

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ
ΤΟΜΕΑΣ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΕΠΙΧΕΙΡΗΣΙΑΚΗΣ
ΕΡΕΥΝΑΣ**

**ΣΥΓΚΡΙΣΗ 24 ΕΚΤΙΜΗΤΩΝ
ΕΤΕΡΟΓΕΝΕΙΑΣ ΣΤΗ ΜΕΤΑ-
ΑΝΑΛΥΣΗ ΤΥΧΑΙΩΝ ΕΠΙΔΡΑΣΕΩΝ:
ΜΙΑ ΕΡΕΥΝΑ ΠΡΟΣΟΜΟΙΩΣΗΣ**

ΠΕΤΡΟΠΟΥΛΟΥ ΜΑΡΙΑ

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ

ΙΩΑΝΝΙΝΑ 2015

Η παρούσα Μεταπτυχιακή Διατριβή εκπονήθηκε στο πλαίσιο των σπουδών για την απόκτηση του Μεταπτυχιακού Διπλώματος Ειδίκευσης στη

ΣΤΑΤΙΣΤΙΚΗ ΚΑΙ ΕΠΙΧΕΙΡΗΣΙΑΚΗ ΕΡΕΥΝΑ

που απονέμει το Τμήμα Μαθηματικών του Πανεπιστημίου Ιωαννίνων.

Εγκρίθηκε την 16/09/2015 από την Τριμελή Επιτροπή Εξέτασης:

Μαυρίδης Δημήτριος, Επίκουρος Καθηγητής του Παιδαγωγικού Τμήματος Δημοτικής Εκπαίδευσης του Πανεπιστημίου Ιωαννίνων
(Επιβλέπων Καθηγητής)

Λουκάς Σωτήριος, Καθηγητής του Τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων

Μπατσίδης Απόστολος, Επίκουρος Καθηγητής του Τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων

ΥΠΕΥΘΥΝΗ ΔΗΛΩΣΗ

"Δηλώνω υπεύθυνα ότι η παρούσα διατριβή εκπονήθηκε κάτω από τους διεθνείς ηθικούς και ακαδημαϊκούς κανόνες δεοντολογίας και προστασίας της πνευματικής ιδιοκτησίας. Σύμφωνα με τους κανόνες αυτούς, δεν έχω προβεί σε ιδιοποίηση ξένου επιστημονικού έργου και έχω πλήρως αναφέρει τις πηγές που χρησιμοποίησα στην εργασία αυτή."

Μαρία Πετροπούλου

Αφιερώνεται στους γονείς μου
Αθανάσιο και Δήμητρα.

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου κ. Δημήτριο Μαυρίδη για την πολύτιμη βοήθεια και υποστήριξη του καθ' όλη τη διάρκεια της μεταπτυχιακής μου διατριβής καθώς και για την άριστη συνεργασία. Ακόμα, ευχαριστώ τα μέλη της κριτικής μου επιτροπής κ. Σωτήριο Λουκά και κ. Απόστολο Μπατσίδα καθώς και τους καθηγητές του Τομέα Στατιστικής και Επιχειρησιακής Έρευνας κ. Κωνσταντίνο Ζωγράφο και κ. Κωνσταντίνα Σκούρη για τις χρήσιμες συμβουλές τους.

ΠΕΡΙΛΗΨΗ

Στην παρούσα μεταπτυχιακή διατριβή παρουσιάζεται η έννοια της ετερογένειας στη μετα-ανάλυση καθώς και οι μέθοδοι εκτίμησης της. Καθώς η μετα-ανάλυση συμβάλει στη διεξαγωγή συμπερασμάτων και κλινικών αποφάσεων κρίνεται σημαντική η μελέτη της ετερογένειας και κατά συνέπεια των εκτιμητών ετερογένειας οι οποίοι διαδραματίζουν καθοριστικό ρόλο για τη δημιουργία ακριβών και έγκυρων αποτελεσμάτων. Με βάση το γεγονός αυτό σκοπός μας αποτελεί η σύγκριση των εκτιμητών ετερογένειας έπειτα από μια μελέτη προσομοίωσης. Η μεταπτυχιακή διατριβή έχει οργανωθεί σε 5 κεφάλαια τα οποία περιγράφονται ως εξής:

Στο Κεφάλαιο 1 περιγράφεται η ιεράρχηση των αποδείξεων με τους διάφορους τύπους έρευνας που μπορούν να πραγματοποιηθούν καθώς και η πιο ισχυρή μέθοδος έρευνας που αποτελεί την κορυφή της πυραμίδας της ιεράρχησης των αποδείξεων, η συστηματική ανασκόπηση και μετα-ανάλυση. Επιπρόσθετα, περιγράφονται τα κεντρικά σημεία των κεφαλαίων και ο σκοπός της μεταπτυχιακής διατριβής.

Στο Κεφάλαιο 2 περιγράφεται η έννοια της συστηματικής ανασκόπησης και τα στάδια αυτής καθώς και η έννοια της μετα-ανάλυσης. Ακόμα, ορίζονται τα δύο είδη αποτελεσμάτων που αποτελούν αντικείμενο μελέτης μας, τα διχότομα και τα συνεχή αποτελέσματα καθώς και την έννοια του μέτρου σχέσης /μεγέθους επίδρασης (effect size) για τα δύο είδη αποτελεσμάτων.

Αντικείμενο μελέτης του Κεφαλαίου 3 αποτελεί η περιγραφή των δύο μοντέλων μετα-ανάλυσης, το μοντέλο σταθερών και το μοντέλο τυχαίων επιδράσεων καθώς και τις διαφορές τους. Περιγράφεται η έννοια της ετερογένειας, οι τρόποι με τους οποίους μπορούμε να την ποσοτικοποιήσουμε και κάποια βασικά πεδία της μετα-ανάλυσης τα οποία είναι συνδεδεμένα με την έννοια της ετερογένειας όπως το διάγραμμα δάσους, το σφάλμα δημοσίευσης και το διάγραμμα χωνί.

Στο Κεφάλαιο 4 παρουσιάζονται όλες οι μέθοδοι εκτίμησης της ετερογένειας καθώς και τα χαρακτηριστικά που έχει η κάθε μέθοδος. Στο πλαίσιο αυτό, παραθέτουμε έναν συγκεντρωτικό πίνακα με όλους τους

εκτιμητές ετερογένειας καθώς και τα βασικά χαρακτηριστικά τους. Επιπλέον, αναφέρονται για κάθε εκτιμητή ετερογένειας τα αποτελέσματα από μελέτες προσομοίωσης που έχουν πραγματοποιηθεί.

Καθώς στη βιβλιογραφία δεν μας παρέχεται μια ολοκληρωμένη γνώση για τη συμπεριφορά όλων των εκτιμητών ετερογένειας, στο Κεφάλαιο 5 παρουσιάζεται η μελέτη προσομοίωσης που πραγματοποιήσαμε για την σύγκριση των μεθόδων εκτίμησης. Παρουσιάζονται τα βήματα της καθώς και ο τρόπος διεξαγωγής, παρατίθενται διαγράμματα από τα αποτελέσματα, πραγματοποιούνται οι κατάλληλες συγκρίσεις μεταξύ των εκτιμητών και τέλος διεξάγονται συμπεράσματα για τα διχότομα και για τα συνεχή αποτελέσματα.

Τέλος, η διατριβή ολοκληρώνεται με το παράρτημα στο οποίο παρατίθενται οι κώδικες που δημιουργήθηκαν για τη πραγματοποίηση της μελέτης προσομοίωσης στο στατιστικό πακέτο της R και στο WinBUGS.

ABSTRACT

In the present thesis the concept of heterogeneity in meta-analysis and the methods of heterogeneity estimators are presented. When meta-analysis contributes in clinical decision making and to draw conclusions, it is important to study the heterogeneity and therefore the heterogeneity estimators which play a decisive role in the creation of accurate and valid results. Based on this fact, our objective is to compare the heterogeneity estimators after a simulation study. The thesis is organized into five chapters, which are described as follows:

Chapter 1 describes the hierarchy of evidence with the various types of research that can be carried and the most powerful search method which is the apex of the hierarchy of evidence, a systematic review and meta-analysis. Additionally, it states the purpose of postgraduate thesis and the main contents of chapters.

Chapter 2 presents the concept of systematic review, the stages and the concept of meta-analysis. Moreover, two types of outcomes that are in the subject of our study, are presented, dichotomous and continuous outcomes.

This chapter concludes with the concept of the effect size for dichotomous and continuous outcomes.

Chapter 3 contains the description of fixed and random effects models and their main differences. Additionally, Chapter 3 describes the concept of heterogeneity, the ways that we can quantify some basic fields of meta-analysis which are connected with the concept of heterogeneity such as forest plot, publication bias and funnel plot.

Chapter 4 presents all the assessment methods and characteristics that each method has and we present a summary of all heterogeneity estimators and their basic characteristics. Furthermore, results from previous simulation studies have been carried out for each heterogeneity estimator. Since the literature did not provided knowledge about the behavior of all heterogeneity estimators, chapter 5 describes the simulation study of assessment methods, presents the steps, the method that was used in the simulation study, diagrams, results with the appropriate comparisons between estimators and finally the conclusions for dichotomous and continuous outcomes.

Finally, the thesis concludes with the appendix which lists the codes used to perform the simulation study in the statistical packages R and WinBUGS.

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1

ΕΙΣΑΓΩΓΗ

1.1 Ιεράρχηση των αποδείξεων.....	1
1.2 Σκοπός της Μεταπτυχιακής Διατριβής	3

ΚΕΦΑΛΑΙΟ 2

ΣΥΣΤΗΜΑΤΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

ΚΑΙ ΜΕΤΑ-ΑΝΑΛΥΣΗ

2.1 Συστηματική ανασκόπηση (Systematic review).....	5
2.2 Τι είναι η μετα-ανάλυση (meta-analysis).....	12
2.3 Μέτρο σχέσης /μέγεθος επίδρασης (effect size).....	12

ΚΕΦΑΛΑΙΟ 3

ΜΟΝΤΕΛΑ ΜΕΤΑ-ΑΝΑΛΥΣΗΣ

ΚΑΙ ΕΤΕΡΟΓΕΝΕΙΑ

3.1 Μοντέλα μετα-αναλύσεων.....	21
3.2 Μετα-ανάλυση σταθερών επιδράσεων (Fixed Effect model).....	21
3.3 Τι είναι ετερογένεια (heterogeneity).....	23
3.4 Μετα-ανάλυση τυχαίων επιδράσεων (Random Effect model).....	25
3.5 Διάγραμμα δάσους (forest plot).....	27
3.6 Εξέταση και τρόποι υπολογισμού της ετερογένειας.....	28
3.7 Τα μοντέλα μετα-ανάλυσης στο διάγραμμα δάσους.....	30
3.8 Σύγκριση των δύο μοντέλων μετα-ανάλυσης.....	32
3.9 Συστηματικό σφάλμα δημοσίευσης (publication bias).....	33
3.10 Διάγραμμα χωνί (funnel plot).....	33

ΚΕΦΑΛΑΙΟ 4

ΕΚΤΙΜΗΤΕΣ ΕΤΕΡΟΓΕΝΕΙΑΣ

4.1 Εκτίμηση της συνολικής επίδρασης μ του πληθυσμού.....	37
4.2 Μέθοδοι εκτίμησης της ετερογένειας.....	38

4.3 Ιδιότητες ενός καλού εκτιμητή ετερογένειας	39
4.4 Γενική μέθοδος εκτίμησης των ροπών (General method of moments (GMM)) για την ετερογένεια	40
4.4.1 Μέθοδος General Hedges-Olkin (GHO)	43
4.4.2 Μέθοδος DerSimonian and Laird (DL).....	43
4.4.3 Μέθοδος Paule-Mandel (PM).....	44
4.4.4 Μέθοδος δύο βημάτων DerSimonian and Laird (two-step DerSimonian and Laird method (DL2))	47
4.4.5 Μέθοδος δύο βημάτων General Hedges-Olkin (two-step General Hedges-Olkin method (GHO2))	47
4.4.6 Θετικός εκτιμητής DerSimonian and Laird (Positive DerSimonian and Laird (DLp)).....	48
4.4.7 Μέθοδος Hartung-Makambi (HM).....	49
4.4.8 Μέθοδος Hunter-Schmidt (HS)	50
4.5 Μέθοδος Μέγιστης Πιθανοφάνειας (Maximum likelihood method (ML))	50
4.5.1 Εκτιμητής μέγιστης πιθανοφάνειας (Maximum likelihood estimator (ML))	50
4.5.2 Περιορισμένη μέθοδος μέγιστης πιθανοφάνειας (Restricted Maximum likelihood method (REML))	55
4.5.3 Προσεγγισμένη περιορισμένη μέθοδος μέγιστης πιθανοφάνειας (Approximate Restricted Maximum likelihood method (AREML))...	59
4.6 Μέθοδος σταθμισμένων ελαχίστων τετραγώνων (Weighted Least-Squares method)	60
4.6.1 Εκτιμητής Sidik-Jonkman (SJ)	60
4.6.2 Εκτιμητής (SJgho)	63
4.7 Μέθοδος Non-Parametric Bootstrap για τον εκτιμητή DerSimonian and Laird (Non-Parametric bootstrap method (DLb))	63
4.8 Εκτιμητής Empirical Bayes (EB).....	64
4.9 Μέθοδος Bayes.....	64

4.9.1 Εκτιμητής Rukhin Bayes (RB).....	65
4.9.2 Θετικός εκτιμητής Rukhin Bayes (RBp).....	65
4.9.3 Εκτιμητής Full Bayesian (FB).....	66
4.9.4 Εκτιμητής Bayes Modal (BM)	67
4.10 Εκτιμητής Malzahn, Böhning, and Holling (MBH)	68
4.11 Διαστήματα εμπιστοσύνης για την συνολική επίδραση μ	68
4.11.1 Wald-type (Wt) διαστήματα εμπιστοσύνης	68
4.11.2 Knapp and Hartung (KH) διαστήματα εμπιστοσύνης.....	69
4.12 Διαστήματα εμπιστοσύνης για την ετερογένεια με τη μέθοδο non-parametric bootstrap.....	69
4.13 Συγκεντρωτικοί πίνακες εκτιμητών ετερογένειας.....	70
4.14 Συμπεράσματα από μελέτες προσομοίωσης	72
4.14.1 Εκτιμητής DL	72
4.14.2 Εκτιμητής GHO.....	73
4.14.3 Εκτιμητής PM.....	74
4.14.4 Εκτιμητής DLp	74
4.14.5 Εκτιμητής DL2	75
4.14.6 Εκτιμητής GHO2.....	75
4.14.7 Εκτιμητής HM.....	75
4.14.8 Εκτιμητής HS	75
4.14.9 Εκτιμητής ML	75
4.14.10 Εκτιμητής REML	76
4.14.11 Εκτιμητής AREML	76
4.14.12 Εκτιμητής SJ.....	76
4.14.13 Εκτιμητής SJgho.....	77
4.14.14 Εκτιμητής DLb	77
4.14.15 Εκτιμητής EB	77
4.14.16 Εκτιμητής RBo	77

4.14.17 Εκτιμητής RBp	77
4.14.18 Εκτιμητής FB.....	78
4.14.19 Εκτιμητής BM.....	79
4.14.20 Εκτιμητής MBH.....	79

ΚΕΦΑΛΑΙΟ 5

ΠΡΟΣΟΜΟΙΩΣΗ

5.1 Σκοπός της έρευνας προσομοίωσης	81
5.2 Σενάρια προσομοίωσης	82
5.3 Μετρήσεις προσομοίωσης.....	82
5.4 Βήματα προσομοίωσης	85
5.4.1 Δημιουργία δεδομένων	85
5.4.2 Υπολογισμός Μετρήσεων.....	88
5.5 Εξαγωγή συμπερασμάτων για τα δεδομένα με διχότομο αποτέλεσμα	91
5.5.1 Σύγκριση των εκτιμητών ετερογένειας ως προς τη μεροληψία..	91
5.5.2 Σύγκριση των εκτιμητών ετερογένειας ως προς το μέσο τετραγωνικό σφάλμα	93
5.5.3 Σύγκριση των εκτιμητών ετερογένειας ως προς την πιθανότητα κάλυψης μ	94
5.5.4 Σύγκριση των εκτιμητών ετερογένειας ως προς την πιθανότητα κάλυψης τ^2	96
5.5.5 Σύγκριση των εκτιμητών ετερογένειας ως προς το εμπειρικό type error I	98
5.5.6 Σύγκριση των εκτιμητών ετερογένειας ως προς την ισχύ	100
5.5.7 Σύγκριση των εκτιμητών ως προς τα επίπεδα σημαντικότητας	101
5.6 Εξαγωγή συμπερασμάτων για τα δεδομένα με συνεχή αποτέλεσμα	104
5.6.1 Σύγκριση των εκτιμητών ετερογένειας ως προς τη μεροληψία	104

5.6.2 Σύγκριση των εκτιμητών ετερογένειας ως προς το μέσο τετραγωνικό σφάλμα	105
5.6.3 Σύγκριση των εκτιμητών ετερογένειας ως προς την πιθανότητα κάλυψης μ	106
5.6.4 Σύγκριση των εκτιμητών ετερογένειας ως προς την πιθανότητα κάλυψης τ^2	106
5.6.5 Σύγκριση των εκτιμητών ετερογένειας ως προς το εμπειρικό type error I.....	107
5.6.6 Σύγκριση των εκτιμητών ετερογένειας ως προς την ισχύ	108
5.6.7 Σύγκριση των εκτιμητών ως προς τα επίπεδα σημαντικότητας	108
5.7 Συμπεράσματα	110
5.7.1 Συμπεράσματα για τα διχότομα δεδομένα	110
5.7.2 Συμπεράσματα για τα συνεχή δεδομένα	111
5.8 Γενικό συμπέρασμα	112
5.9 Διαγράμματα	118
ΠΑΡΑΡΤΗΜΑ	159
ΒΙΒΛΙΟΓΡΑΦΙΑ	193

ΚΕΦΑΛΑΙΟ 1

ΕΙΣΑΓΩΓΗ

1.1 Ιεράρχηση των αποδείξεων

Η εξαγωγή συμπερασμάτων από τις διάφορες μελέτες που προσεγγίζουν ένα ερευνητικό ερώτημα διαδραματίζει καθοριστικό ρόλο στην εφαρμογή της ιατρικής πρακτικής. Υπάρχουν διάφορα είδη μελετών και τεκμηρίων για την επαλήθευση ή όχι μιας ερευνητικής υπόθεσης. Υπάρχουν μελέτες σε ζώα, μελέτες παρατήρησης, τυχαιοποιημένες μελέτες, προσφυγές σε αυθεντίες και άλλοι πολλοί τρόποι στους οποίους μπορεί να καταφύγει κάποιος για την επαλήθευση ή όχι μιας ερευνητικής υπόθεσης. Όμως δεν έχουν όλα τα τεκμήρια/αποδείξεις το ίδιο βάρος.

Στη βάση της πυραμίδας ιεράρχησης των αποδείξεων όπως φαίνεται στο Διάγραμμα 1.1 βρίσκεται η γνώμη των εμπειρογνώμων (Expert opinion), έρευνες (Surveys), έρευνες σε ζώα (Animal research) και έπειτα οι μελέτες ασθενών μαρτύρων (Case control studies).

Κατά τις μελέτες ασθενών μαρτύρων χωρίζουμε τον πληθυσμό σε μια ομάδα που έχει κάποιο χαρακτηριστικό (ασθένεια) και σε μια άλλη που δεν έχει αυτό το χαρακτηριστικό και παρατηρούμε τους παράγοντες που μπορεί να σχετίζονται με την έκβαση που μας ενδιαφέρει. Αυτού του είδους μελέτες είναι λιγότερο αξιόπιστες από τις μελέτες κοόρτης και τις τυχαιοποιημένες ελεγχόμενες δοκιμές καθώς για την εξαγωγή ενός συμπεράσματος δε μπορούμε να βασιστούμε στη στατιστική σχέση μεταξύ παραγόντων (Isenburg, 2015).

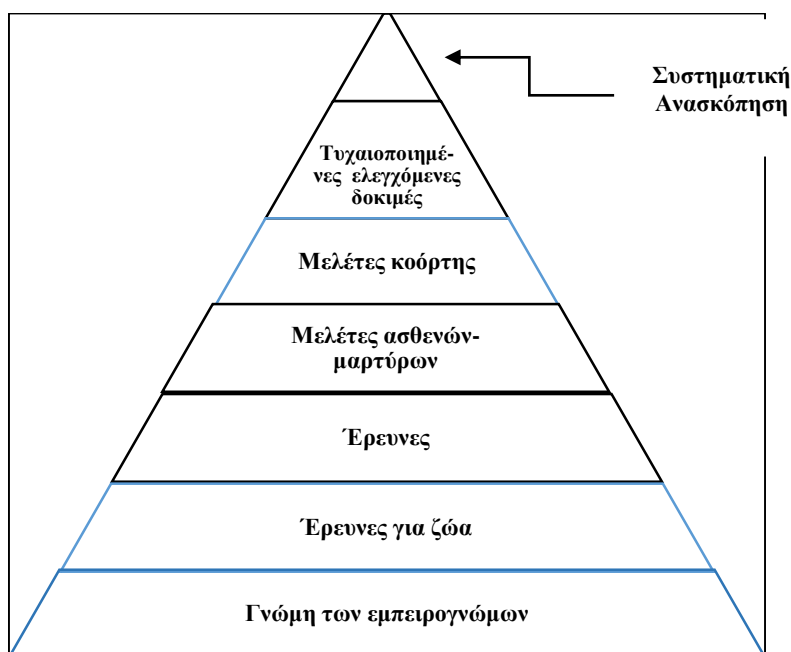
Οι μελέτες κοόρτης (cohort studies) αποτελούν ένα είδος μελετών παρατήρησης. Κατά τις μελέτες κοόρτης ορίζουμε ομάδες ατόμων ανάλογα με την έκβαση τους ή όχι και εξάγουμε συμπεράσματα από την παρατήρηση του αριθμού των ατόμων που μπορεί να εμφανίσουν την έκβαση μέσα σε ένα χρονικό διάστημα. Οι μελέτες κοόρτης δεν είναι τόσο αξιόπιστες όσο οι τυχαιοποιημένες ελεγχόμενες δοκιμές καθώς οι ομάδες μπορεί να διαφέρουν ως προς κάποια χαρακτηριστικά τα οποία μπορεί να επηρεάσουν την έκβαση (Isenburg, 2015).

Παρά το γεγονός ότι δεν υπάρχει παγκοσμίως καθιερωμένη ιεράρχηση των αποδεικτικών στοιχείων υπάρχει ευρεία συμφωνία ότι το πιο έγκυρο (gold standard) τύπο έρευνας αποτελούν οι τυχαιοποιημένες ελεγχόμενες δοκιμές (Randomized controlled trials (RCTs)). Κατά τις τυχαιοποιημένες ελεγχόμενες δοκιμές τα άτομα χωρίζονται τυχαία σε δύο ομάδες όπου στη μια ομάδα (πειραματική ομάδα / experimental group) γίνεται μια παρέμβαση και στην άλλη (ομάδα ελέγχου / control group) συνήθως μια εικονική παρέμβαση (placebo) και οι διαφορές μεταξύ των δύο ομάδων οφείλονται αποκλειστικά στις παρεμβάσεις διότι οι συμμετέχοντες έχουν τυχαιοποιηθεί και στις δύο ομάδες και η θεωρία πιθανοτήτων εξασφαλίζει ότι δε θα διαφέρουν σε κανένα άλλο χαρακτηριστικό πέρα από την παρέμβαση που δέχονται.

Σε κάθε ιατρικό πρόβλημα υπάρχει πληθώρα μεμονωμένων μελετών. Ωστόσο, η εξαγωγή συμπερασμάτων από τις μελέτες μπορεί να έχει αρνητικές συνέπειες καθώς οι μεμονωμένες μελέτες ενδέχεται να δίνουν μεροληπτικά (μελέτες χαμηλής ποιότητας) ή αντικρουόμενα αποτελέσματα χωρίς να υπάρχει η επίγνωση αν οι διαφορές αυτές είναι τυχαίες ή πραγματικές (Higgins & Green, 2008). Για το λόγο αυτό δεν είναι ασφαλές να εξάγει κάποιος συμπεράσματα κοιτάζοντας τα αποτελέσματα από μεμονωμένες μελέτες.

Με βάση το παραπάνω γεγονός, οι μελετητές ανέπτυξαν νέες ερευνητικές μεθόδους για τη εξαγωγή ενός έγκυρου αποτελέσματος με στόχο τη βελτίωση της ιατρικής πρακτικής. Οι νέες ερευνητικές μέθοδοι αποτέλεσαν μεθόδους σύνθεσης των μεμονωμένων μελετών. Οι διαδικασίες αυτές που είναι γνωστές ως συστηματικές ανασκοπήσεις (systematic reviews) αποτελούν την κορυφή της πυραμίδας της ιεράρχησης των αποδεικτικών στοιχείων και αντικείμενο μελέτης του Κεφαλαίου 2. Με την έννοια συστηματική ανασκόπηση θεωρούμε μια εκτενή βιβλιογραφική ανασκόπηση με βάση κάποια συγκεκριμένη ερευνητική υπόθεση. Αναπόσπαστο κομμάτι της συστηματικής ανασκόπησης αποτελεί η διαδικασία της μετα-ανάλυσης (meta-analysis). Με τον όρο μετα-ανάλυση εννοούμε μια στατιστική διαδικασία σύνθεσης έπειτα από την ανασκόπηση της βιβλιογραφίας των μεμονωμένων μελετών με στόχο την παροχή ενός συγκεντρωτικού αποτελέσματος.

Επειδή κατά την μετα-ανάλυση έχουμε σύνθεση διαφορετικών μελετών και η κάθε μελέτη έχει διαφορετικά χαρακτηριστικά δημιουργείται ετερογένεια, την έννοια της οποίας αναλύουμε στο Κεφάλαιο 3. Η ετερογένεια αυτή κατά τη πράξη είναι άγνωστη ποσότητα και οι μέθοδοι της εκτίμησής της αποτελούν αντικείμενο μελέτης του Κεφαλαίου 4. Στο πλαίσιο αυτό, δίνεται μια ανασκόπηση των μελετών προσομοίωσης που έχουν εμφανιστεί στην βιβλιογραφία. Για την σύγκριση των μεθόδων εκτίμησης της ετερογένειας στο Κεφάλαιο 5 παρατίθεται μια ολοκληρωμένη γνώση της συμπεριφοράς των εκτιμητών της ετερογένειας έπειτα από μια πρωτότυπη έρευνα προσομοίωσης.



