ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΒΙΟΙΑΤΡΙΚΗ

Λογισμικό μέτα-ανάλυσης στο STATA

Δεμετσιώτης Κυριάκος

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ
Υπεύθυνος
Παντελεήμων Γ. Μπάγκος
Καθηγητής

Λαμία, 2024

ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΒΙΟΙΑΤΡΙΚΗ

Λογισμικό μέτα-ανάλυσης στο STATA

Δεμετσιώτης Κυριάκος

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Επιβλέπων
Παντελεήμων Γ. Μπάγκος
Καθηγητής

Λαμία, 2024

Με ατομική μου ευθύνη και γνωρίζοντας τις κυρώσεις [1], που προβλέπονται από της διατάξεις της παρ. 6 του άρθρου 22 του Ν. 1599/1986, δηλώνω ότι:

1. *Δεν παραθέτω κομμάτια βιβλίων ή άρθρων ή εργασιών άλλων αυτολεξεί* ***χωρίς να τα περικλείω σε εισαγωγικά*** *και χωρίς να αναφέρω το συγγραφέα, τη χρονολογία, τη σελίδα. Η αυτολεξεί παράθεση χωρίς εισαγωγικά χωρίς αναφορά στην πηγή, είναι λογοκλοπή. Πέραν της αυτολεξεί παράθεσης, λογοκλοπή θεωρείται και η παράφραση εδαφίων από έργα άλλων, συμπεριλαμβανομένων και έργων συμφοιτητών μου, καθώς και η παράθεση στοιχείων που άλλοι συνέλεξαν ή επεξεργάσθηκαν, χωρίς αναφορά στην πηγή. Αναφέρω πάντοτε με πληρότητα την πηγή κάτω από τον πίνακα ή σχέδιο, όπως στα παραθέματα.*

2. *Δέχομαι ότι η αυτολεξεί* ***παράθεση χωρίς εισαγωγικά****, ακόμα κι αν συνοδεύεται από αναφορά στην πηγή σε κάποιο άλλο σημείο του κειμένου ή στο τέλος του, είναι αντιγραφή. Η αναφορά στην πηγή στο τέλος π.χ. μιας παραγράφου ή μιας σελίδας, δεν δικαιολογεί συρραφή εδαφίων έργου άλλου συγγραφέα, έστω και παραφρασμένων, και παρουσίασή τους ως δική μου εργασία.*

3. *Δέχομαι ότι υπάρχει επίσης περιορισμός στο μέγεθος και στη συχνότητα των παραθεμάτων που μπορώ να εντάξω στην εργασία μου εντός εισαγωγικών. Κάθε μεγάλο παράθεμα (π.χ. σε πίνακα ή πλαίσιο, κλπ), προϋποθέτει ειδικές ρυθμίσεις, και όταν δημοσιεύεται προϋποθέτει την άδεια του συγγραφέα ή του εκδότη. Το ίδιο και οι πίνακες και τα σχέδια*

4. *Δέχομαι όλες τις συνέπειες σε περίπτωση λογοκλοπής ή αντιγραφής.*

Ημερομηνία: 16/06/2024

Ο Δηλών

Δεμετσιώτης Κυριάκος

(Υπογραφή)

(1) «Όποιος εν γνώσει του δηλώνει ψευδή γεγονότα ή αρνείται ή αποκρύπτει τα αληθινά με έγγραφη υπεύθυνη δήλωση του άρθρου 8 παρ. 4 Ν. 1599/1986 τιμωρείται με φυλάκιση τουλάχιστον τριών μηνών. Εάν ο υπαίτιος αυτών των πράξεων σκόπευε να προσπορίσει στον εαυτόν του ή σε άλλον περιουσιακό όφελος βλάπτοντας τρίτον ή σκόπευε να βλάψει άλλον, τιμωρείται με κάθειρξη μέχρι 10 ετών.

Λογισμικό μέτα-ανάλυσης στο STATA

Δεμετσιώτης Κυριάκος

Τριμελής Επιτροπή:

Παντελεήμων Γ. Μπάγκος, Καθηγητής

Γεωργία Μπράλιου, Επίκουρη καθηγήτρια

Αθανάσιος Σαχλάς, Επίκουρος καθηγητής

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον κ. **Μπάγκο Παντελεήμονα**, Καθηγητή στο τμήμα Πληροφορικής με Εφαρμογές στη Βιοιατρική, για την εμπιστοσύνη που μου έδειξε κατά την εκπόνηση της συγκεκριμένης εργασίας, τις πολύτιμες συμβουλές του και την άρτια καθοδήγηση του.

Ιδιαίτερα επιθυμώ να ευχαριστήσω την Καλλιόπη Εξάρχου-Κουβέλη, υποψήφια διδάκτορα στο Τμήμα Πληροφορικής με Εφαρμογές στη Βιοιατρική, για την υπομονή που επέδειξε, τις ωφέλιμες γνώσεις της και την ενθάρρυνση της, ώστε να βγει το καλύτερο δυνατό αποτέλεσμα.

Επιπλέον, ένα μεγάλο ευχαριστώ στους γονείς μου για την συμπαράσταση που δείχνουν σε όλα τα στάδια της ζωής μου και για την βοήθεια τους στο να πραγματοποιήσω κάθε μου όνειρο.

Τέλος, οφείλω να ευχαριστήσω την αδερφή μου, που με στηρίζει στις αποφάσεις μου και στην οποία θα είμαι πάντα δίπλα όπου και όποτε χρειαστεί.

"If you are feeling disheartened, that you are somehow not enough, set your heart ablaze."

-Kyojuro Rengoku, from the Demon Slayer Anime

Περιεχόμενα

Ευχαριστίες iv
Περιεχόμενα v
Κατάλογος πινάκων vi
Κατάλογος σχημάτων vii
Περίληψη viii
Abstract ix
1. Εισαγωγή 1
1.1 Εισαγωγικοί όροι 1
1.2 Εισαγωγή στην μετά-ανάλυση 8
1.3 Μοντέλα της μετά-ανάλυσης 11
1.4 Ετερογένεια 15
1.5 Γράφημα forest 17
1.6 Ανάλυση συνεχών δεδομένων 20
1.7 Μέθοδοι μετατροπής συνεχών δεδομένων σε μεγέθη επίδρασης 23
2. Μέθοδοι 25
2.1 Εργαλεία της μετά-ανάλυσης 25
2.2 Μεγέθη επίδρασης 28
2.3 Σχεδιασμός των μεθόδων 30
3. Αποτελέσματα 32
3.1 Υλοποίηση του προγράμματος 32
3.3 Εφαρμογή των μεθόδων σε πραγματικά σύνολα δεδομένων 51
3.3.1 Παράδειγμα 1: Επίδραση της ροσιγλιταζόνης στην συγκέντρωση της HbA1c . 51
3.3.2 Παράδειγμα 2: Επίδραση της μετφορμίνης στην συγκέντρωση της HbA1c 55
3.3.3 Παράδειγμα 3: Επίδραση της μιγλιτόλης στην συγκέντρωση της HbA1c 58
4. Συζήτηση 61
Βιβλιογραφικές Αναφορές 63

Κατάλογος πινάκων

1 Πίνακας αποτελεσμάτων ανά θεραπεία……..6
2 Παράδειγμα συνεχών δεδομένων για 2 ομάδες ατόμων με τιμή αποκοπής, cutoff score = 9 (κόκκινη γραμμή)…….21
3 2x2 πίνακας συνάφειας μετά από διχοτόμηση των συνεχών δεδομένων…….22
4 Πίνακας αποτελεσμάτων ανά ομάδα θεραπείας …..…….....28
5 Στατιστικά μέτρα λόγου σχετικών πιθανοτήτων…….....28
6 Στατιστικά μέτρα διαφοράς κινδύνου……..29
7 Στατιστικά μέτρα τυποποιημένων διαφορών μέσου……29
8 Περιγραφή του σετ δεδομένων Rosiglitazone…….....52
9 Λίστα τιμών των μεταβλητών του σετ δεδομένων Rosiglitazone……52
10 Πίνακας αποτελεσμάτων της μετά-ανάλυσης, χρησιμοποιώντας την metan, το μοντέλο τυχαίων επιδράσεων και την μέθοδο F, στο σετ δεδομένων Rosiglitazone...53
11 Μετά-ανάλυση των μελετών (k = 6) του σετ δεδομένων Rosiglitazone, χρησιμοποιώντας το μοντέλο τυχαίων επιδράσεων…….54
12 Μετά-ανάλυση των μελετών (k = 6) του σετ δεδομένων Rosiglitazone, χρησιμοποιώντας το μοντέλο σταθερών επιδράσεων……..55
13 Λίστα τιμών των μεταβλητών του σετ δεδομένων Metformin……...56
14 Πίνακας αποτελεσμάτων της μετά-ανάλυσης, χρησιμοποιώντας το πρόγραμμα ameta, το μοντέλο τυχαίων επιδράσεων και την μέθοδο S, στο σετ δεδομένων Metformin…….56
15 Μετά-ανάλυση των μελετών (k = 4) του σετ δεδομένων Metformin.dta, χρησιμοποιώντας το μοντέλο τυχαίων επιδράσεων…….57
16 Μετά-ανάλυση των μελετών (k = 4) του σετ δεδομένων Metformin.dta, χρησιμοποιώντας το μοντέλο σταθερών επιδράσεων……..58
17 Λίστα τιμών των μεταβλητών του σετ δεδομένων Miglitol……...58
18 Πίνακας αποτελεσμάτων της μετά-ανάλυσης, χρησιμοποιώντας την metan, το μοντέλο σταθερών επιδράσεων και την μέθοδο HH, στο σετ δεδομένων Miglitol…59
19 Μετά-ανάλυση των μελετών (k = 3) του σετ δεδομένων Miglitol.dta, χρησιμοποιώντας το μοντέλο τυχαίων επιδράσεων……60

20 Μετά-ανάλυση των μελετών (k = 3) του σετ δεδομένων Miglitol.dta, χρησιμοποιώντας το μοντέλο σταθερών επιδράσεων……61

Κατάλογος σχημάτων

1 Γραφική παράσταση κανονικής κατανομής…….....3
2 Γραφική παράσταση τυποποιημένης κανονικής κατανομής…….....4
3 Γράφημα forest για το σετ δεδομένων BCG……...18
4 Διάγραμμα διασποράς ARE – OR για RiskCon = 0.2 και ίσα μεγέθη δειγμάτων...43
5 Διάγραμμα διασποράς ARE – OR για RiskCon = 0.5 και ίσα μεγέθη δειγμάτων...44
6 Διάγραμμα διασποράς ARE – OR για RiskCon = 0.8 και ίσα μεγέθη δειγμάτων...45
7 Διάγραμμα διασποράς ARE – OR για RiskCon = 0.2 και άνισα μεγέθη δειγμάτων (ncon = 100)…….46
8 Διάγραμμα διασποράς ARE – OR για RiskCon = 0.5 και άνισα μεγέθη δειγμάτων (ncon = 100)…….47
9 Διάγραμμα διασποράς ARE – OR για RiskCon = 0.8 και άνισα μεγέθη δειγμάτων (ncon = 100)…….48
10 Διάγραμμα διασποράς ARE – OR για RiskCon = 0.2 και άνισα μεγέθη δειγμάτων (ncon = 10)……...49
11 Διάγραμμα διασποράς ARE – OR για RiskCon = 0.5 και άνισα μεγέθη δειγμάτων (ncon = 10)……...49
12 Διάγραμμα διασποράς ARE – OR για RiskCon = 0.8 και άνισα μεγέθη δειγμάτων (ncon = 10)……...50
13 Γράφημα forest της metan για το σετ δεδομένων Rosiglitazone……54
14 Γράφημα forest της ameta για το σετ δεδομένων Metformin…….57
15 Γράφημα forest της metan για το σετ δεδομένων Miglitol…….60

Περίληψη

Τα αποτελέσματα συνεχούς μέτρησης τυχαιοποιημένων κλινικών δοκιμών συνήθως προκύπτουν από μια πληθώρα τεχνικών και κλιμάκων μέτρησης. Για να προκύψει όμως ένας καθολικός εκτιμητής κατά την μετά-ανάλυση τέτοιων δοκιμών, είναι απαραίτητο να ακολουθηθεί μια κοινή γραμμή τόσο στις μεθόδους όσο και στις μονάδες μέτρησης των δοκιμών αυτών. Σε περιπτώσεις όπου αυτό δεν είναι εφικτό, συνηθέστερα γίνεται χρήση των τυποποιημένων διαφορών των μέσων (Standardized Mean Differences – SMDs). Οι SMDs εμπεριέχουν τις παρατηρούμενες διαφορές των μέσων των 2 ομάδων (ομάδα ελέγχου και πειραματική ομάδα) δια της τυπικής τους απόκλισης σε κάθε δοκιμή. Το πλεονέκτημα των SMDs σε κάθε κλινική δοκιμή είναι ότι παραμένουν ανεξάρτητες από τον τρόπο διεξαγωγής αυτής, αποτελώντας έτσι ένα ιδανικό μέγεθος επίδρασης, στο οποίο οι μονάδες μέτρησης είναι πάντα εκφρασμένες σε τυπικές αποκλίσεις. Σε πρακτικό επίπεδο όμως, οι SMDs δεν είναι εύκολα ερμηνεύσιμες, με αποτέλεσμα οι επιστήμονες να χρησιμοποιούν εκ νέου μεθόδους για να εξάγουν χρήσιμα στατιστικά συμπεράσματα. Αυτές περιλαμβάνουν τη διχοτόμηση των συνεχών δεδομένων με βάση ένα ορισμένο κατώφλι με σκοπό την κατηγοριοποίηση των ατόμων σε 2 ομάδες και την μετέπειτα σύγκρισή τους. Σε αυτή την εργασία, ασχολούμαστε με την υλοποίηση των μεθόδων αυτών και πιο συγκεκριμένα των μεθόδων των Hasselbad and Hedges (HH), των Cox and Snell (CS), του Furukawa (FU) και του Suissa (SU), οι οποίες χρησιμοποιούν τις SMDs για την εξαγωγή άλλων μεγεθών επίδρασης με μεγαλύτερη ερμηνευσιμότητα (odds ratios, risk ratios, risk differences, numbers needed to treat). Για την υλοποίηση αυτή, έγινε χρήση του στατιστικού πακέτου Stata και των συναρτήσεων που διαθέτει.

Λέξεις - κλειδιά: μετά-ανάλυση τυχαίων επιδράσεων, συνεχή δεδομένα, ανάλυση απόκρισης, διχοτόμηση, τυποποιημένες διαφορές των μέσων, Stata

Abstract

Continuous measurement results of randomized clinical trials are usually derived from a variety of measurement techniques and scales. However, to obtain a universal estimator during the meta-analysis of such trials, it is necessary to follow a common line both in the methods and in the measurement units of these trials. In cases where this is not possible, Standardized Mean Differences (SMDs) are more commonly used. SMDs contain the observed differences in means of the 2 groups (control group and experimental group) divided by their standard deviation in each trial. The advantage of SMDs in any clinical trial is that they remain independent of the way the latter is conducted, thus constituting an ideal effect size, in which the units of measurement are always expressed in standard deviations. On a practical level, however, SMDs are not easily interpretable, resulting in scientists re-using methods to draw useful statistical conclusions. These methods include splitting continuous data based on a certain threshold with the purpose of categorizing individuals into 2 groups and their subsequent comparison. In this thesis, we deal with the implementation of these methods, and more specifically with these of Hasselblad and Hedges (HH), Cox and Snell (CS), Furukawa (FU) και Suissa (SU), which use SMDs to derive other effect sizes with greater interpretability (odds ratios, risk ratios, risk differences, numbers needed to treat). For this implementation, the statistical package Stata and its functions were used.

Keywords: random-effects meta-analysis, continuous data, responder analysis, dichotomization, standardized mean differences, Stata

1. Εισαγωγή

1.1 Εισαγωγικοί όροι

Παρακάτω θα παρουσιαστούν κάποιοι βασικοί στατιστικοί όροι, καθώς και μια σύντομη ανάλυση αυτών, με σκοπό την καλύτερη κατανόησή τους κατά την αναφορά τους παρακάτω στην εργασία. Εδώ πρέπει να σημειωθεί ότι δεν θα γίνει αναφορά σε όλες τις στατιστικές έννοιες, αλλά μόνο σε αυτές που αφορούν το ζητούμενο της παρούσας εργασίας. Πιο συγκεκριμένα, στην περίπτωση μας, θα ασχοληθούμε με **αριθμητικές, συνεχείς μεταβλητές** (π.χ. αρτηριακή πίεση) και σε μέτρα που αφορούν τέτοιου είδους μεταβλητές.

1. Δειγματικός μέσος (Sample mean) ν παρατηρήσεων ενός δείγματος του πληθυσμού, είναι το άθροισμα των αριθμητικών τιμών a_i κάθε παρατήρησης διά το πλήθος των παρατηρήσεων.

$$\mu = \frac{1}{\nu}\sum_{i=1}^{\nu} a_i \tag{1.1}$$

2. Διακύμανση (Variance) του δειγματικού μέσου, είναι η μέση τετραγωνική απόσταση των a_i παρατηρήσεων από τον δειγματικό μέσο.

$$s^2 = \frac{\sum_{i=1}^{\nu}(a_i - \mu)^2}{\nu - 1} \tag{1.2}$$

Σε αυτό το σημείο, αξίζει να αναφέρουμε ότι παρ' όλη την σημαντικότητα του εκτιμητή της διακύμανσης, δεν είναι διαισθητικά ερμηνεύσιμος από στους ερευνητές. Επίσης, όπως φαίνεται και από τον μαθηματικό τύπο παραπάνω, η διακύμανση δίνει μεγάλη βαρύτητα σε ακραίες τιμές (δηλαδή αριθμητικές τιμές των παρατηρήσεων που βρίσκονται μακριά από τον μέσο).

3. Τυπική απόκλιση (Standard deviation) του δειγματικού μέσου, είναι η τετραγωνική ρίζα της διακύμανσης του.

$$SD = \sqrt{s^2} \tag{1.3}$$

Η τυπική απόκλιση δηλώνει το πόσο αναμένεται να απέχει μία αριθμητική τιμή μιας παρατήρησης από τον αριθμητικό μέσο του πληθυσμού. Υψηλή τυπική απόκλιση υποδηλώνει ότι τα δεδομένα βρίσκονται διασκορπισμένα, ενώ χαμηλή τυπική απόκλιση δείχνει ότι τα δεδομένα βρίσκονται γύρω από τον αριθμητικό μέσο του πληθυσμού.

Τα παραπάνω περιγραφικά στατιστικά, όπως ονομάζονται, αναφέρονται σε ένα δείγμα του πληθυσμού, και όχι σε ολόκληρο τον πληθυσμό. Σκοπός τους είναι να εξάγουν συμπεράσματα για τις **παραμέτρους** ολόκληρου του πληθυσμού που τον περιγράφουν και είναι πρακτικά άγνωστες.

4. Κατανομή Πιθανότητας μιας μεταβλητής ενός πληθυσμού, είναι το σύνολο όλων των πιθανών αποτελεσμάτων της μεταβλητής, καθώς και των σχετικών συχνοτήτων εμφάνισης αυτών. Οι κατανομές πιθανοτήτων χωρίζονται σε συνεχείς και διακριτές, ανάλογα με το είδος της μεταβλητής, και οι μεταβλητές που ακολουθούν την εκάστοτε κατανομή ονομάζονται **τυχαίες μεταβλητές.** Για να περιγραφεί μαθηματικά η κατανομή πιθανότητας μιας συνεχούς, τυχαίας μεταβλητής, χρησιμοποιείται η **συνάρτηση πυκνότητας πιθανότητας.** Δύο από τις σημαντικότερες συνεχείς κατανομές πιθανοτήτων είναι:

- **Η Κανονική (ή Γκαουσιανή) Κατανομή**
- **Η Λογιστική Κατανομή**

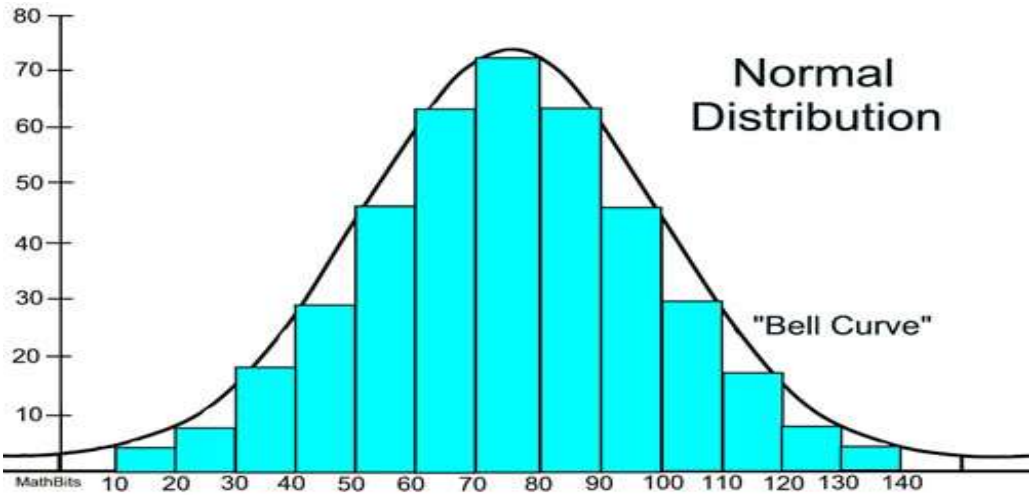
4.1. Ως **Κανονική Κατανομή (Normal Distribution)** μιας συνεχούς τυχαίας μεταβλητής X, χαρακτηρίζουμε τη συμμετρική κατανομή όπου η πλειοψηφία των αριθμητικών τιμών της μεταβλητής X βρίσκονται γύρω από την τιμή του μέσου. Η κατανομή αυτή έχει ως παραμέτρους την μέση τιμή και την τυπική απόκλιση, ενώ το σχήμα της μοιάζει με καμπάνα. Όσον αφορά τον συμβολισμό, ότι δηλαδή η συνεχής τυχαία μεταβλητή X ακολουθεί την κανονική κατανομή με μέσο μ και διακύμανση σ^2, μπορεί να περιγραφεί με τον ακόλουθο τρόπο:

$$X \sim N(\mu, \sigma^2) \tag{1.4}$$

Η συνάρτηση πυκνότητας πιθανότητάς της είναι:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \qquad (1.5)$$

Και η γραφική της παράσταση:



Σχήμα 1. Γραφική παράσταση κανονικής κατανομής.

