70% of data were missing than with 30% of missingness.

Three-Way MANOVA for Effect Size outcomes ES = .7 and ES = .5

A three-way MANOVA with two dependent variables (ES = .7 and ES = .5) and three independent variables (method x sample size x percent of missingness) was performed to gain a meaningful comparison between the two methods. MANOVA was used primarily to determine whether the mean values for ES = .7 and ES = .5 outcomes differed between the two methods. The effects of sample size and percent missingness were also noted. As seen above, the between-studies factor had no interpretable effects in any of the ANOVAs. The multivariate test statistic, Wilks' Lambda (Λ), is the most widely used multivariate test statistic and was used in this test (Bray et al., 1985). Box's

M test of homogeneity of variance-covariance matrices was evaluated before interpreting the MANOVA results.

As in ANOVA, partial eta squared (η_p^2) rules of thumb for small =.01, medium =.06, and large =.14 were used to interpret the main effects and interactions. The Box's M test result is displayed in Table 25. Box's M suggested that the assumption of homogeneity of variances and covariances was violated. However, since the sample sizes were equal, this test result is not crucial because the MANOVA test statistic in this case is robust to violations of the assumption of homogeneity of variance and covariance matrices (Field & Miles, 2010).

Table 25
Box's Test of Equality of Variance/Covariance Matrices

Box's M	24418.13
F	478.42
df1	51.00
df2	424543627.78
p	<.001

The multivariate test of ES = .7 and ES = .5 showed that there was no significant effect of percent of missing data and method on the effect size estimates for ES = .7 and ES = .5 (Table 26). Sample size was significant, but it had a negligible partial eta square (p = .039, partial η_p^2 = .000). Furthermore, the two-way and three-way interactions were also nonsignificant. Next, a multivariate test was conducted for standard error. Table 27 shows Box' M test result which again suggested that homogeneity of variance/covariance matrices cannot be assumed. The MANOVA revealed that sample size had a large effect on the standard error estimates of ES = .7 and ES = .5 (p <.001, partial η_p^2 = .335).

73

Percent of missing data and method were also significant, p <.001, partial η_p^2 = .041, and

p <.001, partial η_p^2 = .043, respectively, and had a small effect (Table 28).

Table 26
Manova Summary Table for the Effect Size Estimate

Effect	Value	F	df	Error df	p	Partial Eta Squared η_p^2
SZ	1.00	2.52	4.00	35962.00	.039	.000
Miss	1.00	1.80	4.00	35962.00	.127	.000
Method	1.00	.89	2.00	17981.00	.412	.000
SZ * Miss	1.00	1.77	8.00	35962.00	.077	.000
SZ * Method	1.00	.14	4.00	35962.00	.967	.000
Miss * Method	1.00	.72	4.00	35962.00	.578	.000
SZ * Miss * Method	1.00	.19	8.00	35962.00	.992	.000

Table 27
Box's Test of Equality of Variance/Covariance Matrices

Box's M	56519.83
F	1107.39
df1	51.00
df2	424543627.78
p	<.001

Table 28
Manova Summary Table for the Standard Error Estimate

Effect	Value	F	df	Error df	p	Partial Eta Squared η_p^2
SZ	.44	4536.92	4.00	35962.00	<.001	.335
Miss	.92	382.79	4.00	35962.00	<.001	.041
Method	.96	408.07	2.00	17981.00	<.001	.043
Method * SZ	1.00	8.96	4.00	35962.00	<.001	.001
Method * Miss	.97	137.44	4.00	35962.00	<.001	.015
SZ * Miss	.94	147.12	8.00	35962.00	<.001	.032
Method * SZ * Miss	.99	26.78	8.00	35962.00	<.001	.006

Since multivariate significance was detected, univariate ANOVA results are presented and interpreted next as a follow up to MANOVA to identify any significant group differences for each of the effect sizes (.7 and .5).

The ANOVA revealed that the difference by method was only significant for ES = .7, with a small effect, (p <.001, partial η_p^2 = .043) (Table 29). Sample size differences were also significant for both ES = .7 and ES = .5 and had a large effect (p <.001, partial η_p^2 = .433) and (p <.001, partial η_p^2 = .324), respectively. The percent of missing data was found to be significant with a small effect on the standard error estimates for ES = .7 and ES = .5 (p <.001, partial η_p^2 = .033) and (p <.001, partial η_p^2 = .049), respectively. Even though the interaction effect of method and sample size was found to be statistically significant for standard error estimates, the value of η_p^2 was very small and ignorable (p <.001, partial η_p^2 = .001) and (p <.001, partial η_p^2 = .001), respectively. The method by sample size two-way interaction was only significant with a small effect on ES = .7) standard error estimates (p <.001, partial η_p^2 = .030). The last two-way interaction between sample size and percent of missing data was significant and had a moderate effect for ES = .5 standard error estimates (p <.001, partial η_p^2 = .055). However, sample size and percent of missingness interaction had an ignorable effect on ES = .7 standard error estimates. Finally, the three way- interaction between method, sample size, and percent of missingness was found to have an ignorable effect on both ESs standard error estimates.

Table 29
Univariate Anova Summary

Source	Dependent Variable	Sum of Squares	df	Mean Square	F	p	Partial Eta Squared η_p^2
Method	(ES = .7)	.57	1.00	.57	814.85	<.00	.043
	(ES = .5)	.00	1.00	.00	.51	.474	.000
SZ	(ES = .7)	9.65	2.00	4.82	6854.4	<.00	.433
	(ES = .5)	7.36	2.00	3.68	4315.1	<.00	.324
Miss	(ES = .7)	.43	2.00	.22	306.03	<.00	.033
	(ES = .5)	.79	2.00	.40	463.94	<.00	.049
Method * SZ	(ES = .7)	.01	2.00	.01	10.61	<.00	.001
	(ES = .5)	.01	2.00	.01	7.20	<.00	.001
Method * Miss	(ES = .7)	.39	2.00	.19	275.22	<.00	.030
	(ES = .5)	.00	2.00	.00	1.86	.155	.000
SZ * Miss	(ES = .7)	.10	4.00	.03	35.63	<.00	.008
	(ES = .5)	.89	4.00	.22	260.24	<.00	.055
Method * SZ * Miss	(ES = .7)	.11	4.00	.03	38.84	<.00	.009
	(ES = .5)	.05	4.00	.01	14.16	<.00	.003
Error	(ES = .7)	12.65	17982.0	.00			
	(ES =. 5)	15.34	17982.0	.00			
Total	(ES = .7)	23.92	17999.0				
	(ES = .5)	24.44	17999.0				

Scheffé's post-hoc test was performed to find out which pairs of means were significantly different (Table 30). For ES = .7, the mean standard error was .037 and .056 higher when the sample size was small than when the sample size was medium and large, respectively. The mean standard error was .018 higher when sample size was medium than when it was large. Regarding ES = .5, the mean standard error was .029 and .049 higher when the sample size was small than when the sample size was medium and large, respectively. The mean standard error was also .020 higher when the sample size was medium than when the sample size was large.

Table 30
Pairwise Comparison Results for ES = .7 and ES = 0.5

	(I) Sample Size	(J) Sample Size	Mean Difference (I-J)	SD	p
ES = .7 SE	Small = 5	Medium = 20	.037	.000	<.001
		Large = 50	.056	.000	<.001
	Medium = 20	Large = 50	.018	.000	<.001
ES = .5 SE	Small = 5	Medium = 20	.029	.001	<.001
		Large = 50	.049	.001	<.001
	Medium = 20	Large = 50	.020	.001	<.001

Note: ES = effect size; SE = standard error; SD = standard deviation

Table 31
Pairwise Comparison Results for ES = .7 and ES = 0.5

Dependent Variable	(I) Percentage of Missingness	(J) Percentage of Missingness	Mean Difference (I-J)	SD	p
ES = .7 SE	No Missing data	30% missing data	-.006	.000	<.001
		70% missing data	-.012	.000	<.001
	30% missing data	70% missing data	-.006	.000	<.001
ES = .5 SE	No Missing data	30% missing data	-.008	.001	<.001
		70% missing data	-.016	.001	<.001
	30% missing data	70% missing data	-.008	.001	<.001

Tables 32, 33, and 34 present the mean and standard deviation of standard error estimates by method, sample size, and percent of missing data for ES = .7 and ES = .5.

Table 32
Means and Standard Deviations for MVMA and UVMA Estimates

	(ES = .7) SE		(ES = .5) SE	
	Mean	SD	Mean	SD
UVMA	.064	.041	.061	.037
MVMA	.052	.030	.061	.036

Table 33

Means and Standard Deviations for MVMA and UVMA Standard Error Estimates by Percent of Missing Data

		MVMA		UVMA	
Percent of missing data		ES = .7 SE	ES = .5 SE	ES = .7 SE	ES = .5 SE
No Missing data	Mean	.052	.053	.052	.053
	SD	.030	.030	.030	.031
30% missing data	Mean	.053	.060	.064	.061
	SD	.030	.040	.044	.041
70% missing data	Mean	.053	.069	.075	.069
	SD	.030	.036	.044	.037

Table 34

Means and Standard Deviations for MVMA and UVMA Standard Error Estimates by Sample Size

		MVMA		UVMA	
Sample Size		ES = .7 SE	ES = .5 SE	ES = .7 SE	ES = .5 SE
Small = 5	Mean	.083	.088	.095	.086
	SD	.033	.046	.053	.049
Medium = 20	Mean	.045	.056	.058	.059
	SD	.008	.021	.022	.022
Large = 50	Mean	.029	.037	.038	.038
	SD	.003	.011	.010	.012

Results of the univariate and multivariate meta-analysis of CDSSs

Univariate Meta-Analysis of the Effect of CDSSs on LDL Levels

The first of three meta-analyses estimated the mean effect of CDSS's on reducing LDL levels in diabetic patients. A random-effects model and restricted maximum likelihood estimation (REML) were used in the three meta-analyses. The number of studies were included in this meta-analysis was 16 studies and a total number of 10,603 patients. The estimate of the mean effect size of CDSS's using a random-effects model was significant SMD = -0.07, 95% CI [-0.12, -0.02]. According to Cohen's (1988) guidelines, defined SMD effect sizes as small (0.2), medium (0.5), and large (0.8)., the SMD of -0.07 is considered as a very small. The estimate for τ^2, the total heterogeneity, was $\tau^2 = 0.001$, and the estimated between-studies standard deviation τ was $\tau = 0.03$. The result of the test of heterogeneity was $(df=15) = 20.81$, $p = .007$, and $I^2 = 12.88\%$. According to Higgins and Thompson's (2002), guidelines for interpreting the descriptive statistic I^2, the I^2 value of about approximately 13% for this meta-analysis indicated the heterogeneity might not be important.

Figure 6 shows the forest plot, the estimated mean difference, and its 95% confidence interval. The squares close to the center of the estimated mean difference will be larger for studies with smaller variances and so more precision. The dashed vertical line down the middle represents the line of no effect. Each square has a horizontal line extending through it that represents the study's 95% confidence interval. The shorter the line, the more precise the estimate of that study's effect size (Borenstein et al., 2011; Card, 2015). The black diamond on the summary line represents two things: the center of

this black diamond represents the mean effect size for this sample of studies and its width indicates the 95% confidence interval for the mean effect size.

Publication Bias

The funnel plot is usually used to assess bias in a meta-analysis (Ferrer, 1998; Song et al., 2002; Tang & Liu, 2000). Therefore, the first step was to generate a funnel plot in order to get an initial understanding of the risk of bias if present. In this plot, the y-axis represents the standard error and the x-axis represents the effect size. Each dot represents a study. If publication bias is present, the funnel plot will be asymmetrical. Figure 7 shows the funnel plot of the studies included and as it appears from the plot, no evidence of bias is detected as the dots seem to be randomly scattered around the funnel. In order to discover the number of missing studies that would lead the overall effect to a non-zero level, the Fail-safe N method was used. The result of Fail-safe N indicated that 17 studies with effect size zero could be added to the meta-analysis before the result lost statistical significance. However, looking at the forest plot in figure 8, there are only three studies that do not include the line of no effect. Therefore, it might not be difficult to get 17 studies. Finally, and as a way to visualize the studies that might be missing, a Trim and Fill algorithm was used and then the funnel plot in Figure 8 was generated.

The hollow dots indicate studies that need to be added in order to have a more symmetric plot. AS shown in Figure 10, only one study was added to the funnel plot which is evidence that the risk of publication bias was low if not present at all.