Διάγραμμα 1.1: Ενδεικτική πυραμίδα ιεράρχησης των αποδεικτικών στοιχείων

1.2 Σκοπός της Μεταπτυχιακής Διατριβής

Σκοπός της μεταπτυχιακής διατριβής αποτελεί η παρουσίαση της έννοιας της ετερογένειας που δημιουργείται κατά τη στατιστική μέθοδο της μετα-ανάλυσης και των μεθόδων εκτίμησής της. Έπειτα από την κριτική ανασκόπηση στόχος αποτελεί η δημιουργία μιας πρωτοποριακής για την υπάρχουσα βιβλιογραφία μελέτης προσομοίωσης η οποία θα

συμπεριλαμβάνει όλους τους εκτιμητές που χρησιμοποιούνται για τον υπολογισμό της ετερογένειας. Με πρόφαση την ανάγκη για τη σύγκριση όλων των εκτιμητών ετερογένειας το ενδιαφέρον μας επικεντρώνεται στη μελέτη της συμπεριφοράς των εκτιμητών ετερογένειας ως προς τις ιδιότητες που μπορεί να έχει ένας καλός εκτιμητής. Έπειτα από την επίγνωση αυτή, η συμβολή των καλών εκτιμητών θα οδηγήσουν στη βελτίωση της κλινικής πρακτικής.

ΚΕΦΑΛΑΙΟ 2

ΣΥΣΤΗΜΑΤΙΚΗ ΑΝΑΣΚΟΠΗΣΗ ΚΑΙ ΜΕΤΑ-ΑΝΑΛΥΣΗ

2.1 Συστηματική ανασκόπηση (Systematic review)

Στον τομέα της υγείας είναι αναγκαία η ολοένα συλλογή πληροφοριών για την έρευνα της υγειονομικής περίθαλψης (Γαλάνης, 2009). Η εποχή μας χαρακτηρίζεται από πληθώρα πληροφοριών και ταχύτητα μετάδοσης της γνώσης γεγονός που καθιστά δύσκολη τη θέση των ερευνητών ώστε να αξιολογήσουν τις πληροφορίες που αποτελούν αντικείμενο του ενδιαφέροντος τους (Πατελάρου & Μπροκαλάκη, 2010). Για τη διεξαγωγή μιας μελέτης κρίνεται λοιπόν απαραίτητη η συγκέντρωση των κατάλληλων και αξιόπιστων πληροφοριών (Γαλάνης, 2009). Επιπρόσθετα, οι μελετητές οφείλουν να έχουν το χρόνο και τους πόρους για την εύρεση πληροφοριών καθώς και τις δεξιότητες για τη σωστή ερμηνεία τους (Higgins & Green, 2008). Επομένως μια αξιόπιστη συλλογή μελετών θα οδηγήσει σε έγκυρες αποφάσεις που θα βελτιώσουν την πρακτική κλινική. Βασικά εργαλεία για τη συλλογή και ανάλυση των μελετών αποτελούν οι ανασκοπήσεις (Καράσσα, 2006).

Οι αφηγηματικές ανασκοπήσεις (narrative reviews) αποτελούν μη αξιόπιστες μεθόδους καθώς καθίστανται υποκειμενικές δεδομένου ότι αγνοούν σημαντικές παραμέτρους όπως τον σχεδιασμό, το αποτέλεσμα της μελέτης ή το μέγεθος του δείγματος. Με τη συμβολή της συστηματικής ανασκόπησης, η έρευνα εμπλουτίζεται με έγκυρα και έγκαιρα αποτελέσματα αποσαφηνίζοντας αβέβαια ή ελλιπή θέματα. Πιο συγκεκριμένα, η συστηματική ανασκόπηση αποτελεί μία ερευνητική μέθοδο κατά την οποία πραγματοποιείται ανασκόπηση της βιβλιογραφίας με βάση μια ερευνητική υπόθεση (Γαλάνης, 2009). Στόχος της συστηματικής ανασκόπησης αποτελεί η συγκέντρωση όλων των απαραίτητων αποδεικτικών στοιχείων με την επιλογή των καλύτερων

μεθοδολογικών μελετών για την εξαγωγή συμπερασμάτων (Higgins & Green, 2008). Η συστηματική ανασκόπηση βασίζεται σε ένα πρωτόκολλο διεξαγωγής, το οποίο περιλαμβάνει καθορισμένα στάδια βασισμένα σε επιστημονικές αρχές που οφείλει να τηρεί κάθε μελετητής ώστε να αποφεύγονται τυχαία σφάλματα. Πιο αναλυτικά, τα στάδια της συστηματικής ανασκόπησης περιγράφονται παρακάτω:

1) Διατύπωση επιστημονικής υπόθεσης (Defining the review question)

Το πρώτο και πιο σημαντικό στάδιο της συστηματικής ανασκόπησης αποτελεί ο καθορισμός του αντικειμένου της μελέτης. Βασική αρχή είναι η διατύπωση της ερευνητικής υπόθεσης. Η ερευνητική υπόθεση πρέπει να είναι σαφής και επιστημονικά τεκμηριωμένη (Πατελάρου & Μπροκαλάκη, 2010). Ακόμα, η επιστημονική υπόθεση πρέπει να λαμβάνεται έτσι ώστε να προσδιορίζει τον τύπο των συμμετεχόντων, τα είδη των παρεμβάσεων καθώς και τα είδη των αποτελεσμάτων (Higgins & Green, 2008). Όσον αφορά τον τύπο των συμμετεχόντων, το εύρος του πληθυσμού πρέπει να λαμβάνει υπόψη την ποικιλομορφία των μελετών αλλά παράλληλα να περιορίζεται σε μια συγκεκριμένη ομάδα ατόμων (Higgins & Green, 2008). Ως προς τα είδη των παρεμβάσεων, οι μελετητές οφείλουν να ορίζουν τις θεραπείες για το αντικείμενο μελέτης τους καθώς και αυτές να μπορούν να είναι συγκρίσιμες (Higgins & Green, 2008). Οι θεραπείες μπορεί να είναι μη δραστικές όπως η χορήγηση εικονικού φάρμακου (placebo) ή δραστικές με τη χορήγηση διαφορετικών ειδών φαρμάκων ή θεραπειών (Higgins & Green, 2008). Ως προς τα είδη των αποτελεσμάτων, οφείλεται να καθορίζονται τα αποτελέσματα που αποτελούν αντικείμενο του ενδιαφέροντος. Αποτέλεσμα μπορεί να θεωρηθεί η θνησιμότητα, τα καρδιακά προβλήματα, ο καρκίνος ή κάποια ανεπιθύμητα γεγονότα (adverse events) (Higgins & Green, 2008). Το αποτέλεσμα μπορεί να είναι ωφέλιμο ή επιβλαβές (Higgins & Green, 2008). Ως ωφέλιμο θεωρούμε για παράδειγμα ένα μη θανατηφόρο καρδιακό επεισόδιο. Τέλος, κατά τη διατύπωση της επιστημονικής υπόθεσης πρέπει να ορίζεται ο σκοπός για τον οποίο γίνεται η μελέτη με το εύρος του να περιορίζεται κατά πολύ από τους παράγοντες και τα

αποτελέσματα που μπορούν να αποτελούν αντικείμενο μελέτης είτε να ορίζεται πιο γενικά (Higgins & Green, 2008).

2) Καθορισμός των κριτηρίων εισόδου και αποκλεισμού των μελετών (Defining eligibility criteria)

Πριν από την αναζήτηση της βιβλιογραφίας θα πρέπει να καθοριστούν τα κριτήρια εισόδου ή αποκλεισμού των μελετών με σκοπό τη χρησιμοποίηση των πιο κατάλληλων μελετών στη συστηματική ανασκόπηση. Τα κριτήρια ένταξης και αποκλεισμού των μελετών αναφέρονται στο πρωτόκολλο μιας συστηματικής ανασκόπησης, οφείλουν να είναι αντικειμενικά και να μην αφήνονται στην υποκειμενικότητα των μελετητών (Γαλάνης, 2009). Ως κριτήρια αποκλεισμού συνήθως ορίζονται οι μελέτες που κατά τον χρόνο μεταβλήθηκε η καθιερωμένη θεραπευτική παρέμβαση, μελέτες που δεν είναι στα αγγλικά λόγω αδυναμίας μετάφρασής τους ή αδημοσίευτες μελέτες (Γαλάνης, 2009). Με τον όρο αδημοσίευτες μελέτες θεωρούμε διδακτορικές διατριβές, μελέτες συνεδρίων ή μελέτες που εξήγαγαν συμπεράσματα με αρνητική έκβαση (Γαλάνης, 2009).

3) Αναζήτηση της βιβλιογραφίας (Searching for studies)

Είναι απαραίτητο να καθορίζεται αναλυτικά και με σαφήνεια η στρατηγική αναζήτησης των κατάλληλων μελετών στη συστηματική ανασκόπηση ώστε να είναι εφικτή η επανάληψή της ή ο έλεγχος της αξιοπιστίας της (Γαλάνης, 2009). Για την εύρεση μελετών οι ερευνητές στρέφονται στην αναζήτηση της πληροφορίας στο διαδίκτυο χρησιμοποιώντας σχετικές λέξεις κλειδιά, σε γραπτές εργασίες, σε μελέτες ακαδημαϊκών, κρατικών ή ιδιωτικών οργανισμών (Γαλάνης, 2009). Για τη διευκόλυνση της αναζήτησης της πληροφορίας έχει συμβάλει ο διεθνής μη κερδοσκοπικός οργανισμός Cochrane Collaboration. Η Cochrane Collaboration έχει αναπτύξει μια βάση δεδομένων, την CENTRAL (The Cochrane Central Register of Controlled Trials), η οποία αποτελεί την πιο σημαντική πηγή αναζήτησης ελεγχόμενων δοκιμών (Higgins & Green, 2008). Η CENTRAL αποτελεί μέρος της Cochrane Library και στοχεύει στον εντοπισμό, στην αξιολόγηση και στην αξιόπιστη σύνθεση των

αποδεικτικών στοιχείων ώστε να δημιουργηθούν συστηματικές ανασκοπήσεις βάση αυστηρών μεθοδολογικών κριτηρίων (Higgins & Green, 2008). Πέρα από την προετοιμασία προωθεί την προσβασιμότητα των συστηματικών ανασκοπήσεων σε μελετητές και επιστήμονες υγείας έτσι ώστε να είναι καλά πληροφορημένοι για να προβούν στις σωστές λήψεις αποφάσεων υγειονομικής περίθαλψης (Higgins & Green, 2008). Ο οργανισμός παρέχει πληθώρα από συστηματικές ανασκοπήσεις και ερευνητικά πρωτόκολλα. Αποτελείται από ομάδες Cochrane Review Groups (CRGs) στις οποίες συνεργάζονται εθελοντικά ερευνητές από όλον τον κόσμο, επαγγελματίες υγείας και άνθρωποι που χρησιμοποιούν τις υπηρεσίες υγείας οι οποίοι συμβάλλουν στη συνεχή εξέλιξη της.

Πέρα από την CENTRAL, το διαδίκτυο διευκολύνει τη διαδικασία αναζήτησης με τις σημαντικότερες μηχανές αναζήτησης της επιστήμης υγείας όπως η MEDLINE (Medical Literature Analysis and Retrieval System Online), EMBASE (The Excerpta Medical Database), CANCERLIT (Cancer Literature), TOXLINE (Toxicology Literature Online), DAI (Dissertation Abstracts International) και το Εθνικό Κέντρο Τεκμηρίωσης (National Documentation Centre) (Γαλάνης, 2009). Άλλες πηγές αναζήτησης αποτελούν οι εθνικές βάσεις δεδομένων (National and regional databases) στις οποίες περιέχονται βιβλιογραφία που παράχθηκε από την τοπική κοινωνία και δεν είναι διαθέσιμη κάπου αλλού. Η αναζήτηση σε ειδικές βιβλιογραφίες (Subject-specific electronic bibliographic databases) με βάση το αντικείμενο του ενδιαφέροντος όπως σε θέματα υγείας, φαρμακολογίας, οικονομίας, ψυχολογίας, η αναζήτηση σε βάση δεδομένων διδακτορικών διατριβών (Theses databases) καθώς και της γκρίζας βιβλιογραφίας (Grey literature), δηλαδή της βιβλιογραφίας που δεν έχει δημοσιευθεί επίσημα (Higgins & Green, 2008).

4) Αξιολόγηση και επιλογή των μελετών (Selecting studies)

Κατά την αξιολόγηση και επιλογή των μελετών σκοπός είναι η επιλογή των καλύτερων μεθοδολογικών μελετών που θα οδηγήσουν σε μια πιο αξιόπιστη συστηματική ανασκόπηση. Το γεγονός αυτό σημαίνει ότι αν βάλεις κακής ποιότητας μελέτες στη μετα-ανάλυση θα πάρεις λανθασμένα

αποτελέσματα (garbage in-garbage out). Τα κύρια στάδια επιλογής των μελετών αποτελούν η σύνθεση των αναφορών που μπορεί να προέρχονται από το ίδιο άρθρο, η απόρριψη μη σχετικών άρθρων με βάση τον τίτλο και την περίληψή τους, η ανάκτηση του πλήρες κειμένου, η σύνθεση των άρθρων από την ίδια μελέτη, ο έλεγχος για την ικανοποίηση των κριτηρίων καταλληλότητας της μελέτης και η αιτιολόγηση για την επιλογή της ώστε να προβεί στη συλλογή των δεδομένων της (Higgins & Green, 2008). Όταν μια μελέτη δεν ικανοποιεί τα κριτήρια καταλληλότητας τότε απορρίπτεται από την περαιτέρω ανάλυση. Η Cochrane περιλαμβάνει μια λίστα με τις μελέτες που απορρίφθηκαν καθώς και τον κύριο λόγο απόρριψής τους (Higgins & Green, 2008).

5) Καταγραφή των δεδομένων (Collecting data)

Στο στάδιο αυτό οι μελετητές πρέπει να καθορίσουν τα στοιχεία που πρέπει να αντλήσουν από τη μελέτη έτσι ώστε να μπορούν να προχωρήσουν στην ανάλυση τους με όλες τις πληροφορίες που κρίνονται απαραίτητες σύμφωνα με το θέμα που μελετούν. Κατά την καταγραφή των δεδομένων κρίνεται απαραίτητη η άντληση των βασικών χαρακτηριστικών των συμπεριλαμβανόμενων μελετών καθώς και η αξιολόγηση της ποιότητας των μελετών με πληροφορίες που διευκολύνουν την αξιολόγηση του κινδύνου μεροληψίας μιας μελέτης (Assessments of the risk of bias) (Higgins & Green, 2008). Τα βασικά χαρακτηριστικά των μελετών αποτελούν τα χαρακτηριστικά των συμμετεχόντων όπως το φύλο, η ηλικία καθώς και πληροφορίες για το αντικείμενο μελέτης καθώς και το είδος των θεραπειών που μπορεί να είναι φαρμακολογικές ή μη (Higgins & Green, 2008). Κατά τις φαρμακολογικές θεραπείες μας ενδιαφέρει η δοσολογία των φαρμάκων, η χρονική διάρκεια μεταξύ των δοσολογιών και η διάρκεια της θεραπείας (Higgins & Green, 2008). Μη φαρμακολογική δόση μπορεί να αποτελεί κάποιου είδους ψυχικής θεραπείας ή γυμναστικής με το ενδιαφέρον να επικεντρώνεται στο περιεχόμενο της θεραπείας και στον τρόπο διεξαγωγής της (Higgins & Green, 2008). Επιπρόσθετα, βασικά χαρακτηριστικά των μελετών αποτελούν η ακεραιότητα των παρεμβάσεων οι οποίες πρέπει να είναι σύνθετες και όχι προληπτικές, οι

μετρήσεις των αποτελεσμάτων, η παρουσία ανεπιθύμητων αποτελεσμάτων (adverse outcomes), τα οποία μπορούν να φανούν χρήσιμα κατά τη μελέτη επιβλαβών αποτελεσμάτων, καθώς και τα αποτελέσματα των μελετών (Higgins & Green, 2008). Για τον έλεγχο της ύπαρξης μιας τυχαιοποιημένης μελέτης, οι ερευνητές πρέπει να προβούν στην άντληση πληροφοριών με σκοπό την αξιολόγηση του κινδύνου μεροληψίας. Οι μελετητές πρέπει να εξετάζουν μια μελέτη ως προς την ύπαρξη ενός μη ολοκληρωμένου αποτελέσματος (incomplete outcome), την τύφλωση των συμμετεχόντων (χρήση ερωτηματολογίου ή μη - bleeding), την κατανομή μιας σειράς απόκρυψης (allocation sequence concealment) και την επιλεκτική αναφορά του αποτελέσματος (selective outcome reporting) (Higgins & Green, 2008). Οι παραπάνω πληροφορίες για την αξιολόγηση του κινδύνου μεροληψίας διεξάγονται με τη χρήση του εργαλείου της Cochrane Collaboration (Cochrane Collaboration tool) (Higgins & Green, 2008). Έπειτα από την άντληση των απαραίτητων πληροφοριών, οι μελετητές προβαίνουν στην καταγραφή τους σε προκαθορισμένες πλατφόρμες με στόχο την ανάλυση των δεδομένων.

6) Στατιστική ανάλυση (Statistical analysis)

Επόμενο βήμα μετά από τη συλλογή και την καταγραφή των δεδομένων όλων των μελετών μιας συστηματικής ανασκόπησης αποτελεί η στατιστική ανάλυση. Σκοπός της στατιστικής ανάλυσης είναι η σύνθεση των αποτελεσμάτων των μεμονωμένων μελετών σ' ένα αποτέλεσμα. Η στατιστική διαδικασία αυτή πραγματοποιείται με τη διαδικασία της μετα-ανάλυσης. Η μετα-ανάλυση είναι μια στατιστική μέθοδος κατά την οποία συντίθενται τα αποτελέσματα των μεμονωμένων μελετών με βάση μια ερευνητική υπόθεση με σκοπό τη δημιουργία ενός συγκεντρωτικού αποτελέσματος (Higgins & Green, 2008). Με τη βοήθεια της μετα-ανάλυσης αυξάνεται η ισχύς και η ακρίβεια καθώς συνδέουμε τα αποτελέσματα των μελετών (Higgins & Green, 2008). Ακόμα, οι μεμονωμένες μελέτες μπορεί να δημιουργούν αντικρουόμενα αποτελέσματα και για το λόγο αυτό με την μετα-ανάλυση γίνεται εφικτό να εξερευνηθούν και να ποσοτικοποιηθούν οι λόγοι που οδηγούν σε αυτήν τη διαφορά (Higgins & Green, 2008). Επιπρόσθετα, μπορούμε να

μελετήσουμε θέματα που δεν θα μπορούσαμε να έχουμε την πληροφορία από μεμονωμένες μελέτες (Higgins & Green, 2008). Η μετα-ανάλυση αποτελεί προαιρετικό στάδιο της συστηματικής ανασκόπησης καθώς πέρα από τα θετικά χαρακτηριστικά που μπορεί να έχει, υπάρχουν περιπτώσεις που η πραγματοποίηση της μπορεί να προσφέρει αρνητικές συνέπειες. Τέτοιες περιπτώσεις μπορεί να αποτελεί το ενδεχόμενο μιας συστηματικής ανασκόπησης με μεγάλη ποικιλομορφία μελετών (Higgins & Green, 2008). Ακόμα όταν σε μια συστηματική ανασκόπηση έχουμε μελέτες με κίνδυνο μεροληψίας τα συγκεντρωτικά αποτελέσματα μπορεί να είναι παραπλανητικά (Higgins & Green, 2008). Σε αυτές τις περιπτώσεις είναι καλύτερη η αποφυγή της μετα-ανάλυσης κατά τη στατιστική ανάλυση. Όταν δεν πραγματοποιείται η μετα-ανάλυση τότε οι μελετητές παρουσιάζουν τα αποτελέσματα των μεμονωμένων μελετών και όχι ενός συγκεντρωτικού αποτελέσματος.

7) Παρουσίαση αποτελεσμάτων (Presenting results)

Στο στάδιο της παρουσίασης των αποτελεσμάτων παρατίθενται ένας πίνακας με τα βασικά χαρακτηριστικά των συμπεριλαμβανόμενων μελετών, μέθοδοι διεξαγωγής, συμμετέχοντες, θεραπείες και αποτελέσματα καθώς και η παρουσίαση ενός πίνακα με τη μεροληψία αναφέροντας το είδος μεροληψίας που μπορεί να παρουσιάζει η κάθε μελέτη (Higgins & Green, 2008). Κατά την παρουσίαση των αποτελεσμάτων πρέπει να αναφερθεί ο αριθμός των ανακτηθέντων μελετών που προέκυψε κατά την αναζήτηση και οι αιτίες αποκλεισμού των μελετών που τελικά δεν συμπεριλήφθηκαν στην ανασκόπηση (Πατελάρου & Μπροκαλάκη, 2010). Ακόμα, εάν η συστηματική ανασκόπηση περιείχε στατιστική ανάλυση οφείλεται να παρουσιαστούν τα αποτελέσματα της μετα-ανάλυσης καθώς και κάποια διαγράμματα της ανάλυσης όπως το διάγραμμα δάσους (forest plot) ή διαγράμματα της ανάλυσης ευαισθησίας ή του σφάλματος δημοσίευσης τα οποία αποτελούν μια επιπρόσθετη διερεύνηση κατά την διεξαγωγή της μετα-ανάλυσης (Higgins & Green, 2008). Για την ερμηνεία αυτών των διαγραμμάτων παραπέμπουμε στο Κεφάλαιο 3.

8) Ερμηνεία αποτελεσμάτων και εξαγωγή συμπερασμάτων (Interpreting results and drawing conclusions)

Το τελευταίο βήμα αποτελεί η ερμηνεία των αποτελεσμάτων και η εξαγωγή των συμπερασμάτων ως προς το αντικείμενο μελέτης της συστηματικής ανασκόπησης. Τα ευρήματα της συστηματικής ανασκόπησης αποσκοπούν στην καθοδήγηση των επιστημόνων υγείας έτσι ώστε να λαμβάνουν σωστές αποφάσεις κατά την εφαρμογή της κλινικής πρακτικής. Για το λόγο αυτό κρίνεται απαραίτητη η σωστή παρουσίαση των αποτελεσμάτων. Για τη σαφή διατύπωση των ευρημάτων πρέπει να παρουσιάζονται πληροφορίες για όλα τα σημαντικά αποτελέσματα, ενώ η διατύπωση των αποτελεσμάτων να μην αφήνεται στην υποκειμενικότητα του μελετητή (Higgins & Green, 2008). Ως προς την ανάλυση, η σαφής παρουσίαση πραγματοποιείται με τη χρήση διαστημάτων εμπιστοσύνης και p-τιμών ώστε να παρουσιάζονται τα επίπεδα της στατιστικής σημαντικότητας με αντικειμενικό τρόπο (Higgins & Green, 2008).

2.2 Τι είναι η μετα-ανάλυση (meta-analysis)

Πολλές συστηματικές ανασκοπήσεις κατά το στάδιο της στατιστικής ανάλυσης περιέχουν μετα-αναλύσεις (Higgins & Green, 2008). Η μετα-ανάλυση είναι μια στατιστική μέθοδος η οποία συνθέτει τις μελέτες που έχουν επιλεγεί από τη συστηματική ανασκόπηση με σκοπό τη δημιουργία ενός συγκεντρωτικού αποτελέσματος (Γαλάνης, 2009). Επομένως οι μονάδες σε μια μετά-ανάλυση δεν είναι άτομα αλλά μελέτες οι οποίες συμβάλουν στο αποτέλεσμα με βάση την πληροφορία που διαθέτουν. Έτσι μεγάλες μελέτες συμβάλλουν περισσότερο στη μετα-ανάλυση απ' ότι οι μικρές μελέτες. Η σύνθεση των μελετών μέσω της διαδικασίας της μετα-ανάλυσης αυξάνει την ακρίβεια και την εγκυρότητα των αποτελεσμάτων.

2.3 Μέτρο σχέσης /μέγεθος επίδρασης (effect size)

Οι μελέτες συγκρίνουν εκβάσεις μεταξύ δυο ομάδων, την ομάδα ελέγχου και την πειραματική ομάδα. Στην ομάδα ελέγχου χορηγείται συνήθως ένα εικονικό φάρμακο (placebo) ή η συνηθισμένη θεραπεία έτσι ώστε να συγκριθεί με την πειραματική ομάδα στην οποία χορηγείται η θεραπεία

της οποίας θέλουμε να ελέγξουμε την αποτελεσματικότητα της. Έχοντας χωρίσει το δείγμα της μελέτης στις δύο ομάδες ορίζουμε ως μέτρο σχέσης /μέγεθος επίδρασης (effect size) τους τρόπους με τους οποίους μπορούμε να συγκρίνουμε τις δύο ομάδες και να ποσοτικοποιήσουμε τη σχετική αποτελεσματικότητα των θεραπειών. Αρχικό βήμα για την εξαγωγή της μετα-ανάλυσης αποτελεί ο υπολογισμός του μεγέθους επίδρασης καθώς από κάθε μελέτη απαιτείται το μέγεθος επίδρασης (effect size) και το τυπικό της σφάλμα (standard error). Κατά την εφαρμογή μιας μελέτης μπορούμε να εξάγουμε διάφορα είδη δεδομένων. Αυτά αποτελούν τα διχότομα (dichotomous), τα συνεχή (continuous), τα διατεταγμένα (ordinal), οι μετρήσεις και τα ποσοστά (counts and rates) καθώς και τα time-to-event δεδομένα (Higgins & Green, 2008):

- Διχότομα (dichotomous): Το αποτέλεσμα του κάθε ατόμου είναι ένα από τις δύο κατηγορίες ανταπόκρισης (γεγονός ή μη γεγονός).
- Συνεχή (continuous): Το αποτέλεσμα του κάθε ατόμου μπορεί να πάρει οποιαδήποτε συνεχή τιμή και επομένως αποτελεί μέτρηση μιας αριθμητικής ποσότητας.
- Διατεταγμένα (ordinal): Κατά τα διατεταγμένα δεδομένα οι ασθενείς κατατάσσονται σε διατεταγμένες κατηγορίες ανάλογα με το αποτέλεσμα που εμφανίζουν.
- Μετρήσεις και ρυθμοί (counts and rates): Τα counts data μετράνε τον αριθμό των φορών που εκδηλώθηκε το γεγονός. Τα counts data χωρίζονται σε σπάνια γεγονότα (rare events) και κοινά γεγονότα (common events). Ο ρυθμός (rate) σχετίζει τις φορές που έχουν συμβεί τα γεγονότα προς τον χρόνο που θα μπορούσαν να έχουν συμβεί.
- Time-to-event: Το αποτέλεσμα λαμβάνεται ως το χρονικό διάστημα μέχρι το γεγονός. Στην κατηγορία αυτή ανήκουν τα δεδομένα επιβίωσης (survival) καθώς συνήθως τα δεδομένα αυτά εστιάζουν στο θάνατο άρα στο χρονικό διάστημα μέχρι την εμφάνιση θανάτου.