Για να θεωρηθεί ότι η κατανομή μίας τυχαίας μεταβλητής X «προσεγγίζει» την κανονική κατανομή θα πρέπει να πληρούνται 4 βασικά κριτήρια:

1. Περίπου 50% των παρατηρήσεων να βρίσκεται πάνω από τον μέσο και 50% να βρίσκεται κάτω από τον μέσο.
2. Περίπου το 68% των παρατηρήσεων να βρίσκεται σε απόσταση το πολύ μιας τυπικής απόκλισης από τον μέσο και άρα στο εύρος (μ-σ, μ+σ).
3. Περίπου το 95% των παρατηρήσεων να βρίσκεται σε απόσταση το πολύ 2 τυπικών αποκλίσεων από τον μέσο και άρα στο εύρος (μ-2σ, μ+2σ).
4. Περίπου το 99,7% των παρατηρήσεων να βρίσκεται στο διάστημα (μ-3σ, μ+3σ). Παρατηρήσεις που βρίσκονται εκτός του συγκεκριμένου εύρους έχουν πολύ μικρή συχνότητα εμφάνισης, χωρίς αυτό να σημαίνει ότι δεν ανήκουν απαραίτητα στην κατανομή.

Για τα διάφορα μ και σ προκύπτουν αντίστοιχα, διαφορετικές κανονικές κατανομές**.** Στην **τυπική κανονική κατανομή**, ο μέσος ισούται με 0 και η τυπική της απόκλιση

ισούται με 1. Για να μετατρέψουμε μια κανονική τυχαία μεταβλητή X σε μία **τυπική κανονική τυχαία μεταβλητή Z,** χρησιμοποιείται ο ακόλουθος τύπος:

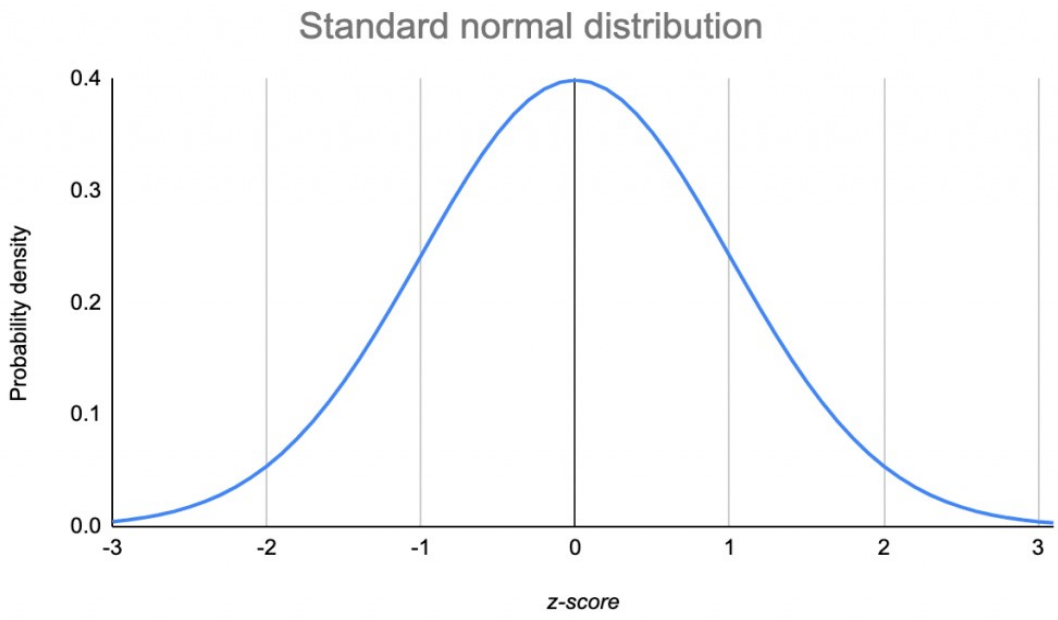
$$Z = \frac{X - \mu}{\sigma} \tag{1.6}$$

όπου μ και σ, ο μέσος και η τυπική απόκλιση της κανονικής κατανομής αντίστοιχα.

Η συνάρτηση πυκνότητας πιθανότητας θα μεταβληθεί ανάλογα, ως εξής:

$$\Phi(Z) = \frac{1}{\sqrt{2\pi}} e^{\frac{Z^2}{2}} \tag{1.7}$$

Και η γραφική παράσταση θα είναι η παρακάτω:



Σχήμα 2. Γραφική παράσταση τυπικής κανονικής κατανομής.

Η μεταβλητή Z εκφράζει το πόσες τυπικές αποκλίσεις απέχει η τιμή μιας παρατήρησης από τον μέσο της κανονικής κατανομής.

4.2. Ως **Λογιστική κατανομή (Logistic Distribution)** μιας συνεχούς τυχαίας μεταβλητής X, χαρακτηρίζουμε την συνεχή κατανομή που σχηματικά μοιάζει με την κανονική κατανομή, αλλά με πιο πλατιές «ουρές» στην γραφική παράσταση (μεγαλύτερη διακύμανση). Σημαντική για την εργασία είναι η **τυπική λογιστική**

κατανομή, όπου για μ = 0 και παράμετρο κλίμακας β = 1 η μεταβλητή έχει συνάρτηση πυκνότητας πιθανότητας:

$$f(x) = \frac{e^{-x}}{(1-e^{-x})^2} \qquad (1.8)$$

με διακύμανση $\pi^2 / 3$.

5. Μελέτη ασθενών μαρτύρων (case-control study) ονομάζεται το είδος επιδημιολογικής μελέτης κατά το οποίο εξετάζεται η επίδραση ενός παράγοντα σε 2 ομάδες ατόμων. Η 1η ομάδα ονομάζεται **ομάδα ελέγχου (controls)** και στα άτομα αυτής απουσιάζει η ύπαρξη της εξεταζόμενης π.χ. ασθένειας, ενώ ισχύει το αντίθετο για τη 2η **ομάδα των περιπτώσεων (cases)**. Στόχος είναι να βρεθεί αν η παρουσία (και αντίστοιχα η απουσία) του παράγοντα επηρεάζει κάποια από τις 2 ομάδες.

6. Τυχαιοποιημένη κλινική δοκιμή (Randomized Clinical Trial) ονομάζεται το είδος της μελέτης, κατά το οποίο εξετάζεται η επίδραση των εξεταζόμενων θεραπειών στην υγεία των ατόμων. Η λέξη «τυχαιοποιημένη» αναφέρεται στην ομαδοποίηση των ατόμων στις επιμέρους ομάδες θεραπειών, και γίνεται τυχαία. Η ομάδα που δεν λαμβάνει την θεραπεία ή λαμβάνει μια συμβατική θεραπεία, ονομάζεται **ομάδα ελέγχου (control group).** Στην περίπτωση που δεν λαμβάνει κάποια θεραπεία η ομάδα ελέγχου, εξετάζεται η επίδραση του φαινομένου placebo (placebo effect) στους ασθενείς. Η ομάδα (ή ομάδες) που δέχεται κάποια θεραπεία ονομάζεται **ομάδα θεραπείας (treatment group ή experimental group)**. Μετά το πέρας της κλινικής δοκιμής, τα άτομα κατηγοριοποιούνται ανάλογα με το αποτέλεσμα της εκάστοτε θεραπείας, σε αυτούς που **ανταποκρίθηκαν (treatment responders)** και σε αυτούς που **δεν ανταποκρίθηκαν στην θεραπεία (treatment non responders).**

7. Μέγεθος επίδρασης (Effect size) θεωρείται οποιοδήποτε στατιστικό μέτρο ικανό να ποσοτικοποιήσει την αιτιώδη σχέση μεταξύ 2 μεταβλητών ή ομάδων (π.χ. ομάδα ελέγχου σε σχέση με την ομάδα θεραπείας). Τα μεγέθη επίδρασης παίζουν πολύ σημαντικό ρόλο στην μετά-ανάλυση, καθώς συμβάλλουν στην εξαγωγή συμπερασμάτων. Τα μεγέθη επίδρασης που θα χρησιμοποιηθούν στην εργασία αυτή είναι:

- Ο λόγος σχετικών πιθανοτήτων (Odds ratio, OR)

- Η διαφορά κινδύνου (Risk Difference)
- Οι τυποποιημένες διαφορές μέσων (Standardized Mean Differences, Cohen's d, SMD).

Για τα πρώτα 2 μεγέθη επίδρασης χρειάζεται πρώτα να δημιουργηθεί ένας 2x2 πίνακας, γνωστός και ως **πίνακας συνάφειας**. Θα θέσουμε ως παράδειγμα μια τυχαιοποιημένη κλινική δοκιμή και θα καταγράψουμε τα αποτελέσματα της. Η ομάδα ελέγχου και η ομάδα θεραπείας θα αποτελέσουν τις γραμμές του πίνακα, ενώ οι ανταποκρίνοντες και μη στην θεραπεία, τις στήλες αυτού.

Πίνακας 1. Πίνακας αποτελεσμάτων ανά θεραπεία.

	Treatment Responders	**Treatment Non-Responders**	**Total**
Treatment Group	**a**	**b**	**a+b**
Control Group	**c**	**d**	**c+d**
Total	**a+c**	**b+d**	**a+b+c+d**

Ο λόγος σχετικών πιθανοτήτων θα υπολογιστεί ως εξής:

$$OR = \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a * d}{b * c} \quad (1.9)$$

Ως εξεταζόμενο γεγονός θα ορίσουμε την ανταπόκριση στην θεραπεία. Ο αριθμητής του μη απλοποιημένου κλάσματος δείχνει την αναλογία αυτών που ανταποκρίθηκαν στην θεραπεία σε σχέση με αυτούς που δεν ανταποκρίθηκαν, ενώ η αντίστοιχη αναλογία βρίσκεται στον παρονομαστή. Με βάση τα συγκεκριμένα δεδομένα, τα αποτελέσματα του OR ερμηνεύονται ως εξής:

- Αν OR > 1, η πιθανότητα να ανταποκριθεί στην θεραπεία ο ασθενής είναι **μεγαλύτερη** στην ομάδα θεραπείας απ' ότι στην ομάδα ελέγχου. Άρα η θεραπεία που δόθηκε στην ομάδα θεραπείας φαίνεται να επιδρά καλύτερα στην υγεία των ατόμων απ' ότι η αντίστοιχη της ομάδας ελέγχου.

- Αν OR = 1, η πιθανότητα ανταπόκρισης στην θεραπεία είναι **ίδια** και στις 2 ομάδες, συνεπώς δεν παρατηρείται σημαντική διαφορά μεταξύ των 2 θεραπειών.
- Αν OR < 1, η πιθανότητα ανταπόκρισης στην θεραπεία είναι **μεγαλύτερη** στην ομάδα ελέγχου απ' ότι στην ομάδα θεραπείας, άρα η συμβατική (ή καθόλου) θεραπεία φαίνεται να είναι πιο αποτελεσματική.

Η **διαφορά κινδύνου** υπολογίζεται από τον ακόλουθο τύπο:

$$RD = \frac{\alpha}{a+b} - \frac{c}{c+d} = risk_{exp} - risk_{con} \qquad (1.10)$$

Το καθένα από τα 2 κλάσματα εκφράζει τον **κίνδυνο,** δηλαδή την πιθανότητα, εμφάνισης ενός επιλεγμένου συμβάντος, για την καθεμία ομάδα αντίστοιχα. Σαν συμβάν ορίζεται πάλι η ανταπόκριση στην θεραπεία. Ανάλογα με το αποτέλεσμα η διαφορά κινδύνου ερμηνεύεται ως εξής:

- Αν RD < 0, ο κίνδυνος εμφάνισης του συμβάντος είναι **μειωμένος** στην ομάδα θεραπείας, σε σχέση με την ομάδα ελέγχου. Η θεραπεία της ομάδας ελέγχου φαίνεται να είναι αποτελεσματικότερη.
- Αν RD = 0, ο κίνδυνος είναι **ίδιος** και στις 2 ομάδες. Καμία σημαντική διαφορά όσον αφορά την επίδραση των 2 θεραπειών.
- Αν RD > 0, ο κίνδυνος είναι **αυξημένος** στην ομάδα θεραπείας, σε σχέση με την ομάδα ελέγχου. Η θεραπεία της ομάδας θεραπείας δείχνει πιο αποτελεσματική.

Οι τυποποιημένες διαφορές μέσων (ή Cohen's d) μπορούν να υπολογιστούν με βάση τον τύπο:

$$\frac{mean_{exp} - mean_{con}}{sd_{pooled}} \qquad (1.11)$$

Ο αριθμητής αναφέρεται στην διαφορά ανάμεσα στους μέσους των 2 ομάδων θεραπείας, και διαιρείται με την συγκεντρωτική διακύμανση των 2 ομάδων η οποία ισούται με:

$$sd_{pooled} = \sqrt{\frac{(n_{exp} - 1) * sd^2{}_{exp} + (n_{con} - 1) * sd^2{}_{con}}{(n_{exp} + n_{con}) - 2}} \qquad (1.12)$$

όπου n_{exp} και n_{con} είναι ο αριθμός των ατόμων που συμμετέχουν στην ομάδα θεραπείας και την ομάδα ελέγχου, αντίστοιχα.

Τέλος, το **p-value** είναι ένα στατιστικό μέτρο, ευρέως χρησιμοποιούμενο στις στατιστικές δοκιμασίες και ιδιαίτερα στον **έλεγχο υποθέσεων**, όπως επίσης και η **τιμή άλφα (alpha value)** που αποτελεί το **επίπεδο σημαντικότητας** (**significance level)** και ορίζεται από τον εκάστοτε ερευνητή. Κατά τον έλεγχο υποθέσεων διατυπώνονται «νοητά» 2 υποθέσεις σχετικά με την αιτιώδη σχέση 2 μεταβλητών ή ομάδων (π.χ. ομάδα ελέγχου και ομάδα θεραπείας). Η υπόθεση ότι τα δεδομένα των 2 ομάδων δεν παρουσιάζουν στατιστικά σημαντικές διαφορές ονομάζεται **μηδενική υπόθεση (H_0),** ενώ αντίστοιχα ορίζεται και η **εναλλακτική υπόθεση (H_1)** που δηλώνει την ύπαρξη στατιστικά σημαντικών διαφορών. Για να καθοριστεί ποια υπόθεση τείνει να ισχύει για τα δεδομένα γίνεται χρήση στατιστικών μεθόδων (όπως το t-test), οι οποίες μεταξύ άλλων υπολογίζουν και το p-value. Αν το p-value της μεθόδου είναι μικρότερο από το alpha value που έχει οριστεί τότε **απορρίπτεται η H_0,** ενώ στην αντίθετη περίπτωση **δεν απορρίπτεται η H_0**. Στατιστικά μιλώντας, το p-value εκφράζει την **πιθανότητα** να προκύψουν αποτελέσματα τουλάχιστον τόσο ακραία όσο το αποτέλεσμα που παρατηρήθηκε στην πραγματικότητα υπό την μηδενική υπόθεση**,** άρα οι διαφορές που θα παρατηρηθούν να οφείλονται στην τύχη. Τις περισσότερες φορές χρησιμοποιείται alpha value ίσο με 0.05.

1.2 Εισαγωγή στην μετά-ανάλυση

Ως μετά-ανάλυση ορίζεται η μέθοδος συνδυασμού προγενέστερων αναλύσεων (εξ' ου και το όνομα μετά-ανάλυση) με σκοπό την εξαγωγή ενός στατιστικά πιο ορθού συμπεράσματος, σε σχέση με αυτό που θα έδινε κάθε ανάλυση ξεχωριστά (Borenstein et al., 2009, σελ. xxi). Βασική προϋπόθεση για την αποφυγή λανθασμένων συμπερασμάτων είναι η σύνθεση των μελετών να γίνεται μεθοδικά. Αυτό με την

σειρά του σημαίνει, ότι πληροφορίες σχετικά με τα δεδομένα των αναλύσεων, τον τρόπο διεξαγωγής αυτών, την διατύπωση των αντίστοιχων υποθέσεων (βλ. έλεγχο υποθέσεων παραπάνω), πρέπει να είναι κοινές σε όλες τι επιμέρους αναλύσεις που συμμετέχουν στην μετά-ανάλυση.

Σαν παράδειγμα θα μπορούσε να χρησιμοποιηθεί μια μετά-ανάλυση που εξετάζει την επίδραση ενός φαρμάκου στους εξεταζόμενους ασθενείς. Στην περίπτωση αυτή, γίνεται σύνθεση τυχαιοποιημένων κλινικών δοκιμών, όπου στην κάθε μία αναφέρεται και το αντίστοιχο μέγεθος επίδρασης. Η σημασία του μεγέθους επίδρασης αποδεικνύεται καθοριστική, καθώς συντελεί στην εξαγωγή πολύτιμων συμπερασμάτων. Αν ακολουθεί ίδια πορεία σε όλες τις μελέτες (για παράδειγμα OR>1 σε κάθε μελέτη), πιθανόν να προκύψει ένα στατιστικά ισχυρό αποτέλεσμα, ενώ στην περίπτωση που παρουσιάζει διαφορές, να μαρτυρά την ύπαρξη παραγόντων που επηρεάζουν τις αριθμητικές τιμές του (Borenstein et al., 2009, σελ. xxii).

Παρ' όλο που στο παραπάνω παράδειγμα αναφέρθηκε μόνο το μέγεθος επίδρασης, βασικά στοιχεία της μετά-ανάλυσης αποτελούν επίσης το **συγκεντρωτικό αποτέλεσμα (summary effect)** και το **p-value** (Borenstein et al., 2009, σελ. 3). Επιπρόσθετα, ανατίθενται διαφορετικά **βάρη** σε κάθε μελέτη, αναλογικά με την επίδραση που έχει η τελευταία στην τελική μετά-ανάλυση. Η ανάθεση των βαρών γίνεται με την χρήση μαθηματικών τεχνικών (Borenstein et al., 2009, σελ. 4).

Τέλος, θα πρέπει να αναφερθούμε και στην ύπαρξη της **μεταβλητότητας** (μέτρο της οποίας είναι τα είδη διακυμάνσεων που θα αναφερθούν παρακάτω) που υπάρχει στις μελέτες μιας μετά-ανάλυσης και κατά πόσο αυτή οφείλεται στην τύχη (Borenstein et al., 2009)

Παρακάτω θα περιγραφεί το μαθηματικό και θεωρητικό υπόβαθρο όλων των δομικών στοιχείων της μετά-ανάλυσης αλλά και των στατιστικών μεθόδων που αναφέρθηκαν παραπάνω.

Πριν εμβαθύνουμε όμως περισσότερο σε ό,τι έχει σχέση με το μέγεθος επίδρασης, θα ήταν καλό να αναφέρουμε σαν όρο την **επίδραση θεραπείας (treatment effect)**, καθώς είναι πολύ πιθανόν να γίνει αναφορά σε αυτήν παρακάτω στην εργασία. Σαν επίδραση θεραπείας αποκαλείται το μέγεθος επίδρασης όταν η εκάστοτε μελέτη εξετάζει την επίδραση μιας ιατρικής παρέμβασης. Επειδή ως γνωστόν το μέγεθος

επίδρασης εξετάζει την σχέση μεταξύ 2 ομάδων, είναι απαραίτητο οι ομάδες να διαφέρουν ως προς την μέθοδο ιατρικής παρέμβασης που έλαβαν, για να μπορέσει να χρησιμοποιηθεί σωστά ο όρος (Borenstein et al., 2009, σελ. 17). Διαφορετικά, γίνεται λόγος απλά για μέγεθος επίδρασης.

Επιπλέον, θα γίνει αναφορά στις ορολογίες **εκτίμηση του μεγέθους επίδρασης** και **του μεγέθους επίδρασης ως παράμετρο.** Όταν αναφέρεται ως παράμετρος (συμβολίζεται με το γράμμα θ) εννοείται το πραγματικό μέγεθος επίδρασης που θα ήταν γνωστό αν είχαμε στην διάθεση μας ολόκληρο τον εξεταζόμενο πληθυσμό (ή έστω ένα μεγάλο κλάσμα αυτού). Καθώς αυτό δεν είναι πρακτικά δυνατό, χρησιμοποιείται η εκτίμηση της παραμέτρου αυτής (συμβολίζεται με το γράμμα Υ) η οποία είναι λογικό να διαφέρει από την παράμετρο θ ως προς την αριθμητική της τιμή.

Ένα από τα πρώτα στάδια μιας μετά-ανάλυσης είναι, μεταξύ άλλων, η επιλογή του μεγέθους επίδρασης. Σε αρχικό στάδιο, αυτή η επιλογή βασίζεται σε ορισμένα κριτήρια, μερικά εκ των οποίων είναι τα εξής (Borenstein et al., 2009, σελ. 18):

1. Το μέγεθος επίδρασης που θα επιλεχθεί πρέπει να είναι διαισθητικά ερμηνεύσιμο στην επιστημονική κοινότητα που ασχολείται με την μετά-ανάλυση αυτή. Επιπλέον για τους ίδιους λόγους, μπορεί να χρειαστεί μετατροπή της αριθμητικής τιμής του κατά την παρουσίαση (π.χ. μετατροπή του νεπέριου λογαρίθμου του Odds Ratio στην κανονική μορφή αυτού).