Figure 10
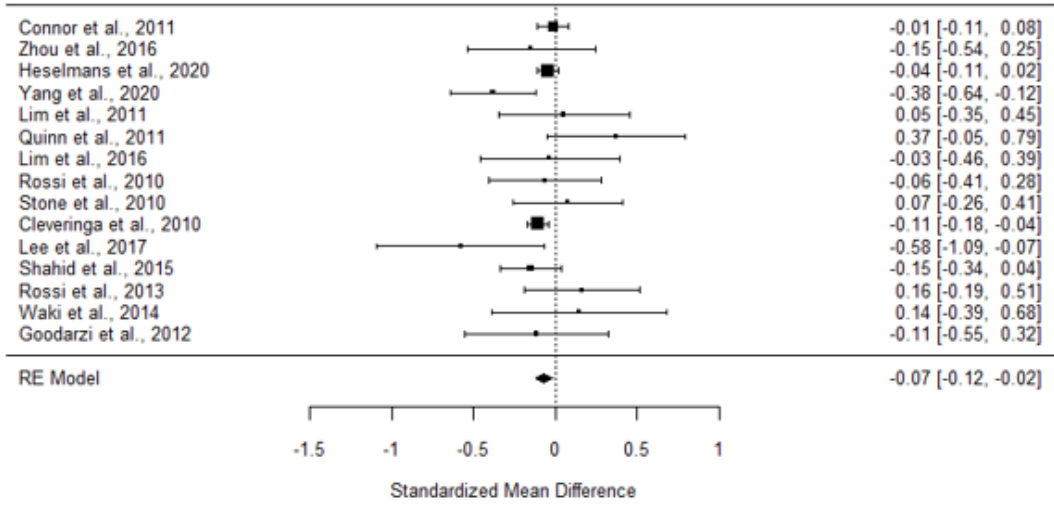*Forest Plot for the LDL Random-Effects Model*

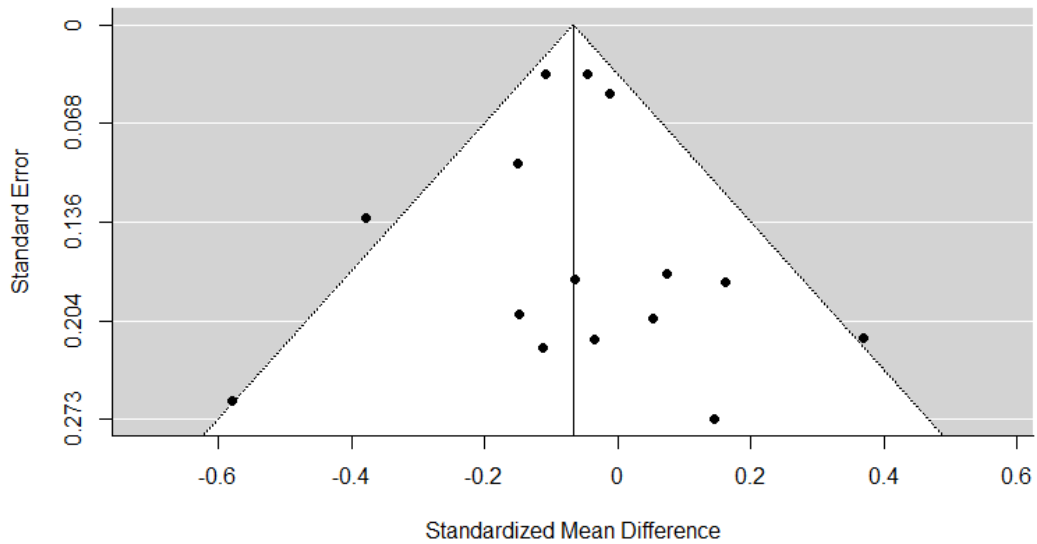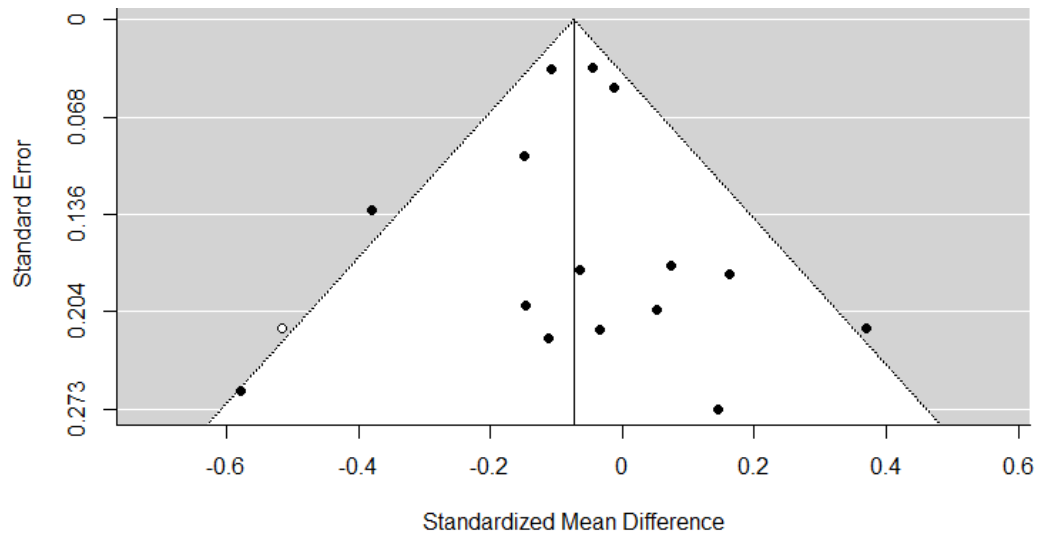

Figure 12
Funnel Plot for LDL Random-Effects Model

Figure 15

Funnel Plot for LDL after Trim & Fill Random-Effects Model

Univariate Meta-Analysis of the Effect of CDSSs on HbA1c levels

The number of studies included in this meta-analysis was 38 studies with a total number of 18,144 patients. The estimate of the mean effect size pf CDSS's using a random-effects model was statistically significant *SMD* = -0.31, 95% *CI* [-0.48, -0.13]. SMD of -0.31 for the effect of CDSSs on the level of HbA1is considered as a small effect size. The estimate for τ^2, the total heterogeneity, was $\tau^2 = 0.26$, and the estimated between-studies standard deviation τ was $\tau = 0.51$. The result of the test of heterogeneity was Q $(df = 37) = 1774.13$, $p < .001$, and $I^2 = 96.28\%$. The I^2 value of about approximately 96% indicated considerable heterogeneity. Figure 9 shows the forest plot for the HbA1c random-effects model.

Publication Bias

The funnel plot presented in Figure 10 shows the funnel plot of the 38 studies included. It appears from the plot that some degree of asymmetry is present. There were

82

some studies that have very strong effect that appear as outliers on the left and right side of the funnel. Most of the effect sizes are scattered around the mean effect size. Additionally, studies with small or non-significant effects might be missed as the lower base of the funnel plot had almost no studies.

Since significant heterogeneity was detected and a random effects model was used, Fail Safe *N* was not reported per Card's (2015) recommendation. This is because the computation of File Safe N does take into account whether the studies are homogeneous or heterogeneous, which makes the method invalid, especially when the heterogeneity is large (Card, 2015). Figure 11 shows the Funnel Plot for HbA1c after a trim and fill random-effects model was generated. The plot is identical to the original funnel plot meaning that the meta-analysis of the HbA1c is robust to publication bias.

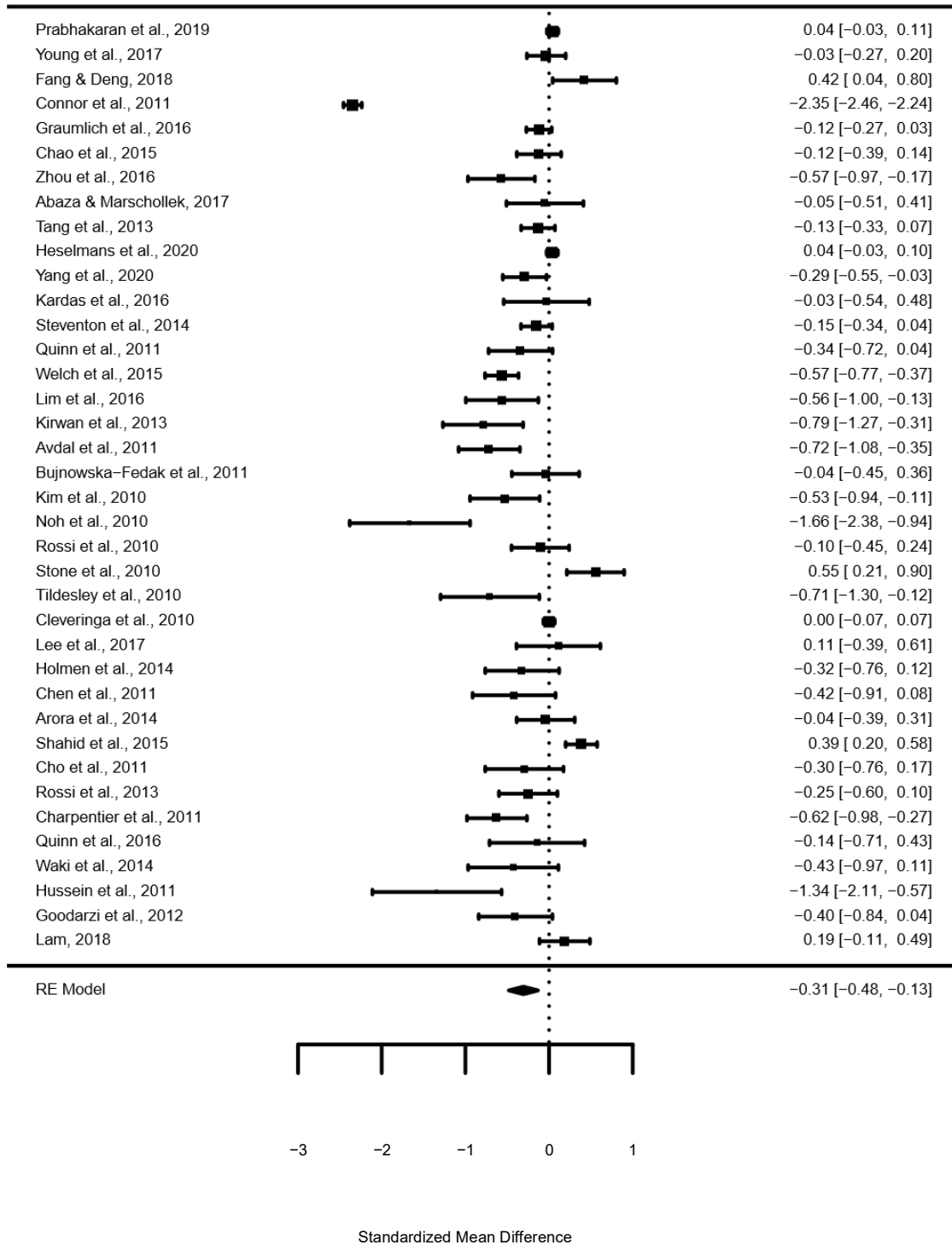Figure 18

Forest Plot for the HbA1c Random-Effects Model

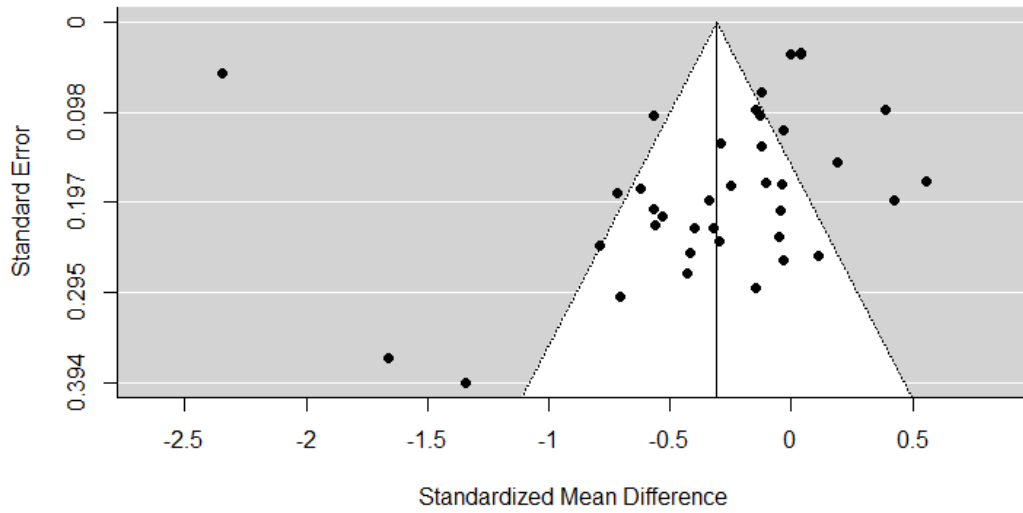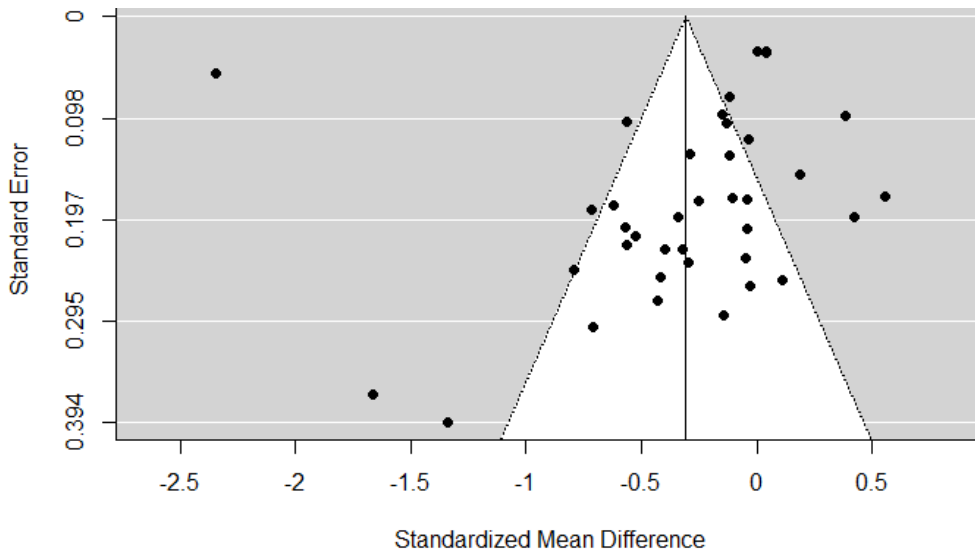Figure 23
Funnel Plot for HbA1c Random-Effects Model