Στη διατριβή θα εστιάσουμε στα δύο είδη δεδομένων, τα διχότομα και τα συνεχή, τα οποία αποτελούν τους πιο συνηθισμένους τύπους

αποτελεσμάτων κατά τη διεξαγωγή μιας μελέτης. Για τα διχότομα δεδομένα τα μέτρα σχέσης /μεγέθη επίδρασης είναι ο λόγος αναλογιών (odds ratio), ο λόγος κινδύνου (risk ratio) και η διαφορά κινδύνου (risk difference). Για τον υπολογισμό των μέτρων σχέσης/μεγεθών επίδρασης και της τυπικής τους απόκλισης χρειαζόμαστε τις ποσότητες που παρατίθενται στον Πίνακα 2.1.

	Γεγονός (event)	Μη γεγονός (Non-event)	Σύνολο
Πειραματική ομάδα	α	β	α+β
Ομάδα ελέγχου	γ	δ	γ+δ

Πίνακας 2.1 Οι ποσότητες για τα διχότομα δεδομένα

Έχοντας τις ποσότητες του Πίνακα 2.1 μπορούμε να υπολογίσουμε το μέγεθος επίδρασης δηλαδή τους διαφορετικούς τρόπους με τους οποίους μπορούμε να συμπεράνουμε πόσο πιθανό είναι να συμβεί ένα γεγονός στην πειραματική ομάδα σε σχέση με την ομάδα ελέγχου με τη βοήθεια των κινδύνων και των αναλογιών. Η **αναλογία** ορίζεται ως το πηλίκο του αριθμού των γεγονότων δια του αριθμού των μη γεγονότων και ο **κίνδυνος** είναι το πηλίκο του αριθμού των ευνοϊκών γεγονότων δια του συνολικού αριθμού γεγονότων.

Για τα διχότομα δεδομένα το μέγεθος επίδρασης δίνεται από τη συνάρτηση $effect\ size = g(p_t) - g(p_c)$, όπου $p_t = \frac{\alpha}{\alpha+\beta}$, $n_1 = \alpha + \beta$ η πιθανότητα ανταπόκρισης και το μέγεθος στην πειραματική ομάδα και $p_c = \frac{\gamma}{\gamma+\delta}$, $n_2 = \gamma + \delta$ η πιθανότητα ανταπόκρισης και το μέγεθος στην ομάδα ελέγχου και τη συνάρτηση $g(\cdot)$ να αποτελεί τη συνάρτηση ένωσης (link function).

Α) Όταν η συνάρτηση $g(\cdot)$ είναι η ταυτοτική, το μέγεθος επίδρασης είναι η διαφορά κινδύνου, δηλαδή σε αυτήν την περίπτωση

$$effect\ size = g(p_t) - g(p_c) = p_t - p_c = \frac{\alpha}{\alpha + \beta} - \frac{\gamma}{\gamma + \delta} = RD$$

Το μέγεθος επίδρασης είναι (βλέπε ορισμούς με bold στη σελίδα 14) η διαφορά μεταξύ του κινδύνου στην πειραματική ομάδα και στην ομάδα ελέγχου, δηλαδή

$$RD = \frac{a}{a + \beta} - \frac{\gamma}{\gamma + \delta}$$

Η ασυμπτωτική διακύμανση δίνεται ως

$$\begin{aligned} \text{Var}(g(p_t) - g(p_c)) &= \\ &= \text{Var}(p_t - p_c) = \text{Var}(p_t) + \text{Var}(p_c) \\ &= \frac{p_t(1-p_t)}{n_1} + \frac{p_c(1-p_c)}{n_2} = \frac{a\beta}{(a+\beta)^3} + \frac{\gamma\delta}{(\gamma+\delta)^3} \end{aligned}$$

με $\text{Var}(p_t) = \frac{p_t(1-p_t)}{n_1}$ καθώς $a \sim \text{binomial}(n_1, p_t)$ με $\text{var}(a) =$

$$n_1 p_t(1-p_t) \text{ και } p_t = \frac{a}{n_1} \leftrightarrow \text{var}(p_t) = \frac{1}{n_1^2} \text{var}(a) = \frac{p_t(1-p_t)}{n_1}.$$

Επομένως $\text{Var}(RD) = \frac{a\beta}{(a+\beta)^3} + \frac{\gamma\delta}{(\gamma+\delta)^3}$, το κατά προσέγγιση τυπικό

σφάλμα $se(RD) = \sqrt{\frac{a\beta}{(a+\beta)^3} + \frac{\gamma\delta}{(\gamma+\delta)^3}}$ και το $(1-\alpha)\%$ διάστημα

εμπιστοσύνης της διαφοράς κινδύνου $RD \pm z_{\alpha/2} \sqrt{\text{Var}(RD)}$.

Όταν η διαφορά κινδύνου ισούται με το μηδέν δεν υπάρχει διαφορά ανάμεσα στην ομάδα ελέγχου και στην πειραματική ομάδα, ενώ όταν είναι μικρότερη από το μηδέν τότε το γεγονός είναι λιγότερο πιθανό να συμβεί στην πειραματική ομάδα σε σχέση με την ομάδα ελέγχου. Αντίθετα, όταν η διαφορά κινδύνου είναι μεγαλύτερη από το μηδέν τότε το γεγονός είναι περισσότερο πιθανό να συμβεί στην πειραματική ομάδα σε σχέση με την ομάδα ελέγχου.

B) Ο λόγος κινδύνου ορίζεται ως

$$RR = \frac{\text{κίνδυνος στην πειραματική ομάδα}}{\text{κίνδυνος στην ομάδα ελέγχου}} = \frac{\frac{a}{a+\beta}}{\frac{\gamma}{\gamma+\delta}} = \frac{\alpha(\gamma+\delta)}{\gamma(\alpha+\beta)}$$

Λόγω κανονικότητας μετράμε το λόγο κινδύνου στη λογαριθμική κλίμακα και επομένως όταν η συνάρτηση $g(\cdot)$ είναι ο λογάριθμος, το μέγεθος επίδρασης είναι ο λόγος κινδύνου

$$\begin{aligned} \text{effect size} &= g(p_t) - g(p_c) = \log(p_t) - \log(p_c) \\ &= \log \frac{\alpha}{\alpha + \beta} - \log \frac{\gamma}{\gamma + \delta} \end{aligned}$$

Η ασυμπτωτική διακύμανση για το μέγεθος επίδρασης κατά την πειραματική ομάδα θα είναι

$$\begin{aligned} \text{Var}(g(p_t)) &= \left(\frac{\partial g(p_t)}{\partial p_t} \right)^2 \text{Var}(p_t) \\ &= \left(\frac{\partial \log p_t}{\partial p_t} \right)^2 \cdot \left(\frac{p_t(1-p_t)}{n_1} \right) \\ &= \left(\frac{1}{p_t} \right)^2 \left(\frac{p_t(1-p_t)}{n_1} \right) = \frac{1-p_t}{n_1 p_t} \end{aligned}$$

Η ασυμπτωτική διακύμανση για το μέγεθος επίδρασης κατά την ομάδα ελέγχου θα είναι

$$\begin{aligned} \text{Var}(g(p_c)) &= \left(\frac{\partial g(p_c)}{\partial p_c} \right)^2 \text{Var}(p_c) \\ &= \left(\frac{\partial \log p_c}{\partial p_c} \right)^2 \cdot \left(\frac{p_c(1-p_c)}{n_2} \right) \\ &= \left(\frac{1}{p_c} \right)^2 \left(\frac{p_c(1-p_c)}{n_2} \right) = \frac{1-p_c}{n_2 p_c} \end{aligned}$$

Επομένως,

$$\text{Var}(g(p_t)) = \frac{1-p_t}{n_1 p_t} = \frac{\beta}{\alpha(\alpha+\beta)} = \frac{\beta+\alpha-\alpha}{\alpha(\alpha+\beta)} = \frac{1}{\alpha} - \frac{1}{\alpha+\beta}$$

και

$$\text{Var}(g(p_c)) = \frac{1-p_c}{n_2 p_c} = \frac{\delta}{\gamma(\gamma+\delta)} = \frac{\delta+\gamma-\gamma}{\gamma(\gamma+\delta)} = \frac{1}{\gamma} - \frac{1}{\gamma+\delta}$$

Συνεπώς η ασυμπτωτική διακύμανση για το λόγο κινδύνου θα είναι

$$\text{Var}(\log RR) = \text{Var}(g(p_t)) + \text{Var}(g(p_c)) = \frac{1}{\alpha} - \frac{1}{\alpha+\beta} + \frac{1}{\gamma} - \frac{1}{\gamma+\delta}$$

και το ασυμπτωτικό τυπικό σφάλμα

$$\text{se}(\log RR) = \sqrt{\frac{1}{\alpha} - \frac{1}{\alpha+\beta} + \frac{1}{\gamma} - \frac{1}{\gamma+\delta}}$$

Το $(1-\alpha)\%$ διάστημα εμπιστοσύνης για το λογάριθμο του λόγου κινδύνου είναι

$$\log RR \pm z_{\alpha/2} \sqrt{\text{Var}(\log RR)} = (L_{\text{riskratio}}, U_{\text{riskratio}})$$

και το $(1 - \alpha)\%$ διάστημα εμπιστοσύνης του λόγου κινδύνου

$$(e^{Lriskratio}, e^{Uriskratio})$$

Όταν ο λόγος κινδύνου ισούται με τη μονάδα τότε δεν υπάρχει διαφορά στον κίνδυνο ανάμεσα στην πειραματική ομάδα και στην ομάδα ελέγχου. Όταν ο λόγος κινδύνου είναι μικρότερος από τη μονάδα σημαίνει ότι το γεγονός είναι λιγότερο πιθανό να συμβεί στην πειραματική ομάδα σε σχέση με την ομάδα ελέγχου. Αντίθετα, όταν είναι μεγαλύτερο από τη μονάδα τότε το γεγονός είναι πιο πιθανό να συμβεί στην πειραματική ομάδα σε σχέση με την ομάδα ελέγχου.

Γ) Ο λόγος αναλογιών ορίζεται ως

$$OR = \frac{\text{αναλογία στην πειραματική ομάδα}}{\text{αναλογία στην ομάδα ελέγχου}} = \frac{\frac{\alpha}{\beta}}{\frac{\gamma}{\delta}} = \frac{\alpha\delta}{\beta\gamma}$$

Λόγω κανονικότητας μετράμε το λόγο αναλογιών στη λογαριθμική κλίμακα και επομένως όταν η συνάρτηση $g(\cdot)$ είναι η $\text{logit}(x) = \log \frac{x}{1-x}$, το μέγεθος επίδρασης είναι ο λόγος αναλογιών

$$\begin{aligned} \text{effect size} &= g(p_t) - g(p_c) = \text{logit}(p_t) - \text{logit}(p_c) \\ &= \log \frac{\alpha}{\beta} - \log \frac{\gamma}{\delta} \end{aligned}$$

και

$$\log OR = \log \frac{\alpha}{\beta} - \log \frac{\gamma}{\delta}$$

Η ασυμπτωτική διακύμανση για το μέγεθος επίδρασης κατά την πειραματική ομάδα θα είναι

$$\begin{aligned} \text{Var}(g(p_t)) &= \left(\frac{\partial g(p_t)}{\partial p_t} \right)^2 \text{Var}(p_t) \\ &= \left(\frac{\partial \log \left(\frac{p_t}{1-p_t} \right)}{\partial p_t} \right)^2 \cdot \left(\frac{p_t(1-p_t)}{n_1} \right) \\ &= \left(\frac{1}{(1-p_t)p_t} \right)^2 \left(\frac{p_t(1-p_t)}{n_1} \right) \\ &= \frac{1}{n_1 p_t (1-p_t)} \quad \text{με} \quad \text{Var}(p_t) = \frac{p_t(1-p_t)}{n_1}. \end{aligned}$$

$$\text{Επομένως } \text{Var}(g(p_t)) = \frac{1}{\alpha} + \frac{1}{\beta}$$

Με παρόμοιο τρόπο, η ασυμπτωτική διακύμανση για το μέγεθος επίδρασης κατά την ομάδα ελέγχου και την πειραματική ομάδα θα είναι

$$\text{Var}(g(p_c)) = \frac{1}{\gamma} + \frac{1}{\delta}$$

και επομένως η ασυμπτωτική διακύμανση για το λόγο αναλογιών είναι

$$\begin{aligned} \text{Var}(\log OR) &= \text{Var}(g(p_t) - g(p_c)) = \text{Var}(g(p_t)) + \text{Var}(g(p_c)) \\ &= \frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} + \frac{1}{\delta} \end{aligned}$$

και το ασυμπτωτικό τυπικό σφάλμα

$$se(\log OR) = \sqrt{\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} + \frac{1}{\delta}}$$

Το $(1 - \alpha)\%$ διάστημα εμπιστοσύνης για το λογάριθμο του λόγου αναλογιών είναι

$$\log OR \pm z_{\alpha/2} \sqrt{\text{Var}(\log OR)} = (L_{\text{oddsratio}}, U_{\text{oddsratio}})$$

και το $(1 - \alpha)\%$ διάστημα εμπιστοσύνης του λόγου αναλογιών
($e^{L_{\text{oddsratio}}}, e^{U_{\text{oddsratio}}}$)

Όταν ο λόγος αναλογιών ισούται με τη μονάδα τότε δεν υπάρχει διαφορά στις αναλογίες ανάμεσα στην πειραματική ομάδα και στην ομάδα ελέγχου. Όταν ο λόγος αναλογιών είναι μικρότερος από τη μονάδα σημαίνει ότι το γεγονός είναι λιγότερο πιθανό να συμβεί στην πειραματική ομάδα σε σχέση με την ομάδα ελέγχου. Αντίθετα, όταν είναι μεγαλύτερος από τη μονάδα τότε το γεγονός είναι πιο πιθανό να συμβεί στην πειραματική ομάδα σε σχέση με την ομάδα ελέγχου. Για παράδειγμα, ένα OR ίσο με δύο σημαίνει ότι το γεγονός είναι δυο φορές πιο πιθανό να συμβεί στην πειραματική ομάδα σε σχέση με την ομάδα ελέγχου.

Συνοπτικά το μέγεθος επίδρασης και η ασυμπτωτική διακύμανση κάθε εκτιμητή παρουσιάζονται στον Πίνακα 2.2.

Συνάρτηση σύνδεσης	Μέγεθος επίδρασης	Εκτιμητής	Ασυμπτωτική διακύμανση
Ταυτότητα	Risk difference	$\frac{a}{a + \beta} - \frac{\gamma}{\gamma + \delta}$	$\frac{a\beta}{(a + \beta)^3} + \frac{\gamma\delta}{(\gamma + \delta)^3}$
Λογάριθμος	Log risk ratio	$\log \frac{a}{\alpha + \beta} - \log \frac{\gamma}{\gamma + \delta}$	$\frac{1}{a} - \frac{1}{a + \beta} + \frac{1}{\gamma} - \frac{1}{\gamma + \delta}$
Logit	Log odds ratio	$\log \frac{a}{\beta} - \log \frac{\gamma}{\delta}$	$\frac{1}{a} + \frac{1}{\beta} + \frac{1}{\gamma} + \frac{1}{\delta}$

Πίνακας 2.2: Συγκεντρωτικός πίνακας για το διχότομο αποτέλεσμα

Για τα συνεχή δεδομένα θεωρούμε μια μελέτη η οποία έχει δύο ανεξάρτητες ομάδες, την πειραματική και την ομάδα ελέγχου. Συμβολίζουμε με m_t και sd_t τη μέση τιμή και την τυπική απόκλιση για την πειραματική ομάδα και m_c και sd_c τη μέση τιμή και την τυπική απόκλιση για την ομάδα ελέγχου όπως περιγράφονται και στον Πίνακα 2.3. Τα μέτρα επίδρασης προκύπτουν από τη σύγκριση των μέσων στις δύο ομάδες καθώς τα πιο γνωστά μέτρα σχέσης είναι η διαφορά μέσων (mean difference) και η τυποποιημένη διαφορά μέσων (standardized mean difference).

	Μέσος (mean)	Τυπική απόκλιση (Standard deviation)	Μέγεθος δείγματος (Sample size)
Πειραματική ομάδα	m_t	sd_t	n_t
Ομάδα ελέγχου	m_c	sd_c	n_c

Πίνακας 2.3: Οι ποσότητες για τα συνεχή δεδομένα

A) Η διαφορά μέσων (mean difference)

Η διαφορά μέσων ορίζεται ως η διαφορά της μέσης τιμής της πειραματικής ομάδας m_t και της ομάδας ελέγχου m_c

$$MD = m_t - m_c$$

με το κατά προσέγγιση τυπικό σφάλμα $se = \sqrt{\frac{sd_t^2}{n_t} + \frac{sd_c^2}{n_c}}$.

B) Η τυποποιημένη διαφορά μέσων (standardized mean difference)

Η τυποποιημένη διαφορά μέσων με βάση τις πραγματικές τιμές της μελέτης όταν οι τυπικές αποκλίσεις των δύο ομάδων είναι ίσες $sd_t = sd_c = s$ δίνεται ως $\delta = \frac{m_t - m_c}{s}$ (Cohen, 1988).

Εκτιμώντας την τυποποιημένη διαφορά μέσων από τις μελέτες χρησιμοποιούμε το μέγεθος επίδρασης γνωστό ως Cohen's d

$$d = \frac{\bar{X}_1 - \bar{X}_2}{S}$$

όπου $S = \sqrt{\frac{(n_t-1)S_t^2 + (n_c-1)S_c^2}{n_t+n_c-2}}$, \bar{X}_1, \bar{X}_2 οι δειγματικές μέσες τιμές και S_t, S_c

οι δειγματικές πληθυσμιακές διακυμάνσεις για τις δύο ομάδες με το κατά

προσέγγιση τυπικό του σφάλμα $se(d) = \sqrt{\frac{n_t+n_c}{n_t n_c} + \frac{d^2}{2(n_t+n_c)}}$ (Cohen,

1988). Ο εκτιμητής Cohen's d δημιουργεί μεροληψία, χρησιμοποιώντας τη διόρθωση $J = 1 - \frac{3}{4(n_t+n_c-2)-1}$ παίρνουμε έναν αμερόληπτο εκτιμητή που προτείνεται από τον Hedges (1981) γνωστό ως Hedge's g

$$g = J \times d$$

με το κατά προσέγγιση τυπικό του σφάλμα $se(g) = \sqrt{J^2 \times se(d)^2}$. Όταν ένα μέτρο σχέσης για τα συνεχή δεδομένα είναι ίσο με το μηδέν τότε δεν υπάρχει διαφορά ανάμεσα στην πειραματική ομάδα και στην ομάδα ελέγχου, ενώ όταν είναι μικρότερο από το μηδέν σημαίνει ότι το γεγονός είναι λιγότερο πιθανό να συμβεί στην πειραματική ομάδα σε σχέση με την ομάδα ελέγχου. Αντίθετα, όταν είναι μεγαλύτερο από το μηδέν τότε το γεγονός είναι πιο πιθανό να συμβεί στην πειραματική ομάδα σε σχέση με την ομάδα ελέγχου.

ΚΕΦΑΛΑΙΟ 3

ΜΟΝΤΕΛΑ ΜΕΤΑ-ΑΝΑΛΥΣΗΣ

ΚΑΙ ΕΤΕΡΟΓΕΝΕΙΑ

3.1 Μοντέλα μετα-αναλύσεων

Η μετα-ανάλυση είναι μια στατιστική διαδικασία που αποτελείται από δυο στάδια. Ο υπολογισμός των μεγεθών επίδρασης και τα αντίστοιχα σφάλματα από κάθε μελέτη αποτελούν το πρώτο στάδιο της διαδικασίας της μετα-ανάλυσης. Έπειτα κατά το δεύτερο στάδιο η μετα-ανάλυση συνθέτει τα μεγέθη επίδρασης έτσι ώστε να προκύψει ένα συγκεντρωτικό αποτέλεσμα. Υπάρχουν δύο προσεγγίσεις για την σύνθεση των αποτελεσμάτων η μετα-ανάλυση σταθερών επιδράσεων (fixed effects) και η μετα-ανάλυση τυχαίων επιδράσεων (random effects).

3.2 Μετα-ανάλυση σταθερών επιδράσεων (Fixed Effect model)

Κατά τη μετα-ανάλυση σταθερών επιδράσεων θεωρούμε ότι όλες οι μελέτες που συμπεριλαμβάνονται στη μετα-ανάλυση εκτιμούν ένα κοινό πραγματικό μέγεθος επίδρασης μ (fixed effect) και η μεταβλητότητα ανάμεσα στα παρατηρούμενα μεγέθη επίδρασης (observed effect sizes) οφείλεται καθαρά και μόνο στο τυχαίο σφάλμα.

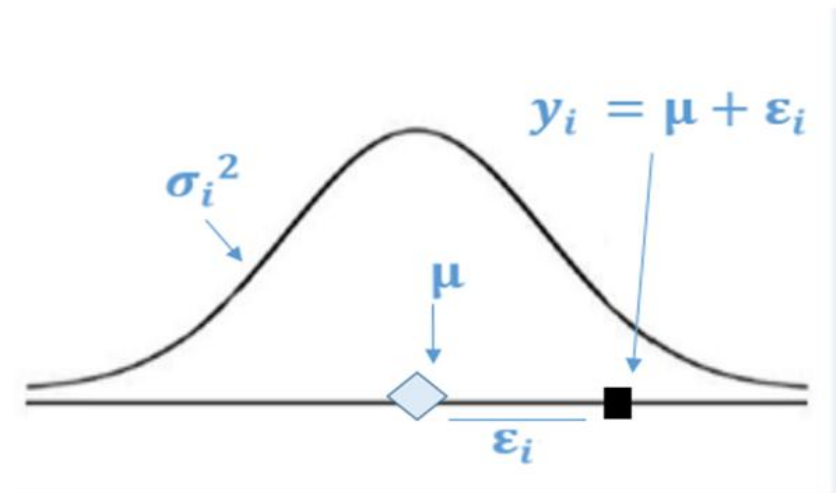
Έστω μία μελέτη με παρατηρούμενο μέγεθος επίδρασης y_1 και μ το πραγματικό μέγεθος επίδρασης (Borenstein et al., 2009). Το παρατηρούμενο μέγεθος επίδρασης θα απέχει από το πραγματικό όσο το τυπικό σφάλμα της εκτίμησης ε_1 που δημιουργείται μέσα σε μια μελέτη (Borenstein et al., 2009). Το παρατηρούμενο μέγεθος επίδρασης υποθέτουμε ότι κατανέμεται γύρω από την κανονική κατανομή με μέση τιμή όσο το πραγματικό μέγεθος επίδρασης μ και πληθυσμιακή διακύμανση σ^2 (Borenstein et al., 2009). Η μετα-ανάλυση σταθερών επιδράσεων στην περίπτωση μιας μελέτης περιγράφεται από το μοντέλο

$$y_1 = \mu + \varepsilon_1 \quad , \quad \varepsilon_1 \sim N(0, \sigma_1^2)$$

Έχοντας k μελέτες που προέρχονται από τον ίδιο πληθυσμό με πληθυσμιακή διακύμανση σ_i^2 και θεωρώντας ως $y_i, i = 1, \dots, k$ τα παρατηρούμενα μεγέθη επιδράσεων το μοντέλο που περιγράφει τη μετα-ανάλυση σταθερών επιδράσεων όπως παρουσιάζεται και στο διάγραμμα 3.1 είναι:

$$y_i = \mu + \varepsilon_i \quad , \quad \varepsilon_i \sim N(0, \sigma_i^2)$$

(Borenstein et al., 2009).



Διάγραμμα 3.1: Μετα-ανάλυση σταθερών επιδράσεων

Στόχος της μετα-ανάλυσης αποτελεί η σύνθεση των μελετών για τη δημιουργία ενός συγκεντρωτικού αποτελέσματος. Για το λόγο αυτό πρέπει να καθοριστεί ο τρόπος με τον οποίο η κάθε μελέτη θα επηρεάσει το αποτέλεσμα με βάση την πληροφορία που διαθέτει. Ο τρόπος συμβολής της κάθε μελέτης αποτελεί μια διαδικασία στη μετα-ανάλυση γνωστή ως ανάθεση βαρών. Για το μοντέλο τυχαίων επιδράσεων έχουν αναπτυχθεί μέθοδοι οι οποίες προσδίδουν η καθεμία διαφορετικές αναθέσεις βαρών. Χρησιμοποιούμε την πιο δημοφιλή μέθοδο ανάθεσης βαρών, την μέθοδο αντίστροφης διασποράς. Σύμφωνα με τη μέθοδο αυτή σε κάθε μελέτη ανατίθεται ως βάρος το αντίστροφο της διακύμανσης της. Επομένως, στην ειδική περίπτωση την μετα-ανάλυση σταθερών επιδράσεων,

λαμβάνοντας υπόψη ότι η προαναφερθείσα διακύμανση είναι ίση με την πληθυσμιακή διακύμανση προκύπτει ότι σε κάθε μελέτη ανατίθενται τα βάρη $w_{i,FE} = \frac{1}{\sigma_i^2}$.