2. Παράγοντες που διαφέρουν ανάμεσα στις μελέτες (π.χ. μέγεθος δείγματος κάθε μελέτης) δεν πρέπει να επηρεάζουν το μέγεθος επίδρασης. Η λογική αυτή αποσκοπεί στο να είναι εφικτή η σύγκριση ανάμεσα στα μεγέθη επίδρασης όλων των μελετών σε μία μετά-ανάλυση.

3. Η εκτίμηση του μεγέθους επίδρασης πρέπει να είναι εφικτή και μετά το πέρας της μετά-ανάλυσης, με βάση τα δεδομένα που αναφέρονται στην κάθε μελέτη ξεχωριστά. Για παράδειγμα, να μπορεί να υπολογιστεί το Cohen’s d από τους μέσους κάθε μελέτης, χωρίς να χρειαστεί ο υπολογισμός των μέσων από την αρχή.

Με βάση τα παραπάνω κριτήρια προκύπτουν μερικά από τα πιο ευρέως διαδεδομένα μεγέθη επίδρασης. Για να επιλεχθεί το καταλληλότερο ωστόσο, πρέπει να εξεταστεί η

φύση των δεδομένων και του αποτελέσματος (outcome). Ανάλογα με την περίπτωση, η ανάθεση μεγέθους επίδρασης γίνεται ως εξής (Borenstein et al., 2009, σελ. 18):

1. Στην περίπτωση που τα συνοπτικά δεδομένα (summary data) βασίζονται στον μέσο και την τυπική απόκλιση 2 τουλάχιστον ομάδων, χρησιμοποιούνται οι διαφορές στους μέσους, το Cohen's d και ο λόγος απόκρισης.

2. Αν τα συνοπτικά δεδομένα βασίζονται σε ένα δυαδικό αποτέλεσμα (π.χ. ανταπόκριση ή μη ανταπόκριση στην θεραπεία), τότε συνήθη μεγέθη επίδρασης είναι ο σχετικός κίνδυνος, η διαφορά κινδύνου και ο λόγος σχετικών πιθανοτήτων.

3. Αν στις επιμέρους μελέτες εξετάζεται η συσχέτιση μεταξύ 2 μεταβλητών, χρησιμοποιείται συνήθως ο συντελεστής συσχέτισης.

Εδώ αξίζει να αναφερθεί ότι από τα συνεχή δεδομένα που θα εξεταστούν στο πλαίσιο της παρούσας εργασίας, θα προκύψει το Cohen's d το οποίο με την σειρά του θα μετατραπεί σε ένα από τα μεγέθη επίδρασης της 2^{ης} περίπτωσης.

1.3 Μοντέλα της μετά-ανάλυσης

Τα στατιστικά μοντέλα χρησιμοποιούνται για τον συνδυασμό των στατιστικών στοιχείων κάθε μελέτης, αλλά και για την εύρεση του καθολικού αποτελέσματος κατά την μετά-ανάλυση. Τα πιο διαδεδομένα από αυτά είναι το **μοντέλο σταθερών επιδράσεων (fixed-effect model)** και το **μοντέλο τυχαίων επιδράσεων (random-effects model)**.

Προτού γίνει η περιγραφή της μαθηματικής τους υπόστασης, πρέπει να τονιστεί ότι η επιλογή του κατάλληλου μαθηματικού μοντέλου δεν είναι μια τυχαία διαδικασία. Αντίθετα, είναι άρρηκτα συνδεδεμένη με την αντίστοιχη υπόθεση που γίνεται για τα δεδομένα (βλ. παρακάτω), ενώ σκοπός της είναι να προάγει τους στόχους της ανάλυσης αλλά και την ερμηνευσιμότητα των αποτελεσμάτων της (Borenstein et al., 2010). Για τον λόγο αυτό, οποιαδήποτε ομοιότητα στους μαθηματικούς τύπους και

στον τρόπο εκτίμησης των παραμέτρων (βλ. παρακάτω), δεν αντανακλάται στο τελικό αποτέλεσμα των 2 μοντέλων (Borenstein et al., 2010).

Εδώ, είναι σκόπιμο να αναφερθούν τα είδη διακυμάνσεων κατά την μετά-ανάλυση για να μην χρειαστεί επεξήγηση αυτών παρακάτω. Σύμφωνα με τους Borenstein et al., (2010), αυτά είναι:

1. Η διακύμανση του πληθυσμού (population variance, βλ. Εισαγωγή).
2. Η διακύμανση του πραγματικού μεγέθους επίδρασης μεταξύ των μελετών (between-studies variance).
3. Η διακύμανση του σφάλματος εντός της μελέτης (within-study error variance).
4. Η συνολική διακύμανση του σφάλματος της μελέτης (overall-study error variance).
5. Η διακύμανση του σφάλματος της μετά-ανάλυσης (meta-analysis error variance)

Το **μοντέλο σταθερών επιδράσεων** κάνει την υπόθεση ότι όλες οι επιμέρους μελέτες έχουν ένα **κοινό πραγματικό μέγεθος επίδρασης** και οποιαδήποτε διαφορά κατά την εκτίμηση αυτού οφείλεται στο δειγματοληπτικό σφάλμα (sampling error) (Borenstein et al., 2010). Το μαθηματικό υπόβαθρο του συγκεκριμένου μοντέλου, είναι το ακόλουθο (Borenstein et al., 2010):

Το παρατηρούμενο μέγεθος επίδρασης (δηλαδή η εκτίμησή του) κάθε μελέτης *i* ορίζεται ως:

$$Y_i = \theta + \varepsilon_i \qquad (1.13)$$

όπου θ είναι το κοινό πραγματικό μέγεθος επίδρασης και ε_i είναι η διαφορά μεταξύ του παρατηρούμενου και του πραγματικού μεγέθους επίδρασης.

Η εντός των μελετών διακύμανση του σφάλματος δίνεται από τον τύπο:

$$V_i = \sigma^2 \qquad (1.14)$$

όπου σ^2 είναι η διακύμανση του μεγέθους επίδρασης κάθε μελέτης.

Όσον αφορά τα βάρη που ανατίθεται στις μελέτες, αυτά υπολογίζονται ως εξής:

$$W_i = \frac{1}{V_i} \tag{1.15}$$

Μερικές σημειώσεις-παρατηρήσεις που αξίζει να αναφερθούν όσον αφορά το μοντέλο σταθερών επιδράσεων:

1. Η διαφορά ε_i αποτελεί την διακύμανση του σφάλματος εντός των μελετών. Καθώς υπάρχει μόνο μια πηγή διακύμανσης στο μοντέλο αυτό, η ε_i ταυτίζεται με την συνολική διακύμανση του σφάλματος της μελέτης (Borenstein et al., 2010).

2. Όσον αφορά την ανάθεση βαρών, προκύπτει ότι ευνοούνται οι μεγαλύτερες (σε μέγεθος δείγματος και με μικρή διακύμανση) μελέτες απ' ότι οι μικρότερες. Αυτό συμβαίνει διότι αυτές συντελούν περισσότερο στον υπολογισμό του **κοινού** μεγέθους επίδρασης (Borenstein et al., 2010).

3. Η ύπαρξη του κοινού μεγέθους επίδρασης προϋποθέτει πολύ στενά πλαίσια όσον αφορά τις συνθήκες διεξαγωγής της μετά-ανάλυσης, τους ερευνητές, το δείγμα, κ.λπ. (Borenstein et al., 2010). Αυτό μπορεί να επιτευχθεί όταν η μετά-ανάλυση αφορά ένα συγκεκριμένο πληθυσμιακό σύνολο, όμως τότε το συμπέρασμα που θα προκύψει δεν γίνεται να επεκταθεί σε κάτι γενικότερο (π.χ. σε όλο τον πληθυσμό) (Borenstein et al., 2010).

Στον αντίποδα, στο **μοντέλο τυχαίων επιδράσεων** γίνεται η υπόθεση ότι υπάρχουν **πολλά πραγματικά μεγέθη επίδρασης** και σκοπός μας είναι η εύρεση του **μέσου της κατανομής τους** (Borenstein et al., 2010). Κάνοντας την υπόθεση ότι σε κάθε μελέτη θα αντιστοιχούσε ένα διαφορετικό πραγματικό μέγεθος επίδρασης, η επιλογή κάποιων εξ αυτών για τους σκοπούς της μετά-ανάλυσης δικαιολογεί την έκφραση «τυχαίες επιδράσεις» (Borenstein et al., 2010).

Από μαθηματικής σκοπιάς, το παρατηρούμενο μέγεθος επίδρασης κάθε μελέτης *i* ορίζεται ως (Borenstein et al., 2010):

$$Y_i = \mu + \xi_i + \varepsilon_i \tag{1.16}$$

όπου μ είναι ο μέσος της κατανομής των πραγματικών μεγεθών επίδρασης, ξ_i είναι η διαφορά ανάμεσα στον μέσο μ και το πραγματικό μέγεθος επίδρασης (θ) της μελέτης, και ε_i είναι η διαφορά ανάμεσα στο πραγματικό και το παρατηρούμενο μέγεθος επίδρασης (Borenstein et al., 2010).

Τα βάρη που ανατίθενται στις μελέτες υπολογίζονται ως εξής:

$$W_i^* = \frac{1}{V_i + \tau^2} \qquad (1.17)$$

όπου V_i είναι η εντός της μελέτης διακύμανση του σφάλματος, ενώ τ^2 είναι η μεταξύ των μελετών διακύμανση - κοινή για όλες τις μελέτες.

Για το μοντέλο τυχαίων επιδράσεων θα πρέπει επίσης να τονιστούν τα εξής:

1. Τα 2 είδη διακυμάνσεων που αναφέρθηκαν προκύπτουν από τα 2 επίπεδα δειγματοληψίας που χρησιμοποιούνται. Από την μία έχουμε ομαδοποίηση των ατόμων για χάρη μιας αυτούσιας μελέτης (1ο επίπεδο) και από την άλλη ομαδοποίηση μελετών που, όπως αναφέρθηκε παραπάνω, αντιπροσωπεύουν διαφορετικό πραγματικό μέγεθος επίδρασης η κάθε μία (Borenstein et al., 2010).

2. Με την χρήση του μοντέλου αυτού δεν αγνοούνται οι μικρότερες μελέτες, καθώς συντελούν στην εύρεση του μέσου των πραγματικών μεγεθών επίδρασης (αφού εκπροσωπούν διαφορετικό μέγεθος επίδρασης οι ίδιες) (Borenstein et al., 2010). Αυτό εν μέρη φαίνεται και στις αριθμητικές τιμές των βαρών, οι οποίες διαθέτουν στον παρονομαστή ένα μέγεθος που είναι κοινό σε όλες τις μελέτες (τ^2) (Borenstein et al., 2010).

3. Εφόσον οι μελέτες αποτελούν ένα τυχαίο δείγμα του συνόλου μελετών (άρα και των διαφορετικών πληθυσμών) και στόχος είναι η εύρεση του μέσου των διαφόρων μεγεθών επίδρασης, τα αποτελέσματα της μετά-ανάλυσης μπορούν να επεκταθούν σε κάτι πιο γενικό (π.χ. ολόκληρο το δείγμα-πληθυσμό) (Borenstein et al., 2010).

1.4 Ετερογένεια

Μετά τους υπολογισμούς που έχουν προηγηθεί με βάση τα 2 μοντέλα, το επόμενο βήμα περιλαμβάνει τον υπολογισμό του συγκεντρωτικού αποτελέσματος που αποτελεί ένα από τα κύρια σημεία σε μία μετά-ανάλυση. Πολύ σημαντικό ρόλο όμως παίζει και η μεταβολή που παρατηρείται στη συμπεριφορά των πραγματικών μεγεθών επίδρασης και η οποία μπορεί να είναι είτε μικρή είτε μεγάλη (Borenstein et al., 2010). Αυτή **η μεταβολή μεταξύ των πραγματικών μεγεθών επίδρασης** αποτελεί την **ετερογένεια** μεταξύ των μελετών και όπως είναι φυσικό παρατηρείται στα μοντέλα τυχαίων επιδράσεων, στα οποία, όπως έχει αναφερθεί, διαφέρουν τα πραγματικά μεγέθη επίδρασης ανά μελέτη (Borenstein et al., 2009, σελ. 105).

Σε πρακτικό επίπεδο παρ' όλα αυτά, προκύπτει πάντα μια εκτίμηση του πραγματικού μεγέθους επίδρασης (το παρατηρούμενο μέγεθος επίδρασης δηλαδή), η οποία διαφέρει τόσο μεταξύ του πραγματικού μεγέθους επίδρασης της μελέτης όσο και μεταξύ των εκτιμήσεων των μεγεθών επίδρασης των υπολοίπων μελετών (Borenstein et al., 2009, σελ. 105). Ο προσδιορισμός της ετερογένειας έγκειται: 1) στην **μεταξύ των μελετών μεταβλητότητα** (μέτρο της οποίας είναι η μεταξύ των μελετών διακύμανση) και 2) στην **διαφορά μεταξύ των παρατηρούμενων μεγεθών επίδρασης**, κάνοντας την υπόθεση ότι η μοναδική πηγή μεταβλητότητας είναι το δειγματοληπτικό σφάλμα (δηλαδή αν ίσχυε το μοντέλο σταθερών επιδράσεων) (Borenstein et al., 2009, σελ. 108).

Εν συνεχεία, γίνεται χρήση στατιστικών μεθόδων για την ποσοτικοποίηση της ετερογένειας, καθώς και δοκιμές που θα εξάγουν συμπεράσματα για το αν η τελευταία είναι στατιστικά σημαντική (δηλαδή η ύπαρξη ή μη αυτής).

Όσον αφορά την υπολογιστική προσέγγιση της ετερογένειας, οι μαθηματικοί τύποι θα περιγραφούν στην επόμενη ενότητα. Η διαδικασία ξεκινάει με τον υπολογισμό **των σταθμισμένων τετραγωνικών αποκλίσεων (Weighted Sum of Squares ή Q)**, ενώ η διαφορά ανάμεσα στην παρατηρούμενη τιμή του Q και την αναμενόμενη τιμή του αντικατοπτρίζει (δεν είναι ίσο με) την ετερογένεια των πραγματικών μεγεθών επίδρασης (Borenstein et al., 2009, σελ. 110).

Το Q λειτουργεί επίσης και ως τεστ σημαντικότητας για την ύπαρξη ετερογένειας. Σαν μηδενική υπόθεση θεωρούμε το γεγονός ότι οι μελέτες μοιράζονται το ίδιο μέγεθος επίδρασης. Αυτό υπολογιστικά σημαίνει ότι η μηδενική υπόθεση είναι ότι το Q ακολουθεί μια κεντρική χ^2 κατανομή με $k - 1$ βαθμούς ελευθερίας, όπου k είναι ο αριθμός των μελετών. Στην περίπτωση που το ***p-value*** **είναι μικρότερο από 0.05 ή 0.1** (ανάλογα με την τιμή που έχει οριστεί από τους ερευνητές), **η μηδενική υπόθεση απορρίπτεται** (Borenstein et al., 2009, σελ.112).

Όσον αφορά τα τεστ σημαντικότητας (όπως το Q), αυτά έχουν από την φύση τους χαμηλή ισχύ κα άρα τα μη σημαντικά *p-value* δεν θα πρέπει να οδηγούν στην υπόθεση ότι υπάρχει **ομοιογένεια** μεταξύ των μεγεθών επίδρασης των μελετών (Hardy & Thompson, 1998). Η στατιστική ισχύς αυτών καθορίζεται τόσο από τον αριθμό όσο και από το μέγεθος των μελετών (Borenstein et al., 2009, σελ. 112-113).

Παρά ταύτα, χρησιμοποιώντας το Q, μπορεί να προκύψει το στατιστικό $\boldsymbol{I^2}$, το οποίο είναι ο λόγος ανάμεσα στην **περίσσεια μεταβλητότητα (excess dispersion)**, που είναι η διαφορά ανάμεσα στο παρατηρούμενο και το αναμενόμενο Q, και την **συνολική μεταβλητότητα ($\boldsymbol{Q}$)** των μελετών (Borenstein et al., 2009, σελ. 117; Higgins et al., 2003). Επομένως, αυτό με την σειρά του εκφράζει **τι ποσοστό της συνολικής διακύμανσης αντιστοιχεί στην ετερογένεια**, χωρίς να υπολογίζεται (όπως και με το Q) η αριθμητική τιμή της (Borenstein et al., 2009, σελ. 120; Higgins et al., 2003). Σε αντίθεση με το Q, το I^2 δεν εξαρτάται από τον αριθμό των μελετών (Borenstein et al., 2009, σελ. 120).

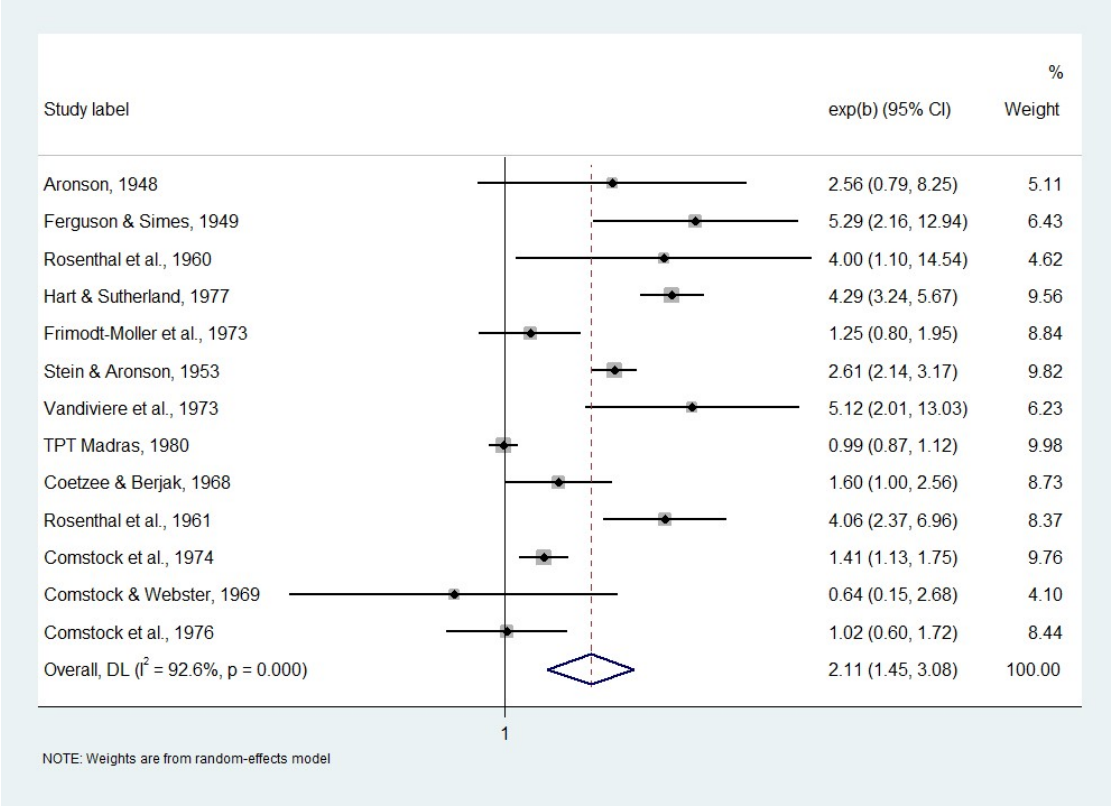
Τέλος, ως $\boldsymbol{\tau^2}$ ορίζεται η διακύμανση των πραγματικών μεγεθών επίδρασης και η εκτίμηση αυτού του μεγέθους συμβολίζεται με $\mathbf{T^2}$ (Borenstein et al., 2009, σελ. 114). Το τ^2, όπως αναφέρθηκε παραπάνω, αποτελεί την μεταξύ των μελετών διακύμανση (μέτρο της μεταξύ των μελετών μεταβλητότητας) και σε αυτή την εργασία ο υπολογισμός του θα γίνει με την βοήθεια του **εκτιμητή Der Simonian-Laird (DL)**. Η μέθοδος των στιγμών, όπως ονομάζεται αλλιώς, δεν βασίζεται σε κάποια υπόθεση σχετικά με την κατανομή των τυχαίων επιδράσεων, παρ' όλο που το ίδιο το μοντέλο κάνει την υπόθεση ότι ακολουθούν την κανονική κατανομή (Thorlund, Wetterslev, et al., 2011). Ο συγκεκριμένος εκτιμητής φαίνεται: 1) να υποεκτιμά την μεταξύ των μελετών διακύμανση, όταν υπάρχει μεγάλη ετερογένεια (υψηλό I^2) και μικρού ή μεσαίου μεγέθους μελέτες, 2) να είναι σχεδόν αμερόληπτος για μετά-αναλύσεις που

εμπεριέχουν σημαντικό αριθμό (10-20 και πάνω) μελετών διαφόρων μεγεθών (μικρών και μεγάλων στην ίδια μετά-ανάλυση) και 3) να υπερεκτιμά την μεταξύ των μελετών διακύμανση για μικρότερο αριθμό μελετών (Langan et al., 2019). Ο DL εκτιμητής είναι ο πιο απλός, από υπολογιστικής πλευράς, ενώ υπάρχουν και άλλοι ίσως και καταλληλότεροι εκτιμητές, οι οποίοι όμως δεν θα αναφερθούν στα πλαίσια της παρούσας εργασίας (Thorlund, Wetterslev, et al., 2011).