Figure 21
Funnel Plot for HbA1c after Trim & Fill Random-Effects Model

Univariate Meta-Analysis of the Effect of CDSSs on Blood Pressure levels

The third UVMA was the meta-analysis of the effect of CDSSs on blood pressure levels. As mentioned in Chapter 2, the outcome of blood pressure was converted to Pulse Pressure (PP) in order to have only one value that represents blood pressure level. The number of studies included in this meta-analysis was 20 studies and a total number of 11,841 patients. The estimate of the mean effect size was non-significant SMD = -0.55, 95% CI [-1.24, 0.15]. The estimate for τ^2, the total heterogeneity, was τ^2 = 2.48, and the estimated between-studies standard deviation τ was τ = 1.58. The result of the test of heterogeneity was Q (df = 19) = 7950.36, $p < .001$, and I^2 = 99.47%. According to the value of I^2, a considerable heterogeneity is present. Figure 12 shows the forest plot for the pulse pressure random-effects model.

Publication Bias

A funnel plot was created for the 20 studies included in the analysis (Figure 13). It can be seen from the plot that a large amount of heterogeneity is present from the spread of the studies. There were some studies that had a very strong effect on the left side of the funnel. Studies with small or non-significant effects could be missed at the lower base of the funnel plot as well as the top left.

Again, Fail Safe N was not reported per Card's (2015) recommendation because of the significant heterogeneity. Figure 14 shows the funnel plot for PP after a trim and fill random-effects model was generated which indicated that 5 studies could be missed on the right side of the funnel which is an indication of publication bias.

Figure 26
Forest Plot for PP Random-Effects Model

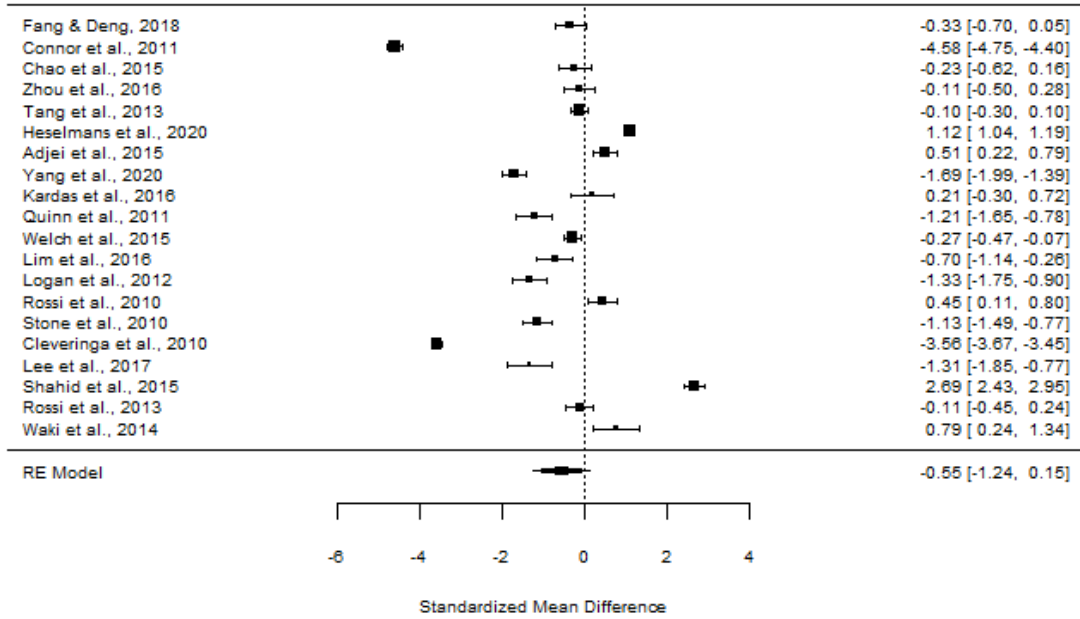|		
Fang & Deng, 2018		-0.33 [-0.70, 0.05]
Connor et al., 2011		-4.58 [-4.75, -4.40]
Chao et al., 2015		-0.23 [-0.62, 0.16]
Zhou et al., 2016		-0.11 [-0.50, 0.28]
Tang et al., 2013		-0.10 [-0.30, 0.10]
Heselmans et al., 2020		1.12 [1.04, 1.19]
Adjei et al., 2015		0.51 [0.22, 0.79]
Yang et al., 2020		-1.69 [-1.99, -1.39]
Kardas et al., 2016		0.21 [-0.30, 0.72]
Quinn et al., 2011		-1.21 [-1.65, -0.78]
Welch et al., 2015		-0.27 [-0.47, -0.07]
Lim et al., 2016		-0.70 [-1.14, -0.26]
Logan et al., 2012		-1.33 [-1.75, -0.90]
Rossi et al., 2010		0.45 [0.11, 0.80]
Stone et al., 2010		-1.13 [-1.49, -0.77]
Cleveringa et al., 2010		-3.56 [-3.67, -3.45]
Lee et al., 2017		-1.31 [-1.85, -0.77]
Shahid et al., 2015		2.69 [2.43, 2.95]
Rossi et al., 2013		-0.11 [-0.45, 0.24]
Waki et al., 2014		0.79 [0.24, 1.34]
RE Model		-0.55 [-1.24, 0.15]

Figure 29
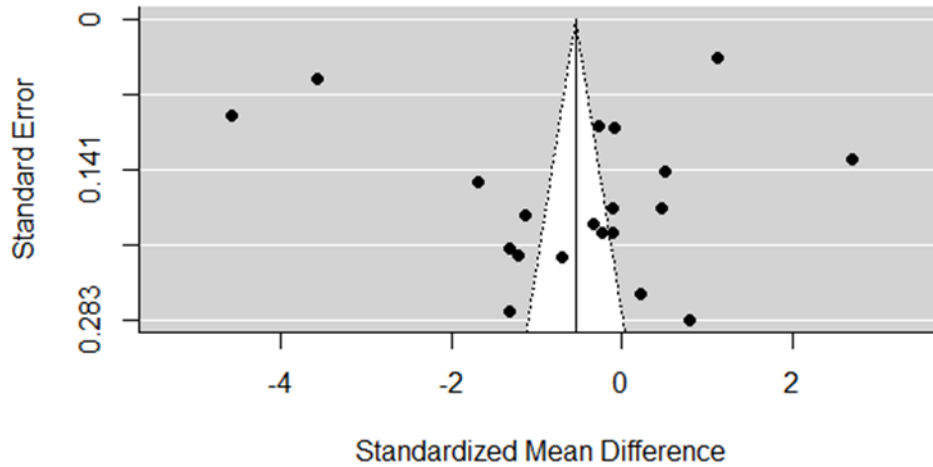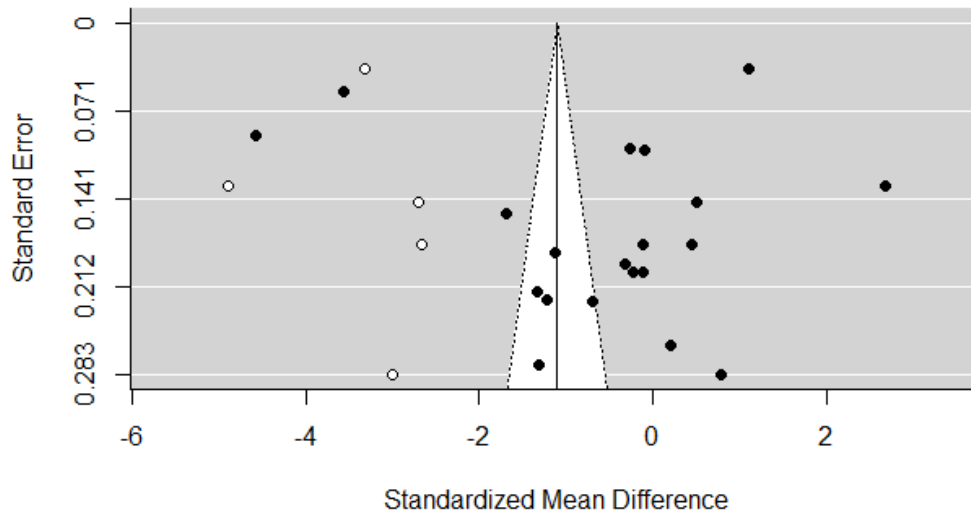*Funnel Plot for PP Random-Effects Model*

87

Figure 32

Funnel Plot for PP after Trim & Fill Random-Effects Model

Multivariate Meta-Analysis for the Effect of CDSSs on the levels of LDL, HbA1C,

and PP

Two multivariate meta-analyses of 41 studies were conducted in order to assess the effect of the CDSSs on the three outcomes under study. The first MVMA was conducted with missing cells deleted which contained no information about the particular outcome. The second MVMA was conducted with missing values imputed, assuming that they were missing completely at random. The result of both MVMAs are presented next.

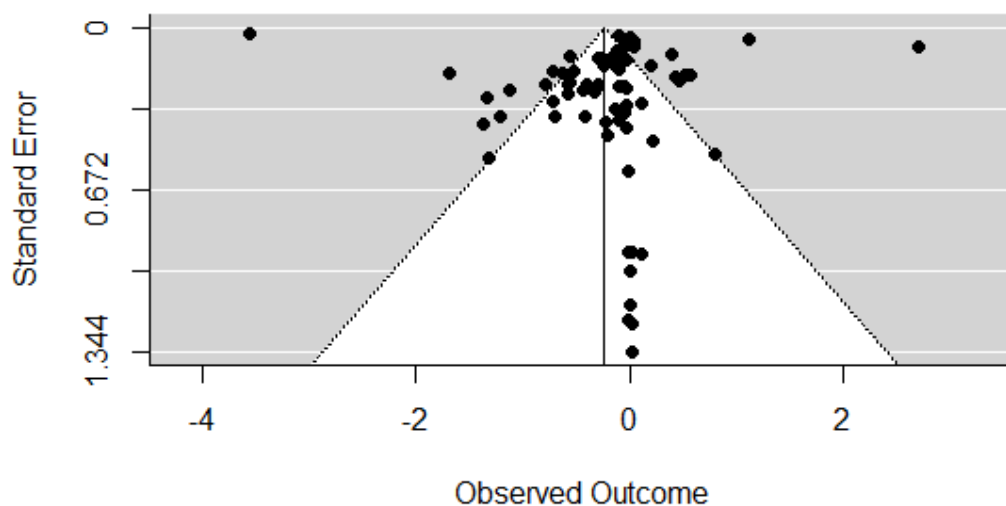**MVMA with Missing Values Deleted**

In this MVMA, the result shows that the estimated effect size of the CDSSs for all of the three outcomes was less than small, LDL = -0.10, 95% *CI* [-0.61, 0.41], HbA1c = -0.27, 95% *CI* [-0.53, -0.01] and PP = -0.26, 95% *CI* [-0.62, 0.11].). However, CDSSs had a significant and less than small effect on HbA1c levels. The estimate for τ^2, the total

88

heterogeneity, was $\tau^2 = 0.58$, and the estimated between-studies standard deviation τ was $\tau = 0.76$. The result of the test of heterogeneity was Q $(df = 67) = 44480.55$, $p < .001$ meaning that heterogeneity was present.

Publication Bias

The funnel plot of the three outcomes shows that some degree of asymmetry might be present (Figure 17). It appears from the plot that there might be some studies that are missing on the left (upper and lower) sides of the funnel. To find out how many studies are missing, a trim and fill approach was used. However, the use of the trim and fill approach is slightly different in MVMA. In order to run a trim and fill random effects model, the function "rma" was used instead of "rma.mv" since the latter does not support trim and fill for MVMA. The result of the trim and fill approach showed that there were 21 missing studies for the funnel plot to be symmetric. Figure 17 shows the funnel plot after trim and fill. It should be mentioned that in Figure 16 and Figure 17, the standard errors were plotted

Figure 35
Funnel Plot for LDL, HbA1c, and PP Random-Effects Model

89

against the observed values instead of the residuals in order to run the trim and fill analysis for MVMA.