Σκοπός της μετα-ανάλυσης αποτελεί η δημιουργία ενός συγκεντρωτικού αποτελέσματος από τη σύνθεση των μελετών. Έπειτα από την ανάθεση βαρών έχουμε την πληροφορία της ποσοτικοποίησης της επίδρασης που θα έχει η κάθε μελέτη. Επειδή όμως θέλουμε ένα συνολικό αποτέλεσμα στόχος μας αποτελεί η ποσοτικοποίησης της συνολικής επίδρασης. Η συνολική επίδραση αποτελεί επομένως την κατεύθυνση των αποτελεσμάτων όλων των μελετών και ονομάζεται συμβατικά διαμάντι. Η συνολική επίδραση ορίζεται ως ο σταθμισμένος μέσος όρος των παρατηρούμενων μεγεθών επίδρασης (Borenstein et al., 2009):

$$\hat{\mu}_{FE} = \frac{\sum_{i=1}^k w_{i,FE} y_i}{\sum_{i=1}^k w_{i,FE}}$$

Για τη σαφή παρουσίαση των αποτελεσμάτων από μια μετα-ανάλυση οι μελετητές οφείλουν να παρουσιάζουν τη συνολική επίδραση αλλά και ένα διάστημα εμπιστοσύνης ώστε να κρίνεται η στατιστική σημαντικότητα με μεγαλύτερη ακρίβεια. Για τη διεξαγωγή ενός διαστήματος εμπιστοσύνης απαραίτητο αποτελεί ο υπολογισμός της διακύμανσης της συνολικής επίδρασης η οποία ορίζεται για το μοντέλο σταθερών επιδράσεων ως το αντίστροφο του αθροίσματος των βαρών της μετα-ανάλυσης $V_M = \frac{1}{\sum_{i=1}^k w_{i,FE}}$ και το τυπικό σφάλμα της συνολικής διακύμανσης $s_M = \sqrt{V_M}$. Υποθέτοντας ότι η συνολική επίδραση-διαμάντι κατανέμεται ασυμπτωτικά κανονικά το κάτω (l) και άνω άκρο (u) του $(1 - \alpha)\%$ διαστήματος εμπιστοσύνης της είναι $l = \hat{\mu}_{FE} - z_{\alpha/2} \sqrt{V_M}$ και $u = \hat{\mu}_{FE} + z_{\alpha/2} \sqrt{V_M}$ (Borenstein et al., 2009).

3.3 Τι είναι ετερογένεια (heterogeneity)

Η μετα-ανάλυση συνθέτει διαφορετικές μελέτες που ικανοποιούν μια κοινή ερευνητική υπόθεση. Οποιαδήποτε διαφορετικότητα παρατηρείται μεταξύ των μελετών αυτών ονομάζεται ετερογένεια. Υπάρχουν τρία είδη ετερογένειας η κλινική (clinical heterogeneity), η μεθοδολογική (methodological heterogeneity) και η στατιστική (statistical heterogeneity)

(Higgins & Green, 2008). Η διαφορετικότητα κάθε κλινικής μελέτης ως προς τα βασικά χαρακτηριστικά διεξαγωγής της αποτελεί την κλινική ετερογένεια (Higgins & Green, 2008). Επομένως η κλινική ετερογένεια δημιουργείται από τη μεταβλητότητα των συμμετεχόντων, των παρεμβάσεων και των εκβάσεων (Higgins & Green, 2008). Ως προς τους συμμετέχοντες, μεταβλητότητα μπορεί να δημιουργεί η διαφορά ηλικίας καθώς η μια μελέτη μπορεί να έχει συμμετέχοντες νέων και η άλλη ενήλικων. Ακόμα, διαφορετική χορήγηση φαρμάκων, δοσολογίας είτε διάρκειας της θεραπείας μπορεί να οδηγήσει σε μεταβλητότητα των παρεμβάσεων. Τέλος, ο διαφορετικός τρόπος μέτρησης ενός αποτελέσματος μπορεί να δημιουργήσει μεταβλητότητα ως προς την έκβαση. Η μεθοδολογική ετερογένεια οφείλεται στη μέθοδο με την οποία θα εφαρμοστεί η κλινική μελέτη καθώς και στην ποιότητα των μελετών (Higgins & Green, 2008). Ουσιαστικά μπορεί να προκύπτει από τη διαφορετικότητα του σχεδιασμού της κάθε μελέτης είτε από τη διαφορετική ποιότητα των μελετών δηλαδή τη διαφορετικότητα στον κίνδυνο μεροληψίας μιας μελέτης (risk of bias) το οποίο σχετίζεται με τη διαφορετικότητα του τρόπου διεξαγωγής της, αν για παράδειγμα θα είναι τυφλή ή θα περιλαμβάνει ερωτηματολόγιο (bleeding) ή θα έχει ένα μη ολοκληρωμένο αποτέλεσμα (incomplete outcome) (Higgins & Green, 2008). Ένα τρίτο είδος μεταβλητότητας αποτελεί η στατιστική ετερογένεια καθώς η δημιουργία της οφείλεται στην παρουσία είτε μεθοδολογικής είτε κλινικής ετερογένειας είτε σε συνδυασμό αυτών (Higgins & Green, 2008). Από την οπτική πλευρά ενός στατιστικού, η στατιστική ετερογένεια οφείλεται στην επιπρόσθετη μεταβλητότητα των παρατηρούμενων μεγεθών επίδρασης πέρα του τυχαίου σφάλματος. Στη μετα-ανάλυση η στατιστική ετερογένεια ονομάζεται ετερογένεια (heterogeneity) και συμβολίζεται με τ^2 . Η ποσοτικοποίηση της ετερογένειας αποτελεί σημαντικό στάδιο κατά τη διεξαγωγή μετα-ανάλυσης καθώς επηρεάζει σημαντικά τα αποτελέσματα της κατά τη σύνθεση των μελετών. Κατά το μοντέλο σταθερών επιδράσεων παρουσιάστηκε ότι η μεταβλητότητα των παρατηρούμενων μεγεθών οφείλεται καθαρά και μόνο στην ύπαρξη του τυχαίου σφάλματος. Επομένως κατά το μοντέλο σταθερών επιδράσεων δεν έχουμε ύπαρξη

πρόσθετης μεταβλητότητας μεταξύ των μελετών και συνεπώς η ετερογένεια είναι μηδενική. Για τη συμβολή της ετερογένειας έχει αναπτυχθεί ένα ακόμα μοντέλο μετα-ανάλυσης από τον DerSimonian and Laird (1986) γνωστό ως μοντέλο τυχαίων επιδράσεων (Random Effect model).

3.4 Μετα-ανάλυση τυχαίων επιδράσεων (Random Effect model)

Κατά τη μετα-ανάλυση τυχαίων επιδράσεων οι διαφορετικές μελέτες εκτιμούν διαφορετικά αποτελέσματα τα οποία σχετίζονται (Higgins & Green, 2008). Οι διαφορές στα παρατηρούμενα μεγέθη επίδρασης αποδίδονται όχι μόνο στο τυχαίο σφάλμα, αλλά και σε διαφορές των πραγματικών επιδράσεων δηλαδή στη στατιστική ετερογένεια. Επομένως με το μοντέλο τυχαίων επιδράσεων η μετα-ανάλυση λαμβάνει υπόψη τη διαφορετικότητα των μελετών, εξάγοντας ένα συγκεντρωτικό αποτέλεσμα το οποίο περιλαμβάνει την ετερογένεια των μελετών.

Έστω μία μελέτη με το παρατηρούμενο μέγεθος επίδρασης y_1 να απέχει από το πραγματικό μέγεθος επίδρασης θ_1 λόγω του τυπικού σφάλματος της εκτίμησης μέσα σε μια μελέτη ε_1 και να κατανέμεται γύρω από μία κατανομή με πραγματικό μέγεθος επίδρασης θ_1 και με διακύμανση σ^2 (Borenstein et al., 2009). Το πραγματικό μέγεθος επίδρασης θ_1 υποθέτουμε ότι κατανέμεται γύρω από μια κανονική κατανομή με μέση τιμή μ και διακύμανση τ^2 (Borenstein et al., 2009). Η μετα-ανάλυση τυχαίων επιδράσεων στην περίπτωση της μιας μελέτης περιγράφεται από το μοντέλο (Borenstein et al., 2009):

$$y_1 = \theta_1 + \varepsilon_1 \quad , \quad \varepsilon_1 \sim N(0, \sigma^2)$$

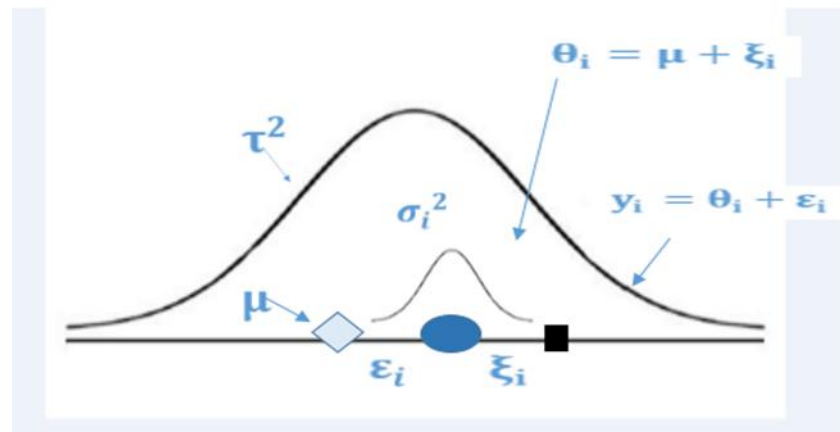
$$\theta_1 = \mu + \xi_1 \quad , \quad \xi_1 \sim N(0, \tau^2)$$

Έχοντας k μελέτες που προέρχονται από τον ίδιο πληθυσμό με πληθυσμιακή διακύμανση σ_i^2 και θεωρώντας ως $y_i, i = 1, \dots, k$ τα παρατηρούμενα μεγέθη επιδράσεων το μοντέλο που περιγράφει τη μετα-ανάλυση τυχαίων επιδράσεων είναι:

$$y_i = \theta_i + \varepsilon_i \quad , \quad \varepsilon_i \sim N(0, \sigma_i^2)$$

$$\theta_i = \mu + \xi_i \quad , \quad \xi_i \sim N(0, \tau^2)$$

Borenstein et al. (2009) με την οπτική παρουσίαση του μοντέλου στο Διάγραμμα 3.2.



Διάγραμμα 3.2 Μετα-ανάλυση τυχαίων επιδράσεων

Κατά το μοντέλο τυχαίων επιδράσεων ως μέθοδο για την ανάθεση βαρών εφαρμόζουμε την πιο δημοφιλή μέθοδο του αντιστρόφου της συνολικής διακύμανσης. Έχοντας ότι η συνολική διακύμανση αποτελείται από την πληθυσμιακή διακύμανση και την ετερογένεια, τα βάρη για το μοντέλο τυχαίων επιδράσεων θα είναι

$$w_{i,RE} = \frac{1}{\sigma_i^2 + \tau^2}$$

(Borenstein et al., 2009). Κατά την εφαρμογή, δεν υπάρχει η επίγνωση της ποσοτικοποίησης της ετερογένειας τ^2 μεταξύ των μελετών. Επομένως για τη διεκπεραίωση της ανάθεσης βαρών κρίνεται απαραίτητη η εκτίμηση της ετερογένειας των μελετών στους υπολογισμούς. Για τον εκτιμητή της ετερογένειας $\hat{\tau}^2$ έχουν αναπτυχθεί πολλές μέθοδοι εκτίμησης οι οποίες θα περιγραφούν αναλυτικά στο Κεφάλαιο 4.

Έχοντας πραγματοποιήσει την ανάθεση βαρών επόμενο βήμα καθίσταται η εξαγωγή της συνολικής επίδρασης που ουσιαστικά αποτελεί ένα σταθμισμένο μέσο όρο των παρατηρούμενων μεγεθών επίδρασης

$$\hat{\mu}_{RE} = \frac{\sum_{i=1}^k w_{i,RE} Y_i}{\sum_{i=1}^k w_{i,RE}}$$

(Borenstein et al., 2009). Η διακύμανση της συνολικής επίδρασης ορίζεται ως το αντίστροφο του αθροίσματος των βαρών της μετα-ανάλυσης $V_M = \frac{1}{\sum_{i=1}^k w_{i,RE}}$ και το τυπικό σφάλμα ως η τετραγωνική ρίζα της συνολικής διακύμανσης $s_M = \sqrt{V_M}$. Υποθέτοντας ότι η συνολική επίδραση κατανέμεται ασυμπτωτικά, το κάτω (l) και άνω άκρο (u) του $(1 - \alpha)\%$ διαστήματος εμπιστοσύνης της είναι $l = \hat{\mu}_{RE} - z_{\alpha/2} \sqrt{V_M}$ και $u = \hat{\mu}_{RE} + z_{\alpha/2} \sqrt{V_M}$ (Borenstein et al., 2009).

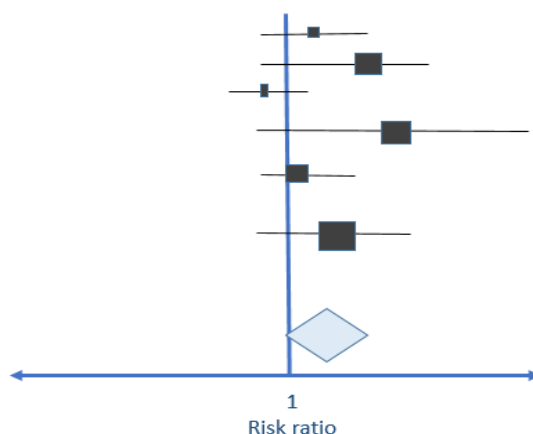
3.5 Διάγραμμα δάσους (forest plot)

Η οπτική παρουσίαση των αποτελεσμάτων στη μετα-ανάλυση μπορεί να πραγματοποιηθεί με το διάγραμμα δάσους (forest plot). Ουσιαστικά, πρόκειται για ένα διάγραμμα απεικόνισης των εκτιμήσεων καθώς και των διαστημάτων εμπιστοσύνης της κάθε μελέτης αλλά και της συνολικής επίδρασης. Στο διάγραμμα η κατακόρυφη αντιπροσωπεύει τη γραμμή χωρίς αποτέλεσμα ενώ η οριζόντια τη γραμμή με το μέγεθος επίδρασης (Nikolakoroulou et al., 2014a). Έχοντας ως μέγεθος επίδρασης τον λόγο κινδύνου παραθέτουμε το Διάγραμμα 3.3 για μια οπτική παρουσίαση του. Όταν το μέγεθος επίδρασης είναι ο λόγος αναλογιών ή ο λόγος κινδύνου τότε η κατακόρυφη γραμμή θα τέμνει την οριζόντια στη μονάδα ενώ όταν το μέγεθος επίδρασης είναι η διαφορά κινδύνου, η μέση διαφορά ή η τυποποιημένη μέση διαφορά τότε η γραμμή χωρίς αποτέλεσμα θα τέμνει τη γραμμή με το μέγεθος επίδρασης στο σημείο μηδέν.

Εκτός από τον οριζόντιο άξονα ο οποίος αποτελεί τη γραμμή με το μέγεθος επίδρασης, το διάγραμμα δάσους αποτελείται από οριζόντιες γραμμές οι οποίες αντιπροσωπεύουν και μια μελέτη. Σε αυτές τις γραμμές το τετράγωνο παριστάνει την εκτίμηση της κάθε μελέτης και το μέγεθός του δηλώνει το μέγεθος της μελέτης, ενώ το μήκος της γραμμής δηλώνει το διάστημα εμπιστοσύνης της εκτίμησης (Nikolakoroulou et al., 2014).

Στο κάτω μέρος του διαγράμματος δάσους απεικονίζεται το αποτέλεσμα της μετα-ανάλυσης με τη συνολική επίδραση το οποίο συμβατικά ονομάζεται διαμάντι και απεικονίζεται με ένα ρόμβο και το διάστημα

εμπιστοσύνης του. Για την εξαγωγή αποτελεσμάτων το ενδιαφέρον επικεντρώνεται στη συμπεριφορά του διαμαντιού. Εάν το διαμάντι βρίσκεται δεξιά από το σημείο τομής των δύο ευθειών η έκβαση είναι υπέρ της πειραματικής ομάδας ενώ αν βρίσκεται αριστερά η έκβαση είναι υπέρ της ομάδας ελέγχου (Nikolakoroulou et al., 2014). Στην περίπτωση του Διαγράμματος 3.3 το διαμάντι βρίσκεται δεξιά από την γραμμή χωρίς αποτέλεσμα και συνεπώς η έκβαση είναι υπέρ της πειραματικής ομάδας.



Διάγραμμα 3.3: Διάγραμμα δάσους (Forest plot)

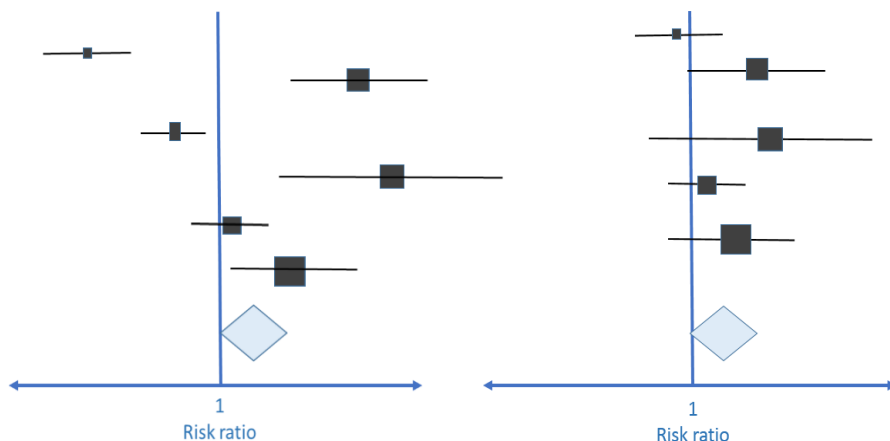
3.6 Εξέταση και τρόποι υπολογισμού της ετερογένειας

Κατά την εφαρμογή της μετα-ανάλυσης είναι σημαντική η εξέταση της ύπαρξης ετερογένειας καθώς και η ποσοτικοποίηση της.

I. Εξέταση ετερογένειας με το διάγραμμα δάσους

Ένας οπτικός τρόπος για την εξέταση της ύπαρξης ετερογένειας γίνεται με τη βοήθεια του διαγράμματος δάσους. Κατά την εξέταση αυτή πρέπει να κριθεί εάν τα διαστήματα εμπιστοσύνης των μελετών στο διάγραμμα δάσους έχουν κοινά σημεία, δηλαδή εάν αλληλεπικαλύπτονται οι οριζόντιες γραμμές των μελετών (Higgins & Green, 2008). Κατά την ύπαρξη ετερογένειας, τα διαστήματα εμπιστοσύνης κάποιων μελετών δεν έχουν κοινά σημεία όπως φαίνεται στο αριστερό διάγραμμα δάσους του Διαγράμματος 3.4, ενώ όταν τα διαστήματα εμπιστοσύνης μεταξύ

κάποιων μελετών έχουν κοινά σημεία τότε δεν υπάρχει ετερογένεια όπως φαίνεται στο δεξί διάγραμμα δάσους του Διαγράμματος 3.4. Όσο περισσότερο δεν αλληλεπικαλύπτονται τα διαστήματα εμπιστοσύνης των μελετών στο διάγραμμα δάσους τόσο μεγαλύτερη ετερογένεια έχουμε.



Διάγραμμα 3.4: Εξέταση ύπαρξης ετερογένειας στο διάγραμμα δάσους

II. Q τεστ

Ένας τρόπος για τον υπολογισμό της ετερογένειας επιτυγχάνεται με το Q τεστ. Θεωρώντας k μελέτες σε μία μετα-ανάλυση ενδιαφερόμαστε για τον έλεγχο της μηδενικής υπόθεσης H_0 έναντι της εναλλακτικής H_a

H_0 : οι μελέτες είναι ομοιογενείς

H_a : οι μελέτες δεν είναι ομοιογενείς

Ο έλεγχος της υπόθεσης πραγματοποιείται με τη χρήση του στατιστικού $Q = \sum_{i=1}^k w_{i,FE} (y_i - \hat{\mu}_{FE})^2$ γνωστό ως το Q –στατιστικό του Cochran το οποίο εντάχθηκε από τον Cochran (1954) όπου $w_{i,FE} = \frac{1}{\sigma_i^2}$, $\hat{\mu}_{FE} = \frac{\sum_{i=1}^k w_i y_i}{\sum_{i=1}^k w_i}$ τα βάρη και η συνολική επίδραση του μοντέλου σταθερών επιδράσεων. Το στατιστικό του Cochran Q ακολουθεί την χ^2 κατανομή με $k - 1$ βαθμούς ελευθερίας (Higgins & Thompson, 2002). Έχοντας ότι

$$y_i \sim N(\mu, \sigma^2) \Rightarrow \frac{y_i - \mu}{\sigma} \sim N(0,1)$$

$$\frac{(y_i - \mu)^2}{\sigma^2} \sim N^2(0,1) \equiv \chi_1^2 \Rightarrow \sum_{i=1}^k w_i (y_i - \mu)^2 \sim \chi_k^2$$

$$Q = \sum_{i=1}^k w_i (y_i - \hat{\mu})^2 \sim \chi_{k-1}^2$$

Εάν η p -τιμή που προκύπτει από τον έλεγχο του στατιστικού είναι μεγαλύτερη από το επίπεδο σημαντικότητας δηλώνει την ύπαρξη ετερογένειας με την τιμή της να αναδεικνύει το μέγεθος της ετερογένειας (Higgins & Thompson, 2002).

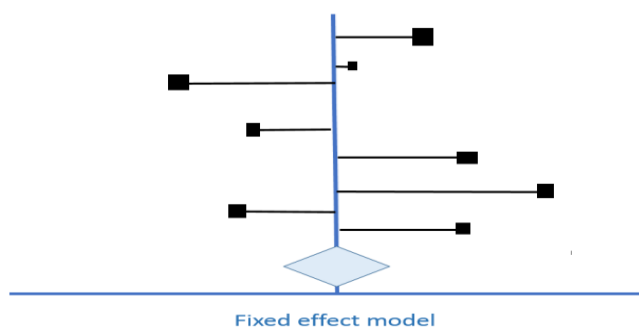
III. I^2

Ένας τρόπος υπολογισμού της ετερογένειας πραγματοποιείται με το μέτρο I^2 το οποίο εντάχθηκε από τους Higgins & Thompson (2002). Ο δείκτης I^2 μετράει το ποσοστό της μεταβλητότητας που οφείλεται στην ετερογένεια παρά στη δειγματική διακύμανση και εκφράζεται από τη σχέση $I^2 = \frac{Q-k+1}{Q} \times 100\%$ (Higgins & Thompson, 2002). Όταν το I^2 είναι ίσο με το μηδέν τότε γίνεται φανερό ότι δεν υπάρχει ετερογένεια δηλαδή το τ^2 είναι ίσο με το μηδέν και συνεπώς τα δύο μοντέλα μετα-αναλύσεων, το μοντέλο σταθερών επιδράσεων και μοντέλο τυχαίων επιδράσεων, ταυτίζονται. Συνήθως το I^2 είναι μεγαλύτερο από το 50% καθώς στις μετα-αναλύσεις το ποσοστό της μεταβλητότητας οφείλεται περισσότερο στην παρουσία ετερογένειας παρά στην δειγματική διακύμανση.

3.7 Τα μοντέλα μετα-ανάλυσης στο διάγραμμα δάσους

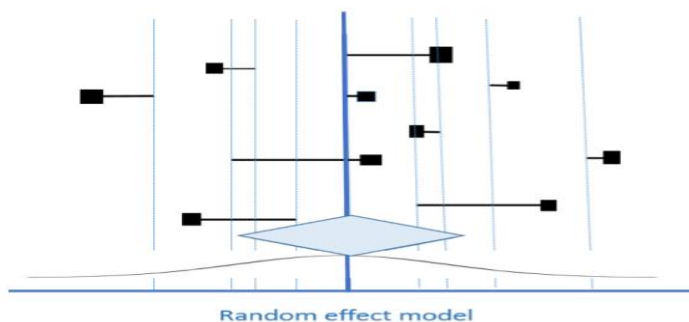
Για το μοντέλο σταθερών επιδράσεων τα παρατηρούμενα μεγέθη επίδρασης είναι τα τετράγωνα όπως φαίνονται στο Διάγραμμα 3.5. Οι οριζόντιες γραμμές από τα παρατηρούμενα μεγέθη επίδρασης μέχρι τη γραμμή χωρίς αποτέλεσμα αποτελεί το δειγματικό σφάλμα (Nikolakoroulou et al., 2014). Όπως παρατηρούμε από το Διάγραμμα 3.5 οι αποκλίσεις από τη συνολική επίδραση-διαμάντι οφείλονται καθαρά και μόνο στο δειγματικό σφάλμα (Nikolakoroulou et al., 2014). Συνεπώς το διαμάντι για το μοντέλο σταθερών επιδράσεων αποτελεί η συνολική

«αληθινή» κοινή επίδραση (Nikolakoroulou et al., 2014).



Διάγραμμα 3.5: Διάγραμμα δάσους για την μετα-ανάλυση σταθερών επιδράσεων

Για το μοντέλο τυχαίων επιδράσεων τα παρατηρούμενα μεγέθη επίδρασης είναι τα τετράγωνα όπως φαίνονται στο Διάγραμμα 3.6 που εκτιμούν με τυχαίο σφάλμα (οριζόντιες γραμμές) τις διαφορετικές πραγματικές επιδράσεις οι οποίες δεν έχουν παρατηρηθεί (κάθετες διακεκομμένες οριζόντιες γραμμές) (Nikolakoroulou et al., 2014). Η κατανομή των τυχαίων επιδράσεων (distribution of underlying effects) είναι μια κοινή κατανομή των πραγματικών πειραματικών επιδράσεων (Nikolakoroulou et al., 2014). Η κατανομή αυτή είναι κανονική κατανομή με μέση τιμή ίση με το άθροισμα της συνολικής επίδρασης και διακύμανση τ^2 (Nikolakoroulou et al., 2014). Συνεπώς το διαμάντι για το μοντέλο τυχαίων επιδράσεων αποτελεί η εκτίμηση της μέσης τιμής της κατανομής των τυχαίων επιδράσεων (Nikolakoroulou et al., 2014).