1.5 Γράφημα forest

Στην υπό-ενότητα αυτή θα παρουσιαστούν τα βασικά στοιχεία ενός γραφήματος forest (forest plot). Σαν παράδειγμα θα χρησιμοποιήσουμε το forest plot που προκύπτει από την ανάλυση του σετ δεδομένων (dataset) bcg. Το BCG είναι ένα εμβόλιο που προστατεύει κατά της φυματίωσης, μόλυνση που προκαλείται από βακτήριο και προσβάλλει κυρίως τα πνευμόνια αλλά και άλλα μέρη του σώματος. Η αποτελεσματικότητα του εμβολίου μπορεί να ποικίλλει, γεγονός που οδήγησε σε υποθέσεις όπως αυτή των Berkey et al. (1995), οι οποίοι αναφέρουν ότι τα διαφορετικά γεωγραφικά πλάτη πιθανόν αιτιολογούν τις μεταβολές αυτές (χωρίς να σημαίνει ότι αποτελούν την πραγματική αιτία).

Σαν μέγεθος επίδρασης έχει επιλεγεί ο λόγος σχετικών πιθανοτήτων (ως συμβάν ορίστηκε η ανταπόκριση στην θεραπεία του εμβολίου) και σαν μοντέλο επιλέχθηκε το μοντέλο των τυχαίων επιδράσεων. Στο συγκεκριμένο παράδειγμα οι επιμέρους μελέτες ανήκουν στην κατηγορία των τυχαιοποιημένων κλινικών δοκιμών.



Σχήμα 3. Γράφημα forest για το σετ δεδομένων BCG

Στα αριστερά του διαγράμματος παρουσιάζονται οι δείκτες των μελετών. Στο παράδειγμα εμφανίζονται τα ονόματα των επιστημόνων που είναι συντελεστές στην μετά-ανάλυση (μπορεί να παρουσιάζονται και επιπλέον δείκτες όπως π.χ. η χώρα διεξαγωγής). Στην δεξιά πλευρά διαφαίνονται οι αριθμητικές τιμές των μεγεθών επίδρασης κάθε μελέτης, ενώ οι τιμές εντός των παρενθέσεων, αποτελούν τα **διαστήματα εμπιστοσύνης** των τελευταίων. Τα διαστήματα εμπιστοσύνης δείχνουν το εύρος τιμών μέσα στο οποίο πιθανότατα βρίσκεται το πραγματικό μέγεθος επίδρασης (ή οποιασδήποτε άλλης πληθυσμιακής παραμέτρου όπως π.χ. του μέσου) και εξαρτάται άμεσα από το **επίπεδο εμπιστοσύνης** που ορίζεται. Το επίπεδο εμπιστοσύνης (εδώ 95% όπως φαίνεται πάνω δεξιά στο γράφημα) εκφράζει αυτήν ακριβώς την πιθανότητα. Τέλος, στην τέρμα δεξιά στήλη εμφανίζονται τα βάρη που ανατίθενται σε κάθε μελέτη ξεχωριστά.

Όσον αφορά τα γραφικά στοιχεία του γραφήματος, οι οριζόντιες γραμμές εκτείνονται σε όλο το εύρος του διαστήματος εμπιστοσύνης του μεγέθους επίδρασης κάθε μελέτης. Οι μικρότερες σε έκταση γραμμές υποδηλώνουν ένα μικρότερο διάστημα

εμπιστοσύνης, ενώ το αντίθετο ισχύει για τις μακρύτερες. Το εύρος των διαστημάτων εμπιστοσύνης υποδηλώνει και την **ακρίβεια (precision)** μιας μελέτης, με τα μικρότερα διαστήματα να ανήκουν στις πιο ακριβείς μελέτες και το αντίστροφο. Στις πιο ακριβείς μελέτες, ακόμη, ανατίθενται μεγαλύτερα βάρη λόγω της μικρότερης διακύμανσης του σφάλματος εντός της μελέτης. Στην μέση ακριβώς των γραμμών βρίσκεται ένα μικρό σημάδι σε σχήμα ρόμβου που δείχνει το μέγεθος επίδρασης, το οποίο περικλείεται από ένα γκρίζο τετράγωνο. Το μέγεθος αυτού είναι αναλογικό με την αριθμητική τιμή του βάρους που έχει ανατεθεί στην συγκεκριμένη μελέτη. Στο τέλος του γραφήματος βρίσκεται το **συγκεντρωτικό αποτέλεσμα,** σε σχήμα ρόμβου, το οποίο προκύπτει από τον συνδυασμό των μεγεθών επίδρασης, αλλά και των βαρών κάθε μελέτης (για υπολογισμό βλ. παρακάτω). Το συγκεντρωτικό αποτέλεσμα είναι από τους κυριότερους στόχους σε μια μετά-ανάλυση καθώς εκπληρώνει τον σκοπό της που είναι η εξαγωγή ενός στατιστικά ισχυρότερου συμπεράσματος από τις επιμέρους μελέτες της.

Προχωρώντας στη συνέχεια, στην ερμηνεία του συγκεκριμένου διαγράμματος forest, το μέγεθος επίδρασης που επιλέχθηκε αποτελεί το **αποτέλεσμα της παρέμβασης (intervention effect ή treatment effect)**, εφόσον οι μελέτες μας είναι τυχαιοποιημένες κλινικές δοκιμές. Με βάση όσα έχουν αναφερθεί παραπάνω στο κεφάλαιο «Εισαγωγή», μπορεί να γίνει ερμηνεία των λόγων σχετικών πιθανοτήτων. Για την κάθε μελέτη ισχύει:

- Αν OR = 1, τότε δεν παρατηρείται διαφορά μεταξύ των 2 θεραπευτικών προσεγγίσεων.
- Αν OR > 1, η θεραπεία με το εμβόλιο BCG φαίνεται αποτελεσματικότερη.
- Αν OR < 1, η συμβατική θεραπεία (ή μη) της ομάδας ελέγχου δείχνει πιο αποτελεσματική από το εμβόλιο BCG.

Καθοριστικό σημείο και για τις 3 περιπτώσεις αποτελεί (για το συγκεκριμένο μέγεθος επίδρασης) η τιμή 1, στην οποία στηρίζεται η μηδενική υπόθεση. Αν δηλαδή στο διάστημα εμπιστοσύνης του λόγου σχετικών πιθανοτήτων εμπεριέχεται η τιμή 1, το p-value είναι μεγαλύτερο του 0.05, με αποτέλεσμα να μην απορρίπτεται η H_0 και άρα η μελέτη να θεωρείται μη στατιστικά σημαντική (Borenstein et al., 2009, σελ.11-12). Για τους ίδιους λόγους όταν δεν ανήκει η τιμή 1 στο διάστημα εμπιστοσύνης, το p-value είναι μικρότερο από 0.05 και άρα η μελέτη θεωρείται στατιστικά σημαντική.

Από τις 13 μελέτες, οι 6 δεν είναι στατιστικά σημαντικές. Παρ' όλα αυτά όμως, το συγκεντρωτικό αποτέλεσμα δείχνει να είναι στατιστικά σημαντικό με OR = 2.11 και χωρίς να περιλαμβάνει το 1 στο διάστημα εμπιστοσύνης του. Τέλος, όσον αφορά το διάστημα εμπιστοσύνης του συγκεντρωτικού αποτελέσματος, με μια πρώτη ματιά φαίνεται να είναι μεγάλου εύρους.

Εκτός από το συγκεντρωτικό αποτέλεσμα, είναι σημαντική η εκτίμηση και η ερμηνεία της ετερογένειας. Για το σκοπό αυτό παρατίθεται κάτω αριστερά στο διάγραμμα το στατιστικό $\boldsymbol{I^2}$. Σύμφωνα με τους Higgins et al. (2003), όταν η αριθμητική τιμή του είναι 75% και πάνω, το $\boldsymbol{I^2}$ φανερώνει ότι ένα μεγάλο μέρος της παρατηρούμενης μεταβλητότητας οφείλεται στην ετερογένεια των πραγματικών μεγεθών επίδρασης και συνεπώς θα πρέπει να προσδιοριστούν τα αίτια της (Borenstein et al., 2009, σελ. 119; Higgins et al., 2003). Ένας τρόπος, εύκολα αντιληπτός, θα ήταν να γίνει ανάλυση υποομάδων (subgroup analysis) με βάση κάποιον παράγοντα που πιθανότατα δικαιολογεί την ετερογένεια (Borenstein et al., 2009, σελ. 119).

1.6 Ανάλυση συνεχών δεδομένων

Η προηγούμενη μετά-ανάλυση βασίστηκε σε δεδομένα από τα οποία θα μπορούσαν να εξαχθούν σχετικά εύκολα οποιαδήποτε από τα γνωστά μεγέθη επίδρασης. Στα πλαίσια αυτής της εργασίας, ωστόσο, θα εξεταστούν σετ δεδομένων τα οποία αναφέρουν μόνο συνεχή δεδομένα και τα μεγέθη επίδρασης θα πρέπει να προκύψουν από εκεί. Πιο συγκεκριμένα, στα πλαίσια μιας τυχαιοποιημένης κλινικής δοκιμής οι ασθενείς θα ομαδοποιούνται σε treatment responders και treatment non-responders ανάλογα με κάποια ποσοτική μεταβολή στην συνεχή μεταβλητή. Η ανάλυση αυτή ονομάζεται **ανάλυση των ανταποκρινόμενων (responder analysis)** (Da Costa et al., 2012).

Το κύριο πρόβλημα των μελετών που αναφέρουν συνεχή δεδομένα είναι ότι μπορεί να διαφέρουν μεταξύ τους ως προς τον τρόπο διεξαγωγής και πιθανότατα ως προς τις μονάδες μέτρησης της συνεχούς μεταβλητής (Da Costa et al., 2012). Για να

μπορέσουν να συνδυαστούν οι μελέτες σε μια μετά-ανάλυση χρησιμοποιούνται οι τυποποιημένες διαφορές μέσων. Το συγκεκριμένο μέγεθος επίδρασης στερείται ερμηνευσιμότητας από τους ερευνητές, παρ' όλο που ανάγει τα δεδομένα σε μια κοινή κλίμακα (ως τυποποιημένο μέγεθος μετριέται σε τυπικές αποκλίσεις) (Da Costa et al., 2012).

Εδώ πρέπει να τονιστεί ότι η συνήθης πρακτική για τα συνεχή δεδομένα περιλαμβάνει τη διχοτόμηση του σετ δεδομένων με βάση μία αριθμητική τιμή (score ή threshold) που ορίζει ο ερευνητής και την κατηγοριοποίηση των ασθενών σε ανταποκρίνοντες ή μη στην θεραπεία (Da Costa et al., 2012). Η τιμή που θα χρησιμοποιηθεί μπορεί να προέρχεται από την εμπειρία του ερευνητή και συνεπώς να έχει κάποια κλινική βάση, ειδάλλως παρατηρείται ο διαχωρισμός γύρω από την **διάμεσο** (την μεσαία σε σειρά παρατήρηση, **median split**) (Nuzzo, 2019). Παραδείγματα τέτοιων score που έχουν κλινική βάση υπάρχουν για πολλές παθολογικές καταστάσεις, όπως υψηλή συστολική πίεση (140 mm Hg και πάνω), υψηλή διαστολική πίεση (90 mm Hg και πάνω), τιμές του δείκτη BMI που κατατάσσουν ένα άτομο ως υπέρβαρο κ.α.. Παρακάτω θα δείξουμε ένα υποθετικό παράδειγμα με 2 ομάδες, συνεχή δεδομένα και μία αυθαίρετη τιμή αποκοπής (την τιμή 9 για το παράδειγμα).

Πίνακας 2. Παράδειγμα συνεχών δεδομένων για 2 ομάδες ατόμων με τιμή αποκοπής, cutoff score = 9 (κόκκινη γραμμή).

Treatment Group	**Control group**
7.1149	7.5530
7.8414	7.8315
…	…
8.9820	8.9739
9.0314	9.0177
…	…
11.7012	11.8082
11.8376	12.1028

Οπότε θα προκύψει ένας 2x2 πίνακας ως εξής (θεωρούμε ότι το εξεταζόμενο ενδεχόμενο είναι να βρίσκεται μια τιμή κάτω από την τιμή αποκοπής):

Πίνακας 3. 2x2 πίνακας συνάφειας μετά από διχοτόμηση των συνεχών δεδομένων.

	Treatment Responders	**Treatment Non-Responders**	**Total**
Treatment Group	21	79	100
Control Group	12	88	100
Total	33	167	200

Παρ' όλη την απλότητα που την διακατέχει ως μέθοδο, η διχοτόμηση φαίνεται να έχει αρνητικό αντίκτυπο στα αποτελέσματα των μελετών καθώς και στην ερμηνεία αυτών. Αρχικά, έχει αποδειχθεί μαθηματικά ότι 158 διχοτομημένες παρατηρήσεις ισοδυναμούν με 100 παρατηρήσεις συνεχών δεδομένων, και αυτό γιατί η διχοτόμηση του σετ δεδομένων προκαλεί **απώλεια πληροφορίας (Fisher information)** (Fedorov et al., 2009). Συγκεκριμένα, η παραπάνω διατύπωση αναφέρεται στην καλύτερη δυνατή περίπτωση όπου η διαδικασία της διχοτόμησης γίνεται χρησιμοποιώντας την διάμεσο των παρατηρήσεων ως τιμή αποκοπής (Fedorov et al., 2009). Κατά συνέπεια, η απώλεια πληροφορίας οδηγεί σε μείωση της **στατιστικής ισχύος** (δηλαδή τον εντοπισμό ενός στατιστικά σημαντικού μεγέθους επίδρασης, όταν αυτό όντως υπάρχει), κάτι το οποίο έχει επαληθευτεί από συγκρίσεις της διχοτόμησης με άλλες στατιστικές μεθόδους (γραμμική παλινδρόμηση, συσχέτιση μεταξύ 2 μεταβλητών) σε προσομοιωμένα δεδομένα (simulated data) (MacCallum et al., 2002; Nuzzo, 2019). Σε 2^{ο} χρόνο, η δημιουργία αυστηρά 2 διακριτών υποομάδων, μειώνει την ερμηνευσιμότητα καθώς οι παρατηρήσεις μπορούν να ανήκουν είτε στην μία είτε στην άλλη ομάδα και δεν γίνεται λόγος για κάποιες πιθανές ενδιάμεσες κατηγορίες (Nuzzo, 2019).

Σε αυτό το σημείο είναι δυνατόν να προκύψουν διάφορα ερωτήματα-προβλήματα όπως:

- Είναι δυνατόν η εκτίμηση του OR να έχει μεγαλύτερη ακρίβεια (άρα μικρότερη διακύμανση);

- Πως αξιοποιούμε τα δεδομένα αν μια μελέτη αναφέρει μόνο δειγματικούς μέσους και τυπικές αποκλίσεις;
- Στις μελέτες που έχει προηγηθεί διχοτόμηση δεν είναι πάντα απαραίτητο να έχει χρησιμοποιηθεί η ίδια τιμή αποκοπής. Τότε όμως πως θα προκύψουν μεγέθη επίδρασης τα οποία να είναι κατάλληλα για μετά-ανάλυση;

Παρακάτω αναφέρονται 4 μέθοδοι, οι οποίες αξιοποιούν μέσους και τυπικές αποκλίσεις (και επακολούθως τις SMD) για να εξάγουν μεγέθη επίδρασης που προσεγγίζουν τα αντίστοιχα μεγέθη επίδρασης διχοτομημένων δεδομένων (approximated effect sizes) (Da Costa et al., 2012).

1.7 Μέθοδοι μετατροπής συνεχών δεδομένων σε μεγέθη επίδρασης

Σύμφωνα με τους **Hasselblad και Hedges**, μπορεί να εξαχθεί ο νεπέριος λογάριθμος του OR και το τυπικό του σφάλμα από τις SMD και το τυπικό σφάλμα αυτών. Βασικές προϋποθέσεις είναι: 1) οι παρατηρήσεις των 2 ομάδων να ακολουθούν την **λογιστική κατανομή** και 2) οι αντίστοιχες διακυμάνσεις των ομάδων να είναι ίσες. Σημαντικό ρόλο παίζει η ποσότητα 1.81, η οποία αποτελεί (προσεγγιστικά) την τυπική απόκλιση της **τυπικής λογιστικής κατανομής** (Chinn, 2000; Da Costa et al., 2012; Hasselblad & Hedges, 1995).

Παρόμοια υπολογιστικά, με την παραπάνω μέθοδο είναι και αυτή των **Cox και Snell**, η οποία χρησιμοποιεί την τιμή 1.65 για τους αντίστοιχους υπολογισμούς, με την προϋπόθεση ότι τα δεδομένα ακολουθούν την κανονική κατανομή και οι διακυμάνσεις μεταξύ των ομάδων είναι ίσες μεταξύ τους (Anzures-Cabrera et al., 2011; Cox and Snell, 1989; Da Costa et al., 2012)

Διαφορετική προσεγγιστικά από τις 2 παραπάνω μεθόδους, είναι η μέθοδος του **Furukawa.** Κύριο συστατικό της μεθόδου είναι ο καθορισμός (και υπολογισμός) του **κινδύνου της ομάδας ελέγχου** (με την χρήση των SMD). Στον υπολογισμό του κινδύνου συμπεριλαμβάνεται και η τιμή αποκοπής (cutoff score) που θα

χρησιμοποιούνταν για μια απευθείας διχοτόμηση του συνόλου δεδομένων των 2 ομάδων. Σαν γεγονός, σε αυτή την εργασία, θα θεωρείται η περίπτωση που οι αριθμητικές τιμές του ασθενή **ξεπεράσουν** την τιμή αποκοπής. Με βάση τον κίνδυνο της ομάδας ελέγχου, υπολογίζεται ο κίνδυνος της ομάδας θεραπείας και έπειτα εξάγονται το Odds Ratio και το τυπικό σφάλμα αυτού (Da Costa et al., 2012; Furukawa et al., 2005; Furukawa & Leucht, 2011).

Βασισμένη στον κίνδυνο των 2 ομάδων είναι και η μέθοδος του **Suissa**. Ο κίνδυνος της ομάδας ελέγχου είναι ίδιος με την προηγούμενη μέθοδο, ενώ ο κίνδυνος της ομάδας θεραπείας διαφέρει από αυτήν, ως προς τον υπολογισμό του. Επιπλέον, η συγκεκριμένη μέθοδος διαφέρει από την προηγούμενη ως προς τον υπολογισμό του τυπικού σφάλματος του Odds Ratio. Αυτό γίνεται με την εισαγωγή της **διακύμανσης του κινδύνου** των 2 ομάδων στην διαδικασία υπολογισμού του. Με την σειρά της η διακύμανση του κινδύνου χρησιμοποιείται για να δικαιολογήσει την διαφορά ανάμεσα στον αναμενόμενο και παρατηρούμενο αριθμό συμβάντων (**number of events**), σε κάθε ομάδα (Anzures-Cabrera et al., 2011; Da Costa et al., 2012; Suissa, 1991).

Παρακάτω, στο κεφάλαιο «Αποτελέσματα» θα εξετάσουμε την σχέση μεταξύ των 4 μεθόδων και της διχοτόμησης χρησιμοποιώντας το στατιστικό της **Ασυμπτωτικής Σχετικής Αποτελεσματικότητας** (**Asymptotic Relative Efficiency, ARE)**, όπου θα συγκριθούν οι διακυμάνσεις των εκτιμητών (των logOR δηλαδή) που προκύπτουν από κάθε μέθοδο, πάντα σε σχέση με την μέθοδο της διχοτόμησης (Bagos et al., 2011; Riffenburgh, 2012).

2. Μέθοδοι

2.1 Εργαλεία της μετά-ανάλυσης

Σε πρώτο χρόνο, θα παρουσιαστούν οι μαθηματικές μέθοδοι που αφορούν την μετά-ανάλυση και έπειτα οι μαθηματικοί τύποι από τις 5 ομάδες επιστημόνων, για την μετατροπή των SMD σε μεγέθη επίδρασης.

Αρχικά, για τα μοντέλα τυχαίων επιδράσεων, το **συγκεντρωτικό αποτέλεσμα** σύμφωνα με τους Borenstein et al. (2010), υπολογίζεται ως εξής:

$$\widehat{M_w} = \frac{\sum_{i=1}^{k} \widehat{W}_i^* Y_i}{\sum_{i=1}^{k} \widehat{W}_i^*} \quad (2.1)$$

όπου $\widehat{W}_i^*$ είναι το βάρος της i-οστής μελέτης (όπως έχει οριστεί στο προηγούμενο κεφάλαιο για το μοντέλο τυχαίων επιδράσεων) και Y_i το παρατηρούμενο μέγεθος επίδρασης.

Για να υπολογιστεί το διάστημα εμπιστοσύνης του συγκεντρωτικού αποτελέσματος, πρέπει πρώτα να υπολογιστεί η **διακύμανση του σφάλματος της μετά-ανάλυσης,** σύμφωνα με τον ακόλουθο τύπο (Borenstein et al., 2010):

$$V_M = \frac{1}{\sum_{i=1}^{k} \widehat{W}_i^*} \quad (2.2)$$

Το **τυπικό σφάλμα** της εκτίμησης μιας παραμέτρου είναι η αναμενόμενη διαφορά του από την αντίστοιχη παράμετρο (π.χ. διαφορά δειγματικού από τον πραγματικό μέσο του πληθυσμού). Για το συγκεντρωτικό αποτέλεσμα, το τελευταίο ορίζεται ως (Borenstein et al., 2010):

$$SE_{\widehat{M_w}} = \sqrt{V_M} \quad (2.3)$$

Και τελικά το **διάστημα εμπιστοσύνης** του συγκεντρωτικού αποτελέσματος θα έχει την μορφή (Borenstein et al., 2010):

$$\left(\widehat{M_w} - 1.96 * SE_{\widehat{M_w}}\,,\ \widehat{M_w} + 1.96 * SE_{\widehat{M_w}}\right) \tag{2.4}$$

Η τιμή 1.96 προκύπτει για επίπεδο εμπιστοσύνης 95% (και άρα για alpha value ίσο με 1 – 0.95 = 0.05). Για διαφορετικά επίπεδα σημαντικότητας οι τιμές υπολογίζονται από το $Z_{1-a/2}$, που είναι το (1– α/2)-οστό εκατοστημόριο της τυπικής κανονικής κατανομής και υπολογίζεται από τον σχετικό πίνακα (δεν περιέχεται στην εργασία) (Borenstein et al., 2010).