MVMA with Missing Values Imputed

In this MVMA, instead of deleting the cells that had missing values, the MICE Package (Multivariate Imputation via Chained Equations) was used (Buuren, S. V., & Groothuis-Oudshoorn, K, 2010). MICE is one of the commonly used packages for R users (Z. Zhang, 2016). The result of this MVMA showed that the estimated effect size for the three outcomes were , LDL = -0.19, 95% CI [-0.39,0.00], HbA1c = -0.21, 95% CI [-0.43, -0.01] and PP = -0.19, 95% CI [-0.37, -0.01].). CDSSs had a significant effect on the levels of HbA1c and PP. CDSSs had a small effect on HbA1c, and a less than small effect on PP. The estimate for τ^2, the total heterogeneity, was $\tau^2 = 0.43$, and the estimated between-studies standard deviation τ was $\tau = 0.65$. The result of the test of heterogeneity was Q ($df = 120$) = 65478.70, $p < .001$ indicating that heterogeneity was present.

Publication Bias

The funnel plot of LDL, HbA1c, and PP shows that there is still some degree of asymmetry (Figure 18). It can be seen from the funnel plot in Figure 18 that there might be some missing studies on the left (upper and lower) sides of the funnel. The result of the trim and fill approach showed that 36 studies should be added to the funnel plot to be symmetric. Figure 19 shows the funnel plot after trim and fill.

Figure 39
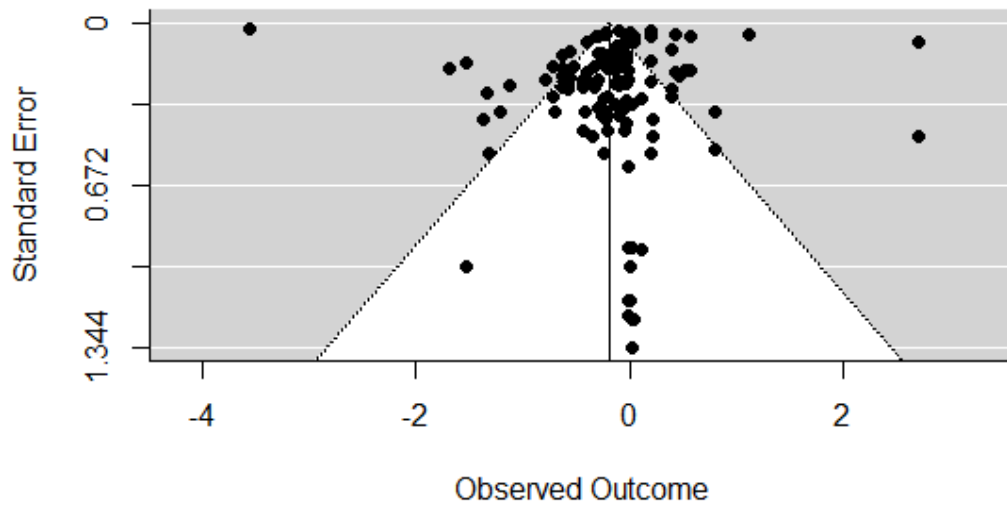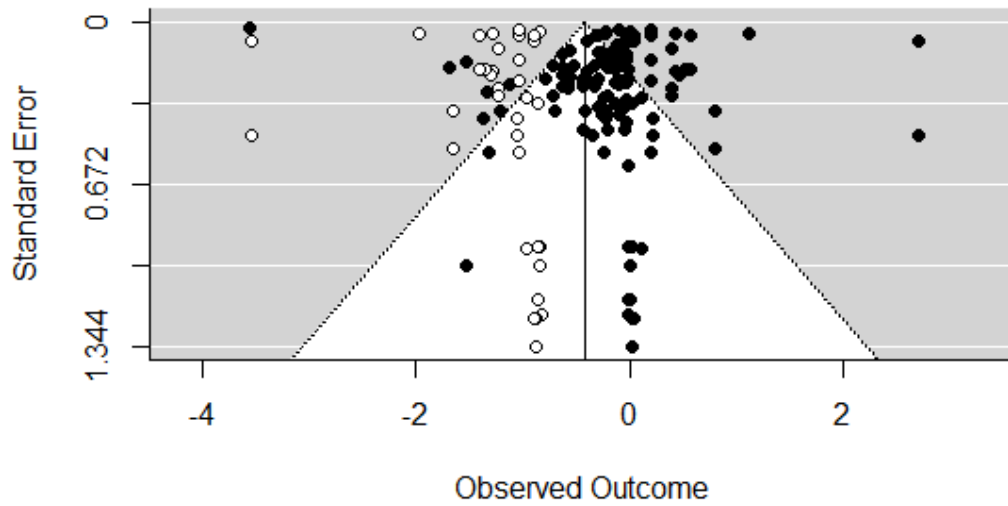*Funnel Plot for LDL, HbA1c, and PP Random-Effects Model*

Figure 38
Funnel Plot for LDL, HbA1c, and PP After Trim & Fill Random-Effects Model

The results of the three UVMAs and the two MVMAs are summarized in Table 35.

Table 35

Summary of the result of the two MVMAs

Outcome	UVMA			MVMA Imputed missing values			MVMA Deleted missing values		
	ES	SE	95% CI	ES	SE	95% CI	ES	SE	95% CI
LDL	-0.07	0.03	-0.12 - -0.02	-0.19	0.10	-0.39 - 0.00	-0.10	0.26	-0.61 - +0.41
HbA1c	-0.31	0.09	-0.48 - -0.13	-0.21	0.11	-0.43 - -0.01	-0.27	0.13	-0.53 - -0.01
PP	-0.55	0.35	-1.24 - +0.15	-0.19	0.09	-0.37 - -0.01	-0.26	0.18	-0.62 - +0.11

The aim of this dissertation was twofold. First, simulations were conducted to understand whether sample size, percent of missing data, and between-study correlation had an effect on meta-analysis estimates. Secondly, an empirical study was conducted to evaluate the effect of CDSSs on diabetes management. The idea of conducting these two studies was to first understand what effect the factors would have on the two outcomes (ES = .7 and ES = .5) and then compare the findings with the results of the empirical study of the CDSSs. However, several factors affected this comparison with the empirical study. The simulations had two outcomes while the empirical study had three. In the empirical study, due to the lack of information reported in the selected studies, several imputations were conducted in the MVMA in order to estimate the variance-covariance matrix and the missing values. As a result of these imputations, MVMA estimates was slightly affected. Thus, no direct comparisons between the simulation and the empirical study were sound.

Chapter Four

Discussion

This chapter summarizes the primary findings with links to the literature and concludes with a review of the study limitations and recommendations for future study.

Summary of the Primary Findings of the Simulation Study

To get an initial understanding of what impact the three factors (sample size, percent of missing data, and between-study correlation) had on both UVMA and MVMA methods, two-way ANOVAs and three-way ANOVAs were conducted in order to evaluate the impact of these factors on the UVMA effect size and standard error estimates. The between-study correlation was not included in this the UVMA analysis since a between-study correlation does not exist in the case of UVMA. This exclusion was because the between-study correlation showed no impact on the effect size estimation and their standard errors. Subsequently 3-way MANOVAs were used that included method (UVMA/MVMA) as a factor in addition to sample size and percent of missing data to directly compare the effects of method and its interaction with the other two factors. In this subsequent analysis, ES = .7 and ES = .5 were treated simultaneously as the dependent variables.

ES = .7. The two-way ANOVA of the UVMA simulation found that there was no interpretable main or interactive effect of sample size and percent of missingness on the effect size estimate. However, the analysis showed that different levels of these factors might have an impact on the quality of the effect size estimates (standard errors). The influence of the percent of missingness in general was moderate. However, different levels of sample size were found to have a large impact on the standard error estimates for ES= .7. The effect of the percent of missingness on the ES = .7 standard error estimate was small when the sample size was either small or large and this effect was moderate when sample size was medium. The interaction effect of sample size and percent of missingness was in general small. To summarize, standard errors were lower when there were no missing data and were incrementally larger when there were higher levels of missing data. The effect was most pronounced for small samples sizes and differences were smaller for large sample sizes. This result can be understood as the influence of sample size on any standard error estimate: as sample size increases, the standard error decreases.

ES = .5. Again, the two-way ANOVA of the UVMA simulation yielded no significant main or interactive effect of sample size and percent of missingness on the effect size estimate. However, sample size again played a role in the estimation of the standard errors. The analysis showed that sample size had a large effect in estimating the standard errors. Moreover, percent of missing data was found to have a moderate effect as in the case of ES = .7. The only difference was in the joint effect of sample size and percent of missingness. The interaction of these two factors was moderate which is

94

different than for ES = .7. The effect of missingness was small in the case of small and large sample size. Percent of missing data showed a moderate effect when the sample size was medium. However, as in the case of ES =.7, the difference in standard error estimates tends to increase as the percent of missing data increases in both the medium and large sample size.

It should be noted that results showed that 30% of missing data across the three levels of sample size had a small difference when compared to the situation when no missing data was present. Therefore, according to this finding, researchers should not be overly concerned as long as the percent of missing data is 30% or less in order to get effect size estimates that have almost the quality (small standard errors) of the estimates that would be produced if no missing data were present.

Regarding MVMA, the ANOVA had one more factor added which was the between-study correlation. The analysis yielded essentially the same results as the two-way ANOVA of the UVMA in case of sample size and percent of missing data for the two effect size estimates (ES = .7 and ES = .5). No interpretable effects of these factors were detected. In addition, there was no interpretable effect of the between study-correlation, which supports the findings of previous literature that examined the effect of the between-study correlation and whether it makes a difference in estimating the effect size compared to UVMA (Boca et al., 2017; Price et al., 2019). The two- and three-way interactions were both too small to be interpretable and had little effect on the effect size estimates.

On the other hand, the sample size had a large effect on the estimate of the standard error for ES = .7. The results showed that the standard error estimates decreased

95

as the sample size increased, which was to be expected. Larger sample sizes increase precision and produce smaller standard errors for estimates (Seaman et al., 1999). The analysis also showed that the percent of missing data had little influence on the standard error estimates of ES = .7. In MVMA, the study does not have to have all the outcomes of interest to be included in the analysis. One outcome is enough for a study to be included and that inclusion reduces the effect of missing information. That is considered as an advantage of MVMA over UVMA. The between study correlation effect again was too small to be interpretable and had no effect on standard errors estimates of ES = .7. All interaction terms had no interpretable effects as well.

The situation when ES = .5 was somewhat different. The effect of sample size was still large, but it was smaller than in the case of ES = .7. The effect of percent of missing data was interpretable, though with a small effect on the standard error of ES = .5. Recalling the effect of missingness on the standard error estimates of the ES's in the case of UVMA, percent of missing data had an interpretable effect on both estimates, not only on the one that had missing data. Including studies that have at least one of the estimates under study allowed the MVMA to reduce this effect of missingness and hence keep more information in order to get a better estimate with better statistical properties. The interaction of sample size with percent of missing data was interpretable with a small effect.

The two-way ANOVA of the UVMA and the three-way ANOVA of the MVMA revealed some of the differences between the two methods. In UVMA, the idea of excluding all studies that do not have all the outcomes of interest had an effect on the standard error estimates of all effect sizes. On the other hand, in MVMA, the outcome

with missing data was the only outcome affected by the missingness. Therefore, using MVMA might reduce the risk of the missingness and hence produce estimates with better statistical quality.

In order to better understand these differences, establish a more accurate conclusion, and gain a meaningful comparison, a MANOVA was conducted which allowed the researcher to analyze the effects of both methods simultaneously. Conducting the MANOVA allowed inclusion of two dependent variables, ES = .7 and ES = .5, and three independent variables (method x sample size x percent of missingness). The between-study correlation was excluded from the analysis since it showed no impact on the effect size estimation and their standard errors. The method factor was added which had two levels (UVMA and MVMA) in order to see if the use of the method would have an effect on the estimates. The multivariate analysis of ES = .7 and ES = .5 showed that no interpretable effect was detected from method, sample size, and percent of missing data on effect size estimates which supports the conclusions previously gained from the ANOVAs.

The MANOVA for ES =.7 and ES = .5 standard error estimates showed that sample size had a great effect in estimating the standard errors. Again, as sample size gets larger, the standard errors get smaller and the estimates of ES = .7 and ES = .5 get closer to the true values.
The MANOVA also showed that standard error estimates were slightly affected by the percent of missing data. The type of method used, either UVMA or MVMA, was found to have a small effect in general. An ANOVA was essential to conduct as a follow up to identify any significant group differences for each of the effect sizes (.7 and .5).

For the method factor, the ANOVA showed that using UVMA or MVMA might give a different result in the case of ES = .7. The mean standard errors of MVMA were smaller compared to UVMA in the case of ES = .7. Using UVMA when there were missing studies resulted in standard error inflation for both ES's (ES = .5 and ES = .7) which is a result of excluding the studies that did not have enough data. However, mean standard error estimates for ES = .5 were almost identical for both methods (Table 32).