Διάγραμμα 3.6 Διάγραμμα δάσους για την μετα-ανάλυση τυχαίων επιδράσεων

3.8 Σύγκριση των δύο μοντέλων μετα-ανάλυσης

Τα δύο μοντέλα μετα-ανάλυσης έχουν διαφορετική πηγή μεταβλητότητας. Κατά τη μετα-ανάλυση σταθερών επιδράσεων οι διαφορές στα παρατηρούμενα μεγέθη επίδρασης αποδίδονται στην ύπαρξη τυχαίου σφάλματος, ενώ κατά τη μετα-ανάλυση τυχαίων επιδράσεων οι διαφορές στα παρατηρούμενα μεγέθη επίδρασης αποδίδονται όχι μόνο σε δειγματικό σφάλμα, αλλά και σε πραγματικές διαφορές των πραγματικών επιδράσεων. Επομένως το μοντέλο τυχαίων επιδράσεων λαμβάνει υπόψη την ετερογένεια των μελετών, το οποίο αποτελεί την ουσιαστική διαφορά του από το μοντέλο σταθερών επιδράσεων. Η παραπάνω διαφοροποίηση έχει ως αποτέλεσμα την ύπαρξη διαφορετικών βαρών για τα δύο μοντέλα καθώς έχουν διαφορετικές πηγές μεταβλητότητας και η ανάθεση βαρών πραγματοποιείται με τη μέθοδο της αντίστροφης διασποράς (Nikolakoroulou et al., 2014). Καθώς ο σταθμισμένος μέσος όρος των παρατηρούμενων μεγεθών αποτελεί την συνολική επίδραση, το συνολικό αποτέλεσμα θα είναι διαφορετικό για τα δύο μοντέλα.

Η ετερογένεια είναι σταθερά και τα βάρη τυχαίων επιδράσεων ανατίθενται ως το αντίστροφο της ετερογένειας και της πληθυσμιακής διακύμανσης. Επομένως η ύπαρξη της ετερογένειας στα βάρη μειώνει τις σχετικές διαφορές τους. Συνεπώς τα βάρη να είναι πιο ισορροπημένα στο μοντέλο τυχαίων επιδράσεων (Nikolakoroulou et al., 2014). Όσο αυξάνεται η ετερογένεια τόσο τα βάρη για το μοντέλο τυχαίων επιδράσεων είναι διαφορετικά με τα βάρη για το μοντέλο σταθερών επιδράσεων, ενώ όσο μειώνεται η ετερογένεια τόσο τα βάρη γίνονται παρόμοια για τα δύο μοντέλα. Συνεπώς, η έκταση της ετερογένειας καθορίζει το ποσό που τα βάρη των μοντέλων θα διαφέρουν.

Ακόμα, οι μεγάλες μελέτες έχουν μεγάλη ακρίβεια και επομένως μικρή διασπορά (Nikolakoroulou et al., 2014). Κατά την ανάθεση βαρών με τη μέθοδο αντίστροφης διασποράς οι μεγάλες μελέτες στο μοντέλο σταθερών επιδράσεων έχουν μεγάλα βάρη άρα και μεγαλύτερη επίδραση απ' ό,τι οι μικρές μελέτες. Στο μοντέλο τυχαίων επιδράσεων οι μεγάλες μελέτες δεν έχουν τόσο μεγάλα βάρη γιατί τα βάρη ανατίθενται ως το αντίστροφο της διασποράς και της ετερογένειας. Λόγω λοιπόν της ύπαρξης ετερογένειας τα βάρη είναι πιο ισορροπημένα στο μοντέλο τυχαίων επιδράσεων με τη

μεγαλύτερη επίδραση να έχουν οι μικρές μελέτες (Nikolakopoulou et al., 2014).

Τα διαστήματα εμπιστοσύνης για την συνολική εκτίμηση στο μοντέλο τυχαίων επιδράσεων είναι συχνά μεγαλύτερα από το αντίστοιχο διάστημα του μοντέλου σταθερών επιδράσεων (Nikolakopoulou et al., 2014). Στις περισσότερες περιπτώσεις το μοντέλο τυχαίων επιδράσεων αποδίδει πιο συντηρητικά αποτελέσματα (Nikolakopoulou et al., 2014).

Επομένως, η σύγκριση των δύο μοντέλων μετα-αναλύσεων είναι μια μέθοδος για την εξέταση της ύπαρξης επιδράσεων των μικρών μελετών (Nikolakopoulou et al., 2014). Οι δύο προσεγγίσεις δίνουν παρόμοια εκτίμηση άλλα διαφορετικά διαστήματα εμπιστοσύνης (Nikolakopoulou et al., 2014). Τα μοντέλα σταθερών επιδράσεων και τυχαίων επιδράσεων δίνουν ίδια αποτελέσματα όταν η ετερογένεια είναι ίση με το μηδέν, όμως ακόμα και αν τα αποτελέσματα είναι ίδια οι συνολικές εκτιμήσεις πρέπει να ερμηνεύονται με διαφορετικό τρόπο.

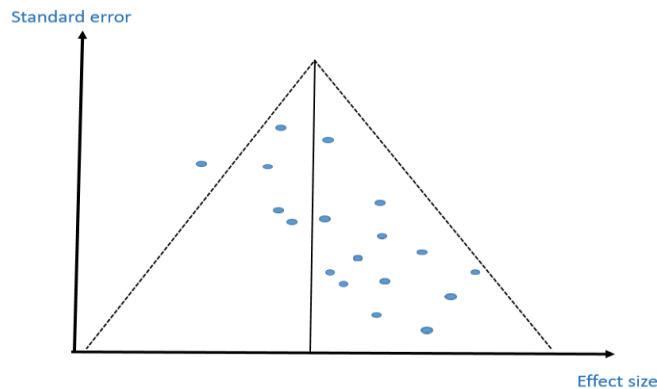
3.9 Συστηματικό σφάλμα δημοσίευσης (publication bias)

Με σκοπό τη διερεύνηση μιας ερευνητικής υπόθεσης διεξάγονται πολλές μελέτες όμως υπάρχει η τάση να δημοσιεύονται μερικές από τις μελέτες αυτές. Η τάση αυτή ονομάζεται συστηματικό σφάλμα δημοσίευσης δημιουργώντας παραπλάνηση και μειώνοντας την αξιοπιστία της μετα-ανάλυσης καθώς δεν έχουμε στη διάθεση μας όλες τις μελέτες αλλά ένα μέρος των μελετών. Μεγαλύτερη πιθανότητα για δημοσίευση έχουν οι μεγάλες μελέτες, μελέτες που δείχνουν σημαντικές επιδράσεις, οι μελέτες που έχουν χορηγούς και οι μικρές μελέτες που δείχνουν σημαντικές επιδράσεις (small-study effects) (Higgins & Green, 2008).

3.10 Διάγραμμα χωνί (funnel plot)

Το διάγραμμα χωνί (funnel plot) εισάχθηκε από τους Light & Pillemer (1984). Αποτελεί ένα διάγραμμα διασποράς των μελετών του μεγέθους επίδρασης (effect size) και ενός μέτρου ακρίβειας των μελετών, συνήθως του τυπικού σφάλματος (Mavridis & Salanti, 2014). Το διάγραμμα χωνί αποτελεί ένα γραφικό τρόπο εξέτασης της ύπαρξης μικρών μελετών οι οποίες δείχνουν σημαντικές επιδράσεις (small-study effects). Για την οπτι-

κή κατανόηση του παραθέτουμε το Διάγραμμα 3.7.

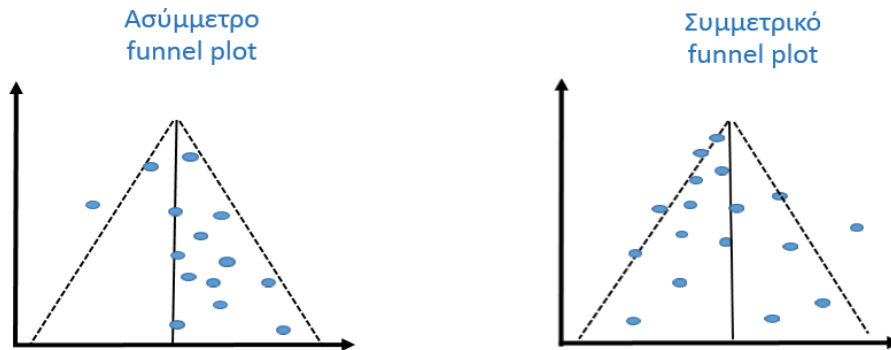


Διάγραμμα 3.7: Διάγραμμα χωνί (funnel plot)

Στο επάνω μέρος του funnel plot, η διασπορά είναι μικρή ενώ στο κάτω μέρος η διασπορά αυξάνει συνεχώς όπως γίνεται αντιληπτό και από το Διάγραμμα 3.7. Οι μεγάλες μελέτες έχουν μεγάλη ακρίβεια άρα μικρή διασπορά και επομένως βρίσκονται στο πάνω μέρος του funnel plot. Οι μικρές μελέτες έχουν μικρή ακρίβεια δηλαδή μεγάλη διασπορά και βρίσκονται στο κάτω μέρος του funnel plot.

Η ασυμμετρία είναι ένας τρόπος για να εξετάσουμε εάν υπάρχει η επίδραση των μικρών μελετών και είναι λάθος να ταυτίζεται με την εξέταση ύπαρξης συστηματικού σφάλματος δημοσίευσης καθώς μπορεί να υπάρχουν και άλλοι λόγοι δημιουργίας σημαντικών επιδράσεων από μικρές μελέτες (Mavridis & Salanti, 2014). Λόγοι ασυμμετρίας μπορεί να δημιουργούνται από τη μεθοδολογική ποιότητα των μελετών καθώς οι μικρές μελέτες συνήθως δεν έχουν καλή μεθοδολογική ποιότητα (Higgins & Thompson, 2002). Ακόμα, μπορεί να προκαλείται και από την ετερογένεια μεταξύ των μικρών και μεγάλων μελετών (Mavridis & Salanti, 2014). Η ασυμμετρία σε μελέτες που αντιστοιχούν σε μη σημαντικά αποτελέσματα μπορεί να υποδεικνύουν ότι μελέτες με αρνητικές εκβάσεις μπορεί να μην έχουν δημοσιευθεί (Mavridis & Salanti, 2014). Η οπτική παρουσίαση της ύπαρξης ασυμμετρίας σε ένα διάγραμμα χωνί γίνεται αντιληπτή με το Διάγραμμα 3.8. Στο αριστερό διάγραμμα έχουμε το funnel plot με τις δημοσιευμένες μελέτες και την ασυμμετρία που δημιουργούν οι αδημοσίευτες μελέτες, ενώ στο δεξί funnel plot

υπάρχει συμμετρία καθώς έχουμε τις δημοσιευμένες μελέτες μαζί με τις αδημοσίευτες.



Διάγραμμα 3.8: Η ασυμμετρία στο funnel plot

ΚΕΦΑΛΑΙΟ 4

ΕΚΤΙΜΗΤΕΣ

ΕΤΕΡΟΓΕΝΕΙΑΣ

4.1 Εκτίμηση της συνολικής επίδρασης μ του πληθυσμού

Θεωρούμε ότι έχουμε k μελέτες με τα παρατηρούμενα μεγέθη επιδράσεων y_1, y_2, \dots, y_k και τις πληθυσμιακές διακυμάνσεις $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$ (DerSimonian & Kacker, 2007). Στο μοντέλο τυχαίων επιδράσεων (random effect model) στόχος μας αποτελεί η εκτίμηση της συνολικής επίδρασης μ του πληθυσμού. Ο υπολογισμός της ετερογένειας τ^2 αποτελεί το ενδιάμεσο βήμα για την εκτίμηση της συνολικής επίδρασης μ (DerSimonian & Kacker, 2007). Εάν οι πληθυσμιακές διακυμάνσεις $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$ και η ετερογένεια τ^2 είναι γνωστές ποσότητες τότε η εκτίμηση της συνολικής επίδρασης υπολογίζεται από τον τύπο $\hat{\mu}_{RE} = \frac{\sum_{i=1}^k w_{i,RE} y_i}{\sum_{i=1}^k w_{i,RE}}$ με τα βάρη τυχαίων επιδράσεων $w_{i,RE} = \frac{1}{\sigma_i^2 + \tau^2}$ (DerSimonian & Kacker, 2007).

Στην πράξη όμως, δεν είναι γνωστή η ετερογένεια ούτε οι πληθυσμιακές διακυμάνσεις. Για το λόγο αυτό στους υπολογισμούς χρησιμοποιούμε τον εκτιμητή $\hat{\tau}^2$ για την ετερογένεια τ^2 και τους εκτιμητές $\hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_k^2$ για τις πληθυσμιακές διακυμάνσεις $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$, θεωρώντας ως $\hat{\sigma}_i^2$ τον εκτιμητή της πληθυσμιακής διακύμανσης για την i -οστή μελέτη (DerSimonian & Kacker, 2007). Θέλοντας να εξάγουμε τον εκτιμητή ετερογένειας η αβεβαιότητα των εκτιμητών $\hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_k^2$ μπορεί να οδηγήσει σε ανακρίβειες και για το λόγο αυτό συνηθίζεται οι εκτιμητές $\hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_k^2$ να θεωρούνται σταθερές και γνωστές τιμές (Sidik & Jonkman, 2007). Επομένως, η εκτίμηση της συνολικής επίδρασης στο μοντέλο τυχαίων επιδράσεων δίνεται ως $\hat{\mu}_{RE} = \frac{\sum_{i=1}^k w_{i,RE} y_i}{\sum_{i=1}^k w_{i,RE}}$ με τα βάρη

$$w_{i,RE} = \frac{1}{\sigma_1^2 + \hat{\tau}^2} \text{ (DerSimonian \& Kacker, 2007).}$$

4.2 Μέθοδοι εκτίμησης της ετερογένειας

Η εκτίμηση της ετερογένειας αποτελεί σημαντικό βήμα στη διαδικασία της μετα-ανάλυσης. Σημαντικό βήμα αποτελεί καθώς το μέγεθος της ετερογένειας των μελετών επηρεάζει την εκτίμηση της συνολικής επίδρασης και κατά επέκταση το αποτέλεσμα της μετα-ανάλυσης. Αφού στην πράξη δεν έχουμε επίγνωση για το μέγεθος της ετερογένειας για να συμπεράνουμε εάν τα αποτελέσματα είναι ακριβή, η εκτίμηση της ετερογένειας αποτελεί αναπόσπαστο κομμάτι του μοντέλου τυχαίων επιδράσεων. Στην προσπάθεια εκτίμησης της ετερογένειας σε μια μετα-ανάλυση τυχαίων επιδράσεων χρησιμοποιήθηκαν αρκετές μέθοδοι εκτίμησης ώστε να παρέχονται στη βιβλιογραφία διάφοροι εκτιμητές ετερογένειας.

Για την εκτίμηση της ετερογένειας υπάρχουν οι επαναληπτικές (iterative methods) και οι μη επαναληπτικές μέθοδοι (non-iterative methods) (Veroniki et al (submitted)). Ένας εκτιμητής ετερογένειας είναι μη επαναληπτικός όταν επιτυγχάνεται μετά από προκαθορισμένο αριθμό βημάτων. Από την άλλη μεριά, ένας εκτιμητής ετερογένειας είναι επαναληπτικός όταν επιτυγχάνεται από τη σύγκλιση μιας επαναληπτικής μεθόδου όπως του Newton-Raphson, της μεθόδου Scoring και της μεθόδου Fisher. Για τη σύγκλιση της επαναληπτικής μεθόδου χρησιμοποιούμε ένα κριτήριο τερματισμού όπως $|\hat{\tau}_{(j+1)}^2 - \hat{\tau}_{(j)}^2| < 0.00001$ στην j -οστή επανάληψη. Εάν ικανοποιείται το κριτήριο τερματισμού τότε η επαναληπτική μέθοδος συγκλίνει και έχουμε ως λύση τον εκτιμητή που προκύπτει από την επανάληψη στην οποία σταματάει η επαναληπτική μέθοδος. Εάν στην j -οστή επανάληψη για παράδειγμα το παραπάνω κριτήριο τερματισμού $|\hat{\tau}_{(j+1)}^2 - \hat{\tau}_{(j)}^2| < 0.00001$ ικανοποιηθεί τότε ο εκτιμητής ετερογένειας θα είναι ο $\hat{\tau}_{(j)}^2$. Οι επαναληπτικές μέθοδοι δεν εξάγουν πάντα έναν εκτιμητή ετερογένειας γιατί μπορεί η επαναληπτική μέθοδος να μην συγκλίνει.

Οι μέθοδοι εκτίμησης της ετερογένειας μπορεί να είναι είτε θετικοί (positive) είτε μη αρνητικοί (non-negative) (Veroniki et al (submitted)).

Ένας μη αρνητικός εκτιμητής μπορεί να είναι θετικός ή μηδέν ($\hat{\tau}^2 \geq 0$), επομένως εάν κατά την εκτίμηση εξάγουμε αρνητικές τιμές αυτές πρέπει να αποκόπτονται στο μηδέν χρησιμοποιώντας την συνάρτηση $\max(0, \hat{\tau}^2)$ (Veroniki et al (submitted)). Η ανάγκη αυτή να αποκόπτουμε τις αρνητικές τιμές στο μηδέν σ' έναν εκτιμητή οδηγεί σε μεροληπτικούς εκτιμητές (bias) (Veroniki et al (submitted)). Ένας θετικός εκτιμητής ($\hat{\tau}^2 > 0$) είναι πάντα μεγαλύτερος από το μηδέν και γι' αυτό δεν χρειάζεται αποκοπή (Veroniki et al (submitted)).

Από τα παραπάνω γίνεται αντιληπτό ότι ένας εκτιμητής ετερογένειας μπορεί να είναι είτε επαναληπτικός είτε μη επαναληπτικός ή μπορεί να είναι είτε θετικός είτε μη αρνητικός (Veroniki et al (submitted)). Οι γενικές μέθοδοι από τις οποίες μπορούν να δημιουργηθούν οι εκτιμητές ετερογένειας είναι η γενική μέθοδος των ροπών (General method of moments (GMM)), η μέθοδος μέγιστης πιθανοφάνειας (Maximum likelihood method (ML)), η μέθοδος Sidik-Jonkman (SJ), η μέθοδος Non-Parametric Bootstrap, η μέθοδος Empirical Bayes (EB), μέθοδος Bayes και η μέθοδος Malzahn, Böhning, and Holling (MBH).

4.3 Ιδιότητες ενός καλού εκτιμητή ετερογένειας

Ένας εκτιμητής λέμε ότι έχει καλές ιδιότητες εάν ικανοποιεί τα ακόλουθα κριτήρια:

Αμεροληψία (bias)

Η μεροληψία του εκτιμητή ετερογένειας είναι η διαφορά μεταξύ της αναμενόμενης τιμής του εκτιμητή και της πραγματικής του τιμής και ορίζεται ως

$$Bias(\hat{\tau}^2) = E(\hat{\tau}^2) - \tau^2 = E(\hat{\tau}^2 - \tau^2)$$

Θετική ή αρνητική μεροληψία των εκτιμητών οδηγεί σε σφάλμα στην εκτίμηση της πραγματικής ετερογένειας είτε αυξημένο είτε μειωμένο. Επομένως για την καλύτερη εκτίμηση απαιτούμε ο εκτιμητής ετερογένειας να είναι αμερόληπτος, δηλαδή $E(\hat{\tau}^2) = \tau^2$.

Αποδοτικότητα (efficiency)

Σε μια καλή εκτίμηση ο εκτιμητής πρέπει να είναι αποδοτικός, δηλαδή να

παραμένει ανεπηρέαστος από τις πιθανές πληθυσμιακές διακυμάνσεις.

Ένα μέτρο αποδοτικότητας αποτελεί το μέσο τετραγωνικό σφάλμα (Mean Square Error (MSE)) το οποίο ορίζεται ως

$$MSE(\hat{\tau}^2) = E(\hat{\tau}^2 - \tau^2) = Var(\hat{\tau}^2) + (Bias(\hat{\tau}^2))^2$$

Όσο πιο μικρό είναι το μέσο τετραγωνικό σφάλμα τόσο πιο αποδοτικός είναι ο εκτιμητής. Έχοντας δύο εκτιμητές ετερογένειας $\hat{\tau}_1^2$ και $\hat{\tau}_2^2$, ο $\hat{\tau}_1^2$ είναι πιο αποδοτικός από τον $\hat{\tau}_2^2$ όταν $MSE(\hat{\tau}_1^2) < MSE(\hat{\tau}_2^2)$.

Υψηλή ισχύ (power)

Στο τεστ ελέγχου της μηδενικής υπόθεσης $H_0 : \tau^2 = \tau_o^2$ με εναλλακτική την $H_a : \tau^2 \neq \tau_o^2$ η ισχύς αποτελεί $P(\text{απορρίπτω } H_0 / H_a \text{ αληθής})$. Επομένως μια καλή εκτίμηση παρέχει στο τεστ υψηλή ισχύ.

Σφάλμα τύπου I (type error I) κοντά στα θεωρητικά επίπεδα (nominal levels)

Στο τεστ ελέγχου της μηδενικής υπόθεσης $H_0 : \tau^2 = \tau_o^2$ το σφάλμα τύπου I αποτελεί $P(\text{απορρίπτω } H_0 / H_0 \text{ αληθής})$. Ένας καλός εκτιμητής έχει την τάση να βρίσκεται κοντά στα θεωρητικά επίπεδα, δηλαδή μια καλή εκτίμηση πρέπει να έχει σφάλμα τύπου I κοντά στο επίπεδο σημαντικότητας α .

Πιθανότητες κάλυψης για τη συνολική επίδραση και της ετερογένειας κοντά στα θεωρητικά επίπεδα (nominal levels)

Μια καλή εκτίμηση πρέπει να έχει πιθανότητες κάλυψης κοντά στα θεωρητικά επίπεδα, έχοντας δηλαδή επίπεδο σημαντικότητας α , η πιθανότητα κάλυψης του εκτιμητή πρέπει να βρίσκεται κοντά στην τιμή $p = 1 - \alpha$.

4.4 Γενική μέθοδος εκτίμησης των ροπών (General method of moments (GMM)) για την ετερογένεια

Μια μέθοδος εκτίμησης αποτελεί η γενική μέθοδος εκτίμησης των ροπών από την οποία προκύπτουν οκτώ εκτιμητές ετερογένειας. Κατά τη γενική

μέθοδο των ροπών εξισώνεται η αναμενόμενη τιμή με την παρατηρούμενη.

Γενικά έχοντας k μελέτες ορίζουμε την ποσότητα $y_w = \frac{\sum_{i=1}^k a_i y_i}{\sum_{i=1}^k a_i}$, με τις ποσότητες a_1, a_2, \dots, a_k να αποτελούν θετικές τιμές που αντιπροσωπεύουν τα βάρη των k μελετών (DerSimonian & Kacker, 2007).

Από Kacker (2004) ισχύει ότι

$$E(y_w) = E\left[\frac{\sum_{i=1}^k a_i y_i}{\sum_{i=1}^k a_i}\right] = \frac{\sum_{i=1}^k a_i E(y_i)}{\sum_{i=1}^k a_i} = E(y_i) \frac{\sum_{i=1}^k a_i}{\sum_{i=1}^k a_i} = E(y_i)$$

$$E(y_i - y_w) = 0$$

Επιπλέον,

$$E(y_i - y_w)^2 - (E(y_i - y_w))^2 = Var(y_i - y_w)$$

$$\begin{aligned} E(y_i - y_w)^2 &= Var(y_i - y_w) \\ &= Var(y_i) + Var(y_w) - 2Cov(y_i, y_w) \\ &= Var(y_i) + \frac{\sum_{i=1}^k a_i^2 Var(y_i)}{\sum_{i=1}^k a_i} - \frac{2a_i Var(y_i)}{\sum_{i=1}^k a_i} \\ E(y_i - y_w)^2 &= Var(y_i) + \frac{\sum_{i=1}^k a_i^2 Var(y_i)}{\sum_{i=1}^k a_i} - \frac{2a_i Var(y_i)}{\sum_{i=1}^k a_i} \end{aligned}$$

Η αναμενόμενη τιμή της έκφρασης $\sum_{i=1}^k a_i (y_i - y_w)^2$ είναι

$$\begin{aligned} E\left[\sum_{i=1}^k a_i (y_i - y_w)^2\right] &= \sum_{i=1}^k a_i E(y_i - y_w)^2 \\ &= \sum_{i=1}^k a_i Var(y_i) + \frac{\sum_{i=1}^k a_i^2 Var(y_i)}{\sum_{i=1}^k a_i} - \frac{2 \sum_{i=1}^k a_i^2 Var(y_i)}{\sum_{i=1}^k a_i} \\ &= \sum_{i=1}^k a_i Var(y_i) - \frac{\sum_{i=1}^k a_i^2 Var(y_i)}{\sum_{i=1}^k a_i} \end{aligned}$$

Και επειδή $Var(y_i) = \tau^2 + \sigma_i^2$ η αναμενόμενη τιμή θα είναι

$$E\left[\sum_{i=1}^k a_i (y_i - y_w)^2\right] = \sum_{i=1}^k a_i (\tau^2 + \sigma_i^2) - \frac{\sum_{i=1}^k a_i^2 (\tau^2 + \sigma_i^2)}{\sum_{i=1}^k a_i}$$

το οποίο μπορεί να γραφεί ως

$$E\left[\sum_{i=1}^k a_i (y_i - y_w)^2\right] = \tau^2 \left[\sum_{i=1}^k a_i - \frac{\sum_{i=1}^k a_i^2}{\sum_{i=1}^k a_i} \right] + \left[\sum_{i=1}^k a_i \sigma_i^2 - \frac{\sum_{i=1}^k a_i^2 \sigma_i^2}{\sum_{i=1}^k a_i} \right]$$

Εφαρμόζοντας τη γενική μέθοδο των ροπών για την έκφραση $\sum_{i=1}^k a_i (y_i - y_w)^2$, εξισώνουμε την έκφραση με την αναμενόμενη τιμή της και άρα προκύπτει ότι

$$\begin{aligned} E\left[\sum_{i=1}^k a_i (y_i - y_w)^2\right] &= \\ &= \sum_{i=1}^k a_i (y_i - y_w)^2 \\ &= \tau^2 \left[\sum_{i=1}^k a_i - \frac{\sum_{i=1}^k a_i^2}{\sum_{i=1}^k a_i} \right] + \left[\sum_{i=1}^k a_i \sigma_i^2 - \frac{\sum_{i=1}^k a_i^2 \sigma_i^2}{\sum_{i=1}^k a_i} \right] \\ &= \sum_{i=1}^k a_i (y_i - y_w)^2 \end{aligned}$$

Λύνοντας ως προς τ^2 και αντικαθιστώντας τα $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$ με $\hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_k^2$ και το τ^2 με $\hat{\tau}^2$ έχουμε τον εκτιμητή της γενικής μεθόδου των ροπών, ο οποίος συμβολίζεται με GMM

$$\hat{\tau}_{GMM}^2 = \frac{\sum_{i=1}^k a_i (y_i - y_w)^2 - \left[\sum_{i=1}^k a_i \hat{\sigma}_i^2 - \frac{\sum_{i=1}^k a_i^2 \hat{\sigma}_i^2}{\sum_{i=1}^k a_i} \right]}{\sum_{i=1}^k a_i - \frac{\sum_{i=1}^k a_i^2}{\sum_{i=1}^k a_i}}$$

(DerSimonian & Kacker, 2007).