Για τον υπολογισμό του *p-value*, θα χρειαστεί πρώτα να οριστεί η **αθροιστική συνάρτηση κατανομής** της τυπικής κατανομής:

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx \tag{2.5}$$

όπου, όσον αφορά την τιμή του ολοκληρώματος, συνήθως υπολογίζεται προσεγγιστικά και με την βοήθεια κατάλληλων προγραμμάτων (Yerukala & Kumar Boiroju, 2015).

Τελικά το *p-value* για τον έλεγχο υποθέσεων είναι ίσο με (Borenstein et al., 2010):

$$p = 2 * [1 - \Phi(|Z|)] \tag{2.6}$$

με το στατιστικό Z (δεν πρέπει να συγχέεται με το Z που αναφέρεται στην τυπική κανονική κατανομή) να υπολογίζεται από τον ακόλουθο τύπο (Borenstein et al., 2010):

$$Z = \frac{\widehat{M_w} - X_0}{SE_{\widehat{M_w}}} \tag{2.7}$$

όπου X_0 μια ορισμένη τιμή για το πραγματικό μέγεθος επίδρασης υπό την μηδενική υπόθεση.

Συνεχίζοντας στα μεγέθη που αφορούν την ετερογένεια, η παρατηρούμενη τιμή του στατιστικού μεγέθους Q ισούται με (Higgins & Thompson, 2002):

$$Q = \sum_{i=1}^{k} \widehat{W}_i * \left(Y_i - \widehat{M_w}\right)^2 \tag{2.8}$$

ενώ με μετασχηματισμούς μπορεί να προκύψει και ο ακόλουθος υπολογισμός (Borenstein et al., 2009, σελ. 109)

$$Q = \sum_{i=1}^{k} \widehat{W}_i * Y_i - \frac{\left(\sum_{i=1}^{k} \widehat{W}_i * Y_i\right)^2}{\sum_{i=1}^{k} \widehat{W}_i} \qquad (2.9)$$

όπου $\widehat{W}_i$ είναι το βάρος της i-οστής μελέτης (όπως έχει οριστεί στο προηγούμενο κεφάλαιο για το μοντέλο σταθερών επιδράσεων) και η αναμενόμενη τιμή του Q να είναι οι **βαθμοί ελευθερίας** (degrees of freedom) (Borenstein et al., 2009, σελ.110):

$$df = k - 1 \qquad (2.10)$$

όπου k ο αριθμός των μελετών.

Για το I^2, χρησιμοποιούνται οι αναμενόμενες και παρατηρούμενες τιμές του Q ως εξής (Borenstein et al., 2009, σελ. 117):

$$I^2 = \max\left\{0, \left(\frac{Q - df}{Q}\right)\right\} * (100)\% \qquad (2.11)$$

Τέλος, με την χρήση του Der Simonian-Laird εκτιμητή, η εκτίμηση της μεταξύ των μελετών διακύμανσης των πραγματικών μεγεθών επίδρασης ορίζεται ως (Dersimonian & Laird, 1986):

$$\hat{\tau}_{DL}{}^2 = \max\left\{0, \frac{Q - (k-1)}{\sum \widehat{W}_i - \frac{\sum \widehat{W}_i{}^2}{\sum \widehat{W}_i}}\right\} \qquad (2.12)$$

2.2 Μεγέθη επίδρασης

Τα μεγέθη επίδρασης που αναφέρθηκαν στην ενότητα «Εισαγωγή» συνοψίζονται εδώ μαζί με τις διακυμάνσεις και τις τυπικές αποκλίσεις τους. Επιπλέον, ακολουθεί μια γενικευμένη εκδοχή του πίνακα συνάφειας για τα μεγέθη του λόγου σχετικών πιθανοτήτων και της διαφοράς κινδύνου (Borenstein et al., 2009, σελ. 33).

Πίνακας 4. Πίνακας αποτελεσμάτων ανά ομάδα θεραπείας.

	Events	***Non-Events***	***Total***
Treated	*A*	*B*	*A+B*
Control	*C*	*D*	*C+D*
Total	*A+C*	*B+D*	*A+B+C+D*

Για τον **λόγο σχετικών πιθανοτήτων** ισχύουν τα παρακάτω (Borenstein et al., 2009 σελ. 36-37):

Πίνακας 5. Στατιστικά μέτρα λόγου σχετικών πιθανοτήτων.

Log of Odds ratio	***Variance***	***Standard* error**
$\ln \boldsymbol{OR} = \ln \frac{\boldsymbol{A} * \boldsymbol{D}}{\boldsymbol{B} * \boldsymbol{C}}$	$V_{LogOR} = \frac{1}{A} + \frac{1}{B} + \frac{1}{C} + \frac{1}{D}$	$SE_{LogOR} = \sqrt{V_{LogOR}}$

Όπως γίνεται άμεσα αντιληπτό, για τους σκοπούς της μετά-ανάλυσης γίνεται χρήση του νεπερίου λογαρίθμου του OR, ενώ στο τέλος της διαδικασίας ανάλογα με το αν έχει επιλεγεί τα δεδομένα να εμφανίζονται σε εκθετική μορφή, απεικονίζονται οι αυτούσιες τιμές των ORs των μελετών (όπως στο forest plot που δημιουργήθηκε πιο πάνω).

Για την **διαφορά κινδύνου** δεν χρειάζεται μετατροπή σε λογαριθμική κλίμακα και επομένως ισχύουν τα ακόλουθα (Borenstein et al., 2009, σελ. 37-38):

Πίνακας 6. Στατιστικά μέτρα διαφοράς κινδύνου.

Risk Difference	***Variance***	***Standard Error***
$\boldsymbol{RD = \frac{A}{A+B} - \frac{C}{C+D}}$	$V_{RD} = \frac{A*B}{(A+B)^3} + \frac{C*D}{(C+D)^3}$	$SE_{RD} = \sqrt{V_{RD}}$

Για λόγους απλοποίησης, θα θέσουμε A+B = n_1 και C+D = n_2. Επίσης, ο δείκτης 1 θα συμβολίζει την ομάδα θεραπείας και ο δείκτης 2 την ομάδα ελέγχου.

Για τον υπολογισμό των τυποποιημένων διαφορών μέσου, επαναλαμβάνεται ότι πρέπει πρώτα να υπολογιστεί η συγκεντρωτική διακύμανση (Da Costa et al., 2012):

$$sd_{pooled} = \sqrt{\frac{(n_1 - 1) * sd^2{}_1 + (n_2 - 1) * sd^2{}_2}{(n_1 + n_2) - 2}} \qquad (2.13)$$

οπότε οι τυποποιημένες διαφορές των μέσων ισούνται (Da Costa et al., 2012):

$$SMD = \frac{mean_1 - mean_2}{sd_{pooled}} \qquad (2.14)$$

όπου $mean_1$ και $mean_2$ οι δειγματικοί μέσοι των αντίστοιχων ομάδων.

Η διακύμανση και το τυπικό σφάλμα εν συνεχεία. ορίζονται αντίστοιχα (Borenstein et al., 2009, σελ. 27):

Πίνακας 7. Στατιστικά μέτρα τυποποιημένων διαφορών μέσου.

Variance	***Standard Error***
$\boldsymbol{V_{SMD} = \frac{n_1 + n_2}{n_1 * n_2} + \frac{SMD^2}{2 * (n_1 + n_2)}}$	$SE_{SMD} = \sqrt{V_{SMD}}$

2.3 Σχεδιασμός των μεθόδων

Οι υπολογισμοί των ζητούμενων μεθόδων της εργασίας παρατίθενται παρακάτω:

1. Μέθοδος των Hasselblad και Hedges (HH)

Σύμφωνα με τους Hasselblad & Hedges (1995):

- $\ln OR = -1.81 * SMD$ (2.15)
- $SE_{\ln OR} = 1.81 * SE_{SMD}$ (2.16)

2. Μέθοδος των Cox και Snell (CS)

Σύμφωνα με τους Cox και Snell (1986):

- $\ln OR = -1.65 * SMD$ (2.17)
- $SE_{\ln OR} = 1.65 * SE_{SMD}$ (2.18)

3. Μέθοδος του Furukawa (F)

Σύμφωνα με τους Da Costa et al. (2012); Furukawa et al. (2005); Thorlund et al. (2011), αρχικά υπολογίζονται οι τύποι για τους κινδύνους των 2 ομάδων, της ομάδας ελέγχου και της ομάδας θεραπείας, αντίστοιχα ως:

$$risk_{con} = \Phi\left(\frac{C - mean_{con}}{sd_{con}}\right) \quad (2.19)$$

$$risk_{exp} = 1 - \Phi\left(SMD - \Phi^{-1}(risk_{con})\right) \quad (2.20)$$

όπου C το σκορ αποκοπής (cut-off score), Φ η αθροιστική κατανομή της τυπικής κανονικής κατανομής και Φ^{-1} η αντίστροφη συνάρτηση της Φ. Να σημειωθεί ότι εξετάζεται το ενδεχόμενο όπου η συνεχή μεταβλητή είναι **μικρότερη** από το cutoff score που έχει οριστεί. Σε περίπτωση που εξετάζεται το αντίθετο ενδεχόμενο θα πρέπει να υπολογιστεί και η ανάλογη πιθανότητα (Thorlund, Walter, et al., 2011).

Για τον υπολογισμό των σχετικών πιθανοτήτων (odds) της κάθε ομάδας ισχύει:

$$odds_g = \frac{risk_g}{1 - risk_g} \tag{2.21}$$

όπου g ο δείκτης της κάθε ομάδας. Έπειτα μπορεί να υπολογιστεί το $\ln OR$ με την γνωστή μεθοδολογία. Όσο για το τυπικό του σφάλμα, υπολογίζεται ως εξής:

$$SE_{\ln OR} = \sqrt{\frac{1}{e_{exp}} + \frac{1}{e_{con}} + \frac{1}{n_{exp}} + \frac{1}{n_{con}}} \tag{2.22}$$

όπου n_{exp} και n_{con} αναπαριστούν τους αριθμούς των ατόμων της κάθε ομάδας και e_{exp} και e_{con} τους εκτιμώμενους αριθμούς των συμβάντων. Οι τελευταίοι υπολογίζονται για κάθε ομάδα με βάση τον αντίστοιχο κίνδυνο ως:

$$e_g = risk_g * n_g \tag{2.23}$$

όπου g ο δείκτης για την κάθε ομάδα ξεχωριστά

4. Η μέθοδος του Suissa (S)

Σύμφωνα με τον Suissa (1991):

Ο μαθηματικός τύπος για τον υπολογισμό του κινδύνου της ομάδας ελέγχου είναι ο ίδιος με την μέθοδο του Furukawa. Όσον αφορά τον κίνδυνο της ομάδας θεραπείας, ο υπολογισμός του είναι παρόμοιος με την ομάδα ελέγχου, δηλαδή:

$$risk_{exp} = \Phi\left(\frac{C - mean_{exp}}{sd_{exp}}\right) \tag{2.24}$$

Για τον υπολογισμό του τυπικού σφάλματος του $\ln OR$ - όπως υπολογίζεται από την μέθοδο του Furukawa - πρέπει πρώτα να υπολογιστούν οι διακυμάνσεις των αντίστοιχων κινδύνων. Οπότε έχουμε:

$$var(risk_g) = \frac{\left(1 + \frac{1}{2}\left(\frac{C - mean_g}{sd_g}\right)^2\right)\left(\exp\left(-\frac{1}{2}\left(\frac{C - mean_g}{sd_g}\right)^2\right)\right)^2}{2\pi n_g} \tag{2.25}$$

Και στη συνέχεια υπολογίζεται το $SE_{\ln OR}$:

$$SE_{\ln OR} = \sqrt{\frac{var(risk_{exp})}{\left(risk_{exp}(1-risk_{exp})\right)^2} + \frac{var(risk_{con})}{\left(risk_{con}(1-risk_{con})\right)^2}} \quad (2.26)$$

Τέλος, η Ασυμπτωτική Σχετική Αποτελεσματικότητα (ARE) θα υπολογιστεί στα πλαίσια της εργασίας ως εξής:

$$ARE = \frac{VarLogOR(F|S|HH|CS)}{VarLogOR(D)} \quad (2.27)$$

όπου F, S, HH, CS συντομογραφίες μίας εκ των μεθόδων που θα εξετάζεται σε σχέση με την μέθοδο της διχοτόμησης (D).

3. Αποτελέσματα

3.1 Υλοποίηση του προγράμματος

3.1.1 Κυρίως πρόγραμμα - κώδικας

Αρχικά, για την δημιουργία του προγράμματος χρησιμοποιήθηκε το στατιστικό πακέτο Stata (έκδοση 13.1). Με την βοήθεια της προγραμματιστικής του γλώσσας και των μαθηματικών του συναρτήσεων, προχωρήσαμε στην υλοποίηση των 4 μεθόδων που αναφέρθηκαν παραπάνω (F, S, CS, HH) καθώς και της συμβατικής διχοτόμησης (συμβολίζεται με D). Το πρόγραμμα θα δέχεται δεδομένα (μέσο, τυπική απόκλιση, αριθμό ατόμων) από κάθε ομάδα (ομάδα ελέγχου και ομάδα θεραπείας), έπειτα θα εφαρμόζεται στα δεδομένα η μέθοδος που έχει επιλέξει ο χρήστης, για την εξαγωγή των ORs, και τέλος θα γίνεται μετά-ανάλυση των μελετών. Όσον αφορά την μετά-ανάλυση, αυτή θα γίνεται με την βοήθεια των ήδη υπαρχουσών υλοποιήσεων της

metan και της ameta. Επιπλέον, οι εντολές αυτές θα χρησιμοποιούν το μοντέλο τυχαίων επιδράσεων και τον εκτιμητή DerSimonian και Laird.

Παρακάτω παρατίθεται ο κώδικας του προγράμματος μαζί με κάποια βοηθητικά σχόλια:

```
*********************************************************
********************
* Methods to convert continuous outcomes into OR and
subsequent meta-analysis
*********************************************************
********************

program define metasmd, rclass
          version 13.1

     syntax varlist(min = 6 max = 6 numeric) [if] [in],
method(string) choice(string) value(real) [adjSMD]
[meta(string)] [study(string)] [model(string)] [eform]

     tempvar  meanexp meancon ncon nexp riskcon riskexp
sdpooled sdcon sdexp SMD SESMD varriskcon varriskexp lnOR
seLnOR r1 r2 nr1 nr2 nr3 nr4 score

     set more off

     qui gen `meancon' = `1'
     qui gen `sdcon' = `2'
     qui gen `ncon' = `3'
     qui gen `meanexp' = `4'
     qui gen `sdexp' = `5'
     qui gen `nexp' = `6'
```

```
     if ("`choice'" == "risk") {
          if (`value' >= 0 & `value' <= 1){
               qui gen `score' = `meancon' +
invnormal(`value')*`sdcon'
          }
          else {
               di "Choose a value meaningful for our
noble cause"
               exit(111)
          }
     }
     else if ("`choice'" == "score") {
          qui gen `score' = `value'
     }

     di ""
     di
"++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
++++"
     di "You must be truly desperate to come to me for
help...      +"
     di "Anyway, guess I have to help you with your
complex task... +"
     di "Let's convert some SMD's to Odds Ratios, shall
we?         +"
     di "Analyzing...
+"
     di
"++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
++++"
     di ""

     qui gen  `sdpooled' = sqrt(((`nexp'-1)*(`sdexp'^2) +
(`ncon'-1)*(`sdcon'^2))/((`nexp'+`ncon') - 2))
```

```
     if ("`adjSMD'" ~= "") {
          qui gen `SMD' = ((`meanexp' -
`meancon')/`sdpooled')*(1 - 3/(4*(`ncon' + `nexp') - 9))
          qui gen `SESMD' = sqrt((`ncon' +
`nexp')/(`ncon' * `nexp') + (`SMD'^2)/(2*(`ncon' + `nexp'
- 3.94)))
     }
     else {
          qui gen  `SMD' = (`meanexp' -
`meancon')/`sdpooled'
          qui gen  `SESMD' = sqrt((`ncon' +
`nexp')/(`ncon' * `nexp') + (`SMD'^2)/(2*(`ncon' +
`nexp')))
     }

     if ("`method'" == "d") {
          if ("`choice'" == "risk") {
               qui gen `r1' = `value'
          }
          else if ("`choice'" == "score") {
               qui gen `r1' = normal((`score' -
`meancon') / `sdcon')
         }

          qui gen `r2' = normal((`score' - `meanexp') /
`sdexp')

          qui gen `nr1' = (`r1' * `ncon') //events
control
          qui gen `nr2' = (`r2' * `nexp') //events
experimental
          qui gen `nr3' = `ncon' - `nr1'
          qui gen `nr4' = `nexp' - `nr2'
```

```
          qui gen `lnOR' = log((`r2'*(1-`r1'))/(`r1'*(1-
`r2')))
          qui gen `seLnOR' = sqrt(1/`nr2' + 1/`nr1' +
1/`nr3' + 1/`nr4')

     }
*****************************Furukawa's
Method*****************************************************

     else if ("`method'" == "f") {

          if ("`choice'" == "risk") {
               qui gen `riskcon' = `value'
          }
          else if ("`choice'" == "score") {
               qui gen `riskcon' = normal((`score' -
`meancon') / `sdcon')
         }
          qui gen `riskexp' = 1 - normal(`SMD' -
invnormal(`riskcon'))

          qui gen `lnOR' = log(((`riskexp')/(1 -
`riskexp')) / ((`riskcon') / (1 - `riskcon')))
          qui gen `seLnOR' = sqrt((1/(`riskexp'*`nexp'))
+ (1/(`riskcon'*`ncon')) + (1/`nexp') + (1/`ncon'))

     }

****************************Suissa's
Method*****************************************************
**

     else if ("`method'" == "s") {
```

```
          if ("`choice'" == "risk") {
               qui gen `riskcon' = `value'
          }
          else if ("`choice'" == "score") {
               qui gen `riskcon' = normal((`score' -
`meancon') / `sdcon')
         }
          qui gen `riskexp' = normal((`score' -
`meanexp') / `sdexp')

          qui gen  `varriskcon' = (1 + 0.5*(((`score' -
`meancon') / `sdcon')^2))*((exp(-0.5*(((`score' -
`meancon') / `sdcon')^2)))^2)/(2*c(pi)*`ncon')
          qui gen  `varriskexp' = (1 + 0.5*(((`score' -
`meanexp') / `sdexp')^2))*((exp(-0.5*(((`score' -
`meanexp') / `sdexp')^2)))^2)/(2*c(pi)*`nexp')

          qui gen `lnOR' = log(((`riskexp')/(1 -
`riskexp')) / ((`riskcon') / (1 - `riskcon')))
          qui gen `seLnOR' =
sqrt((`varriskexp'/((`riskexp'*(1 - `riskexp'))^2) +
(`varriskcon'/((`riskcon'*(1 - `riskcon'))^2))))

     }

**********Hasselblad & Hedges'
Method**************************************************

     else if ("`method'" == "hh") {

          qui gen `lnOR' = `SMD' * -1.81
          qui gen `seLnOR' = `SESMD' * 1.81
```

```
     }


     //Mean scores of each group should follow logistic
distribution and variances have to be equal between
groups

**********Cox & Snell's
Method**************************************************
**

     else if ("`method'" == "cs") {

          qui gen `lnOR' = `SMD' * -1.65
          qui gen `seLnOR' = `SESMD' * 1.65

     }


     //Mean scores of each group should follow normal
distribution and variances have to be equal between
groups

**********Meta-Analysis
Section*************************************************
**

     if ("`model'" == "fixed") {
          if ("`eform'" ~= "") {
               if ("`study'" ~= "") {
                    metan `lnOR' `seLnOR', fixed graph
eform lcols(`study')
               }
               else {
```

```
                    metan `lnOR' `seLnOR', fixed graph
eform
               }
          }
          else {
               if ("`study'" ~= "") {
                    metan `lnOR' `seLnOR', fixed graph
lcols(`study')
               }
               else {
                    metan `lnOR' `seLnOR', fixed graph
               }
          }
     }
     else if ("`model'" == "random") {
          if ("`eform'" ~= "") {
               if ("`meta'" == "ameta") {
                    if ("`study'" ~= "") {
                         ameta `lnOR' `seLnOR',
method(dl) graph eform study(`study')
                    }
                    else {
                         ameta `lnOR' `seLnOR',
method(dl) graph eform
                    }
               }
               else if ("`meta'" == "metan" | "`meta'" ==
"") {
                    if ("`study'" ~= "") {
                         metan `lnOR' `seLnOR', random
eform graph lcols(`study')
                    }
                    else {
```

```
                        metan `lnOR' `seLnOR', random
eform graph
                    }
               }
          }
          else {
               if ("`meta'" == "ameta") {
                    if ("`study'" ~= "") {
                        ameta `lnOR' `seLnOR',
method(dl) graph study(`study')
                    }
                    else {
                        ameta `lnOR' `seLnOR',
method(dl) graph
                    }
               }
               else if ("`meta'" == "metan" | "`meta'" ==
"") {
                    if ("`study'" ~= "") {
                        metan `lnOR' `seLnOR', random
graph lcols(`study')
                    }
                    else {
                        metan `lnOR' `seLnOR', random
graph
                    }
               }
          }
     }

end
```

3.1.2 Περιγραφή του κυρίως προγράμματος

Το πρόγραμμα ξεκινάει αναφέροντας την έκδοση του Stata (13.1) που χρησιμοποιήθηκε κατά την συγγραφή του κώδικα και την εντολή "define, rclass" που δείχνει ότι το πρόγραμμα επιστρέφει τιμές με το που τελειώσει η εκτέλεση του. Συντακτικά μιλώντας, ο χρήστης πρέπει να εισάγει υποχρεωτικά 6 μεταβλητές (μέσους, τυπικές αποκλίσεις και αριθμό ατόμων κάθε ομάδας) σε συνδυασμό με 3 υποχρεωτικές επιλογές (method, choice, value). Η επιλογή **method** αναφέρεται στην μέθοδο που θα επιλέξει ο χρήστης να χρησιμοποιήσει (συμβολισμοί επιλογών: d, f, s, hh, cs). Η επιλογή **choice** έχει να κάνει με το αν ο χρήστης επιλέγει να χρησιμοποιήσει ένα κοινό cutoff score (επιλογή "score") ή κοινή τιμή κινδύνου (επιλογή "risk"). Τέλος, η επιλογή **value** αναφέρεται στην αριθμητική τιμή της εκάστοτε επιλογής (βλ. choice).