Having missing data in ES = .5 was again interpretable with a small effect. In the case of no missing data, both mean standard error estimates produced by UVMA and MVMA were almost identical. However, mean standard error estimates of ES = .7 that were produced by MVMA were smaller than the ones produced by UVMA with 30% and 70% of missingness with ES = .5.

MVMA was also superior to UVMA in the three levels of sample size for ES = .7. Throughout the three levels of sample size, mean standard error estimates of ES = .7 were smaller than the ones from the UVMA. However, mean standard error estimates of ES = .5 were almost identical for both methods.

As a summary of the results of the simulation, between-study correlation showed no significant effect on estimates when it was considered in MVMA simulation. This result has been supported by the literature and still there is no clear explanation for why between-study correlation shows no effect (D.-G. D. Chen & Peace, 2013; Price et al., 2019). Between-study correlation might be meaningful if the studies that would be included in a meta-analysis contain related data. That means different studies might be included that have the same set of patients and/or same characteristics. In this case, considering individual patient data (IPD) would be meaningful. IPD from completed

clinical trials are usually unavailable. However, IPD should be responsibly shared in order to enhance the reliability and efficiency of any comparative meta-analyses (Tudur Smith et al., 2015).

MVMA performed better compared to UVMA when missing data were present. In UVMA, all outcomes of interest would be affected by missing data and would result in estimates with less statistical quality. Therefore, MVMA is more robust to missing data and should be preferred over UVMA in this case. However, when no missing data are present, both methods performed equally and the researcher should choose the most convenient method for his/her study. Increasing the sample size was effective in reducing the standard errors of the estimates in both methods with preference to MVMA. The more studies included in a meta-analysis, the more information about the variables under study will be gained. MVMA offers an easier way to reach a bigger sample size compared to UVMA. As explained before, in UVMA the possibility of a study to be included when there is more than one outcome of interest is low compared to MVMA where only one outcome present in the study is enough for the study to be included (Jackson et al., 2011).

Summary of the Finding of the Univariate Meta-Analyses of Effects of CDSS

In order to evaluate the effect of CDSSs on the management of LDL, HbA1c, and PP levels in diabetic patients, three UVMAs were conducted separately. The first UVMA of 16 studies was conducted to evaluate the effect of CDSSs on LDL levels. The result showed that CDSSs had a significant effect that was less than small on managing LDL levels. This result concurs with some RCT's and systematic reviews found that showed the effect of the CDSSs were very small or not effective at all in managing LDL levels (Ali et al., 2016; P. J. O'Connor et al., 2011). The measures of heterogeneity (Q and I^2)

both indicated heterogeneity was very small and could be ignored. Based on that result, the studies that were included in this meta-analysis shared the same true effect size. This claim could be supported by looking at the forest plot (Figure 8), where it can be seen that the effect did not vary much across studies and hence the pooled effect size should be a good reflection of the true effect size. Regarding publication bias, the funnel plot of the studies presented in Figure 9 showed a symmetric spread of the studies which is a sign of unbiasedness. Fail-safe N revealed that a total of 17 studies with zero effect size could be added before the result lost statistical significance. However, as mentioned before, there were only three studies that do not include the line of no effect meaning that 17 studies wouldn't be difficult to get. Finally, Trim and Fill algorithm indicated that only one study needs to be added to the meta-analysis which again support what the funnel plot showed (Figure 10).

The second UVMA was to examine the effect of CDSSs on managing the levels of HbA1c. In this random-effects meta-analysis, 38 studies were included. The effect of CDSSs was significant and had a small effect in reducing the levels of HbA1c. The literature reported inconsistent results of the effect of CDSS in reducing levels of HbA1c. Some reported non-significant effects and others reported small to moderate effects (Ali et al., 2016; Jeffery et al., 2013). A high amount of heterogeneity was found based on the values of (Q and I^2) meaning that there were differences in sample characteristics of the studies. For instance, and from a clinical standpoint, while one study could have included old people, other studies have recruited young participants. Another factor that could be a reason behind the heterogeneity is the kind of intervention (CDSS) that was used which varies from one study to another. Some CDSS might have a stronger effect and some

100

might have small or no effect. As a result, the pooled effect size could be misleading and not reflect the true effect size. The funnel plot of the 38 studies (Figure 12) shows that there is a possibility of publication bias. However, as can be seen in Figure 13, the trim and fill random-effects model Funnel plot for HbA1c is identical to the original funnel plot which mean the little or no publication bias is present. This could be because the number of the studies included, which is considered as a large sample size in meta-analysis, reduced the possibility of missing related studies. Another possibility might be related to the effectiveness of CDSSs on HbA1c levels. As mentioned earlier, and based on the literature, the effect on HbA1c was found either non-significant or had a small to moderate effect. As a result, most of the studies were published even with non-significant results since no large effect was reported in the literature to the best of the researcher's knowledge.

The last UVMA was conducted for the purpose of evaluating the effectiveness of CDSSs on blood pressure levels which was converted during the coding process to pulse pressure. A total of 20 studies were included in this analysis. The effect of the CDSSs on reducing levels of PP was nonsignificant. CDSSs had been found effective in reducing blood pressure; however, the result of the meta-analysis did not support what was found in the literature (Ali et al., 2016; Holbrook et al., 2009; O'Connor & Sperl-Hillen, 2019). Again, both Q and I^2 indicated that there was considerable heterogeneity. It can be noticed from the forest plot that some studies had strong effects and some did not (Figure 14). As a result, the pooled effect in this case was not reflective of the true effect of the CDSSs. The funnel plot shows that some degree of publication bias was present. The right side of the plot could be missing some studies. Based on the result of Trim and Fill, the number of studies that are missing is 5 studies to make the funnel plot symmetric (Figure 15).

Summary of the Results of the Multivariate Meta-Analyses of Effects of CDSS

Two MVMAs were conducted due to the fact that there were missing values. The first MVMA was conducted with missing values deleted and an imputation technique was used in the second MVMA instead of deleting the missing values. A total of 41 studies were included in these MVMAs.

The first MVMA with missing values deleted yielded a nonsignificant very small effect of CDSSs on the levels of LDL and a significant small effect on the levels of HbA1c. the effect of CDSSs was non-significant small effect on PP. The Q test of homogeneity indicated that heterogeneity was detected. That could be explained by the inconsistent effect of CDSS throughout the studies. This result is different than that found using separate UVMAs. The first thing to notice is that effect size estimates of the three outcomes are slightly different and the effect of CDSS on PP dropped from medium to less than small. Secondly, the standard errors of UVMA were smaller than the ones produced by MVMA for LDL and HbA1c. However, MVMA had a smaller standard error for the PP outcome. MVMA with missing values deleted showed that CDSSs had only a significant effect on HbA1c. However, UVMAs yielded the finding that CDSSs had a significant effect on LDL and HbA1c levels. Therefore, utilizing either UVMA or MVMA could result in different conclusions. The funnel plot in Figure 16 shows some degree of asymmetry. Studies could be missing on the top and the bottom of the left side of the plot. Trim and Fill indicate that 21 studies missing studies would be needed for the plot to be symmetric (Figure 17).

Deleting the missing values was one way to successfully run the R code. However, one of the advantages of MVMA is that it can include more information about

the outcomes. In order to see the performance of MVMA with no missing data, imputation was applied in the second MVMA. The effect of CDSSs on LDL was still nonsignificant. CDSSs were found to have a significant small effect on the levels of HbA1c and a non-significant less-than small effect on PP. Heterogeneity was found to be significant and publication bias was present with 36 studies need to be added to gain symmetry (Figure 19).

When comparing the two MVMAs, it can be seen that MVMA with missing values imputed had smaller standard errors for the estimates. However, effect size estimates produced by MVMA with missing values deleted were closer to the ones produced by the UVMAs. The three meta-analyses results were close to each other. However, conducting three separate meta-analyses was more feasible compared to conducting an MVMA with deleted missing values and an MVMA with imputed missing values. MVMA appeared to be more practical to utilize when all data needed for the analysis are provided or could be computed.

In conclusion, the results of UVMA and the two MVMAs of CDSSs' effect on LDL did not agree. CDSSs' effects on LDL was found to be significant only in the UVMA result. The conclusion was consistent in the case of HbA1c. UVMA and MVMAs of HbA1c yielded a small effect of CDSSs on HbA1c levels. And, UVMA and MVMA with missing data deleted agreed in which both concluded that CDSSs had a nonsignificant effect in reducing levels of PP. The two MVMAs resulted in different conclusions. MVMA with missing values imputed indicated that CDSSs had a significant (less than small) effect on PP while MVMA with deleted missing values resulted in nonsignificant effect of CDSSs on PP levels. This difference between the two MVMAs

could be a result of extra information provided by imputing the missing values rather than just deleting them. It can also be concluded that this result might change based on what method was used--UVMA or MVMA (Jackson et al., 2011).

Overall Summary of the Findings

Looking back at the results of the simulation and the empirical meta-analyses together, a number of conclusions can be drawn. First, the between-study correlation appeared to have no effect on MVMA results. However, further investigation is need since including the between-study correlation should have some effect on the outcome estimates. Based on that null finding, the differences between the UVMAs and the MVMAs estimates were a result of the other factors (sample size, percent of missingness). Second, the simulation study and the meta-analyses of CDSSs showed that MVMAs could give different conclusions compared to UVMAs as they include more studies and more information about the outcomes. Third, it is important to state that effect sizes did not differ to any interpretable extent across all of the analyses conducted, but standard errors differed between UVMA and MVMA. As was found and as was expected, standard errors decreased as sample size increased, and so, standard errors decreased with no missing data and increased with a higher proportion of missing data. Fourth, CDSSs showed a small effect overall on the dependent measures combined. Finally, based on the findings taken together, MVMA is considered more precise since standard errors were smaller.

Limitations of the Study

This study had limitations on two parts (simulation study and the empirical study of CDSSs). In the simulation study, two outcomes were considered in order to make the

process easier to explain. However, if three outcomes were included, the comparisons of the findings of the simulation and the empirical meta-analyses of CDSSs would have more strength and extend the logical conclusions. Another limitation was when missing values were generated. In the simulation, the missing values were only placed on the second outcome (ES = 0.5). The reason behind that was to see if the missingness would only affect ES = 0.5 estimates or it would also affect ES = 0.7 estimates. However, another possibility would be to assign missing values to all outcomes since it is usually the case in real meta-analysis data. A final limitation is related to individual patient data. In the simulation, it would be more accurate if individual patient level data were generated and then meta-analyses were calculated.

Several limitations were also present in UVMAs and MVMAs for the CDSSs. In the study search procedure, not all studies that were evaluating CDSSs on diabetes care management were accessible. In studying the effect of CDSSs, there was no focus on a specific type of CDSS. CDSSs differ in their effect according to their type and mechanism. As a result, the effect sizes of the studies were varying and inconsistent and hence heterogeneity was significant in most of the conducted meta-analyses. Nevertheless, limiting the interest to only one type of CDSSs would severely decrease the number of the studies included in a meta-analysis.

A critical issue that is more likely than not to be present in every MVMA is the lack of reporting of the information needed for the analysis. In MVMA of CDSSs, no within-study correlation was reported in the included studies. As a result, the researcher computed the correlation between the three outcomes based on the literature. This

correlation was then used to impute the covariances in order to run the analysis. Having the actual correlations would result in more accurate findings.

Another common issue in MVMA is missing values. Two MVMAs were performed—one with missing data deleted and one with missing data imputed. Deleting the missing values was an option to get "rma.mv" to run since it does not deal with missingingness. The other option was by imputing the missing values using multiple imputation (MICE). This method assumes that missing values are missing at random which is not always the case. If the possibility of being ignored directly relates with the value of the data, then it is defined as missing not at random. For instance, some studies might report the variances because they are large. Also, smaller studies, are more likely ignore reporting variances compared to larger studies (Idris, 2011). Using multiple imputations increased the standard errors of effect size estimates as can be seen in Table 34 (Jakobsen et al., 2017). Imputing the missing effect sizes of CDSSs relied on the available information of the other studies. This could be misleading since not all studies have the same CDSS and not all CDSSs had the same effect.

Recommendations for Future Study

Several recommendations for future study that could help in furthering our understanding of UVMA and MVMA are provided below.

1- The simulation could be conducted with three or more outcomes with missing values placed randomly on all of them. This could give more insight into conclusions when comparing a simulation with an empirical MVMA.

2- The number of the replications used in the simulation was 1000 replications. Fewer replications could yield much the same accuracy and could be

examined with respect to making the simulation easier to conduct. It is recommended that future study examine use of 100 and 500 replications.

3- Considering the individual patient data when generating meta-analysis data would allow the researcher to adequately calculate within-study and between-study correlations.

4- Even though the consideration of the between-study correlation showed had no effect on the effect size estimates, further investigation using different effect sizes (very small or very large) might reveal some effect of the correlation on the effect size estimates.

5- Meta-analysis could be conducted to evaluate one type of CDSS or at least evaluate CDSSs sharing the same characteristics.

6- Due to the intent of this study, the effects of moderator variables were not examined. Conducting this study with the consideration of moderators that might affect the variation between the selected studies could be beneficial.