Επειδή η ετερογένεια είναι μη-αρνητική $\tau^2 \geq 0$ κατά αποδοτικότητα ο εκτιμητής της γενικής μεθόδου των ροπών είναι μη-αρνητικός. Ακόμα υποθέτοντας ότι οι πληθυσμιακές διακυμάνσεις $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$ είναι γνωστές και κάτω από τις συνθήκες του μοντέλου τυχαίων επιδράσεων ο εκτιμητής GMM είναι αμερόληπτος (Veroniki et al (submitted)). Από τον εκτιμητή της γενικής μεθόδου των ροπών $\hat{\tau}_{GMM}^2$ για τις διάφορες θετικές τιμές $\alpha_1, \alpha_2, \dots, \alpha_k$ μπορούμε να εξάγουμε διαφορετικούς εκτιμητές ετερογένειας.

4.4.1 Μέθοδος General Hedges-Olkin (GHO)

Ο εκτιμητής ετερογένειας General Hedges-Olkin είναι γνωστός και ως εκτιμητής Cochran ANOVA (CA) ή ως τύπος εκτίμησης της διακύμανσης (variance component type estimator (VC)) και εντάχθηκε από τον Cochran (1954). Προκύπτει από τη γενική μέθοδο των ροπών για $a_i = \frac{1}{k}$, $i = 1, \dots, k$

και $y_w = \frac{\sum_{i=1}^k \frac{1}{k} y_i}{\sum_{i=1}^k \frac{1}{k}} = \bar{y}$ (DerSimonian & Kacker, 2007).

Η έκφραση $\sum_{i=1}^k a_i (y_i - y_w)^2$ για $a_i = \frac{1}{k}$ και $y_w = \bar{y}$ γίνεται $\frac{1}{k} \sum_{i=1}^k (y_i - \bar{y})^2 = S_y^2$ η οποία αποτελεί τον εκτιμητή της πληθυσμιακής διακύμανσης. Θέτουμε την αναμενόμενη τιμή του εκτιμητή της πληθυσμιακής διακύμανσης ίση με την παρατηρούμενη τιμή του.

$$E\left[\sum_{i=1}^k \frac{1}{k} (y_i - \bar{y})^2\right] = \sum_{i=1}^k \frac{1}{k} (y_i - \bar{y})^2$$

$$\tau^2 \left[\sum_{i=1}^k \frac{1}{k} - \frac{\sum_{i=1}^k \frac{1}{k^2}}{\sum_{i=1}^k \frac{1}{k}} \right] + \left[\frac{\sum_{i=1}^k \frac{1}{k} \sigma_i^2 - \sum_{i=1}^k \frac{1}{k^2} \sigma_i^2}{\sum_{i=1}^k \frac{1}{k}} \right] = \sum_{i=1}^k \frac{1}{k} (y_i - \bar{y})^2$$

Λύνοντας ως προς τ^2 και αντικαθιστώντας το τ^2 με $\hat{\tau}^2$ και τα σ_i^2 με τα $\hat{\sigma}_i^2$ εξάγουμε τον εκτιμητή General Hedges-Olkin (GHO)

$$\hat{\tau}_{GHO}^2 = \frac{1}{k-1} \sum_{i=1}^k (y_i - \bar{y})^2 - \frac{1}{k} \sum_{i=1}^k \hat{\sigma}_i^2$$

(DerSimonian & Kacker, 2007). Ο εκτιμητής General Hedges-Olkin είναι εύκολο να υπολογιστεί αλλά δεν χρησιμοποιείται ευρέως. Είναι ένας μη επαναληπτικός καθώς δεν απαιτείται κάποια επαναληπτική μέθοδο για να προκύψει η λύση και μη αρνητικός (Veroniki & Salanti, 2014).

4.4.2 Μέθοδος DerSimonian and Laird (DL)

Αποτελεί την πιο δημοφιλή μέθοδο εκτίμησης της ετερογένειας καθώς χρησιμοποιείται πιο συχνά από τις υπόλοιπες μεθόδους (DerSimonian & Laird, 1986). Σύμφωνα με τους DerSimonian & Kacker, (2007) ο εκτιμητής DerSimonian and Laird (1986) προκύπτει από τη γενική μέθοδο

των ρομών για $a_i = w_{i,FE} = \frac{1}{\sigma_i^2}$, $i = 1, \dots, k$ και $y_w = \hat{\mu}_{FE} = \frac{\sum_{i=1}^k w_{i,FE} y_i}{\sum_{i=1}^k w_{i,FE}}$. Η έκφραση $\sum_{i=1}^k a_i (y_i - y_w)^2$ για $a_i = w_{i,FE}$ και $y_w = \hat{\mu}_{FE}$ γίνεται $\sum_{i=1}^k w_{i,FE} (y_i - \hat{\mu}_{FE})^2 = Q$ γνωστή ως το Q -στατιστικό του Cochran.

Η αναμενόμενη τιμή του Q είναι

$$E(Q) = E\left(\sum_{i=1}^k w_{i,FE} (y_i - \hat{\mu}_{FE})^2\right) = \tau^2 \left[\sum_{i=1}^k w_{i,FE} - \frac{\sum_{i=1}^k w_{i,FE}^2}{\sum_{i=1}^k w_{i,FE}} \right] + k - 1$$

Εφαρμόζοντας τη γενική μέθοδο των ρομών για το στατιστικό του Cochran Q έχουμε

$$E(Q) = Q$$

$$Q = \tau^2 \left[\sum_{i=1}^k w_{i,FE} - \frac{\sum_{i=1}^k w_{i,FE}^2}{\sum_{i=1}^k w_{i,FE}} \right] + k - 1$$

Λύνοντας ως προς τ^2 και αντικαθιστώντας το τ^2 με $\hat{\tau}^2$ και τα $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$ με τις εκτιμώμενες $\hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_k^2$ προκύπτει ο εκτιμητής DerSimonian and Laird (DL)

$$\hat{\tau}_{DL}^2 = \frac{Q - (k - 1)}{\sum_{i=1}^k w_{i,FE} - \frac{\sum_{i=1}^k w_{i,FE}^2}{\sum_{i=1}^k w_{i,FE}}}$$

με $w_{i,FE} = \frac{1}{\sigma_i^2}$, $i = 1, \dots, k$.

Ο εκτιμητής DerSimonian and Laird (DL) είναι μη επαναληπτικός και μη αρνητικός (Veroniki et al (submitted)). Αποτελεί έναν αμερόληπτο εκτιμητή όταν οι πληθυσμιακές διακυμάνσεις $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$ είναι γνωστές και όχι εκτιμώμενες (Veroniki et al (submitted)).

4.4.3 Μέθοδος Paule-Mandel (PM)

Η μέθοδος από τους Paule Mandel (1982) προκύπτει από τη γενική μέθοδο

των ρομών για $a_i = w_{i,RE} = \frac{1}{\sigma_i^2 + \tau^2}$ και $y_w = \hat{\mu}_{RE} = \frac{\sum_{i=1}^k w_{i,RE} y_i}{\sum_{i=1}^k w_{i,RE}}$.

Η έκφραση $\sum_{i=1}^k a_i (y_i - y_w)^2$ για $a_i = w_{i,RE}$ και $y_w = \hat{\mu}_{RE}$ γίνεται $\sum_{i=1}^k w_{i,RE} (y_i - \hat{\mu}_{RE})^2 = Q_{gen}$ γνωστή ως Generalised Q -στατιστικό.

Ισχύει ότι το Generalised Q -στατιστικό ακολουθεί την χ^2 κατανομή με $k - 1$ βαθμούς ελευθερίας $Q_{gen} = \sum_{i=1}^k w_{i,RE} (y_i - \hat{\mu}_{RE})^2 \sim \chi_{k-1}^2$ και επειδή η αναμενόμενη τιμή της χ^2 κατανομής είναι $E(\chi_{k-1}^2) = k - 1$, η αναμενόμενη τιμή του Generalised Q -στατιστικού είναι:

$$E(Q_{gen}) = k - 1$$

Εφαρμόζοντας τη γενική μέθοδο των ροπών έχουμε

$$E(Q_{gen}) = Q_{gen}$$

$$\sum_{i=1}^k w_{i,RE} (y_i - \hat{\mu}_{RE})^2 = k - 1$$

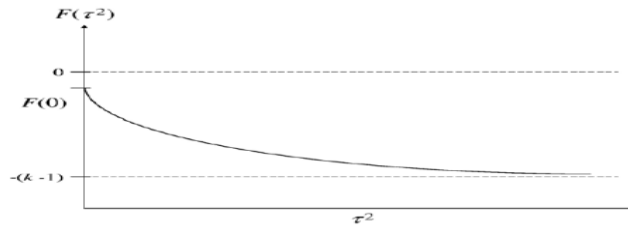
$$\sum_{i=1}^k w_{i,RE} (y_i - \hat{\mu}_{RE})^2 - (k - 1) = 0$$

Αντικαθιστώντας τα $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$ με τα $\hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_k^2$ παίρνουμε την εξίσωση

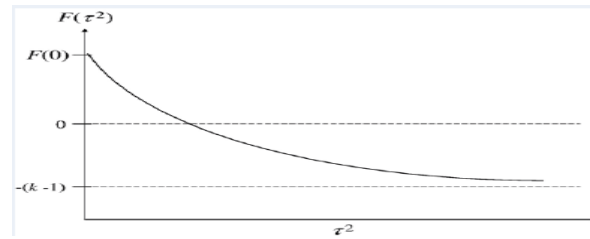
$$F(\tau^2) = \sum_{i=1}^k w_{i,RE} (y_i - \hat{\mu}_{RE})^2 - (k - 1) = 0 \quad (4.4.3.1)$$

όπου $w_{i,RE} = \frac{1}{\hat{\sigma}_i^2 + \tau^2}$ και $\hat{\mu}_{RE} = \frac{\sum_{i=1}^k w_{i,RE} y_i}{\sum_{i=1}^k w_{i,RE}}$ στην εξίσωση (4.4.3.1). Η λύση της εξίσωσης $F(\tau^2) = 0$ αποτελεί τον εκτιμητή Paule-Mandel (PM) ο οποίος συμβολίζεται με $\hat{\tau}_{PM}^2$ (DerSimonian & Kacker, 2007). Σύμφωνα με τους DerSimonian & Kacker (2007) η λύση της εξίσωσης είναι μοναδική και μπορεί να καθοριστεί από μια επαναληπτική μέθοδο ξεκινώντας με αρχικό βήμα $\tau^2 = 0$ ή ελαφρώς μεγαλύτερο από το μηδέν. Ακόμα, η συνάρτησή $F(\tau^2)$ είναι αυστηρά φθίνουσα και επομένως το μέγιστό της είναι μοναδικό. Το σχήμα της παρουσιάζεται από το Διάγραμμα 4.1 όταν $F(0) < 0$ και το Διάγραμμα 4.2 όταν $F(0) > 0$. Για $\tau^2 \rightarrow \infty$ έχουμε ότι $F(\tau^2) = -(k - 1)$, ενώ για $\tau^2 = 0$ έχουμε το μέγιστο της συνάρτησης, $F(0)$ όπως βλέπουμε και από τα διαγράμματα. Εάν το $F(0) < 0$, τότε συμπεραίνουμε ότι η συνάρτηση $F(\tau^2)$ δεν έχει θετική τιμή και συνεπώς ο εκτιμητής $\hat{\tau}_{PM}^2$ είναι μηδέν. Ακόμα όταν $F(0) = 0$ καταλήγουμε στο ίδιο συμπέρασμα ότι δηλαδή ο εκτιμητής $\hat{\tau}_{PM}^2$ είναι μηδέν. Από την άλλη μεριά, όταν $F(0) > 0$ κάνουμε διόρθωση ώστε να επιτύχουμε τον μηδενισμό της συνάρτησης $F(\tau^2) = 0$. Επομένως από τον ορισμό της παραπάνω συνάρτησης καταλαβαίνουμε ότι ο εκτιμητής Paule-Mandel είναι

μη αρνητικός και παράλληλα επαναληπτικός.



Διάγραμμα 4.1 Σχήμα της συνάρτησης $F(\tau^2)$ όταν $F(0) < 0^1$



Διάγραμμα 4.2 Σχήμα της συνάρτησης $F(\tau^2)$ όταν $F(0) > 0^2$

Σύμφωνα με τον Kacker (2004) ο αλγόριθμος για τον υπολογισμό του εκτιμητή Paule-Mandel, ο οποίος είναι απλός ως προς την υλοποίησή του σε κάποιο πρόγραμμα λογισμικού, είναι ο ακόλουθος:

1. Έχοντας το αρχικό βήμα τ^2 (αρχικό) υπολογίζουμε τα βάρη $w_{i,RE} = \frac{1}{\hat{\sigma}_i^2 + \tau^2(\text{αρχικό})}$ και τη συνάρτηση $F(\tau^2(\text{αρχικό}))$.

2. Εάν $F(\tau^2(\text{αρχικό})) < 0$, θέτουμε $\hat{\tau}_{PM}^2 = \tau^2(\text{αρχικό})$.
Εάν $F(\tau^2(\text{προηγούμενο})) = 0$, τότε $\hat{\tau}_{PM}^2 = \tau^2(\text{προηγούμενο})$.
Εάν $F(\tau^2(\text{προηγούμενο})) > 0$, τότε προχωράμε στη διόρθωση

$$\Delta\tau^2 = \frac{\sum_{i=1}^k w_{i,RE} (y_i - \hat{\mu}_{RE})^2 - (k-1)}{\sum_{i=1}^k w_{i,RE}^2 (y_i - \hat{\mu}_{RE})^2}$$

3. Η τιμή στην επόμενη επανάληψη είναι $\tau^2(\text{επόμενο}) = \tau^2(\text{προηγούμενο}) + \Delta\tau^2$.
4. Επαναλαμβάνουμε τα βήματα 2 και 3 μέχρι $F(\tau^2(\text{προηγούμενο})) = 0$. Η τελική τιμή του τ^2 θα είναι το $\hat{\tau}_{PM}^2$.

¹ Πηγή: (DerSimonian & Kacker, 2007)

² Πηγή: (DerSimonian & Kacker, 2007)

Οι Paule και Mandel υποστήριζαν ότι ως αρχικό βήμα είναι καλύτερη η επιλογή μίας τιμής ελαφρώς πιο πάνω από το μηδέν (Kacker, 2004), ενώ ο Kacker πρότεινε ως αρχική τιμή τον εκτιμητή GHO εάν αυτός είναι θετικός, καθώς έτσι επιτυγχάνεται πιο γρήγορη σύγκλιση (Kacker, 2004).

4.4.4 Μέθοδος δύο βημάτων DerSimonian and Laird (two-step DerSimonian and Laird method (DL2))

Οι DerSimonian & Kacker (2007) πρότειναν τη μέθοδο δύο βημάτων με την οποία καταλήγουμε στη δημιουργία ενός νέου εκτιμητή ετερογένειας έπειτα από δύο βήματα. Κατά το πρώτο βήμα εκτιμούμε τον DerSimonian and Laird εκτιμητή. Έχοντας εκτιμήσει με την γενική μέθοδο των ροπών το $\hat{\tau}_{DL}^2$ κατά επέκταση θα έχουμε τα βάρη του μοντέλου τυχαίων επιδράσεων για τον εκτιμητή DerSimonian and Laird $w_{i,RE}(\hat{\tau}_{DL}^2) = \frac{1}{\hat{\sigma}_i^2 + \hat{\tau}_{DL}^2}$, $i = 1, \dots, k$ και την εκτίμηση της συνολικής επίδρασης-διαμάντι

$\hat{\mu}_{RE}(\hat{\tau}_{DL}^2) = \frac{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{DL}^2) y_i}{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{DL}^2)}$. Στο δεύτερο βήμα εφαρμόζουμε τη γενική

μέθοδο των ροπών με σκοπό να βρούμε έναν νέο εκτιμητή για $\alpha_i = w_{i,RE}(\hat{\tau}_{DL}^2) = \frac{1}{\hat{\sigma}_i^2 + \hat{\tau}_{DL}^2}$, $i = 1, \dots, k$, $y_w = \hat{\mu}_{RE}(\hat{\tau}_{DL}^2) = \frac{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{DL}^2) y_i}{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{DL}^2)}$

και παίρνουμε τον εκτιμητή δύο βημάτων DerSimonian and Laird (DL2)

$$\hat{\tau}_{DL2}^2 = \frac{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{DL}^2) (y_i - \hat{\mu}_{RE}(\hat{\tau}_{DL}^2))^2 - \left[\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{DL}^2) \hat{\sigma}_i^2 - \frac{\sum_{i=1}^k w_{i,RE}^2(\hat{\tau}_{DL}^2) \hat{\sigma}_i^2}{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{DL}^2)} \right]}{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{DL}^2) - \frac{\sum_{i=1}^k w_{i,RE}^2(\hat{\tau}_{DL}^2)}{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{DL}^2)}}$$

Ο εκτιμητής δυο βημάτων DerSimonian and Laird δεν είναι επαναληπτικός και είναι μη-αρνητικός.

4.4.5 Μέθοδος δύο βημάτων General Hedges-Olkin (two-step General Hedges-Olkin method (GHO2))

Οι DerSimonian & Kacker (2007) πέρα από τον εκτιμητή δύο βημάτων DerSimonian and Laird πρότειναν και τη δημιουργία του εκτιμητή δύο βημάτων για τον General Hedges-Olkin. Στο πρώτο βήμα εξάγουμε τον General Hedges-Olkin εκτιμητή. Έχοντας εκτιμήσει με τη γενική μέθοδο

των ροπών το $\hat{\tau}_{GHO}^2$ κατά επέκταση θα έχουμε τα βάρη του μοντέλου τυχαίων επιδράσεων για τον εκτιμητή General Hedges-Olkin $w_{i,RE}(\hat{\tau}_{GHO}^2) = \frac{1}{\hat{\sigma}_i^2 + \hat{\tau}_{GHO}^2}$, $i = 1, \dots, k$ και την εκτίμηση της συνολικής επίδρασης-διαμάντι $\hat{\mu}_{RE}(\hat{\tau}_{GHO}^2) = \frac{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{GHO}^2) y_i}{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{GHO}^2)}$. Στο δεύτερο βήμα εφαρμόζουμε τη γενική μέθοδο των ροπών με σκοπό να βρούμε έναν νέο εκτιμητή για $a_i = w_{i,RE}(\hat{\tau}_{GHO}^2) = \frac{1}{\hat{\sigma}_i^2 + \hat{\tau}_{GHO}^2}$, $i = 1, \dots, k$, $\hat{\mu}_{RE}(\hat{\tau}_{GHO}^2) = \frac{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{GHO}^2) y_i}{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{GHO}^2)}$ παίρνουμε τον εκτιμητή δύο βημάτων General Hedges-Olkin (GHO2).

$$\hat{\tau}_{GHO2}^2 = \frac{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{GHO}^2) (y_i - \hat{\mu}_{RE}(\hat{\tau}_{GHO}^2))^2 - \left[\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{GHO}^2) \hat{\sigma}_i^2 - \frac{\sum_{i=1}^k w_{i,RE}^2(\hat{\tau}_{GHO}^2) \hat{\sigma}_i^2}{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{GHO}^2)} \right]}{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{GHO}^2) - \frac{\sum_{i=1}^k w_{i,RE}^2(\hat{\tau}_{GHO}^2)}{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{GHO}^2)}}$$

Από τον ορισμό του εκτιμητή General Hedges-Olkin συμπεραίνουμε ότι δεν είναι επαναληπτικός και είναι μη-αρνητικός.

4.4.6 Θετικός εκτιμητής DerSimonian and Laird (Positive DerSimonian and Laird (DLp))

Οι Kontopantelis et al. (2013) πρότειναν το θετικό εκτιμητή DerSimonian and Laird ο οποίος προκύπτει από τον εκτιμητή DerSimonian and Laird και άρα είναι ένας εκτιμητής που προέρχεται από τη γενική μέθοδο των ροπών. Όταν ο εκτιμητής DerSimonian and Laird, ο οποίος είναι μη αρνητικός, έχει θετική τιμή τότε ο θετικός εκτιμητής DerSimonian and Laird παίρνει την τιμή του. Ενώ στην αντίθετη περίπτωση παίρνει την τιμή 0.01, δηλαδή

$$\hat{\tau}_{DLp}^2 = \begin{cases} \hat{\tau}_{DL}^2, & \text{αν } \hat{\tau}_{DL}^2 > 0 \\ 0.01, & \text{αν } \hat{\tau}_{DL}^2 \leq 0 \end{cases}$$

Ο θετικός εκτιμητής DerSimonian and Laird είναι μη επαναληπτικός και θετικός όπως φαίνεται από τον ορισμό του.

4.4.7 Μέθοδος Hartung-Makambi (HM)

Οι Hartung & Makambi (2003) δημιούργησαν μία μέθοδο η οποία αποτελεί μια παραλλαγή της μεθόδου DerSimonian and Laird (1986). Οι Hartung-Makambi βασίστηκαν το γεγονός ότι ένας μη αρνητικός εκτιμητής ετερογένειας μπορεί να δώσει μηδενικές τιμές ενώ παράλληλα η μεταβλητότητα μεταξύ των μελετών δεν είναι φυσιολογικό να είναι μηδενική. Ως προς αποφυγή αυτής της αρνητικότητας για μια καλύτερη εκτίμηση της ετερογένειας η μέθοδος Hartung-Makambi δημιουργεί ένα θετικό εκτιμητή από μια παραλλαγή πάνω στη μέθοδο DerSimonian and Laird (1986). Θεωρούμε τον εκτιμητή DerSimonian and Laird (1986)

$$\hat{t}_{DL}^2 = \frac{Q - (k - 1)}{\sum_{i=1}^k w_{i,FE} - \frac{\sum_{i=1}^k w_{i,FE}^2}{\sum_{i=1}^k w_{i,FE}}}$$

όπου Q το στατιστικό του Cochran $Q = \sum_{i=1}^k w_{i,FE} (y_i - \hat{\mu}_{FE})^2$ το οποίο αποτελεί την τετραγωνική μορφή των y_i . Σύμφωνα με τη μέθοδο Hartung-Makambi παίρνουμε το μέλος με την τετραγωνική μορφή από τον εκτιμητή DerSimonian and Laird (1986), δηλαδή την ποσότητα

$$\frac{\sum_{i=1}^k w_{i,FE} (y_i - \hat{\mu}_{FE})^2}{\sum_{i=1}^k w_{i,FE} - \frac{\sum_{i=1}^k w_{i,FE}^2}{\sum_{i=1}^k w_{i,FE}}} = \frac{Q}{\sum_{i=1}^k w_{i,FE} - \frac{\sum_{i=1}^k w_{i,FE}^2}{\sum_{i=1}^k w_{i,FE}}}$$

και την πολλαπλασιάζουμε με την ποσότητα $\frac{Q}{(2(k-1)+Q)}$ με σκοπό την απαίτηση της θετικότητας, άρα προκύπτει η ποσότητα

$$\frac{Q}{(2(k-1)+Q)} \cdot \frac{Q}{\sum_{i=1}^k w_{i,FE} - \frac{\sum_{i=1}^k w_{i,FE}^2}{\sum_{i=1}^k w_{i,FE}}}$$

που αποτελεί το θετικό εκτιμητή ετερογένειας Hartung-Makambi (HM)

$$\hat{t}_{HM}^2 = \frac{Q^2}{(2(k-1)+Q) \left(\sum_{i=1}^k w_{i,FE} - \frac{\sum_{i=1}^k w_{i,FE}^2}{\sum_{i=1}^k w_{i,FE}} \right)}$$

Από την παραπάνω διαδικασία συμπεραίνουμε ότι ο εκτιμητής Hartung-Makambi που προκύπτει από τη γενική μέθοδο των ροπών βάση του

στατιστικού του Cochran Q είναι θετικός εκτιμητής και αποτελεί ένα μη επαναληπτικό εκτιμητή.

4.4.8 Μέθοδος Hunter-Schmidt (HS)

Οι Hunter and Schmidt (2004) προτείνουν τον εκτιμητή Hunter-Schmidt

(HS) ο οποίος έχει τη μορφή $\hat{t}_{HM}^2 = \frac{\sum_{i=1}^k w_i \left(y_i - \frac{\sum_{i=1}^k w_i y_i}{\sum_{i=1}^k w_i} \right)^2}{\sum_{i=1}^k w_i} - \frac{\sum_{i=1}^k w_i \frac{1}{\sigma_i^2}}{\sum_{i=1}^k w_i}$ με τα

βάρη w_i να παίρνουν ως αρχικές τιμές τα βάρη του μοντέλου σταθερών επιδράσεων $w_{i,FE}$ ή να σταθμίζονται από το μέγεθος της κάθε μελέτης, δηλαδή $w_i = n_i$ όπου n_i αποτελεί το μέγεθος της i μελέτης. Θεωρώντας ως μέθοδο στάθμισης τα βάρη σταθερών επιδράσεων ο εκτιμητής Hunter-Schmidt (HS) παίρνει τη μορφή

$$\hat{t}_{HM}^2 = \frac{Q - k}{\sum_{i=1}^k w_{i,FE}}$$

Από τον ορισμό του παρατηρούμε ότι ο εκτιμητής Hunter-Schmidt είναι μη επαναληπτικός και μη αρνητικός.