Επιπλέον υπάρχουν και προαιρετικές επιλογές που δεν επηρεάζουν την ροή του προγράμματος. Η επιλογή **adjSMD**, δίνει στον χρήστη την επιλογή να χρησιμοποιήσει την προσαρμοσμένη εκδοχή των SMD, για μικρό σύνολο ατόμων. Η επιλογή **model** δίνει στον χρήστη την επιλογή να χρησιμοποιήσει είτε το μοντέλο σταθερών επιδράσεων είτε το μοντέλο τυχαίων επιδράσεων (χρησιμοποιείται ως εκτιμητής ο DerSimonian - Laird). Επιπλέον η επιλογή **eform** χρησιμοποιείται σε περίπτωση που είναι επιθυμητό τα logORs να εμφανίζονται με την εκθετική μορφή τους (ως ORs δηλαδή). Για τα γραφήματα forest, η επιλογή **study** χρησιμοποιείται για να εμφανίζονται τα ονόματα των μελετών στο τελικό γράφημα. Τέλος, η επιλογή **meta** αφορά την χρήση των εντολών της **metan** και της **ameta** για την διεξαγωγή της μετά-ανάλυσης (σαν προεπιλογή χρησιμοποιείται η metan).

Στην αρχή του προγράμματος γίνονται οι απαραίτητοι υπολογισμοί και οι απαραίτητες αναθέσεις σε προσωρινές μεταβλητές (**tempvar**), οι οποίες υπάρχουν μόνο για το χρονικό διάστημα που εκτελείται το πρόγραμμα. Σε δεύτερο χρόνο, η ίδια διαδικασία επαναλαμβάνεται και στον τομέα του προγράμματος που αφορά τις μεθόδους (π.χ. if "method" == "f" {}). Στο τέλος κάθε μεθόδου υπάρχουν οι εντολές της μετά-ανάλυσης.

3.2.1 Γραφήματα σύγκρισης Ασυμπτωτικής Σχετικής Αποτελεσματικότητας των μεθόδων

Προκειμένου να συγκρίνουμε τις μεθόδους μεταξύ τους, εστιάσαμε στην διακύμανση του νεπέριου λογαρίθμου του λόγου σχετικών πιθανοτήτων (variance of log odds ratio). Η διακύμανση αυτή προκύπτει αν τετραγωνίσουμε το τυπικό σφάλμα του log odds ratio, που προκύπτει από κάθε μέθοδο. Όσον αφορά την Ασυμπτωτική Σχετική Αποτελεσματικότητα (Asymptotic Relative Efficiency, ARE), ο αριθμητής θα περιλαμβάνει την διακύμανση της συμβατικής διχοτόμησης και ο παρονομαστής την διακύμανση των υπόλοιπων τεσσάρων μεθόδων. Συνεπώς θα έχουμε τέσσερις διαφορετικές ARE (ARE_F, ARE_S, ARE_CS, ARE_HH).

Για να εξεταστεί πως μεταβάλλονται οι παραπάνω ARE, θα πρέπει να ελέγξουμε την συμπεριφορά τους ενώ μεταβάλλουμε άλλες παραμέτρους που έχουν να κάνουν με τον υπολογισμό τους. Συγκεκριμένα, θα μεταβάλλουμε το Odds Ratio, τον κίνδυνο της ομάδας ελέγχου, και τον αριθμό ατόμων των 2 ομάδων (ελέγχου και θεραπείας).

Ο κίνδυνος θα λάβει διακριτές τιμές από 0.05 έως 0.95, το Odds Ratio επίσης διακριτές τιμές από 1.1 έως 4, και η καθεμία ομάδα θα αποτελείται από είτε 10 είτε 100 άτομα. Συνολικά, για κάθε συνδυασμό αριθμού ατόμων των 2 ομάδων (4 συνδυασμοί σε πρώτη φάση), έχουμε 440 συνδυασμούς Odds Ratio και risk.

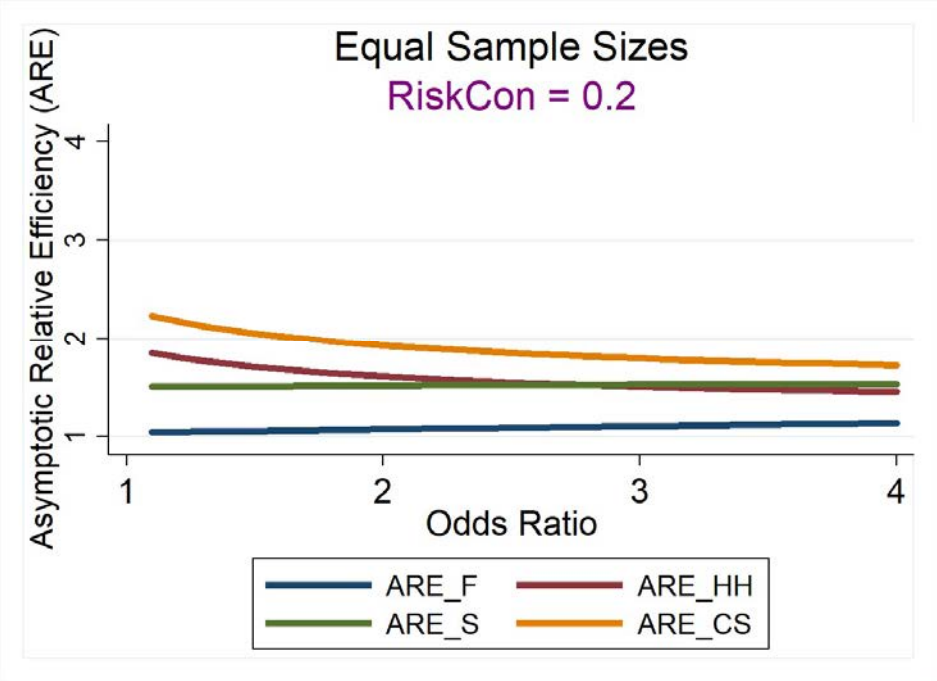
Κάποια στατιστικά στοιχεία όπως π.χ. ο κίνδυνος της ομάδας ελέγχου, η διακύμανση του κινδύνου στην μέθοδο του Suissa κ.α., υπολογίζονται κανονικά από τις ήδη υπάρχουσες παραμέτρους που αναφέρθηκαν παραπάνω.

Από τους παραπάνω συνδυασμούς κρατήσαμε αυτούς που προκύπτουν για $OR \geq 1.1$ και $OR \leq 4$, και risk ομάδας ελέγχου 0.2, 0.5 και 0.8. Όσον αφορά τους συνδυασμούς των αριθμών ατόμων κάθε ομάδας, καταλήξαμε σε 3 από τους 4 λογικούς συνδυασμούς που προκύπτουν, καθώς για τον ίδιο αριθμό ατόμων σε κάθε ομάδα προκύπτουν ακριβώς ίδιες τιμές για το ARE (οι αριθμητικές τιμές των υπόλοιπων παραμέτρων μπορεί να αλλάζουν, αλλά δεν εξετάζονται στην εργασία).

Τα 9 συνολικά διαγράμματα που προκύπτουν παρουσιάζονται παρακάτω:

Για ίσο αριθμό δείγματος και στις 2 ομάδες (Equal Sample Sizes):

Για κίνδυνο ομάδας ελέγχου = 0.2:

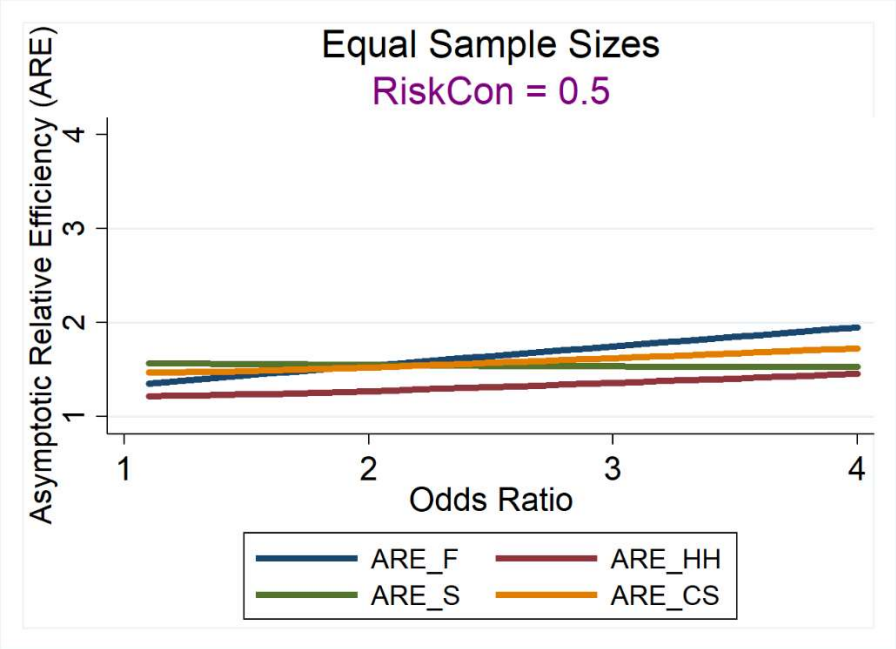


Σχήμα 4. Διάγραμμα διασποράς ARE – OR για RiskCon = 0.2 και ίσα μεγέθη δειγμάτων.

Παρατηρούμε ότι:

- Η ARE_CS υπερισχύει των υπολοίπων σε όλο το εύρος του OR.
- Η ARE_F μειονεκτεί των υπολοίπων σε όλο το εύρος του OR.
- Η ARE_S > ARE_S για OR > 2.6, ενώ ισχύει το αντίθετο για OR ≤ 2.6
-

Για κίνδυνο ομάδας ελέγχου = 0.5:

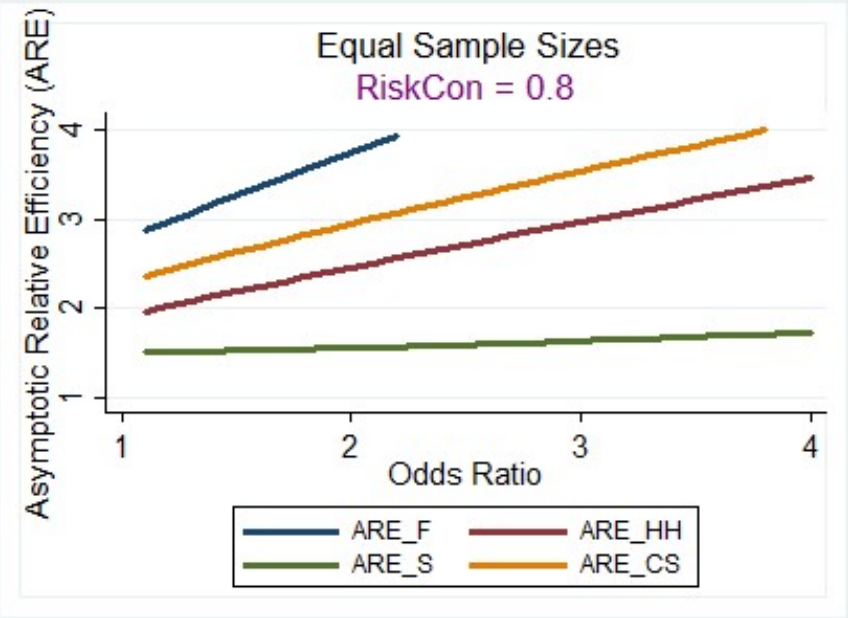


Σχήμα 5. Διάγραμμα διασποράς ARE – OR για RiskCon = 0.5 και ίσα μεγέθη δειγμάτων.

Παρατηρούμε ότι:

- Η ARE_S υπερισχύει έως OR ≤ 2, ενώ στο υπόλοιπο εύρος του OR υπερισχύει η ARE_F.
- Η ARE_HH μειονεκτεί των υπολοίπων σε όλο το εύρος του OR
- Η ARE_CS >ARE_F για OR ≤ 1.8, ενώ ισχύει το αντίθετο για OR >1.8

Για κίνδυνο ομάδας ελέγχου = 0.8:



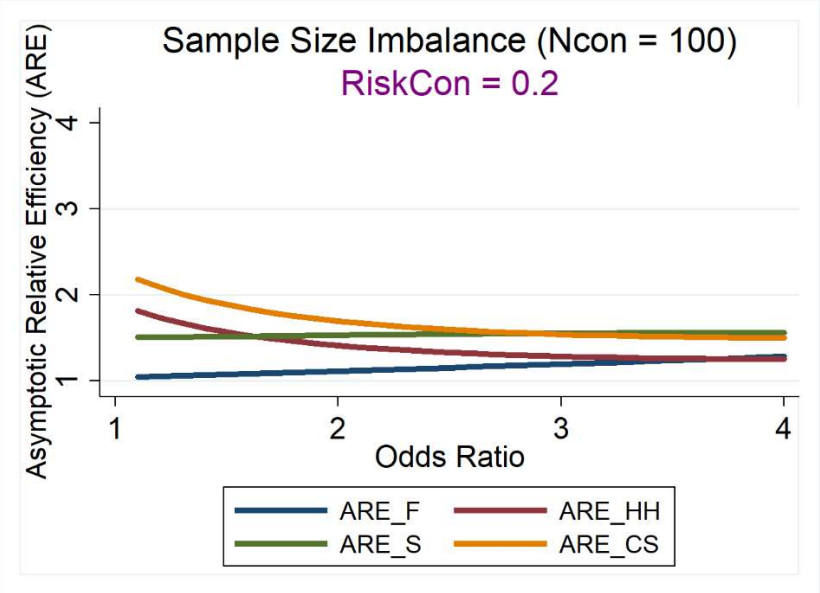
Σχήμα 6. Διάγραμμα διασποράς ARE – OR για RiskCon = 0.8 και ίσα μεγέθη δειγμάτων.

Παρατηρούμε ότι:

- Η ARE_F υπερισχύει των υπολοίπων σε όλο το εύρος του OR.
- Η ARE_S μειονεκτεί των υπολοίπων σε όλο το εύρος του OR.
- Η ARE_CS > ARE_HH σε όλο το εύρος του OR.

Για ανισορροπία στο μέγεθος δείγματος (αριθμός ατόμων στην ομάδα ελέγχου, ncon=100, αριθμός ατόμων στην ομάδα θεραπείας, nexp= 10), έχουμε:

Για κίνδυνο ομάδας ελέγχου = 0.2:

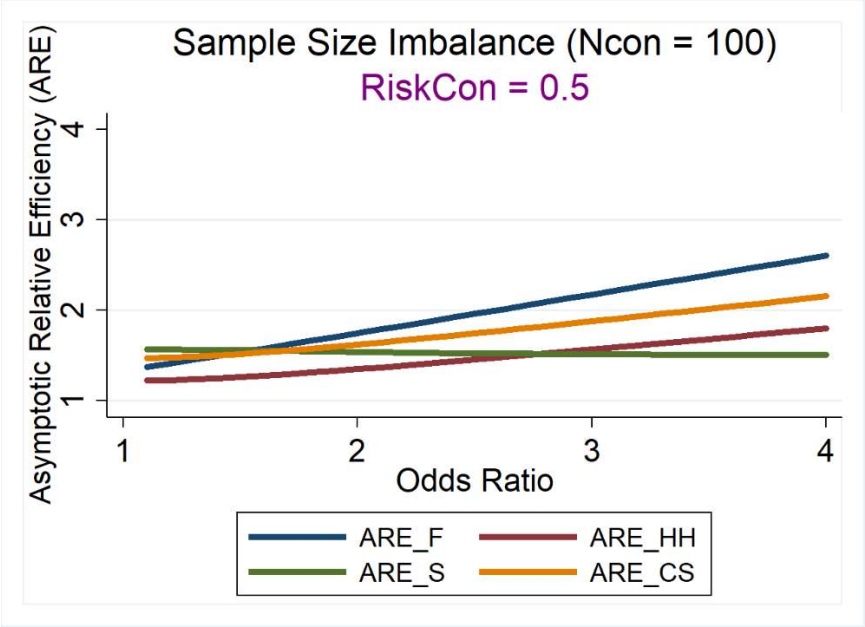


Σχήμα 7. Διάγραμμα διασποράς ARE – OR για RiskCon = 0.2 και άνισα μεγέθη δειγμάτων (ncon = 100).

Παρατηρούμε ότι:

- Η ARE_CS υπερισχύει των υπολοίπων μέχρι OR ≤ 2.8, έπειτα, υπερισχύει η ARE_S.
- Η ARE_F μειονεκτεί των υπολοίπων έως OR ≤ 3.6, έπειτα η ARE_HH μειονεκτεί των υπολοίπων.
- Η ARE_HH > ARE_S έως OR ≤ 1.6, ενώ το αντίθετο ισχύει για OR > 1.6.

Για κίνδυνο ομάδας ελέγχου = 0.5:

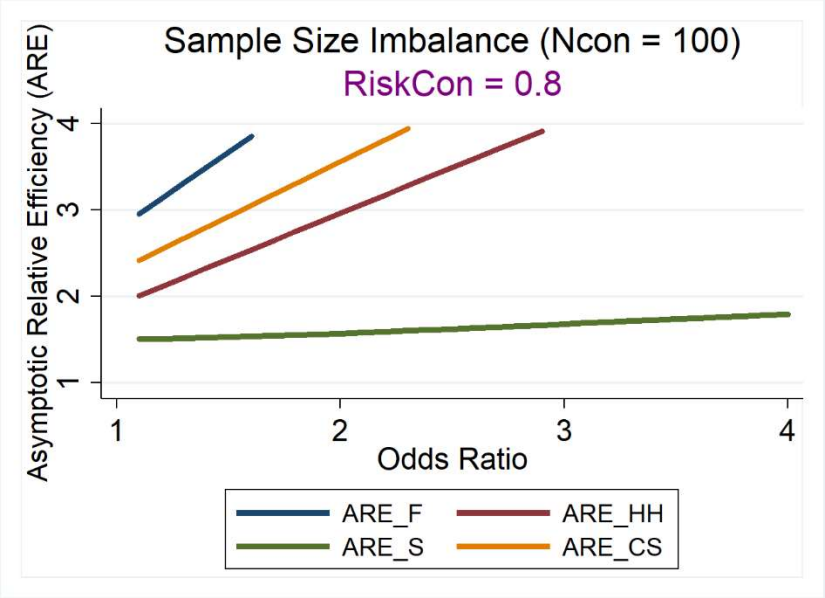


Σχήμα 8. Διάγραμμα διασποράς ARE – OR για RiskCon = 0.5 και άνισα μεγέθη δειγμάτων (ncon = 100).

Παρατηρούμε ότι:

- Η ARE_S υπερισχύει των υπολοίπων έως OR ≤ 1.5, έπειτα υπερισχύει η ARE_F.
- Η ARE_HH μειονεκτεί των υπολοίπων έως OR ≤ 2.6, έπειτα μειονεκτεί των υπολοίπων η ARE_S.
- Η ARE_CS > ARE_F για OR ≤ 1.4, ισχύει το αντίθετο για OR > 1.4.

Για κίνδυνο ομάδας ελέγχου = 0.8:



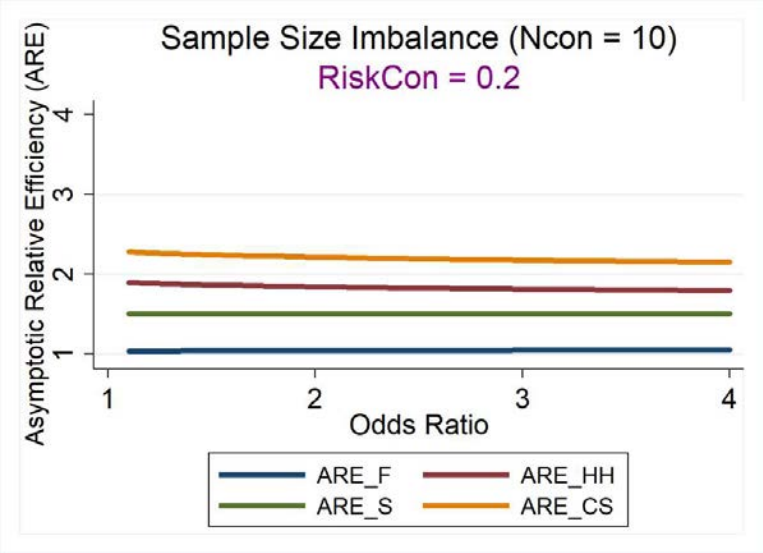
Σχήμα 9. Διάγραμμα διασποράς ARE – OR για RiskCon = 0.8 και άνισα μεγέθη δειγμάτων (ncon = 100).

Παρατηρούμε ότι:

- Η ARE_F υπερισχύει των υπολοίπων σε όλο το εύρος του OR.
- Η ARE_S μειονεκτεί των υπολοίπων σε όλο το εύρος του OR.
- Η ARE_CS > ARE_HH σε όλο το εύρος του OR.