7- Researchers are encouraged to report vital information that could help in doing a complete meta-analysis and avoid deleting or imputing values that are missing. Alternatively, researchers could provide both solutions—with and without imputed values.

8- R Packages that are designed to conduct MVMA need to be developed to handle missing values and efficiently perform publication bias evaluation.

As an assistant lecturer at the University of Benghazi, faculty of public health/department of health informatics, one of my interests was to see how CDSSs would help diabetes physicians and patients in managing diabetes. After I took a meta-

analysis course, I thought that conducting meta-analysis would give a more precise conclusion about how effective CDSSs are. Since evaluating CDSSs was based on three outcomes (LDL, HbA1c, and BP), MVMA was the only efficient way to do such a meta-analysis. However, the use of the MVMA was limited in the literature and the effect of some factors on MVMA estimates were unknown. Therefore, and after consulting my professors, I decided to conduct simulation study first to and then conduct the study of CDSSs. This study was important to the field as relatively little is known about factors influencing MVMA and to me personally because it would help decision makers and health officials in Libya in making decisions on whether or not CDSSs would help improving health care.

References

Ahn, E., & Kang, H. (2018). Introduction to systematic review and meta-analysis. *Korean Journal of Anesthesiology*, *71*(2), 103–112. https://doi.org/10.4097/kjae.2018.71.2.103

Ali, S. M., Giordano, R., Lakhani, S., & Walker, D. M. (2016). A review of randomized controlled trials of medical record powered clinical decision support system to improve quality of diabetes care. *International Journal of Medical Informatics*, *87*, 91–100. https://doi.org/10.1016/j.ijmedinf.2015.12.017

Aloe, A. M., Amo, L. C., & Shanahan, M. E. (2014). Classroom management self-efficacy and burnout: A multivariate meta-analysis. *Educational Psychology Review*, *26*(1), 101–126. https://doi.org/10.1007/s10648-013-9244-0

Aloe, A. M., Shisler, S. M., Norris, B. D., Nickerson, A. B., & Rinker, T. W. (2014). A multivariate meta-analysis of student misbehavior and teacher burnout. *Educational Research Review*, *12*, 30–44. https://doi.org/10.1016/j.edurev.2014.05.003

Aronson, D., & Edelman, E. R. (2014). Coronary artery disease and diabetes mellitus. *Cardiology Clinics*, *32*(3), 439–455. https://doi.org/10.1016/j.ccl.2014.04.001

Bakris, G. L. (2001). A practical approach to achieving recommended blood pressure goals in diabetic patients. *Archives of Internal Medicine*, *161*(22), 2661–2667. https://doi.org/10.1001/archinte.161.22.2661

Becker, B. J. (2005). Failsafe N or file-drawer number. *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*, 111–125.

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, 1088–1101.

Belias, M., Rovers, M. M., Reitsma, J. B., Debray, T. P., & IntHout, J. (2019). Statistical approaches to identify subgroups in meta-analysis of individual participant data: A simulation study. *BMC Medical Research Methodology*, *19*(1), 183.

Boca, S. M., Pfeiffer, R. M., & Sampson, J. N. (2017). Multivariate meta-analysis with an increasing number of parameters. *Biometrische Zeitschrift*, *59*(3), 496–510. https://doi.org/10.1002/bimj.201600013

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. John Wiley & Sons.

Bray, J. H., Maxwell, S. E., & Maxwell, S. E. (1985). *Multivariate analysis of variance*. Sage.

Brown, S. A., Upchurch, S. L., & Acton, G. J. (2003). A framework for developing a coding scheme for meta-aAnalysis. *Western Journal of Nursing Research*, *25*(2), 205–222. https://doi.org/10.1177/0193945902250038

Brown, S. A., Becker, B. J., García, A. A., Brown, A., & Ramírez, G. (2015). Model-driven meta-analyses for informing health care: A diabetes meta-analysis as an exemplar. *Western Journal of Nursing Research*, *37*(4), 517–535. https://doi.org/10.1177/0193945914548229

Buuren, S. van, & Groothuis-Oudshoorn, K. (2010). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 1–68.

Card, N. A. (2015). *Applied meta-analysis for social science research*. Guilford Publications.

Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). Correcting for bias

in psychology: A comparison of meta-analytic methods. *Advances in Methods and*

Practices in Psychological Science, *2*(2), 115–144.

CDC. (2020, January 31). *LDL and HDL Cholesterol: "Bad" and "Good" Cholesterol*.

Centers for Disease Control and Prevention.

https://www.cdc.gov/cholesterol/ldl_hdl.htm

Centers for Disease Control and Prevention (CDC). (2000). Levels of diabetes-related

preventive-care practices—United States, 1997-1999. *MMWR. Morbidity and*

Mortality Weekly Report, *49*(42), 954–958.

Chen, D.-G. D., & Peace, K. E. (2013). *Applied meta-analysis with R*. Chapman and

Hall/CRC.

Chen, Y., Cai, Y., Hong, C., & Jackson, D. (2016). Inference for correlated effect sizes

using multiple univariate meta-analyses. *Statistics in Medicine*, *35*(9), 1405–1422.

https://doi.org/10.1002/sim.6789

Cheung, M. W.-L. (2013). Multivariate meta-analysis as structural equation models.

Structural Equation Modeling: A Multidisciplinary Journal, *20*(3), 429–454.

Crosson, J. C., Ohman-Strickland, P. A., Cohen, D. J., Clark, E. C., & Crabtree, B. F.

(2012). Typical electronic health record use in primary care practices and the

quality of diabetes care. *The Annals of Family Medicine*, *10*(3), 221–227.

https://doi.org/10.1370/afm.1370

Del Re, A. C., Maisel, N., Blodgett, J., & Finney, J. (2013). The declining efficacy of

naltrexone pharmacotherapy for alcohol use disorders over time: A multivariate

meta-analysis. *Alcoholism: Clinical and Experimental Research*, *37*(6), 1064–1068. https://doi.org/10.1111/acer.12067

Delgado-Rodríguez, M. (2001). Glossary on meta-analysis. *Journal of Epidemiology & Community Health*, *55*(8), 534–536. https://doi.org/10.1136/jech.55.8.534

Dickersin, K. (2005). Publication bias: Recognizing the problem, understanding its origins and scope, and preventing harm. *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*, 11–33.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634.

Ellington, E. H., Bastille-Rousseau, G., Austin, C., Landolt, K. N., Pond, B. A., Rees, E. E., … Murray, D. L. (2015). Using multiple imputation to estimate missing data in meta-regression. *Methods in Ecology and Evolution*, *6*(2), 153–163. https://doi.org/10.1111/2041-210X.12322

Ferrer, R. L. (1998). Graphical methods for detecting bias in meta-analysis. *Family Medicine, 30*, 579–583.

Field, A. P., & Miles, J. (2010). *Discovering statistics using SAS: And sex and drugs and rock "n" roll*. SAGE.

French, A., Macedo, M., Poulsen, J., Waterson, T., & Yu, A. (2008, June 06). *Multivariate analysis of variance (MANOVA)*. San Francisco State University. http://userwww.sfsu.edu/efc/classes/biol710/manova/manovanewest.htm

Garg, A. X., Adhikari, N. K., McDonald, H., Rosas-Arellano, M. P., Devereaux, P. J., Beyene, J., … Haynes, R. B. (2005). Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: A systematic review. *Journal of the American Medical Association*, *293*(10), 1223–1238.

Gill, J., Kucharski, K., Turk, B., Pan, C., & Wei, W. (2019). Using electronic clinical decision support in patient-centered medical homes to improve management of diabetes in primary care: The DECIDE study. *Journal of Ambulatory Care Management*, *42*(2), 105–115. https://doi.org/10.1097/JAC.0000000000000267

Gøtzsche, P. C., & Johansen, H. K. (1997). Meta-analysis of prophylactic or empirical antifungal treatment versus placebo or no treatment in patients with cancer complicated by neutropenia. *BMJ*, *314*(7089), 1238.

Haidich, A. B. (2010). Meta-analysis in medical research. *Hippokratia, 14*(Suppl 1), 29-37.

Ho, P. M., Rumsfeld, J. S., Masoudi, F. A., McClure, D. L., Plomondon, M. E., Steiner, J. F., & Magid, D. J. (2006). Effect of medication nonadherence on hospitalization and mortality among patients with diabetes mellitus. *Archives of Internal Medicine*, *166*(17), 1836–1841. https://doi.org/10.1001/archinte.166.17.1836

Holbrook, A., Thabane, L., Keshavjee, K., Dolovich, L., Bernstein, B., Chan, D., … Gerstein, H. (2009). Individualized electronic decision support and reminders to improve diabetes care in the community: COMPETE II randomized trial. *CMAJ : Canadian Medical Association Journal = Journal de l'Association Medicale Canadienne*, *181*(1–2), 37–44. https://doi.org/10.1503/cmaj.081272

Hu, H., Hori, A., Nishiura, C., Sasaki, N., Okazaki, H., Nakagawa, T., … Dohi, S. (2016a). Hba1c, Blood pressure, and lipid control in people with diabetes: Japan epidemiology collaboration on occupational health study. *PLoS ONE*, *11*(7). https://doi.org/10.1371/journal.pone.0159071

Huang, E. P., Wang, X.-F., Choudhury, K. R., McShane, L. M., Gönen, M., Ye, J., … Jackson, E. F. (2015). Meta-analysis of the technical performance of an imaging procedure: Guidelines and statistical methodology. *Statistical Methods in Medical Research*, *24*(1), 141–174.

Idris, N. R. N. (2011). Estimating the bias in meta analysis estimates for continuous data with non-random missing study variance. *MATEMATIKA: Malaysian Journal of Industrial and Applied Mathematics*, 27, 121–128.

Ishak, K. J., Platt, R. W., Joseph, L., & Hanley, J. A. (2008). Impact of approximating or ignoring within-study covariances in multivariate meta-analyses. *Statistics in Medicine*, *27*(5), 670–686. https://doi.org/10.1002/sim.2913

Jackson, D., Riley, R., & White, I. R. (2011). Multivariate meta-analysis: Potential and promise. *Statistics in Medicine*, *30*(20), 2481–2498. https://doi.org/10.1002/sim.4172

Jackson, D., & Turner, R. (2017). Power analysis for random-effects meta-analysis. *Research Synthesis Methods*, *8*(3), 290–302.

Jagannathan, R., Sevick, M. A., Fink, D., Dankner, R., Chetrit, A., Roth, J., … Bergman, M. (2016). The 1-hour post-load glucose level is more effective than HbA1c for screening dysglycemia. *Acta Diabetologica*, *53*(4), 543–550. https://doi.org/10.1007/s00592-015-0829-6

Jakobsen, J. C., Gluud, C., Wetterslev, J., & Winkel, P. (2017). When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Medical Research Methodology*, *17*(1), 162. https://doi.org/10.1186/s12874-017-0442-1

Jeffery, R., Iserman, E., Haynes, R. B., & CDSS Systematic Review Team. (2013). Can computerized clinical decision support systems improve diabetes management? A systematic review and meta-analysis. *Diabetic Medicine: A Journal of the British Diabetic Association*, *30*(6), 739–745. https://doi.org/10.1111/dme.12087

Jia, P., Zhao, P., Chen, J., & Zhang, M. (2019a). Evaluation of clinical decision support systems for diabetes care: An overview of current evidence. *Journal of Evaluation in Clinical Practice*, *25*(1), 66–77. https://doi.org/10.1111/jep.12968

Jia, P., Zhao, P., Chen, J., & Zhang, M. (2019b). Evaluation of clinical decision support systems for diabetes care: An overview of current evidence. *Journal of Evaluation in Clinical Practice*, *25*(1), 66–77. https://doi.org/10.1111/jep.12968

Jüni, P., Holenstein, F., Sterne, J., Bartlett, C., & Egger, M. (2002). Direction and impact of language bias in meta-analyses of controlled trials: Empirical study. *International Journal of Epidemiology*, *31*(1), 115–123.

Kawamoto, K., Houlihan, C. A., Balas, E. A., & Lobach, D. F. (2005). Improving clinical practice using clinical decision support systems: A systematic review of trials to identify features critical to success. *BMJ (Clinical Research Ed.)*, *330*(7494), 765. https://doi.org/10.1136/bmj.38398.500764.8F

Kittler, J. E., Menard, W., & Phillips, K. A. (2007). Weight concerns in individuals with body dysmorphic disorder. *Eating Behaviors*, *8*(1), 115–120. https://doi.org/10.1016/j.eatbeh.2006.02.006

Light, R., & Pillemer, D. (1984). Summing up: The science of reviewing research. *MA: Harvard University Press.*

Lin, L., & Chu, H. (2018). Bayesian multivariate meta-analysis of multiple factors. *Research Synthesis Methods*, *9*(2), 261–272. https://doi.org/10.1002/jrsm.1293

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* SAGE publications, Inc.