4.5 Μέθοδος Μέγιστης Πιθανοφάνειας (Maximum likelihood method (ML))

Ένας τρόπος εύρεσης εκτιμητών πέρα από τη γενική μέθοδο των ροπών αποτελεί η αρχή της μέγιστης πιθανοφάνειας. Ο εκτιμητής ετερογένειας με την αρχή της μέγιστης πιθανοφάνειας είναι ουσιαστικά ο εκτιμητής μέγιστης πιθανοφάνειας. Για την εξαγωγή των εκτιμητών ετερογένειας στην μετα-ανάλυση εφαρμόζεται η μέθοδος μέγιστης πιθανοφάνειας και η περιορισμένη μέθοδος μέγιστης πιθανοφάνειας.

4.5.1 Εκτιμητής μέγιστης πιθανοφάνειας (Maximum likelihood estimator (ML))

Σύμφωνα με τον Viechtbauer (2005) έχοντας k μελέτες με τα παρατηρούμενα μεγέθη επιδράσεων y_1, y_2, \dots, y_k , το γενικό γραμμικό μοντέλο μικτών επιδράσεων (general linear mixed effects model) ορίζεται ως $y = X\beta + Z\gamma + e$ $e \sim N(0, R)$, $\gamma \sim N(0, D)$ με $y \sim N(X\beta, V)$ όπου

y οι $(k \times 1)$ τυχαίες μεταβλητές, X ο $(k \times p)$ πίνακας σταθερών επιδράσεων, β η $(p \times 1)$ παράμετρος σταθερών επιδράσεων, Z ο $(k \times q)$ πίνακας τυχαίων επιδράσεων, γ η $(q \times 1)$ παράμετρος τυχαίων επιδράσεων, e τα $(k \times 1)$ τυχαία σφάλματα με $E(\gamma) = 0, E(e) = 0, Cov(\gamma, e) = 0$. Για το μικτό μοντέλο οι πίνακες συνδιακυμάνσεων είναι ο $(q \times q)$ πίνακας συνδιακυμάνσεων D των τυχαίων επιδράσεων γ , ο $(k \times k)$ πίνακας συνδιακυμάνσεων R των τυχαίων σφαλμάτων e και ο $(k \times k)$ πίνακας V συνδιακυμάνσεων του y , ο οποίος ορίζεται ως $V = ZDZ' + R$. Ο λογάριθμος της πιθανοφάνειας για το γενικό μοντέλο μικτών επιδράσεων είναι

$$\ln L = -\frac{k}{2} \ln(2\pi) - \frac{1}{2} \ln|V| - \frac{1}{2} (y - X\beta)' V^{-1} (y - X\beta)$$

Το μοντέλο τυχαίων επιδράσεων αποτελεί ειδική περίπτωση του γενικού μοντέλου μικτών επιδράσεων με

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix}_{k \times 1}, X = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{k \times 1}, \mu = \beta, Z = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}_{k \times k}$$

$$\gamma = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_k \end{pmatrix}_{k \times 1}, e = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_k \end{pmatrix}_{k \times 1}, V = \begin{pmatrix} \tau^2 + \hat{\sigma}_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \tau^2 + \hat{\sigma}_k^2 \end{pmatrix}_{k \times k}$$

θεωρώντας τα $\hat{\sigma}_i^2$ ως γνωστά. Επομένως θα έχουμε $y_i = \mu + \xi_i + \varepsilon_i$ με $\varepsilon_i \sim N(0, \hat{\sigma}_i^2), \xi_i \sim N(0, \tau^2)$, $\hat{\sigma}_i^2$ γνωστά. Τα παρατηρούμενα μεγέθη κατανέμονται ασυμπτωτικά στην κανονική κατανομή με μέση τιμή μ και διακύμανση $\hat{\sigma}_i^2 + \tau^2$, δηλαδή $y_i \sim N(\mu, \hat{\sigma}_i^2 + \tau^2), i = 1, 2, \dots, k$

Επίσης, ο λογάριθμος της συνάρτησης πιθανοφάνειας των παρατηρούμενων επιδράσεων που αποτελεί την ειδική περίπτωση του γενικού μικτών επιδράσεων μοντέλου ορίζεται ως:

$$\begin{aligned} l(\mu, \tau^2) &= \ln L(\mu, \tau^2) \\ &= -\frac{k}{2} \ln(2\pi) - \frac{1}{2} \ln \prod_{i=1}^k (\hat{\sigma}_i^2 + \tau^2) - \frac{1}{2} \sum_{i=1}^k \frac{(y_i - \mu)^2}{(\hat{\sigma}_i^2 + \tau^2)} \end{aligned}$$

$$= -\frac{k}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^k \ln(\hat{\sigma}_i^2 + \tau^2) - \frac{1}{2} \sum_{i=1}^k \frac{(y_i - \mu)^2}{(\hat{\sigma}_i^2 + \tau^2)}$$

$$\mu \in R, \tau^2 \geq 0 \quad (4.5.1.1)$$

Έχοντας ότι ο λογάριθμος της συνάρτησης πιθανοφάνειας είναι μερικώς διαφορίσιμος, παίρνουμε τις μερικές παραγώγους ως προς μ και ως προς τ^2 και τις θέτουμε ίσες με μηδέν.

$$\frac{\partial \ln L(\mu, \tau^2)}{\partial \mu} = 0 \quad \text{και} \quad \frac{\partial \ln L(\mu, \tau^2)}{\partial \tau^2} = 0$$

Η μερική παράγωγος της πιθανοφάνειας ως προς μ είναι:

$$\frac{\partial \ln L(\mu, \tau^2)}{\partial \mu} = \sum_{i=1}^k \frac{(y_i - \mu)}{(\hat{\sigma}_i^2 + \tau^2)} = \sum_{i=1}^k \frac{y_i}{(\hat{\sigma}_i^2 + \tau^2)} - \mu \sum_{i=1}^k \frac{1}{(\hat{\sigma}_i^2 + \tau^2)}$$

Θέτουμε τη μερική παράγωγο ως προς μ ίση με το μηδέν

$$\frac{\partial \ln L(\mu, \tau^2)}{\partial \mu} = 0$$

και αντικαθιστούμε το τ^2 με τον εκτιμητή του $\hat{\tau}^2$

$$\sum_{i=1}^k \frac{y_i}{(\hat{\sigma}_i^2 + \hat{\tau}^2)} - \mu \sum_{i=1}^k \frac{1}{(\hat{\sigma}_i^2 + \hat{\tau}^2)} = 0$$

Επομένως γνωρίζοντας ότι $w_{i,RE} = \frac{1}{(\hat{\sigma}_i^2 + \hat{\tau}^2)}$ εξάγουμε τον εκτιμητή του μ

$$\hat{\mu}_{RE} = \frac{\sum_{i=1}^k w_{i,RE} y_i}{\sum_{i=1}^k w_{i,RE}} \quad (4.5.1.2)$$

Η μερική παράγωγος της πιθανοφάνειας ως προς τ^2 είναι:

$$\frac{\partial \ln L(\mu, \tau^2)}{\partial \tau^2} = -\frac{1}{2} \sum_{i=1}^k \frac{\frac{\partial (\hat{\sigma}_i^2 + \tau^2)}{\partial \tau^2}}{(\hat{\sigma}_i^2 + \tau^2)} + \frac{1}{2} \sum_{i=1}^k \frac{(y_i - \mu)^2 \cdot \frac{\partial (\hat{\sigma}_i^2 + \tau^2)}{\partial \tau^2}}{(\hat{\sigma}_i^2 + \tau^2)^2}$$

$$\begin{aligned}
 &= -\frac{1}{2} \sum_{i=1}^k \frac{2\tau}{(\hat{\sigma}_i^2 + \tau^2)} + \frac{1}{2} \sum_{i=1}^k \frac{(y_i - \mu)^2 \cdot 2\tau}{(\hat{\sigma}_i^2 + \tau^2)^2} \\
 &= -\sum_{i=1}^k \frac{\tau}{(\hat{\sigma}_i^2 + \tau^2)} + \sum_{i=1}^k \frac{(y_i - \mu)^2 \cdot \tau}{(\hat{\sigma}_i^2 + \tau^2)^2}
 \end{aligned}$$

Θέτουμε τη μερική παράγωγο ως προς τ^2 ίση με το μηδέν

$$\frac{\partial \ln L(\mu, \tau^2)}{\partial \tau^2} = 0$$

δηλαδή

$$\sum_{i=1}^k \frac{1}{(\hat{\sigma}_i^2 + \tau^2)} = \sum_{i=1}^k \frac{(y_i - \mu)^2}{(\hat{\sigma}_i^2 + \tau^2)^2}$$

Αντικαθιστώντας όπου τ^2 τον εκτιμητή του, το $\hat{\tau}^2$ και όπου μ το $\hat{\mu}_{RE}$ έχουμε

$$\begin{aligned}
 \sum_{i=1}^k \frac{1}{(\hat{\sigma}_i^2 + \hat{\tau}^2)} &= \sum_{i=1}^k \frac{\left(y_i - \frac{\sum_{i=1}^k \frac{1}{(\hat{\sigma}_i^2 + \hat{\tau}^2)} y_i}{\sum_{i=1}^k \frac{1}{(\hat{\sigma}_i^2 + \hat{\tau}^2)}} \right)^2}{(\hat{\sigma}_i^2 + \hat{\tau}^2)^2} \\
 \sum_{i=1}^k \frac{1}{(\hat{\sigma}_i^2 + \hat{\tau}^2)} &= \sum_{i=1}^k \frac{(y_i - \hat{\mu}_{RE})^2}{(\hat{\sigma}_i^2 + \hat{\tau}^2)^2}
 \end{aligned}$$

Πολλαπλασιάζουμε και διαιρούμε μέσα στο άθροισμα $\sum_{i=1}^k \frac{1}{(\hat{\sigma}_i^2 + \hat{\tau}^2)}$ την ποσότητα $(\hat{\sigma}_i^2 + \hat{\tau}^2)$

$$\sum_{i=1}^k \left(\frac{1}{(\hat{\sigma}_i^2 + \hat{\tau}^2)} \frac{(\hat{\sigma}_i^2 + \hat{\tau}^2)}{(\hat{\sigma}_i^2 + \hat{\tau}^2)} \right) = \sum_{i=1}^k \frac{(y_i - \hat{\mu}_{RE})^2}{(\hat{\sigma}_i^2 + \hat{\tau}^2)^2}$$

$$\sum_{i=1}^k \left(\frac{\hat{\sigma}_i^2 + \hat{\tau}^2}{(\hat{\sigma}_i^2 + \hat{\tau}^2)^2} \right) = \sum_{i=1}^k \frac{(y_i - \hat{\mu}_{RE})^2}{(\hat{\sigma}_i^2 + \hat{\tau}^2)^2}$$

$$\sum_{i=1}^k \frac{\hat{\sigma}_i^2}{(\hat{\sigma}_i^2 + \hat{\tau}^2)^2} + \hat{\tau}^2 \sum_{i=1}^k \frac{1}{(\hat{\sigma}_i^2 + \hat{\tau}^2)^2} = \sum_{i=1}^k \frac{(y_i - \hat{\mu}_{RE})^2}{(\hat{\sigma}_i^2 + \hat{\tau}^2)^2}$$

Αντικαθιστούμε τις ποσότητες $w_{i,RE} = \frac{1}{(\hat{\sigma}_i^2 + \hat{\tau}^2)}$

$$\sum_{i=1}^k w_{i,RE}^2 \hat{\sigma}_i^2 + \hat{\tau}^2 \sum_{i=1}^k w_{i,RE}^2 = \sum_{i=1}^k w_{i,RE}^2 (y_i - \hat{\mu}_{RE})^2$$

Λύνοντας ως προς $\hat{\tau}^2$ έχουμε τον εκτιμητή μέγιστης πιθανοφάνειας (ML)

$$\hat{\tau}_{ML}^2 = \frac{\sum_i w_{i,RE}^2 (\hat{\tau}_{ML}^2) ((y_i - \hat{\mu}_{RE}(\hat{\tau}_{ML}^2))^2 - \hat{\sigma}_i^2)}{\sum_i w_{i,RE}^2 (\hat{\tau}_{ML}^2)}$$

όπου $\hat{\mu}_{RE}(\hat{\tau}_{ML}^2)$ ο εκτιμητής του μ από την εξίσωση (4.5.1.2) για $\hat{\tau}^2 = \hat{\tau}_{ML}^2$

$$\hat{\mu}_{RE}(\hat{\tau}_{ML}^2) = \frac{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{ML}^2) y_i}{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{ML}^2)}$$

με τα βάρη τυχαίων επιδράσεων για $\hat{\tau}^2 = \hat{\tau}_{ML}^2$

$$w_{i,RE}(\hat{\tau}_{ML}^2) = \frac{1}{\hat{\sigma}_i^2 + \hat{\tau}_{ML}^2}$$

Επομένως οι εκτιμητές μέγιστης πιθανοφάνειας που προκύπτουν είναι

$$\hat{\tau}_{ML}^2 = \frac{\sum_i w_{i,RE}^2(\hat{\tau}_{ML}^2) ((y_i - \hat{\mu}_{RE}(\hat{\tau}_{ML}^2))^2 - \hat{\sigma}_i^2)}{\sum_i w_{i,RE}^2(\hat{\tau}_{ML}^2)} \quad (4.5.1.3)$$

και η εξίσωση (4.5.1.2) για τον εκτιμητή $\hat{\tau}_{REML}^2$ αποτελεί τον εκτιμητή της συνολικής επίδρασης για τον εκτιμητή για την περιορισμένη μέθοδο μέγιστης πιθανοφάνειας

$$\hat{\mu}_{RE}(\hat{\tau}_{ML}^2) = \frac{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{ML}^2) y_i}{w_{i,RE}(\hat{\tau}_{ML}^2)}$$

με τα βάρη τυχαίων επιδράσεων

$$w_{i,RE}(\hat{\tau}_{ML}^2) = \frac{1}{\hat{\sigma}_i^2 + \hat{\tau}_{ML}^2}$$

Οι λύσεις που βρέθηκαν αποτελούν όντως τους εκτιμητές μέγιστης πιθανοφάνειας καθώς ικανοποιούν ότι $A < 0$ ή $\Gamma < 0$ και $A\Gamma - B^2 > 0$ με A, B, Γ να είναι οι ποσότητες

$$A = \frac{\partial^2 \ln L(\mu, \tau^2)}{\partial \mu^2}, B = \frac{\partial^2 \ln L(\mu, \tau^2)}{\partial \mu \partial \tau^2}, \Gamma = \frac{\partial^2 \ln L(\mu, \tau^2)}{\partial (\tau^2)^2}$$

στα σημεία $\hat{\tau}_{ML}^2$ και $\hat{\mu}_{RE}(\hat{\tau}_{ML}^2)$ (Παπαϊωάννου, Γ., Φερεντίνος, Κ., 2000). Καθώς ο εκτιμητής μέγιστης πιθανοφάνειας από την εξίσωση (4.5.1.3) έχει και δεξιά από την ισότητα ποσότητες $\hat{\tau}_{ML}^2$ αυτό δηλώνει ότι για την εύρεση του χρειαζόμαστε μια επαναληπτική μέθοδο. Για την πρώτη επανάληψη οι εκτιμητές $\hat{\tau}_{ML}^2$ που βρίσκονται δεξιά της εξίσωσης (4.5.1.3) μπορούν να πάρουν ως αρχική τιμή το μηδέν, μια τιμή μη επαναληπτικού εκτιμητή ή οποιαδήποτε αποδεκτή τιμή ετερογένειας (Veroniki et al (submitted)). Σε κάθε επανάληψη εκτιμάμε τη συνολική επίδραση $\hat{\mu}_{RE}(\hat{\tau}_{ML}^2)$ και τον εκτιμητή ετερογένειας $\hat{\tau}_{ML}^2$ για τον οποίο απαιτούμε να παίρνει τις μη αρνητικές τιμές (Veroniki et al (submitted)). Χαρακτηριστικές επαναληπτικές μέθοδοι αποτελούν η μέθοδος Newton-Raphson, μέθοδος Scoring, μέθοδος Simplex, Fisher ή ο αλγόριθμος EM (expectation - maximisation). Ο εκτιμητής ετερογένειας μέγιστης πιθανοφάνειας $\hat{\tau}_{ML}^2$ προκύπτει όταν έχουμε σύγκλιση στην επαναληπτική μέθοδο, όταν δηλαδή το κριτήριο τερματισμού ικανοποιηθεί (Veroniki et al (submitted)).

4.5.2 Περιορισμένη μέθοδος μέγιστης πιθανοφάνειας (Restricted Maximum likelihood method (REML))

Η περιορισμένη μέθοδος μέγιστης πιθανοφάνειας (Restricted Maximum likelihood method) είναι γνωστή και ως Residual Maximum likelihood και εφαρμόζεται με στόχο να διορθώσει την αρνητική μεροληψία που δημιουργείται από τη μέθοδο μέγιστης πιθανοφάνειας (Viechtbauer, 2005). Σύμφωνα με τον Viechtbauer (2005) ο λογάριθμος της συνάρτησης μέγιστης πιθανοφάνειας για το γενικό μοντέλο μικτών επιδράσεων είναι

$$\ln L(y) = -\frac{k}{2} \ln(2\pi) - \frac{1}{2} \ln|V| - \frac{1}{2} (y - X\beta)' V^{-1} (y - X\beta)$$

Η μεροληψία που δημιουργεί η μέθοδος μέγιστης πιθανοφάνειας μπορεί να διορθωθεί με ένα γραμμικό μετασχηματισμό του y . Επειδή στόχος μας είναι η δημιουργία ενός αμερόληπτου εκτιμητή και γνωρίζοντας ότι στο μοντέλο μικτών επιδράσεων $E(y) = X\beta$, θέλουμε ένα μετασχηματισμό ο οποίος θα είναι ανεξάρτητος του β . Επομένως πολλαπλασιάζοντας οποιονδήποτε πίνακα K' με το y έτσι ώστε $K'X = 0$ θα δημιουργήσει την ανεξαρτησία που επιθυμούμε. Ένας πίνακας που ικανοποιεί την υπόθεση αυτή είναι ο πίνακας $K' = I - X(X'X)^{-1}X'$, με το ταυτοτικό $p \times p$ πίνακα I και K' να έχει βαθμίδα $k - r(X)$. Πολλαπλασιάζοντας τον πίνακα με το y προκύπτουν τα υπόλοιπα $\tilde{y} = K'y = (I - X(X'X)^{-1}X')y = y - X(X'X)^{-1}X'y = X\mu - X\hat{\mu} = e$. Με τον παραπάνω μετασχηματισμό ισχύει ότι $E(K'y) = K'X\mu = (I - X(X'X)^{-1}X')X\mu = 0$ δημιουργώντας έτσι έναν αμερόληπτο εκτιμητή. Έτσι ο μετασχηματισμός $K'y$ είναι ανεξάρτητος του β με $K'y \sim N(0, K'VK)$. Η περιορισμένη μέθοδος πιθανοφάνειας προκύπτει από τον παραπάνω μετασχηματισμό μεγιστοποιώντας την πιθανοφάνεια του μοντέλου υπολοίπων $\tilde{y} = K'y = K'X\beta + K'(Z\xi + \varepsilon)$.

Επομένως η πιθανοφάνεια θα είναι

$$L(y) = (2\pi)^{-\frac{(k-r(X))}{2}} |K'VK|^{-\frac{1}{2}} e^{-\frac{1}{2}(K'y)'(K'VK)^{-1}(K'y)}$$

και ο λογάριθμος της μέγιστης πιθανοφάνειας

$$l = -\frac{1}{2}(k - r(X)) \ln(2\pi) - \frac{1}{2} \ln|K'VK| - \frac{1}{2} y'K(K'VK)^{-1}K'y$$

Κατά τον Searle (1979) ισχύει ότι

$$\ln|K'VK| = \ln|V| + \ln|X'V^{-1}X| + \ln(|X'X|^{-1}) \quad (4.5.2.1)$$

και

$$y'K(K'VK)^{-1}K'y = (y - X\hat{\mu})'V^{-1}(y - X\hat{\mu}) \quad (4.5.2.2)$$

Χρησιμοποιώντας τις εξισώσεις (4.5.2.1) και (4.5.2.2) και αφαιρώντας το σταθερό όρο από την πιθανοφάνεια καταλήγουμε στην περιορισμένη μέθοδο μέγιστης πιθανοφάνειας του γενικού μοντέλου μικτών επιδράσεων

$$l(\tau^2) = -\frac{1}{2} \ln|V| - \frac{1}{2} \ln|X'V^{-1}X| - \frac{1}{2} \ln\left(\frac{1}{|X'X|}\right) - \frac{1}{2} (y - X\hat{\mu})'V^{-1}(y - X\hat{\mu})$$

Κατά την περιορισμένη μέθοδο μέγιστης πιθανοφάνειας ως αρχικό βήμα χρησιμοποιούμε τον εκτιμητή ML. Για την ειδική περίπτωση του μοντέλου μικτών επιδράσεων όπως περιεγράφηκε στη μέθοδο μέγιστης πιθανοφάνειας έχουμε την περιορισμένη μέγιστη πιθανοφάνεια για το μοντέλο τυχαίων επιδράσεων

$$\begin{aligned} l(\tau^2) &= \\ &= -\frac{1}{2} \ln \prod_{i=1}^k (\tau^2 + \hat{\sigma}_i^2) - \frac{1}{2} \ln \prod_{i=1}^k \frac{1}{(\tau^2 + \hat{\sigma}_i^2)} - \frac{1}{2} \sum_{i=1}^k \frac{(y_i - \hat{\mu}_{RE}(\hat{t}_{ML}^2))^2}{(\hat{\sigma}_i^2 + \tau^2)^2} \\ &= -\frac{1}{2} \sum_{i=1}^k \ln(\tau^2 + \hat{\sigma}_i^2) - \frac{1}{2} \sum_{i=1}^k \ln \frac{1}{(\tau^2 + \hat{\sigma}_i^2)} - \frac{1}{2} \sum_{i=1}^k \frac{(y_i - \hat{\mu}_{RE}(\hat{t}_{ML}^2))^2}{(\hat{\sigma}_i^2 + \tau^2)^2} \end{aligned}$$

Στη συνάρτηση της πιθανοφάνειας έχουμε μόνο την παράμετρο τ^2 ως άγνωστη με σκοπό την εύρεση του εκτιμητή που θα μεγιστοποιήσει την πιθανοφάνεια. Η μερική παράγωγος της πιθανοφάνειας ως προς τ^2 είναι

$$\begin{aligned} \frac{\partial l(\tau^2)}{\partial \tau^2} &= \\ &= -\frac{1}{2} \sum_{i=1}^k \frac{(\hat{\sigma}_i^2 + \tau^2)'}{(\hat{\sigma}_i^2 + \tau^2)} - \frac{1}{2} \sum_{i=1}^k (\hat{\sigma}_i^2 + \tau^2) \frac{\partial(\hat{\sigma}_i^2 + \tau^2)}{\partial \tau^2} + \frac{1}{2} \sum_{i=1}^k \frac{(y_i - \hat{\mu}_{RE}(\hat{t}_{ML}^2))^2 \left[\frac{\partial(\hat{\sigma}_i^2 + \tau^2)}{\partial \tau^2} \right]}{(\hat{\sigma}_i^2 + \tau^2)^2} \\ &= -\frac{1}{2} \sum_{i=1}^k \frac{2\tau}{(\hat{\sigma}_i^2 + \tau^2)} - \frac{1}{2} \sum_{i=1}^k (\hat{\sigma}_i^2 + \tau^2)(2\tau) + \frac{1}{2} \sum_{i=1}^k \frac{(y_i - \hat{\mu}_{RE}(\hat{t}_{ML}^2))^2 (2\tau)}{(\hat{\sigma}_i^2 + \tau^2)^2} \\ &= -\sum_{i=1}^k \frac{\tau}{(\hat{\sigma}_i^2 + \tau^2)} - \sum_{i=1}^k (\hat{\sigma}_i^2 + \tau^2)\tau + \sum_{i=1}^k \frac{(y_i - \hat{\mu}_{RE}(\hat{t}_{ML}^2))^2 \cdot \tau}{(\hat{\sigma}_i^2 + \tau^2)^2} \end{aligned}$$

Θέτουμε τη μερική παράγωγο ως προς τ^2 ίση με το μηδέν

$$\begin{aligned} \frac{\partial l(\tau^2)}{\partial \tau^2} &= 0 \\ \sum_{i=1}^k \frac{1}{(\hat{\sigma}_i^2 + \tau^2)} - \sum_{i=1}^k (\hat{\sigma}_i^2 + \tau^2) &= \sum_{i=1}^k \frac{(y_i - \hat{\mu}_{RE}(\hat{t}_{ML}^2))^2}{(\hat{\sigma}_i^2 + \tau^2)^2} \end{aligned}$$

Αντικαθιστώντας όπου τ^2 τον εκτιμητή του, το $\hat{\tau}^2$ έχουμε

$$\sum_{i=1}^k \frac{1}{(\hat{\sigma}_i^2 + \hat{\tau}^2)} = \sum_{i=1}^k \frac{(y_i - \hat{\mu}_{RE}(\hat{\tau}_{ML}^2))^2}{(\hat{\sigma}_i^2 + \hat{\tau}^2)^2} + \sum_{i=1}^k (\hat{\sigma}_i^2 + \tau^2)$$