Για ανισορροπία στο μέγεθος δείγματος (αριθμός ατόμων στην ομάδα ελέγχου, ncon=10, αριθμός ατόμων στην ομάδα θεραπείας, nexp=100), έχουμε:

Για κίνδυνο ομάδας θεραπείας = 0.2:

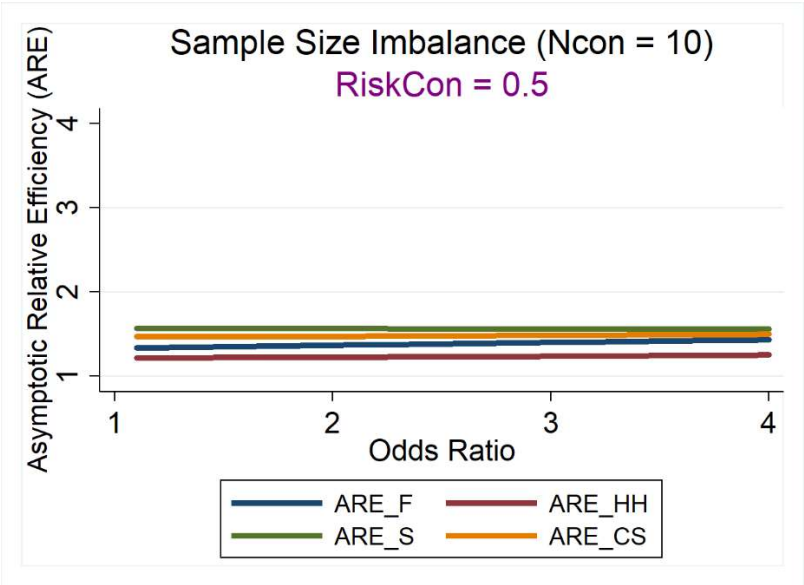


Σχήμα 10. Διάγραμμα διασποράς ARE – OR για RiskCon = 0.2 και άνισα μεγέθη δειγμάτων (ncon = 10).

Παρατηρούμε ότι:

- Η ARE_CS υπερισχύει των υπολοίπων σε όλο το εύρος του OR.
- Η ARE_F μειονεκτεί των υπολοίπων σε όλο το εύρος του OR.
- Η ARE_HH > ARE_S σε όλο το εύρος του OR.

Για κίνδυνο ομάδας ελέγχου = 0.5:

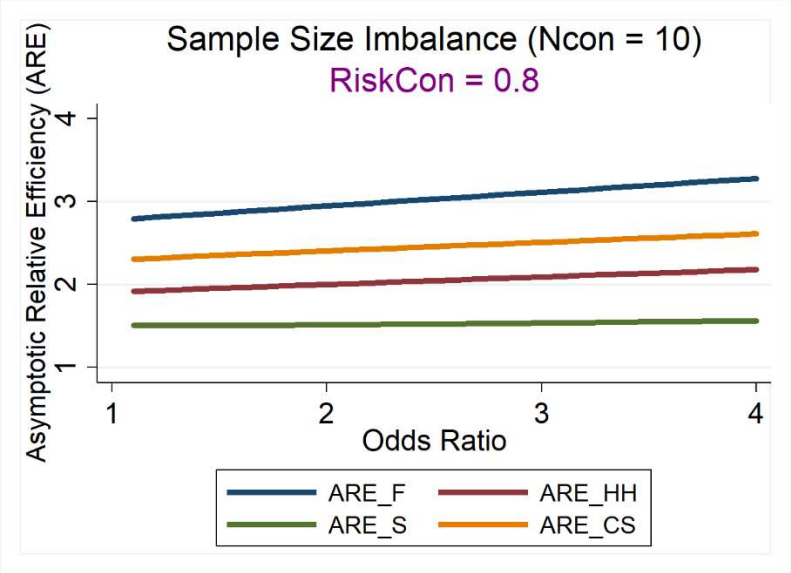


Σχήμα 11. Διάγραμμα διασποράς ARE – OR για RiskCon = 0.5 και άνισα μεγέθη δειγμάτων (ncon = 10).

Παρατηρούμε ότι:

- Η ARE_S υπερισχύει των υπολοίπων σε όλο το εύρος του OR.
- Η ARE_HH μειονεκτεί των υπολοίπων σε όλο το εύρος του OR.
- Η ARE_CS > ARE_F σε όλο το εύρος του OR.

Για κίνδυνο ομάδας ελέγχου = 0.8:



Σχήμα 12. Διάγραμμα διασποράς ARE – OR για RiskCon = 0.8 και άνισα μεγέθη δειγμάτων (ncon = 10).

Παρατηρούμε ότι:

- Η ARE_F υπερισχύει των υπολοίπων σε όλο το εύρος του OR.
- Η ARE_S μειονεκτεί των υπολοίπων σε όλο το εύρος του OR.
- Η ARE_CS > ARE_HH σε όλο το εύρος του OR.

Θα γίνει αναφορά στα διαγράμματα παρακάτω στην εργασία, στο κεφάλαιο «Συζήτηση»

3.3 Εφαρμογή των μεθόδων σε πραγματικά σύνολα δεδομένων

Σε αυτό το σημείο, θα προχωρήσουμε στην εκτέλεση του προγράμματος και θα παρουσιαστούν τα αποτελέσματα αυτού σε 3 σύνολα δεδομένων. Κάθε σύνολο δεδομένων περιέχει μελέτες που ερευνούν την επίδραση παραγόντων σε ασθενείς με διαβήτη τύπου ΙΙ, που έχουν προηγουμένως υποβληθεί σε θεραπεία με σουλφονυλουρίες. Η συνεχής μεταβλητή που εξετάζεται είναι η επί τοις εκατό μεταβολή της συγκέντρωσης της γλυκοζυλιωμένης αιμοσφαιρίνης (HbA1c). Η διαφορά στα σύνολα δεδομένων έχει να κάνει με την χρήση ενός από τους 3 διαφορετικούς παράγοντες (ροσιγλιταζόνη, μετφορμίνη, μιγλιτόλη). Η εξαγωγή των συνόλων δεδομένων έγινε από ένα μεγαλύτερο σύνολο δεδομένων, το dat.senn2013.rda που περιέχει επιπλέον μελέτες, που εστιάζουν σε μερικούς ακόμη παράγοντες, οι οποίες δεν συμπεριλαμβάνονται στην εργασία (Senn et al., 2013). Το dat.senn2013.rda περιέχεται μαζί με άλλα σετ δεδομένων, που προορίζονται για μετά-ανάλυση, στο πακέτο **metadat** το οποίο έχει δημιουργηθεί για την γλώσσα προγραμματισμού R και είναι αναρτημένο στην σελίδα του Github (https://github.com/wviechtb/metadat). Για την ανάλυση στο Stata δημιουργήθηκαν τα αντίστοιχα σετ δεδομένων με ονομασία ανάλογη με τον παράγοντα που εξετάζεται (Rosiglitazone.dta, Metformin.dta, Miglitol.dta). Για σκορ αποκοπής θα χρησιμοποιήσουμε την τιμή -1, καθώς μελέτες έχουν δείξει μείωση του κινδύνου μικροαγγειακών επιπλοκών έως 25%, οπότε σε πρώτο στάδιο φαίνεται ότι το συγκεκριμένο σκορ είναι κλινικά σημαντικό (Stratton et al., 2000).

3.3.1 Παράδειγμα 1: Επίδραση της ροσιγλιταζόνης στην συγκέντρωση της HbA1c

Προκειμένου να λάβουμε κάποιες πληροφορίες σχετικά με το συγκεκριμένο σετ δεδομένων και τις μεταβλητές που το απαρτίζουν, χρησιμοποιούμε την εντολή **describe**:

Πίνακας 8. Περιγραφή του σετ δεδομένων Rosiglitazone.

```
. describe

Contains data from C:\Users\HP\OneDrive\Υπολογιστής\DIB UTH\Rosiglitazone.dta
  obs:             6
 vars:             8                          7 Jun 2024 14:45
 size:           348
-------------------------------------------------------------------------------
              storage   display    value
variable name   type    format     label      variable label
-------------------------------------------------------------------------------
rownames        byte    %8.0g
study           str21   %21s
nc              int     %8.0g
mc              double  %12.0g
sdc             double  %12.0g
ne              int     %10.0g
me              double  %10.0g
sde             double  %10.0g
-------------------------------------------------------------------------------
Sorted by:
```

Στη συνέχεια, μπορούμε να χρησιμοποιήσουμε την εντολή **list** για να παρουσιαστούν στην οθόνη τα περιεχόμενα των μεταβλητών:

Πίνακας 9. Λίστα τιμών των μεταβλητών του σετ δεδομένων Rosiglitazone.

```
. list

     +-----------------------------------------------------------------------+
     | rownames                 study   nc     mc    sdc   ne     me    sde |
     |-----------------------------------------------------------------------|
  1. |        1          Baksi (2004)  233     .1  1.036  218   -1.2  1.112 |
  2. |        2       Davidson (2007)  116    .14  1.093  117   -1.2  1.097 |
  3. |        3       Kerenyi (2004)   154   -.14    .92  160   -.91    .99 |
  4. |        4     Rosenstock (2008)   57   -.08  1.208   59  -1.17  1.229 |
  5. |        5  Wolffenbuttel (1999)  192    .2   1.11  183   -.9    1.1 |
     |-----------------------------------------------------------------------|
  6. |        6           Zhu (2003)  105   -.4    1.3  210   -1.9   1.47 |
     +-----------------------------------------------------------------------+
```

Όσον αφορά τα περιεχόμενα:

- study: η χρονολογία που δημοσιεύθηκε η μελέτη μαζί με τον κύριο επιστήμονα στην διεξαγωγή της.
- nc, mc, sdc: οι μεταβλητές της ομάδας ελέγχου για τον αριθμό ατόμων, τον δειγματικό μέσο και την τυπική απόκλιση, αντίστοιχα.
- ne, me, sde: οι μεταβλητές της ομάδας θεραπείας για τον αριθμό ατόμων, τον δειγματικό μέσο και την τυπική απόκλιση, αντίστοιχα.

Από την στιγμή που δεν λείπουν δεδομένα για την ομαλή λειτουργία του προγράμματος, μπορούμε να το εκτελέσουμε χρησιμοποιώντας την εντολή **metasmd**. Ένα παράδειγμα εκτέλεσης της εντολής, χρησιμοποιώντας την μέθοδο του Furukawa, το μοντέλο τυχαίων επιδράσεων και την metan, θα έχει αυτή την μορφή:

```
metasmd mc sdc nc me sde ne, method(f) choice(score)
value(-1) model(random) study(Studies)
```

Στην οθόνη εμφανίζεται η μετά-ανάλυση μαζί με κάποιες πληροφορίες, τις οποίες θα συνοψίσουμε για κάθε dataset σε αντίστοιχους πίνακες. Από την εκτέλεση της παραπάνω εντολής, προκύπτει η παρακάτω εικόνα της κονσόλας και το αντίστοιχο γράφημα forest:

Πίνακας 10. Πίνακας αποτελεσμάτων της μετά-ανάλυσης, χρησιμοποιώντας την metan, το μοντέλο τυχαίων επιδράσεων και την μέθοδο F, στο σετ δεδομένων Rosiglitazone.

```
Studies included: 6
Participants included: Unknown

Meta-analysis pooling of aggregate data
using the random-effects inverse-variance model
with DerSimonian-Laird estimate of tau²
```

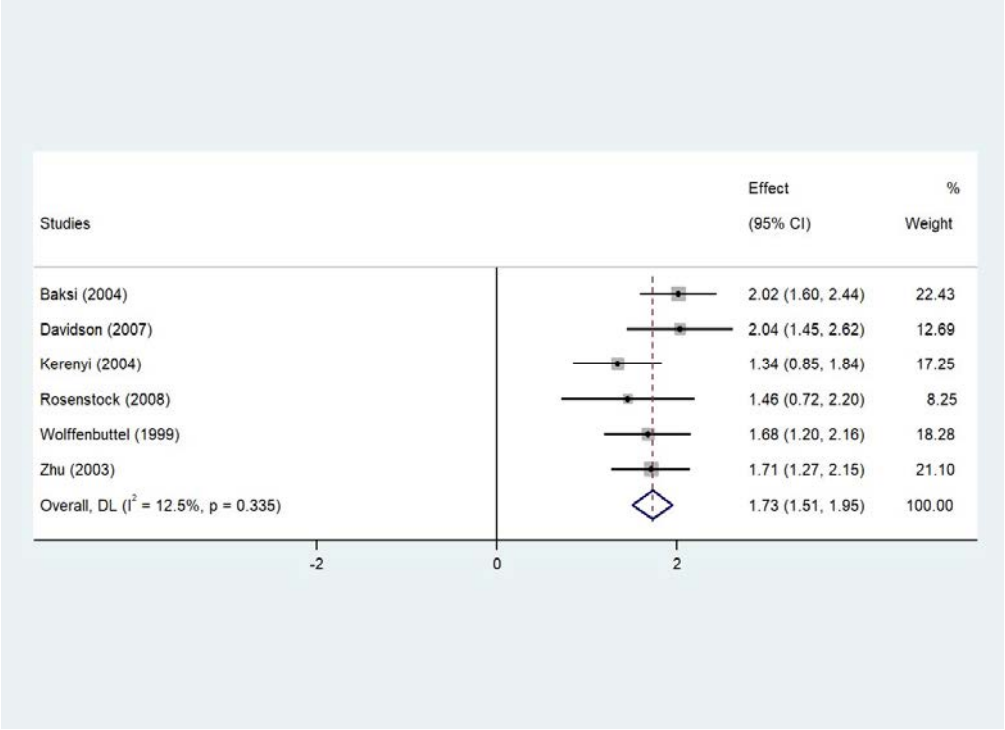
Studies	Effect	[95% Conf.	Interval]	% Weight
Baksi (2004)	2.020	1.595	2.444	22.43
Davidson (2007)	2.035	1.447	2.624	12.69
Kerenyi (2004)	1.344	0.849	1.839	17.25
Rosenstock (2008)	1.459	0.716	2.203	8.25
Wolffenbuttel (1999)	1.680	1.201	2.159	18.28
Zhu (2003)	1.713	1.273	2.153	21.10
Overall, DL	1.732	1.512	1.953	100.00

```
Test of overall effect = 0:  z =  15.398  p = 0.000

Heterogeneity measures, calculated from the data
with Conf. Intervals based on Gamma (random-effects) distribution for Q
```

Measure	Value	df	p-value
Cochran's Q	5.72	5	0.335
	−[95% Conf.	Interval]−	
H	1.069	1.000	1.713
I² (%)	12.5%	0.0%	65.9%

```
H = relative excess in Cochran's Q over its degrees-of-freedom
I² = proportion of total variation in effect estimate due to between-study
>  heterogeneity (based on Q)
```



Σχήμα 13. Γράφημα forest της metan για το σετ δεδομένων Rosiglitazone.

Ακολουθεί ο πίνακας που περιέχει τις συγκεντρωτικές πληροφορίες της μετά-ανάλυσης για κάθε μέθοδο:

Πίνακας 11. Μετά-ανάλυση των μελετών (k = 6) του σετ δεδομένων Rosiglitazone, χρησιμοποιώντας το μοντέλο τυχαίων επιδράσεων.

Method	M	SE(M)	τ^2	Q	I^2	LCI	UCI	p-value
D	1.760	0.1128	0	4.9465	0	1.539	1.981	0.0001
F	1.732	0.1125	0.0096	5.7154	0.1252	1.512	1.953	0.0000
S	1.752	0.1140	0.0264	7.6020	0.3423	1.529	1.975	0.0000
HH	1.883	0.1273	0.0434	9.1854	0.4557	1.633	2.132	0.0000
CS	1.716	0.1160	0.0360	9.1854	0.4557	1.489	1.944	0.0000

Καθώς χρησιμοποιήσαμε το μοντέλο τυχαίων επιδράσεων παρατηρείται η ύπαρξη ετερογένειας σε όλες τις μεθόδους πέραν της διχοτόμησης. Έτσι παρ' όλο που τα τυπικά σφάλματα των logOR των επιμέρους μελετών είναι μικρότερα αυτών της διχοτόμησης (ισχύει για όλες τις μεθόδους), το τυπικό σφάλμα της μετά-ανάλυσης

είναι μεγαλύτερο με την χρήση των μεθόδων (με εξαίρεση την μέθοδο F), παρά με την διχοτόμηση. Αυτό πιθανότατα να οφείλεται και στο μοντέλο που επιλέχθηκε, καθώς μπορεί να μην είναι το κατάλληλο για μικρό αριθμό μελετών (Tufanaru et al., 2015).

Επομένως, παραθέτουμε και το ανάλογο πινακάκι, χρησιμοποιώντας αυτή την φορά το μοντέλο σταθερών επιδράσεων:

Πίνακας 12. Μετά-ανάλυση των μελετών (k = 6) του σετ δεδομένων Rosiglitazone, χρησιμοποιώντας το μοντέλο σταθερών επιδράσεων.

Method	M	SE(M)	Q	I^2	LCI	UCI	p-value
D	1.760	0.1127	4.9465	0	1.539	1.981	0.0000
F	1.735	0.1045	5.7154	0.1252	1.530	1.940	0.0000
S	1.761	0.0910	7.6020	0.3423	1.582	1.939	0.0000
HH	1.891	0.0918	9.1854	0.4557	1.711	2.071	0.0000
CS	1.724	0.0836	9.1854	0.4557	1.560	1.888	0.0000

Εδώ παρατηρείται ξεκάθαρα ότι το τυπικό σφάλμα της μετά-ανάλυσης είναι μικρότερο σε κάθε περίπτωση από την συμβατική διχοτόμηση.

3.3.2 Παράδειγμα 2: Επίδραση της μετφορμίνης στην συγκέντρωση της HbA1c

Το συγκεκριμένο σετ δεδομένων περιέχει 4 μελέτες και έχει την ίδια δομή με το προηγούμενο (ίδιες μεταβλητές, διαφορετικές τιμές). Θα ξαναχρησιμοποιήσουμε την εντολή **list** για να δούμε τα περιεχόμενα των μεταβλητών:

Πίνακας 13. Λίστα τιμών των μεταβλητών του σετ δεδομένων Metformin.

```
. list

     +-------------------------------------------------------------+
     |              Studies   ne     me    sde   nc    mc    sdc |
     |-------------------------------------------------------------|
  1. |      De Fronzo (1995)  213   -1.7  1.459  209    .2  1.446 |
  2. |          Lewin (2007)  431   -.74  1.106  144     0  1.004 |
  3. |         Willms (1999)   29   -2.5   .862    1     1     1 |
  4. | Gonzalez-Ortiz (2004)   34   -1.3  1.861   37   -.9  1.803 |
     +-------------------------------------------------------------+
```

Τα αποτελέσματα του κώδικά μας, χρησιμοποιώντας την μέθοδο S, το μοντέλο τυχαίων επιδράσεων και πρόγραμμα μετά-ανάλυσης το **ameta**, θα είναι τα παρακάτω:

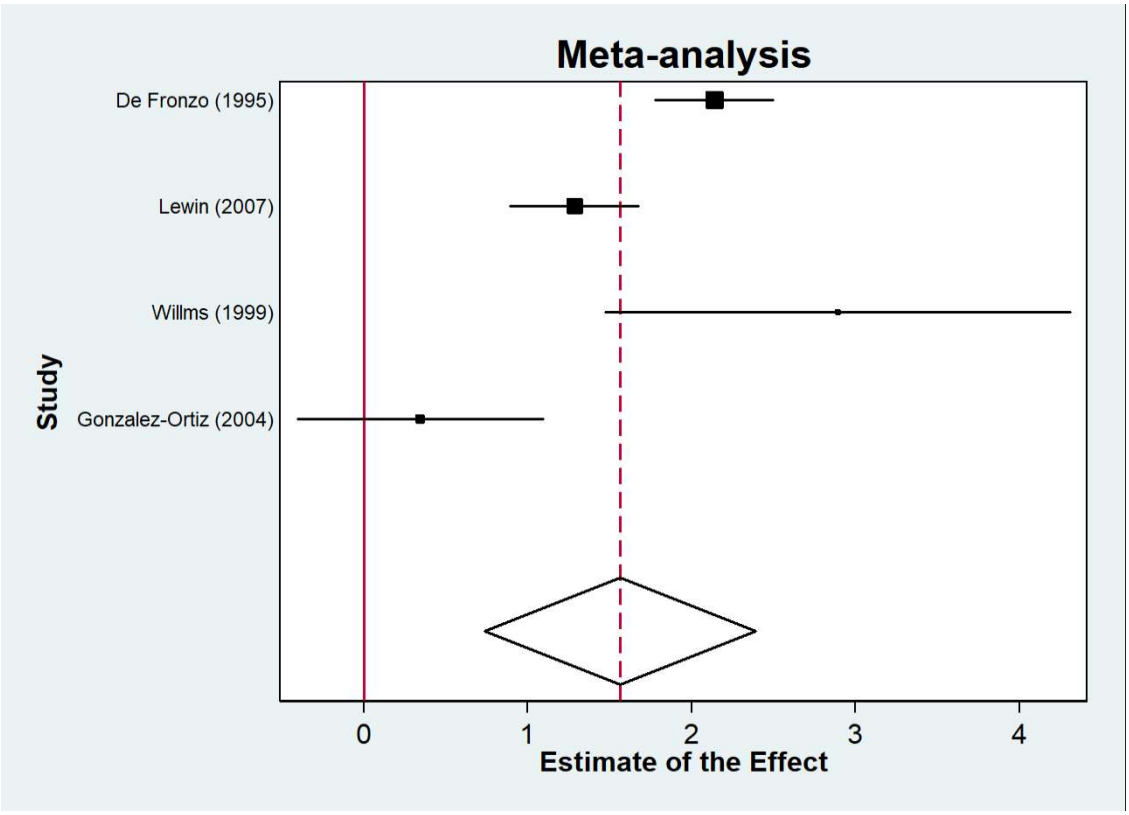
```
metasmd mc sdc nc me sde ne, method(s) choice(score)
value(-1) study(Studies) meta(ameta) model(random)
```

Τα αποτελέσματα της κονσόλας και το διάγραμμα forest φαίνονται παρακάτω:

Πίνακας 14. Πίνακας αποτελεσμάτων της μετά-ανάλυσης, χρησιμοποιώντας το πρόγραμμα ameta, το μοντέλο τυχαίων επιδράσεων και την μέθοδο S, στο σετ δεδομένων Metformin.

```
Performing analysis using DerSimonian-Laird estimation method
The variable __00000D contains the parameter estimate
The variable __00000E contains the standard error of the  estimate

             Study |   Eff. size    [% Conf. Interval]     % Weight
-------------------+-----------------------------------------------------
De Fronzo (1995)   |     2.139        1.779     2.500        29.55
Lewin (2007)       |     1.285        0.892     1.678        29.24
Willms (1999)      |     2.893        1.475     4.311        16.29
Gonzalez-Ortiz (2004)|    0.346       -0.401     1.093        24.92
-------------------+-----------------------------------------------------
 DL pooled ES      |     1.565        0.739     2.391       100.00
-------------------------------------------------------------------------
Number of studies: 4
Cochran's homogeneity test: Q= 25.0245 on 3 degrees of freedom (p= 0.000)
Estimate of between-study variance Tau-squared= 0.5674
Pooled intervention effect m=  1.5653 and its standard error se= 0.4215
I-square= 0.8801   D-square= 0.9110
95% Conf. Interval= [ 0.7392  2.3914]
z=  3.7137   p-value= 0.0002
```



Σχήμα 14. Γράφημα forest της ameta για το σετ δεδομένων Metformin.