Lloyd, S., Bujkiewicz, S., Wailoo, A. J., Sutton, A. J., & Scott, D. (2010). The effectiveness of anti-TNF-α therapies when used sequentially in rheumatoid arthritis patients: A systematic review and meta-analysis. *Rheumatology (Oxford, England)*, *49*(12), 2313–2321. https://doi.org/10.1093/rheumatology/keq169

Mallet, S., Hopewell, S., & Clarke, M. (2002). The use of grey literature in the first 1000 Cochrane reviews. *4th Symposium on Systematic Reviews: Pushing the Boundaries. Oxford, United Kingdom*.

Mavridis, D., & Salanti, G. (2013). A practical introduction to multivariate meta-analysis. *Statistical Methods in Medical Research*, *22*(2), 133–158. https://doi.org/10.1177/0962280211432219

Mittlböck, M., & Heinzl, H. (2006). A simulation study comparing properties of heterogeneity measures in meta-analyses. *Statistics in Medicine*, *25*(24), 4321–4333.

Morrow, D. G., Conner-Garcia, T., Graumlich, J. F., Wolf, M. S., McKeever, S., Madison, A., … Kaiser, D. (2012). An EMR-Based Tool to Support Collaborative Planning

for Medication use among adults with diabetes: Design of a multi-site randomized

control trial. *Contemporary Clinical Trials*, *33*(5), 1023–1032.

https://doi.org/10.1016/j.cct.2012.05.010

Neupane, B., & Beyene, J. (2015). Multivariate Meta-Analysis of Genetic Association

Studies: A Simulation Study. *PLoS ONE*, *10*(7), e0133243.

https://doi.org/10.1371/journal.pone.0133243

O'Connor, P. J., Sperl-Hillen, J. M., Rush, W. A., Johnson, P. E., Amundson, G. H.,

Asche, S. E., … Gilmer, T. P. (2011). Impact of electronic health record clinical

decision support on diabetes care: A randomized trial. *The Annals of Family

Medicine*, *9*(1), 12–21. https://doi.org/10.1370/afm.1196

O'Connor, Patrick J., & Sperl-Hillen, J. M. (2019). Current status and future directions

for electronic point-of-care clinical decision support to improve diabetes

management in primary care. *Diabetes Technology & Therapeutics*, *21*(S2), S2-

26-S2-34. https://doi.org/10.1089/dia.2019.0070

Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of

Educational Statistics*, *8*(2), 157–159.

Paxton, P., Curran, P. J., Bollen, K. A., Kirby, J., & Chen, F. (2001). Monte Carlo

experiments: Design and implementation. *Structural Equation Modeling*, *8*(2),

287–312.

Prabhakaran, D., Jha, D., Prieto-Merino, D., Roy, A., Singh, K., Ajay, V. S., … Sharma,

M. (2019). Effectiveness of an mHealth-Based electronic decision support system

for integrated management of chronic conditions in primary care: The mWellcare

cluster-randomized controlled trial. *Circulation*, *139*(3), 380–391.

https://doi.org/10.1161/CIRCULATIONAHA.118.038192

Price, M. J., Blake, H. A., Kenyon, S., White, I. R., Jackson, D., Kirkham, J. J., … Riley, R. D. (2019). Empirical comparison of univariate and multivariate meta-analyses in Cochrane Pregnancy and Childbirth reviews with multiple binary outcomes. *Research Synthesis Methods*, *10*(3), 440–451. https://doi.org/10.1002/jrsm.1353

Ravnskov, U. (1992). Cholesterol lowering trials in coronary heart disease: Frequency of citation and outcome. *British Medical Journal*, *305*(6844), 15–19.

Riley, R. D. (2009). Multivariate meta-analysis: The effect of ignoring within-study correlation. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, *172*(4), 789–811. JSTOR.

Rosenberg, L. H., Franzén, B., Auer, G., Lehtiö, J., & Forshed, J. (2010). Multivariate meta-analysis of proteomics data from human prostate and colon tumours. *BMC Bioinformatics*, *11*, 468. https://doi.org/10.1186/1471-2105-11-468

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Schnipper, J., Liang, C., Ndumele, C., & Pendergrass, M. (2010). Effects of a computerized order set on the inpatient management of hyperglycemia: A cluster-randomized controlled trial. *Endocrine Practice*, *16*(2), 209–218.

Schwarzer, G., Carpenter, J. R., & Rücker, G. (2015). *Meta-analysis with R*. Springer.

Seaman, D. E., Millspaugh, J. J., Kernohan, B. J., Brundige, G. C., Raedeke, K. J., & Gitzen, R. A. (1999). Effects of sample size on kernel home range estimates. *The*

Journal of Wildlife Management, *63*(2), 739–747.

https://doi.org/10.2307/3802664

Sequist, T. D., Gandhi, T. K., Karson, A. S., Fiskio, J. M., Bugbee, D., Sperling, M., …

Bates, D. W. (2005). A randomized trial of electronic clinical reminders to

improve quality of care for diabetes and coronary artery disease. *Journal of the*

American Medical Informatics Association, *12*(4), 431–437.

https://doi.org/10.1197/jamia.M1788

Sharif, M. O., Sharif, M. O., Ali, H. A., & Ahmed, F. (2013). *Explained: AMSTAR —*

How to tell the good from the bad and the ugly. Oral Health Dent Manag, 12(1),

9-16.

Shorten, A., & Shorten, B. (2013). What is meta-analysis? *Evidence Based Nursing*,

16(1), 3–4. https://doi.org/10.1136/eb-2012-101118

Sim, L. L. W., Ban, K. H. K., Tan, T. W., Sethi, S. K., & Loh, T. P. (2017). Development

of a clinical decision support system for diabetes care: A pilot study. *PloS One*,

12(2), e0173021. https://doi.org/10.1371/journal.pone.0173021

Song, F., Khan, K. S., Dinnes, J., & Sutton, A. J. (2002). Asymmetric funnel plots and

publication bias in meta-analyses of diagnostic accuracy. *International Journal of*

Epidemiology, *31*(1), 88–95.

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., …

Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel

plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*, *343*.

https://doi.org/10.1136/bmj.d4002

Tang, J.-L., & Liu, J. L. (2000). Misleading funnel plot for detection of bias in meta-analysis. *Journal of Clinical Epidemiology*, *53*(5), 477–484.

Tay, T. L., Foo, J. P., Tan, E., Chen, R., Khoo, J., Soh, S. B., … Cho, L. W. (2011). HbA1c may not be a sensitive determinant of diabetic status in the elderly. *Diabetes Research and Clinical Practice*, *92*(2), e31–e33. https://doi.org/10.1016/j.diabres.2011.01.003

Tramèr, M. R., Reynolds, D. J. M., Moore, R. A., & McQuay, H. J. (1997). Impact of covert duplicate publication on meta-analysis: A case study. *BMJ*, *315*(7109), 635–640.

Trikalinos, T. A., Hoaglin, D. C., & Schmid, C. H. (2013). *Empirical and simulation-based comparison of univariate and multivariate meta-analysis for binary outcomes*. Agency for Healthcare Research and Quality (US). http://www.ncbi.nlm.nih.gov/books/NBK132562/

Trikalinos, T. A., Hoaglin, D. C., & Schmid, C. H. (2014). An empirical comparison of univariate and multivariate meta-analyses for categorical outcomes. *Statistics in Medicine*, *33*(9), 1441–1459. https://doi.org/10.1002/sim.6044

Trikalinos, T. A., & Olkin, I. (2012). Meta-analysis of effect sizes reported at multiple time points: A multivariate approach. *Clinical Trials*, *9*(5), 610–620.

Tudur Smith, C., Hopkins, C., Sydes, M. R., Woolfall, K., Clarke, M., Murray, G., & Williamson, P. (2015). How should individual participant data (IPD) from publicly funded clinical trials be shared? *BMC Medicine*, *13*(1), 298. https://doi.org/10.1186/s12916-015-0532-z

Wei, Y., & Higgins, J. P. (2013). Estimating within-study covariances in multivariate

    meta-analysis with multiple outcomes. *Statistics in Medicine*, *32*(7), 1191–1205.

    https://doi.org/10.1002/sim.5679

Maia, J. X., de Sousa, L. A., Marcolino, M. S., Cardoso, C. S., da Silva, J. L., Alkmim,

    M. B., & Ribeiro, A. L. (2016). The impact of a clinical decision support system

    in diabetes primary care patients in a developing country. *Diabetes Technology &*

    Therapeutics, 18(4), 258–263. https://doi.org/10.1089/dia.2015.0253

Zhang, Y. Q., Yu, C. H., & Bao, J. Z. (2017). Impact of daily mean temperature, cold

    spells, and heat waves on stroke mortality a multivariable Meta-analysis from 12

    counties of Hubei province, China. *Zhonghua Liu Xing Bing Xue Za Zhi =*

    Zhonghua Liuxingbingxue Zazhi, *38*(4), 508–513.

    https://doi.org/10.3760/cma.j.issn.0254-6450.2017.04.019

Zhang, Z. (2016). Multiple imputation with multivariate imputation by chained equation

    (MICE) package. *Annals of Translational Medicine*, *4*(2).

    https://doi.org/10.3978/j.issn.2305-5839.2015.12.63

Appendices

Appendix A: 41 Studies Included in The Meta-Analyses

Abaza, H., & Marschollek, M. (2017). SMS education for the promotion of diabetes self-management in low & middle income countries: A pilot randomized controlled trial in Egypt. *BMC Public Health*, *17*(1), 962. https://doi.org/10.1186/s12889-017-4973-5

Adjei, D. N., Agyemang, C., Dasah, J. B., Kuranchie, P., & Amoah, A. G. B. (2015). The effect of electronic reminders on risk management among diabetic patients in low resourced settings. *Journal of Diabetes and Its Complications*, *29*(6), 818–821. https://doi.org/10.1016/j.jdiacomp.2015.05.008

Arora, S., Peters, A. L., Burner, E., Lam, C. N., & Menchine, M. (2014). Trial to examine text message–based mHealth in emergency department patients with diabetes (TExT-MED): A randomized controlled trial. *Annals of Emergency Medicine*, *63*(6), 745-754.e6. https://doi.org/10.1016/j.annemergmed.2013.10.012

Avdal, E. Ü., Kizilci, S., & Demirel, N. (2011). The effects of web-based diabetes education on diabetes care results: A randomized control Study. *CIN: Computers, Informatics, Nursing*, *29*(2), 101–106. https://doi.org/10.1097/NCN.0b013e3181fcbdc6

Bujnowska-Fedak, M. M., Puchała, E., & Steciwko, A. (2011). The impact of telehome care on health status and quality of life among patients with diabetes in a primary care setting in Poland. *Telemedicine and E-Health*, *17*(3), 153–163. https://doi.org/10.1089/tmj.2010.0113

Chao, J., Yang, L., Xu, H., Yu, Q., Jiang, L., & Zong, M. (2015). The effect of integrated health management model on the health of older adults with diabetes in a randomized controlled trial. *Archives of Gerontology and Geriatrics*, *60*(1), 82–88. https://doi.org/10.1016/j.archger.2014.10.006

Charpentier, G., Benhamou, P.-Y., Dardari, D., Clergeot, A., Franc, S., Schaepelynck-Belicar, P., … on behalf of the TeleDiab Study Group. (2011). The Diabeo software enabling individualized insulin dose adjustments combined with telemedicine support improves HbA1c in poorly controlled type 1 diabetic patients: A 6-month, randomized, open-label, parallel-group, multicenter trial (TeleDiab 1 Study). *Diabetes Care*, *34*(3), 533–539. https://doi.org/10.2337/dc10-1259

Chen, S.-Y., Chang, Y.-H., Hsu, H.-C., Lee, Y.-J., Hung, Y.-J., & Hsieh, C.-H. (2011). One-year efficacy and safety of the telehealth system in poorly controlled type 2 diabetic patients receiving insulin therapy. *Telemedicine and E-Health*, *17*(9), 683–687. https://doi.org/10.1089/tmj.2011.0020

Cho, J.-H., Kwon, H.-S., Kim, H.-S., Oh, J.-A., & Yoon, K.-H. (2011). Effects on diabetes management of a health-care provider mediated, remote coaching system via a PDA-type glucometer and the Internet. *Journal of Telemedicine and Telecare*, *17*(7), 365–370. https://doi.org/10.1258/jtt.2011.100913

Cleveringa, F. G. W., Minkman, M. H., Gorter, K. J., van den Donk, M., & Rutten, G. E. H. M. (2010). Diabetes care protocol: Effects on patient-important outcomes. A cluster randomized, non-inferiority trial in primary care. *Diabetic Medicine*, *27*(4), 442–450. https://doi.org/10.1111/j.1464-5491.2010.02968.x

Fang, R., & Deng, X. (2018). Electronic messaging intervention for management of

    cardiovascular risk factors in type 2 diabetes mellitus: A randomised controlled

    trial. *Journal of Clinical Nursing*, *27*(3–4), 612–620.

    https://doi.org/10.1111/jocn.13962

Goodarzi, M., Ebrahimzadeh, I., Rabi, A., Saedipoor, B., & Jafarabadi, M. A. (2012).

    Impact of distance education via mobile phone text messaging on knowledge,

    attitude, practice and self efficacy of patients with type 2 diabetes mellitus in Iran.