Συγκεντρωτικά οι πληροφορίες για την μετά-ανάλυση (μοντέλο τυχαίων επιδράσεων) με κάθε μέθοδο:

Πίνακας 15. Μετά-ανάλυση των μελετών (k = 4) του σετ δεδομένων Metformin.dta, χρησιμοποιώντας το μοντέλο τυχαίων επιδράσεων.

Method	M	SE(M)	τ^2	Q	I^2	LCI	UCI	p-value
D	1.520	0.4255	0.5122	15.7974	0.8101	0.686	2.354	0.001
F	1.307	0.3810	0.4782	20.0472	0.8504	0.560	2.054	0.000
S	1.565	0.4215	0.5674	25.0245	0.8801	0.739	2.391	0.000
HH	1.423	0.4187	0.5879	27.5112	0.8910	0.602	2.243	0.000
CS	1.297	0.3817	0.4886	27.5112	0.8910	0.549	2.045	0.000

Παρατηρούμε την ύπαρξη ετερογένειας, άλλα στην συγκεκριμένη περίπτωση τα τυπικά σφάλματα όλων των μεθόδων είναι μικρότερα της διχοτόμησης. Για το μοντέλο σταθερών επιδράσεων έχουμε:

Πίνακας 16. Μετά-ανάλυση των μελετών (k = 4) του σετ δεδομένων Metformin.dta, χρησιμοποιώντας το μοντέλο σταθερών επιδράσεων.

Method	M	SE(M)	Q	I^2	LCI	UCI	p-value
D	1.628	0.1558	15.80	0.81	1.323	1.934	0.000
F	1.550	0.1359	20.05	0.85	1.284	1.817	0.000
S	1.631	0.1257	25.02	0.88	1.385	1.878	0.000
HH	1.632	0.1216	27.51	0.891	1.393	1.870	0.000
CS	1.487	0.1109	27.51	0.891	1.270	1.705	0.000

Τα συμπεράσματα για το τυπικό σφάλμα της μετά-ανάλυσης είναι όμοια με το μοντέλο τυχαίων επιδράσεων.

3.3.3 Παράδειγμα 3: Επίδραση της μιγλιτόλης στην συγκέντρωση της HbA1c

Το συγκεκριμένο σετ δεδομένων περιέχει 3 μελέτες και με την εντολή **list** τα περιεχόμενα είναι τα παρακάτω:

Πίνακας 17. Λίστα τιμών των μεταβλητών του σετ δεδομένων Miglitol.

```
. list

     +-------------------------------------------------------+
     |          Studies   ne     me    sde   nc   mc    sdc |
     |-------------------------------------------------------|
  1. |  Johnston (1994)   68   -.41  1.072   63  .33  1.032 |
  2. | Johnston (1998a)   91   -.43   .954   43  .98  1.311 |
  3. | Johnston (1998b)   49   -.12    1.4   34  .56  1.166 |
     +-------------------------------------------------------+
```

Στο παρακάτω παράδειγμα θα γίνεται μετά-ανάλυση χρησιμοποιώντας την μέθοδο HH και το μοντέλο σταθερών επιδράσεων (άρα μπορεί να χρησιμοποιηθεί μόνο η metan και όχι το ameta). Συνεπώς η μορφή της εντολής θα είναι η παρακάτω:

```
metasmd mc sdc nc me sde ne, method(hh) choice(score)
value(-1) model(fixed) study(Studies)
```

Τα αποτελέσματα που προκύπτουν και το διάγραμμα forest:

Πίνακας 18. Πίνακας αποτελεσμάτων της μετά-ανάλυσης, χρησιμοποιώντας την metan, το μοντέλο σταθερών επιδράσεων και την μέθοδο HH, στο σετ δεδομένων Miglitol.

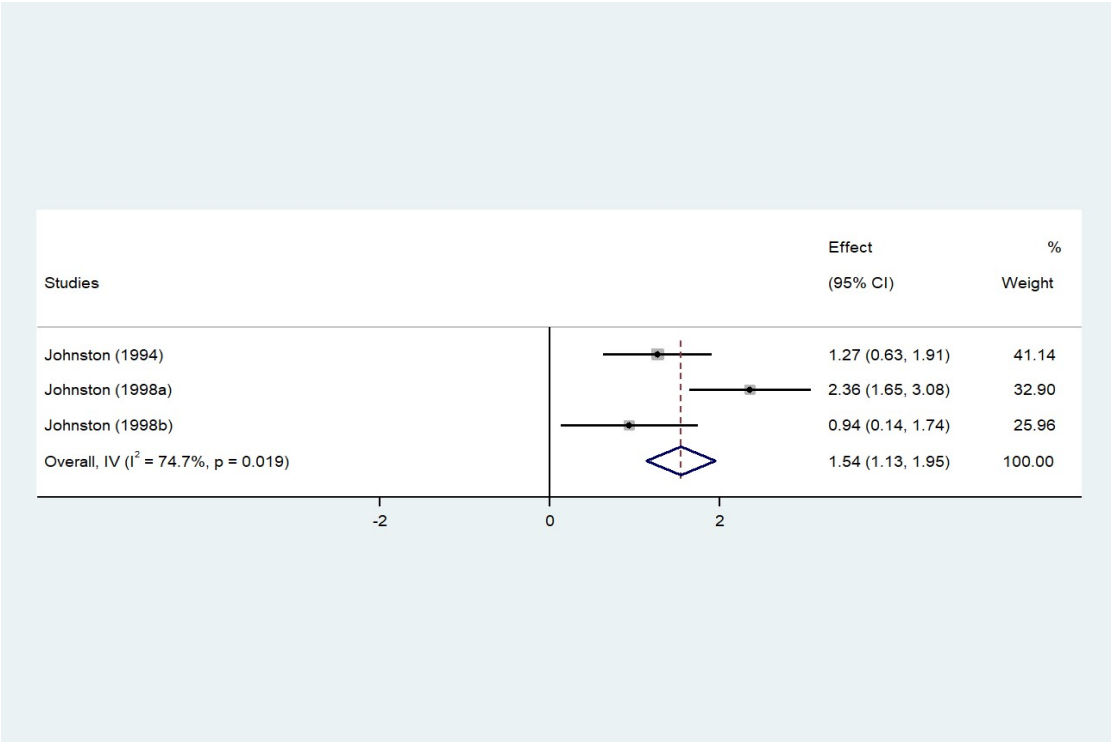
```
Studies included: 3
Participants included: Unknown

Meta-analysis pooling of aggregate data
using the common-effect inverse-variance model

------------------------------------------------------------------------------
Studies                  |   Effect    [95% Conf. Interval]   % Weight
-------------------------+----------------------------------------------------
Johnston (1994)          |      1.272      0.633     1.911      41.14
Johnston (1998a)         |      2.362      1.647     3.077      32.90
Johnston (1998b)         |      0.940      0.135     1.744      25.96
-------------------------+----------------------------------------------------
Overall, IV              |      1.544      1.134     1.954     100.00
------------------------------------------------------------------------------

Test of overall effect = 0:  z =   7.383  p = 0.000


Heterogeneity measures, calculated from the data
with Conf. Intervals based on non-central chi² (common-effect) distribution for Q
------------------------------------------------------------------------------
Measure                  |     Value      df      p-value
-------------------------+----------------------------------------------------
Cochran's Q              |      7.89       2      0.019
                         |          -[95% Conf. Interval]-
H                        |     1.987     1.000     3.209
I² (%)                   |     74.7%      0.0%     90.3%
------------------------------------------------------------------------------
H = relative excess in Cochran's Q over its degrees-of-freedom
I² = proportion of total variation in effect estimate due to between-study heterogeneity (based on Q)
```



Σχήμα 15. Γράφημα forest της metan για το σετ δεδομένων Miglitol.

Συνεχίζουμε με το μοντέλο τυχαίων επιδράσεων και τον πίνακα αποτελεσμάτων του:

Πίνακας 19. Μετά-ανάλυση των μελετών (k = 3) του σετ δεδομένων Miglitol.dta, χρησιμοποιώντας το μοντέλο τυχαίων επιδράσεων.

Method	M	SE(M)	τ^2	Q	I^2	LCI	UCI	p-value
D	1.412	0.3437	0.0000	0.2438	0.0000	0.739	2.086	0.0000
F	1.496	0.3857	0.0914	2.5004	0.2001	0.740	2.252	0.0000
S	1.418	0.2694	0.0000	0.4110	0.0000	0.890	1.946	0.0000
HH	1.532	0.4197	0.3936	7.8932	0.7466	0.710	2.355	0.0000
CS	1.397	1.4660	0.3271	7.8932	0.7466	0.647	2.147	0.0003

Αν εξαιρέσουμε την μέθοδο S, το τυπικό σφάλμα της μετά-ανάλυσης των υπολοίπων μεθόδων ξεπερνάει αυτό της διχοτόμησης. Για το μοντέλο σταθερών επιδράσεων έχουμε αντίστοιχα:

Πίνακας 20. Μετά-ανάλυση των μελετών (k = 3) του σετ δεδομένων Miglitol.dta, χρησιμοποιώντας το μοντέλο σταθερών επιδράσεων.

Method	M	SE(M)	Q	I^2	LCI	UCI	p-value
D	1.412	0.3436	0.24	0.00	0.739	2.086	0.000
F	1.487	0.3405	2.50	0.20	0.819	2.154	0.000
S	1.418	0.2693	0.41	0.00	0.890	1.946	0.000
HH	1.544	0.2091	7.89	0.747	1.134	1.954	0.000
CS	1.408	0.1908	7.89	0.747	1.034	1.782	0.000

Συμπεραίνουμε ότι το τυπικό σφάλμα της μετά-ανάλυσης με την χρήση των μεθόδων είναι μικρότερο απ' ότι με την διχοτόμηση.

4. Συζήτηση

Τα διαγράμματα ARE – OR που αναφέρθηκαν στην Ενότητα 3.2.1 δημιουργήθηκαν με σκοπό να εξεταστεί η ακρίβεια (precision) των μεθόδων και κατά πόσο οι εκτιμήσεις των ORs έχουν μικρότερη διακύμανση απ' την εκτίμηση που προκύπτει από την συμβατική διχοτόμηση. Εκτός από τα ειδικότερα συμπεράσματα που προκύπτουν από το κάθε διάγραμμα ξεχωριστά, επιπλέον 2 συμπεράσματα διαφαίνονται στο σύνολο των διαγραμμάτων. Το πρώτο είναι ότι όλα τα επιμέρους ARE έχουν τιμές μεγαλύτερες της μονάδας, που δείχνει ότι για την ίδια εκτίμηση του OR οι μέθοδοι έχουν μεγαλύτερη ακρίβεια. Το δεύτερο συμπέρασμα έχει να κάνει με τις μεθόδους CS και HH. Όσον αφορά την ακρίβεια, η μέθοδος CS έχει πάντοτε μεγαλύτερη ακρίβεια από την HH, κάτι που προκύπτει από τους μαθηματικούς τους τύπους. Επειδή τα διαγράμματα ασχολούνται μόνο με την ακρίβεια των μεθόδων και υποθέτουν ότι τα δεδομένα ακολουθούν την κανονική κατανομή (ενώ η μέθοδος HH προϋποθέτει ότι τα δεδομένα ακολουθούν λογιστική κατανομή), δεν θα ήταν συνετό να κατηγοριοποιήσουμε τις μεθόδους ως «καλύτερη» ή «χειρότερη», βασιζόμενοι μόνο σε αυτό το μέτρο. Από την στιγμή που το OR των μεθόδων πρέπει να

προσεγγίζει αυτό της διχοτόμησης θα πρέπει να γίνει λόγος για την μεροληψία (bias), δηλαδή την διαφορά ανάμεσα στα 2 ORs, όπως και της κάλυψης (coverage), που εξετάζει αν το προσεγγιστικό OR βρίσκεται εντός του διαστήματος εμπιστοσύνης του παρατηρούμενου OR (π.χ. για 95% διάστημα εμπιστοσύνης) (Anzures-Cabrera et al., 2011). Συνεπώς τα διαγράμματα ARE-OR δεν είναι κατάλληλα για να συνδυάσουν και τις 3 αυτές παραμέτρους και μια καταλληλότερη προσέγγιση θα ήταν μια μελέτη προσομοίωσης δεδομένων όπως αυτή των Anzures-Cabrera et al. (2011). Η συγκεκριμένη μελέτη εξέτασε την απόδοση των μεθόδων (εκτός της F) για διάφορες υποθέσεις κατανομής, risks, SMD και μεγέθη δείγματος. Έπειτα οι ερευνητές, βασισμένοι στα αποτελέσματα, προτείνανε κάτω από ποιες συνθήκες είναι καταλληλότερη για χρήση η κάθε μέθοδος. Όσο δεν παραβιάζεται η συνθήκη για τις ίσες τυπικές αποκλίσεις, τιμές για το risk της ομάδας ελέγχου μεταξύ 0.2 και 0.8 και η κατανομή να είναι συμμετρική (π.χ. κανονική, λογιστική), οι μέθοδοι HH και CS προτείνονται ως κατάλληλες για χρήση. Στην περίπτωση που ισχύουν οι παραπάνω προϋποθέσεις, καθώς επίσης και μικρές τιμές στις SMD και μεγάλο μέγεθος δείγματος (n = 100), προτείνεται η μέθοδος S. Αν οι τυπικές αποκλίσεις διαφέρουν μεταξύ των ομάδων προτείνεται η μέθοδος S μόνο στην περίπτωση που ισχύει: 1) συμμετρική κατανομή, 2) μικρές SMD, 3) μεγάλος αριθμός δείγματος. Σε διαφορετική περίπτωση δεν προτείνεται κάποια από τις μεθόδους. Τέλος, αν τα δεδομένα αντιστοιχούν σε μια λοξή (skewed) κατανομή προτείνονται οι μέθοδοι CS και HH μόνο στην περίπτωση που οι τυπικές τους αποκλίσεις είναι ίσες, το μέγεθος δείγματος μεγάλο και οι SMD έχουν μικρή αριθμητική τιμή.

Εκτός από την αποτελεσματικότητα τους σε σχέση με την διχοτόμηση, οι συγκεκριμένες μέθοδοι μπορούν να επεκταθούν και στην δημιουργία καμπυλών ROC (Receiver Operating Characteristic) για τα διχοτομημένα δεδομένα καθώς και για τον υπολογισμό της ευαισθησίας και της ειδικότητας στα διαγνωστικά τεστ.

Βιβλιογραφικές Αναφορές

Anzures-Cabrera, J., Sarpatwari, A., & Higgins, J. P. (2011). Expressing findings from meta-analyses of continuous outcomes in terms of risks. *Statistics in Medicine*, *30*(25), 2967–2985. https://doi.org/10.1002/sim.4298

Bagos, P. G., Dimou, N. L., Liakopoulos, T. D., & Nikolopoulos, G. K. (2011). Meta-analysis of family-based and case-control genetic association studies that use the same cases. *Statistical Applications in Genetics and Molecular Biology*, *10*(1). https://doi.org/10.2202/1544-6115.1640

Berkey, C. S., Hoaglin, D. C., Mosteller' And, F., & Colditz', G. A. (1995). A RANDOM-EFFECTS REGRESSION MODEL FOR META-ANALY SIS. In *STATISTICS IN MEDICINE* (Vol. 14).

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. (2009). *Introduction to meta-analysis*. John Wiley & Sons.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, *1*(2), 97–111. https://doi.org/10.1002/jrsm.12

Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. In *STATISTICS IN MEDICINE Statist. Med* (Vol. 19).

Cox, D. R., Snell, E. J. (1989). Analysis of Binary Data, 2nd Edition. Chapman & Hall/CRC.

Da Costa, B. R., Rutjes, A. W. S., Johnston, B. C., Reichenbach, S., Nüesch, E., Tonia, T., Gemperli, A., Guyatt, G. H., & Jüni, P. (2012). Methods to convert continuous outcomes into odds ratios of treatment response and numbers needed to treat: Meta-epidemiological study. *International Journal of Epidemiology*, *41*(5), 1445–1459. https://doi.org/10.1093/ije/dys124

Dersimonian, R., & Laird, N. (1986). *Meta-Analysis in Clinical Trials**.

Fedorov, V., Mannino, F., & Zhang, R. (2009). Consequences of dichotomization. *Pharmaceutical Statistics*, *8*(1), 50–61. https://doi.org/10.1002/pst.331

Furukawa, T. A., Cipriani, A., Barbui, C., Brambilla, P., & Watanabe, N. (2005). *Imputing response rates from means and standard deviations in meta-analyses*.

Furukawa, T. A., & Leucht, S. (2011). How to Obtain NNT from Cohen's d: Comparison of two methods. *PLoS ONE*, *6*(4). https://doi.org/10.1371/journal.pone.0019070

Hardy, R. J., & Thompson, S. G. (1998). Detecting and describing heterogeneity in meta-analysis. *Statistics in Medicine*, *17*(8), 841–856. https://doi.org/10.1002/(SICI)1097-0258(19980430)17:8<841::AID-SIM781>3.0.CO;2-D

Hasselblad, V., & Hedges, L. V. (1995). *Meta-Analysis of Screening and Diagnostic Tests*.

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). *Measuring inconsistency in meta-analyses Testing for heterogeneity*.

Langan, D., Higgins, J. P. T., Jackson, D., Bowden, J., Veroniki, A. A., Kontopantelis, E., Viechtbauer, W., & Simmonds, M. (2019). A comparison of

heterogeneity variance estimators in simulated random-effects meta-analyses. *Research Synthesis Methods*, *10*(1), 83–98. https://doi.org/10.1002/jrsm.1316

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, *7*(1), 19–40. https://doi.org/10.1037/1082-989X.7.1.19

Nuzzo, R. L. (2019). Making Continuous Measurements into Dichotomous Variables. *PM and R*, *11*(10), 1132–1134. https://doi.org/10.1002/pmrj.12228

Riffenburgh, R. H. (2012). Sample Size Estimation and Meta-Analysis. In *Statistics in Medicine* (pp. 365–391). Elsevier. https://doi.org/10.1016/b978-0-12-384864-2.00018-4

Senn, S., Gavini, F., Magrez, D., & Scheen, A. (2013). Issues in performing a network meta-analysis. *Statistical Methods in Medical Research*, *22*(2), 169–189. https://doi.org/10.1177/0962280211432220

Stratton, I. M., Adler, A. I., Neil, A. W., Matthews, D. R., Manley, S. E., Cull, C. A., Hadden, D., Turner, R. C., & Holman, R. R. (2000). *Papers Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study*.

Suissa, S. (1991). BINARY METHODS FOR CONTINUOUS OUTCOMES: A PARAMETRIC ALTERNATIVE. In *J Clin Epidemiol* (Vol. 44, Issue 3).

Thorlund, K., Walter, S. D., Johnston, B. C., Furukawa, T. A., & Guyatt, G. H. (2011). Pooling health-related quality of life outcomes in meta-analysis—a tutorial and review of methods for enhancing interpretability. *Research Synthesis Methods*, *2*(3), 188–203. https://doi.org/10.1002/jrsm.46

Thorlund, K., Wetterslev, J., Awad, T., Thabane, L., & Gluud, C. (2011). Comparison of statistical inferences from the DerSimonian–Laird and alternative random-effects model meta-analyses – an empirical assessment of 920 Cochrane primary outcome meta-analyses. *Research Synthesis Methods*, *2*(4), 238–253. https://doi.org/10.1002/jrsm.53

Tufanaru, C., Munn, Z., Stephenson, M., & Aromataris, E. (2015). Fixed or random effects meta-analysis? Common methodological issues in systematic reviews of effectiveness. *International Journal of Evidence-Based Healthcare*, *13*(3), 196–207. https://doi.org/10.1097/XEB.0000000000000065

Yerukala, R., & Kumar Boiroju, N. (2015). Approximations to Standard Normal Distribution Function. *International Journal of Scientific & Engineering Research*, *6*(4). http://www.ijser.org