    Journal of Diabetes & Metabolic Disorders, *11*(1), 10.

    https://doi.org/10.1186/2251-6581-11-10

Graumlich, J. F., Wang, H., Madison, A., Wolf, M. S., Kaiser, D., Dahal, K., & Morrow,

    D. G. (2016). Effects of a patient-provider, collaborative, medication-planning

    tool: A randomized, controlled trial. *Journal of Diabetes Research*, *2016*, 1–16.

    https://doi.org/10.1155/2016/2129838

Heselmans, A., Delvaux, N., Laenen, A., Van de Velde, S., Ramaekers, D., Kunnamo, I.,

    & Aertgeerts, B. (2020). Computerized clinical decision support system for

    diabetes in primary care does not improve quality of care: A cluster-randomized

    controlled trial. *Implementation Science*, *15*(1), 5. https://doi.org/10.1186/s13012-

    019-0955-6

Holmen, H., Torbjørnsen, A., Wahl, A. K., Jenum, A. K., Småstuen, M. C., Årsand, E., &

    Ribu, L. (2014). A mobile health intervention for self-management and lifestyle

    change for persons with type 2 diabetes, part 2: One-year results from the

    Norwegian randomized controlled trial RENEWING HEALTH. *JMIR MHealth*

    and UHealth, *2*(4), e57. https://doi.org/10.2196/mhealth.3882

Hussein, W. I., Hasan, K., & Jaradat, A. A. (2011). Effectiveness of mobile phone short

    message service on diabetes mellitus management; the SMS-DM study. *Diabetes*

    Research and Clinical Practice, *94*(1), e24–e26.

    https://doi.org/10.1016/j.diabres.2011.07.025

Kardas, P., Lewandowski, K., & Bromuri, S. (2016). Type 2 diabetes patients benefit

    from the COMODITY12 mHealth system: Results of a randomised trial. *Journal*

    of Medical Systems, *40*(12), 259. https://doi.org/10.1007/s10916-016-0619-x

Kim, C. S., Park, S. Y., Kang, J. G., Lee, S. J., Ihm, S. H., Choi, M. G., & Yoo, H. J.

    (2010). Insulin dose titration system in diabetes patients using a short messaging

    service automatically produced by a knowledge matrix. *Diabetes Technology &*

    Therapeutics, *12*(8), 663–669. https://doi.org/10.1089/dia.2010.0031

Kirwan, M., Vandelanotte, C., Fenning, A., & Duncan, M. J. (2013). Diabetes self-

    management smartphone application for adults with Type 1 diabetes: Randomized

    rontrolled trial. *Journal of Medical Internet Research*, *15*(11), e235.

    https://doi.org/10.2196/jmir.2588

Lam, Y. Y. (2018). *The effectiveness of EHealth technology on improving self-*

    management of adult patients with type 2 diabetes (T2DM): Randomized

    controlled trial [Unpublished doctoral Thesis]. The Chinese University of Hong

    Kon.

Lee, J. Y., Wong, C. P., Tan, C. S. S., Nasir, N. H., & Lee, S. W. H. (2017).

    Telemonitoring in fasting individuals with Type 2 Diabetes Mellitus during

    Ramadan: A prospective, randomised controlled study. *Scientific Reports*, *7*(1),

    10119. https://doi.org/10.1038/s41598-017-10564-y

Lim, S., Kang, S. M., Shin, H., Lee, H. J., Won Yoon, J., Yu, S. H., … Jang, H. C. (2011). Improved glycemic control without hypoglycemia in elderly diabetic patients using the ubiquitous healthcare service, a new medical information system. *Diabetes Care*, *34*(2), 308–313. https://doi.org/10.2337/dc10-1447

Lim, Soo, Kang, S. M., Kim, K. M., Moon, J. H., Choi, S. H., Hwang, H., … Jang, H. C. (2016). Multifactorial intervention in diabetes care using real-time monitoring and tailored feedback in type 2 diabetes. *Acta Diabetologica*, *53*(2), 189–198. https://doi.org/10.1007/s00592-015-0754-8

Logan, A. G., Irvine, M. J., McIsaac, W. J., Tisler, A., Rossos, P. G., Easty, A., … Cafazzo, J. A. (2012). Effect of home blood pressure telemonitoring with self-care support on uncontrolled systolic hypertension in diabetics. *Hypertension*, *60*(1), 51–57. https://doi.org/10.1161/HYPERTENSIONAHA.111.188409

Noh, J.-H., Cho, Y.-J., Nam, H.-W., Kim, J.-H., Kim, D.-J., Yoo, H.-S., … Yoo, H.-J. (2010). Web-based comprehensive information system for self-Management of diabetes mellitus. *Diabetes Technology & Therapeutics*, *12*(5), 333–337. https://doi.org/10.1089/dia.2009.0122

O'Connor, P. J., Sperl-Hillen, J. M., Rush, W. A., Johnson, P. E., Amundson, G. H., Asche, S. E., … Gilmer, T. P. (2011). Impact of electronic health record clinical decision support on diabetes care: A randomized trial. *The Annals of Family Medicine*, *9*(1), 12–21. https://doi.org/10.1370/afm.1196

Prabhakaran, D., Jha, D., Prieto-Merino, D., Roy, A., Singh, K., Ajay, V. S., Jindal, D., … Sharma, M. (2019). Effectiveness of an mHealth-based electronic decision support system for integrated management of chronic conditions in primary care:

The mWellcare cluster-randomized controlled trial. *Circulation*, *139*(3), 380–391.

https://doi.org/10.1161/CIRCULATIONAHA.118.038192

Quinn, C. C., Shardell, M. D., Terrin, M. L., Barr, E. A., Ballew, S. H., & Gruber-Baldini,

A. L. (2011). Cluster-randomized trial of a mobile phone personalized behavioral

intervention for blood glucose control. *Diabetes Care*, *34*(9), 1934–1942.

https://doi.org/10.2337/dc11-0366

Quinn, C. C., Shardell, M. D., Terrin, M. L., Barr, E. A., Park, D., Shaikh, F., Guralnik, J.

M., & Gruber-Baldini, A. L. (2016). Mobile diabetes intervention for glycemic

control in 45- to 64-year-old persons with type 2 diabetes. *Journal of Applied*

Gerontology, *35*(2), 227–243. https://doi.org/10.1177/0733464814542611

Rossi, M. C. E., Nicolucci, A., Di Bartolo, P., Bruttomesso, D., Girelli, A., Ampudia, F.

J., … Vespasiani, G. (2010). Diabetes interactive diary: a new telemedicine

system enabling flexible diet and insulin therapy while improving quality of life:

An open-label, international, multicenter, randomized study. *Diabetes Care*,

33(1), 109–115. https://doi.org/10.2337/dc09-1327

Rossi, M. C., Nicolucci, A., Lucisano, G., Pellegrini, F., Di Bartolo, P., Miselli, V.,

Anichini, R., & Vespasiani, on behalf of the DID St, G. (2013). Impact of the

"Diabetes Interactive Diary" telemedicine system on metabolic control, risk of

hypoglycemia, and quality of life: a randomized clinical trial in type 1 diabetes.

Diabetes Technology & Therapeutics, *15*(8), 670–679.

https://doi.org/10.1089/dia.2013.0021

Shahid, M., Mahar, S. A., Shaikh, S., & Shaikh, Z. (2015). Mobile phone intervention to improve diabetes care in rural areas of Pakistan: A randomized controlled trial. *J Coll Physicians Surg Pak,* 25(3), 166-171.

Steventon, A., Bardsley, M., Doll, H., Tuckey, E., & Newman, S. P. (2014). Effect of telehealth on glycaemic control: Analysis of patients with type 2 diabetes in the Whole Systems Demonstrator cluster randomised trial. *BMC Health Services Research*, *14*(1), 334. https://doi.org/10.1186/1472-6963-14-334

Stone, R. A., Rao, R. H., Sevick, M. A., Cheng, C., Hough, L. J., Macpherson, D. S., … DeRubertis, F. R. (2010). Active care management supported by home telemonitoring in veterans with type 2 diabetes: The DiaTel randomized controlled trial. *Diabetes Care*, *33*(3), 478–484. https://doi.org/10.2337/dc09-1012

Tang, P. C., Overhage, J. M., Chan, A. S., Brown, N. L., Aghighi, B., Entwistle, M. P., … Young, C. Y. (2013). Online disease management of diabetes: Engaging and motivating patients online with enhanced resources-diabetes (EMPOWER-D), a randomized controlled trial. *Journal of the American Medical Informatics Association*, *20*(3), 526–534. https://doi.org/10.1136/amiajnl-2012-001263

Tildesley, H. D., Mazanderani, A. B., & Ross, S. A. (2010). Effect of internet therapeutic intervention on a1c levels in patients with type 2 diabetes treated with insulin. *Diabetes Care*, *33*(8), 1738–1740. https://doi.org/10.2337/dc09-2256

Waki, K., Fujita, H., Uchimura, Y., Omae, K., Aramaki, E., Kato, S., … Ohe, K. (2014). DialBetics: A novel smartphone-based self-management support system for type 2

diabetes patients. *Journal of Diabetes Science and Technology*, *8*(2), 209–215. https://doi.org/10.1177/1932296814526495

Welch, G., Zagarins, S. E., Santiago-Kelly, P., Rodriguez, Z., Bursell, S.-E., Rosal, M. C., & Gabbay, R. A. (2015). An internet-based diabetes management platform improves team care and outcomes in an urban Latino population. *Diabetes Care*, 38(4), 561-567. https://doi.org/10.2337/dc14-1412

Yang, Y., Lee, E. Y., Kim, H.-S., Lee, S.-H., Yoon, K.-H., & Cho, J.-H. (2020). Effect of a mobile phone–based glucose-monitoring and feedback system for type 2 diabetes management in multiple primary care clinic settings: Cluster randomized controlled trial. *JMIR MHealth and UHealth*, *8*(2), e16266. https://doi.org/10.2196/16266

Young, L. A., Buse, J. B., Weaver, M. A., Vu, M. B., Mitchell, C. M., Blakeney, T., … Monitor Trial Group. (2017). Glucose self-monitoring in non-insulin-treated patients with type 2 diabetes in primary care settings: A randomized trial. *JAMA Internal Medicine*, *177*(7), 920–929. https://doi.org/10.1001/jamainternmed.2017.1233

Zhou, W., Chen, M., Yuan, J., & Sun, Y. (2016). Welltang – A smart phone-based diabetes management application – Improves blood glucose control in Chinese people with diabetes. *Diabetes Research and Clinical Practice*, *116*, 105–110. https://doi.org/10.1016/j.diabres.2016.03.018

Appendix B: IRB Determination Letter

UNIVERSITY of DENVER
OFFICE OF RESEARCH &
SPONSORED PROGRAMS
Research Integrity & Education

March 15, 2021
Abdelfattah Elbarsha
Research Methods and Information Science
Morgridge College of Education
University of Denver

RE: Determination of Proposed Project
Project Title: **Evaluation of the Effect of the Clinical Decision Support Systems on Diabetes Management: A Multivariate Meta-Analysis Comparison with Univariate Meta-Analysis**

Dear Abdelfattah,

Thank you for submitting the IRB Determination Form, to the University of Denver Institutional Review Board for evaluation to determine if the above-referenced project qualifies as human subject research. Based on the information provided, it has been determined that the proposed project does not require IRB review. This determination is based on whether this proposed project is research with human subjects defined by the federal regulations.

The IRB Determination Form was evaluated and it was assessed that the proposed metaanalysis project does not qualify as human subjects research. The planned meta-analysis of publicly available data from published articles does not meet the regulatory definition of research with human subjects.

The Regulatory Definition of Research and Human Subject
Federal research regulations define **research** as "*a systematic investigation, including research development, testing, and evaluation, designed to develop or contribute to generalizable knowledge.*"

During the review of this proposed project, it was noted that the primary intent is to determine how and when estimated produced by multivariate meta-analysis would be different and more statistically precise compared to estimates produced by separate univariate meta-analyses, and to examine the effects of clinical decision support systems on the quality of diabetes care management. This study will analyze de-identified and publicly available datasets found from published articles. This project is designed to develop generalizable knowledge, and therefore this project does qualify as research.

130

Per the regulations, **Human subject** means a living individual about whom an investigator (whether professional or student) conducting research obtains 1) data through intervention or interaction with the individual, or 2) identifiable private information. This project does not involve interaction with living individuals and will utilize de-identified secondary data only, therefore it does not qualify as involving human subjects.

In order for a project to require IRB review, the proposed research must qualify under **both** definitions of being research and involving human subjects. This research project does fulfill the regulatory definition of research but does NOT involve human subjects per the federal regulation definition.

My evaluation, based only on the information provided, determined that the proposed project does not require IRB review.

If you have questions regarding this determination or believe that this proposed project does qualify as human subject research, please feel free to contact me directly at 303-871-4051 or via e-mail at: Ashleigh.Ruehrdanz@du.edu.

Sincerely,

Ashleigh Ruehrdanz
Research Compliance Monitor
Office of Research and Sponsored Programs
University of Denver