Πολλαπλασιάζουμε και διαιρούμε μέσα στο άθροισμα $\sum_{i=1}^k \frac{1}{(\hat{\sigma}_i^2 + \hat{\tau}^2)}$ την ποσότητα $(\hat{\sigma}_i^2 + \hat{\tau}^2)$

$$\sum_{i=1}^k \left(\frac{1}{(\hat{\sigma}_i^2 + \hat{\tau}^2)} \frac{(\hat{\sigma}_i^2 + \hat{\tau}^2)}{(\hat{\sigma}_i^2 + \hat{\tau}^2)} \right) = \sum_{i=1}^k \frac{(y_i - \hat{\mu}_{RE}(\hat{\tau}_{ML}^2))^2}{(\hat{\sigma}_i^2 + \hat{\tau}^2)^2} + \sum_{i=1}^k (\hat{\sigma}_i^2 + \tau^2)$$

$$\sum_{i=1}^k \left(\frac{(\hat{\sigma}_i^2 + \hat{\tau}^2)}{(\hat{\sigma}_i^2 + \hat{\tau}^2)^2} \right) = \sum_{i=1}^k \frac{(y_i - \hat{\mu}_{RE}(\hat{\tau}_{ML}^2))^2}{(\hat{\sigma}_i^2 + \hat{\tau}^2)^2} + \sum_{i=1}^k (\hat{\sigma}_i^2 + \tau^2)$$

$$\sum_{i=1}^k \frac{\hat{\sigma}_i^2}{(\hat{\sigma}_i^2 + \hat{\tau}^2)^2} + \hat{\tau}^2 \sum_{i=1}^k \frac{1}{(\hat{\sigma}_i^2 + \hat{\tau}^2)^2} = \sum_{i=1}^k \frac{(y_i - \hat{\mu}_{RE}(\hat{\tau}_{ML}^2))^2}{(\hat{\sigma}_i^2 + \hat{\tau}^2)^2} + \sum_{i=1}^k (\hat{\sigma}_i^2 + \tau^2)$$

Αντικαθιστούμε τις ποσότητες $w_{i,RE} = \frac{1}{(\hat{\sigma}_i^2 + \hat{\tau}^2)}$

$$\sum_{i=1}^k w_{i,RE}^2 \hat{\sigma}_i^2 + \hat{\tau}^2 \sum_{i=1}^k w_{i,RE}^2 = \sum_{i=1}^k w_{i,RE}^2 (y_i - \hat{\mu}_{RE}(\hat{\tau}_{ML}^2))^2 + \sum_{i=1}^k \frac{1}{w_{i,RE}}$$

Λύνοντας ως προς $\hat{\tau}^2$ και αντικαθιστώντας τον εκτιμητή $\hat{\tau}_{ML}^2$ του πρώτου βήματος με τον εκτιμητή $\hat{\tau}_{REML}^2$ έχουμε τον εκτιμητή περιορισμένης μέγιστης πιθανοφάνειας (REML)

$$\hat{\tau}_{REML}^2 = \frac{\sum_i w_{i,RE}^2 (\hat{\tau}_{REML}^2) ((y_i - \hat{\mu}_{RE}(\hat{\tau}_{REML}^2))^2 - \hat{\sigma}_i^2)}{\sum_i w_{i,RE}^2 (\hat{\tau}_{REML}^2)} + \frac{1}{\sum_{i=1}^k w_{i,RE} (\hat{\tau}_{REML}^2)}$$

όπου $\hat{\mu}_{RE}(\hat{\tau}_{ML}^2)$ ο εκτιμητής του μ για $\hat{\tau}^2 = \hat{\tau}_{REML}^2$

$$\hat{\mu}_{RE}(\hat{\tau}_{REML}^2) = \frac{\sum_{i=1}^k w_{i,RE} (\hat{\tau}_{REML}^2) y_i}{\sum_{i=1}^k w_{i,RE} (\hat{\tau}_{REML}^2)}$$

με τα βάρη τυχαίων επιδράσεων για $\hat{\tau}^2 = \hat{\tau}_{REML}^2$

$$w_{i,RE}(\hat{\tau}_{REML}^2) = \frac{1}{\hat{\sigma}_i^2 + \hat{\tau}_{REML}^2}$$

Επομένως ο εκτιμητής της περιορισμένης μέγιστης πιθανοφάνειας που προκύπτει είναι

$$\hat{\tau}_{REML}^2 = \frac{\sum_i w_{i,RE}^2(\hat{\tau}_{REML}^2) ((y_i - \hat{\mu}_{RE}(\hat{\tau}_{REML}^2))^2 - \hat{\sigma}_i^2)}{\sum_i w_{i,RE}^2(\hat{\tau}_{REML}^2)} + \frac{1}{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{REML}^2)} \quad (4.5.2.3)$$

το οποίο ικανοποιεί την εξίσωση $\frac{\partial^2 l(\tau^2)}{\partial (\tau^2)^2} < 0$ στο σημείο $\hat{\tau}_{REML}^2$ με τα βάρη τυχαίων επιδράσεων

$$w_{i,RE}(\hat{\tau}_{REML}^2) = \frac{1}{\hat{\sigma}_i^2 + \hat{\tau}_{REML}^2}$$

και τη συνολική επίδραση

$$\hat{\mu}_{RE}(\hat{\tau}_{REML}^2) = \frac{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{REML}^2) y_i}{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{REML}^2)}$$

Η μέθοδος περιορισμένης μέγιστης πιθανοφάνειας πρόκειται για μια επαναληπτική μέθοδο όπως τις επαναληπτικές μεθόδους που χρησιμοποιούνται για την μέθοδο μέγιστης πιθανοφάνειας. Για την πρώτη επανάληψη ο εκτιμητής $\hat{\tau}_{REML}^2$ παίρνει ως αρχική τιμή τον εκτιμητή μέγιστης πιθανοφάνειας $\hat{\tau}_{ML}^2$ (Veroniki et al (submitted)). Σε κάθε επανάληψη εκτιμάμε τη συνολική επίδραση και τον εκτιμητή ετερογένειας καθώς απαιτούμε τις μη αρνητικές τιμές στον εκτιμητή ετερογένειας (Veroniki et al (submitted)). Ο εκτιμητής ετερογένειας περιορισμένης μέγιστης πιθανοφάνειας $\hat{\tau}_{REML}^2$ προκύπτει όταν στην επαναληπτική μέθοδο επιτυγχάνεται η σύγκλιση.

4.5.3 Προσεγγισμένη περιορισμένη μέθοδος μέγιστης πιθανοφάνειας (Approximate Restricted Maximum likelihood method (AREML))

Ο εκτιμητής της περιορισμένης μεθόδου μέγιστης πιθανοφάνειας σύμφωνα με τον Morris (1983) και δίνεται και από την σχέση

$$\hat{\tau}_{AREML}^2 = \frac{\sum_i w_{i,RE}^2(\hat{\tau}_{AREML}^2) \left(\left(\frac{k}{k-1} \right) ((y_i - \hat{\mu}_{RE}(\hat{\tau}_{AREML}^2))^2 - \hat{\sigma}_i^2) \right)}{\sum_i w_{i,RE}^2(\hat{\tau}_{AREML}^2)}$$

με τα βάρη τυχαίων επιδράσεων $w_{i,RE}(\hat{\tau}_{AREML}^2)$

$$w_{i,RE}(\hat{\tau}_{AREML}^2) = \frac{1}{\hat{\sigma}_i^2 + \hat{\tau}_{AREML}^2}$$

και τη συνολική επίδραση

$$\hat{\mu}_{RE}(\hat{\tau}_{AREML}^2) = \frac{\sum_{i=1}^k w_{i,RE}(\hat{\tau}_{AREML}^2) y_i}{w_{i,RE}(\hat{\tau}_{AREML}^2)}$$

Ο εκτιμητής ετερογένειας έχει οριστεί ως προσέγγιση του εκτιμητή της περιορισμένης μεθόδου μέγιστης πιθανοφάνειας και συμβολίζεται με AREML (Morris, 1983). Πρόκειται για μια επαναληπτική μέθοδο όπως ο εκτιμητής REML και αποτελεί ένα μη αρνητικό εκτιμητή (Viechtbauer, 2005). Ο εκτιμητής AREML αποδίδει ακριβώς τις ίδιες τιμές με τον εκτιμητή REML όταν οι πληθυσμιακές διακυμάνσεις είναι ίσες και επειδή στην πράξη αυτό είναι σπάνιο, ο εκτιμητής REML προτιμάται (Viechtbauer, 2005).

4.6 Μέθοδος σταθμισμένων ελαχίστων τετραγώνων (Weighted Least-Squares method)

Εκτός από τη γενική μέθοδο των ροπών και της μέγιστης πιθανοφάνειας μια συνηθισμένη μέθοδος εκτίμησης της ετερογένειας αποτελεί η μέθοδος σταθμισμένων ελαχίστων τετραγώνων (Weighted Least-Squares method). Σύμφωνα με τη μέθοδο αυτή ο εκτιμητής ετερογένειας προκύπτει από το σταθμισμένο άθροισμα των τετραγώνων των υπολοίπων. Στη βιβλιογραφία η μέθοδος είναι γνωστή και ως μοντέλο εκτίμησης της διακύμανσης των σφαλμάτων (model error variance estimator (MV)). Με τη μέθοδο των σταθμισμένων ελαχίστων τετραγώνων στη μετα-ανάλυση έχουμε δύο εκτιμητές, τον εκτιμητή Sidik-Jonkman (SJ) και τη βελτιωμένη του προσέγγιση που προτάθηκε από τους ίδιους και είναι γνωστή ως εκτιμητής SJgho.

4.6.1 Εκτιμητής Sidik-Jonkman (SJ)

Σύμφωνα με τους Sidik Jonkman (2005) θεωρούμε ότι έχουμε k μελέτες με τα παρατηρούμενα μεγέθη επιδράσεων y_1, y_2, \dots, y_k για το μοντέλο τυχαίων επιδράσεων

$$\begin{aligned} y_i &= \theta_i + \varepsilon_i, & \varepsilon_i &\sim N(0, \sigma_i^2) \\ \theta_i &= \mu + \xi_i, & \xi_i &\sim N(0, \tau^2) \end{aligned}$$

με σ_i^2 να είναι γνωστά ή αλλιώς

$y_i = \mu + \xi_i + \varepsilon_i$ με $\varepsilon_i \sim N(0, \sigma_i^2)$, $\xi_i \sim N(0, \tau^2)$, σ_i^2 γνωστά ή σε μορφή πινάκων

$$y = X\mu + \gamma + e \quad e \sim N(0, R), \gamma \sim N(0, D) \quad \text{με}$$

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix}_{k \times 1}, X = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{k \times 1}, \gamma = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_k \end{pmatrix}_{k \times 1}, e = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_k \end{pmatrix}_{k \times 1}$$

και πίνακα συνδιακυμάνσεων τον διαγώνιο πίνακα

$$V = \begin{pmatrix} \tau^2 + \sigma_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \tau^2 + \sigma_k^2 \end{pmatrix}_{k \times k}$$

Η αναμενόμενη τιμή και η διακύμανση των παρατηρούμενων μεγεθών επίδρασης είναι

$$E(y_i) = \mu \quad \text{και} \quad \text{Var}(y_i) = \sigma_i^2 + \tau^2, i = 1, 2, \dots, k$$

Η διακύμανση των παρατηρούμενων μεγεθών μπορεί να γραφεί ως

$$\text{Var}(y_i) = \sigma_i^2 + \tau^2 = \tau^2(r_i + 1), i = 1, 2, \dots, k$$

με $r_i = \frac{\sigma_i^2}{\tau^2}$, $\tau^2 \neq 0$ θεωρώντας τα $r_i, i = 1, 2, \dots, k$ ως γνωστά.

Τότε ο πίνακας συνδιακυμάνσεων μπορεί να γραφεί ως

$$V = \tau^2 V_1 = \tau^2 \begin{pmatrix} r_1 + 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & r_k + 1 \end{pmatrix}_{k \times k} \quad \text{με} \quad r_i = \frac{\sigma_i^2}{\tau^2}$$

με την αναμενόμενη τιμή και τη διακύμανση των παρατηρούμενων μεγεθών επίδρασης

$$E(y) = X\mu \\ \text{Var}(y) = \tau^2 V_1$$

Για το μοντέλο ο καλύτερος αμερόληπτος εκτιμητής (best linear unbiased estimator BLUE) δίνεται ως

$$\hat{\theta}_v = (X^T V^{-1} X)^{-1} X^T V^{-1} y = \frac{\sum_{i=1}^k v_i^{-1} y_i}{\sum_{i=1}^k v_i^{-1}}$$

με $v_i = (r_i + 1), i = 1, 2, \dots, k$ και ο εκτιμητής της ετερογένειας προκύπτει από το σταθμισμένο άθροισμα των τετραγώνων των υπολοίπων

$$\hat{t}^2 = \frac{(y - X\hat{\theta}_v)^T V^{-1}(y - X\hat{\theta}_v)}{k-1} = \frac{1}{k-1} \sum_{i=1}^k v_i^{-1} (y_i - \hat{\theta}_v)^2$$

Ο παραπάνω εκτιμητής είναι γνωστός ως εκτιμητής της ετερογένειας Sidik-Jonkman ή αλλιώς μοντέλο εκτίμησης της διακύμανσης των σφαλμάτων (model error variance estimator). Στην πράξη όμως τα $v_i = (r_i + 1), i = 1, 2, \dots, k$ είναι άγνωστα, για το λόγο αυτό θα εκτιμήσουμε τα $r_i = \frac{\sigma_i^2}{\tau^2}$ με τις εκτιμώμενες πληθυσμιακές διακυμάνσεις $\hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_k^2$ και την εκτίμηση ενός εκτιμητή ετερογένειας του τ^2 . Για την εκτίμηση της ετερογένειας χρησιμοποιείται μια διαισθητική εκτίμηση με χρήση της εμπειρικής εκτίμησης της διακύμανσης $\sum_{i=1}^k \frac{(\theta_i - \bar{\theta})^2}{k}$ με $\bar{\theta} = \frac{\sum_{i=1}^k \theta_i}{k}$ έχοντας τα πραγματικά μεγέθη επίδρασης $\theta_i, i = 1, 2, \dots, k$ γνωστά.

Έχοντας τα πραγματικά μεγέθη επίδρασης $\theta_i, i = 1, 2, \dots, k$ άγνωστα θα χρησιμοποιήσουμε τα παρατηρούμενα μεγέθη επίδρασης $y_i, i = 1, 2, \dots, k$ αντί για τα πραγματικά μεγέθη επίδρασης. Επομένως, ο εκτιμητής ετερογένειας είναι

$$\hat{t}_0^2 = \frac{1}{k} \sum_{i=1}^k (y_i - \bar{y})^2 \text{ με } \bar{y} = \frac{\sum_{i=1}^k y_i}{k}$$

Επομένως ο εκτιμητής του r_i είναι $\hat{r}_i = \frac{\hat{\sigma}_i^2}{\hat{t}_0^2}, i = 1, 2, \dots, k$.

Χρησιμοποιώντας τους εκτιμητές \hat{r}_i ο εκτιμητής ετερογένειας Sidik-Jonkman είναι

$$\hat{t}_{SJ}^2 = \frac{1}{k-1} \sum_{i=1}^k \hat{v}_i^{-1} (y_i - \hat{\theta}_v)^2$$

με

$$\hat{\theta}_v = \frac{\sum_{i=1}^k \hat{v}_i^{-1} y_i}{\sum_{i=1}^k \hat{v}_i^{-1}}$$

$$\hat{v}_i = (\hat{r}_i + 1), i = 1, 2, \dots, k$$

$$\hat{r}_i = \frac{\hat{\sigma}_i^2}{\hat{t}_0^2}, i = 1, 2, \dots, k$$

Από την παραπάνω μέθοδο συμπεραίνουμε ότι ο εκτιμητής Sidik-Jonkman είναι μη επαναληπτικός. Ακόμη καταλήγουμε στο γεγονός ότι ο εκτιμητής Sidik-Jonkman είναι θετικός καθώς οποιαδήποτε αρνητική τιμή του δεν χρειάζεται να αποκόπτεται στο μηδέν.

4.6.2 Εκτιμητής (SJgho)

Οι Sidik Jonkman (2007) προτείνουν έναν βελτιωμένο τρόπο εκτίμησης της μεθόδου Sidik-Jonkman δημιουργώντας έτσι ένα νέο εκτιμητή, τον εκτιμητή SJgho. Πρόκειται ουσιαστικά για τον εκτιμητή Sidik-Jonkman με τη διαφορά ότι τα πηλίκια $r_i = \frac{\sigma_i^2}{\tau^2}$ δεν τα εκτιμούμε από τους εκτιμητές των πληθυσμιακών διακυμάνσεων, ούτε από το διαισθητικό εκτιμητή ετερογένειας $\hat{\tau}_0^2$. Στην περίπτωση αυτή η εκτίμηση των πηλίκων r_i βασίζεται στις τιμές του εκτιμητή General Hedges-Olkin, $\hat{\tau}_{GHO}^2$. Στην περίπτωση που ο εκτιμητής $\hat{\tau}_{GHO}^2$ ισούται με μηδέν τότε θεωρούμε ως τιμή του εκτιμητή τη μικρή ποσότητα 0.01 και ο εκτιμητής των πηλίκων r_i γίνεται $\hat{r}_i = \frac{\hat{\sigma}_i^2}{0.01}$.

4.7 Μέθοδος Non-Parametric Bootstrap για τον εκτιμητή DerSimonian and Laird (Non-Parametric bootstrap method (DLb))

Με τον όρο bootstrapping εννοούμε τη δημιουργία ενός τυχαίου δείγματος με επανατοποθέτηση ("Bootstrapping", 2015). Εάν έχουμε για παράδειγμα τέσσερις μελέτες, έστω A,B,Γ,Δ, ένα τυχαίο δείγμα με επανατοποθέτηση μπορεί να αποτελεί το A,A,B,Γ ή το B,Δ,Γ,A.

Κατά τους Kontopantelis et al. (2013), θεωρούμε k μελέτες σε μια μετα-ανάλυση με $y_i \sim N(\hat{\mu}, \hat{\sigma}_i^2 + \hat{\tau}^2)$, όπου $\hat{\tau}^2$ οποιοσδήποτε εκτιμητής ετερογένειας. Η μη παραμετρική μέθοδος bootstrap δημιουργεί ένα τυχαίο δείγμα με επανατοποθέτηση των y_i και των $\hat{\sigma}_i^2, i = 1, \dots, k$. Μετά την παραγωγή του τυχαίου δείγματος εκτιμάμε τον εκτιμητή ετερογένειας γι' αυτό με τη μέθοδο DerSimonian and Laird, $\hat{\tau}_b^2$. Επαναλαμβάνουμε τη διαδικασία για $b = 1, 2, \dots, B$ φορές και σε κάθε επανάληψη εκτιμάμε τον εκτιμητή ετερογένειας DerSimonian and Laird $\hat{\tau}_b^2, b = 1, 2, \dots, B$

τον οποίο αποκόπτουμε στο μηδέν εάν είναι αρνητικός. Έπειτα ο εκτιμητής DerSimonian and Laird της μη παραμετρικής μεθόδου bootstrap προκύπτει από τη μέση τιμή των εκτιμητών $\hat{t}_b^2, b = 1, 2, \dots, B$ και συμβολίζεται με \hat{t}_{DLb}^2 . Όπως δημιουργήσαμε τον εκτιμητή DerSimonian and Laird μπορούμε με τη μη παραμετρική μέθοδο bootstrap να εξάγουμε οποιονδήποτε εκτιμητή ετερογένειας.

4.8 Εκτιμητής Empirical Bayes (EB)

Ο εκτιμητής Empirical Bayes εντάχθηκε από τον Morris (1983), έχει παρόμοια μορφή με τους εκτιμητές από τη μέθοδο πιθανοφάνειας και προκύπτει από τον AREML αντικαθιστώντας τα βάρη $w_{i,RE}^2$ με $w_{i,RE}$ (Sidik & Jonkman, 2007). Επομένως δίνεται από τον τύπο

$$\hat{t}_{EB}^2 = \frac{\sum_i w_{i,RE}(\hat{t}_{EB}^2) \left(\left(\frac{k}{k-1} \right) ((y_i - \hat{\mu}_{RE}(\hat{t}_{EB}^2))^2 - \hat{\sigma}_i^2) \right)}{\sum_i w_{i,RE}(\hat{t}_{EB}^2)}$$

όπου $\hat{\mu}_{RE}(\hat{t}_{EB}^2)$ ο εκτιμητής του μ για $\hat{t}^2 = \hat{t}_{EB}^2$

$$\hat{\mu}_{RE}(\hat{t}_{EB}^2) = \frac{\sum_{i=1}^k w_{i,RE}(\hat{t}_{EB}^2) y_i}{w_{i,RE}(\hat{t}_{EB}^2)}$$

με τα βάρη τυχαίων επιδράσεων για $\hat{t}^2 = \hat{t}_{EB}^2$

$$w_{i,RE}(\hat{t}_{EB}^2) = \frac{1}{\hat{\sigma}_i^2 + \hat{t}_{EB}^2}$$

Από τον ορισμό του παρατηρείται ότι εκτιμητής EB αποτελεί έναν επαναληπτικό και μη αρνητικό εκτιμητή όπως και οι παρόμοιοι εκτιμητές από τη μέθοδο πιθανοφάνειας (Sidik & Jonkman, 2007). Ακόμη ο εκτιμητής EB έχει ακριβώς τις ίδιες τιμές με τον εκτιμητή PM (Paule & Mandel, 1982).

4.9 Μέθοδος Bayes

Η μπεϋζιανή στατιστική ή αλλιώς στατιστική κατά Bayes αποτελεί μια ειδική μέθοδο της στατιστικής βασική υπόθεση της οποίας αποτελεί το γεγονός ότι οι παράμετροι δεν είναι σταθερές αλλά τυχαίες παράμετροι οι οποίες έχουν μία κατανομή που ονομάζουμε εκ των προτέρων κατανομή

(prior distribution). Οι παράμετροι αυτοί μαζί με το τυχαίο δείγμα συνδέονται μέσω της θεωρίας του Bayes και μας δίνουν την από κοινού κατανομή τους, την οποία ονομάζουμε εκ των υστέρων κατανομή (posterior distribution). Κατά τη μέθοδο Bayes οι εκτιμητές ετερογένειας είναι: ο Rukhin Bayes, ο θετικός εκτιμητής Rukhin Bayes, ο Full Bayesian και ο Bayes Modal εκτιμητής.

4.9.1 Εκτιμητής Rukhin Bayes (RB)

Έχοντας το μοντέλο των τυχαίων επιδράσεων $y_i = \mu + \xi_i + \varepsilon_i$ με $\varepsilon_i \sim N(0, \sigma_i^2)$, $\xi_i \sim N(0, \tau^2)$, $\sigma_i^2, i = 1, 2, \dots, k$ γνωστά, θεωρούμε ότι οι παράμετροι σ_i^2 και τ^2 είναι ανεξάρτητοι. Για την παράμετρο τ^2 θεωρούμε την εκ των προτέρων κανονική κατανομή με μέση τιμή β και μεγάλη διακύμανση. Σύμφωνα με τον Rukhin (2013) η γενική μορφή των εκτιμητών Bayes είναι

$$\hat{t}_{RB}^2 = \max \left(0, \frac{\sum_{i=1}^k (y_i - \bar{y})^2}{k+1} + \frac{(\sum_{i=1}^k n_i - k)(2k\beta - (k+1) \sum_{i=1}^k \hat{\sigma}_i^2)}{\sum_{i=1}^k (n_i - k + 2) k(k+1)} \right)$$

με n_i να αποτελεί το μέγεθος του δείγματος στην i μελέτη. Έχοντας τη γενική μορφή των εκτιμητών Bayes για $\beta = 0$ παίρνουμε τον εκτιμητή Rukhin Bayes

$$\hat{t}_{RB0}^2 = \frac{\sum_{i=1}^k (y_i - \bar{y})^2}{k+1} - \frac{(\sum_{i=1}^k n_i - k)(k+1) \sum_{i=1}^k \hat{\sigma}_i^2}{\sum_{i=1}^k (n_i - k + 2) k(k+1)}$$

ο οποίος αποτελεί ένα μη αρνητικό και μη επαναληπτικό εκτιμητή ετερογένειας (Rukhin, 2013).

4.9.2 Θετικός εκτιμητής Rukhin Bayes (RBp)

Από τη γενική μορφή των εκτιμητών Bayes για $\beta = 0.5(k-1) \sum_{i=1}^k \frac{\hat{\sigma}_i^2}{k}$ παίρνουμε το θετικό εκτιμητή Rukhin Bayes ο οποίος είναι μη επαναληπτικός και δίνεται από τη σχέση (Rukhin, 2013):

$$\hat{t}_{RBp}^2 = \frac{\sum_{i=1}^k (y_i - \bar{y})^2}{k+1}$$

4.9.3 Εκτιμητής Full Bayesian (FB)

Στη στατιστική κατά Bayes μπορούμε να εκτιμήσουμε εκτιμητές ετερογένειας με τη μέθοδο Markov Chain Monte Carlo (MCMC). Ως μαρκοβιανή αλυσίδα (markov chain) θεωρείται ένα μαθηματικό σύστημα που μεταβάλλεται από τη μία κατάσταση στην άλλη μέσα σ' ένα πεπερασμένο αριθμό καταστάσεων (“Αλυσίδα Μαρκόφ”, 2015). Με τη μέθοδο Markov Chain Monte Carlo (MCMC) θεωρούμε μια κλάση αλγορίθμων οι οποίοι χρησιμοποιούν ως δείγμα την κατανομή πιθανοτήτων που δημιουργούν τη μαρκοβιανή αλυσίδα (“Markov chain Monte Carlo”, 2015). Η κατάσταση της μαρκοβιανής αλυσίδας μετά από έναν αριθμό σταδίων χρησιμοποιείται ως δείγμα της ζητούμενης κατανομής (“Markov chain Monte Carlo”, 2015). Στη στατιστική κατά Bayes έχουμε τις εκ των προτέρων κατανομές (prior distributions) και με τη βοήθεια αυτών προσπαθούμε να εκτιμήσουμε τις εκ των υστέρων κατανομές (posterior distributions). Επομένως με τη βοήθεια της μεθόδου MCMC μπορούμε να έχουμε τις εκ των προτέρων κατανομές και να παράγουμε τις εκ των υστέρων μετά από ένα πεπερασμένο αριθμό καταστάσεων. Η μέθοδος MCMC μπορεί να εφαρμοστεί σε λογισμικό με μπεϋζιανό περιβάλλον όπως το WinBUGS (Lunn et al., 2000) www.mrc-bsu.cam.ac.uk/bugs/.

Κατά τη μπεϋζιανή στατιστική το μοντέλο τυχαίων επιδράσεων δίνεται ως

$$y_i/\theta_i \sim N(\theta_i, \sigma_i^2)$$

$$\theta_i/\mu, \tau \sim N(\mu, \tau^2)$$

με τις τυχαίες μεταβλητές μ, τ να έχουν τις εκ των προτέρων κατανομές

$$\mu \sim p_1(\cdot)$$

$$\tau \sim p_2(\cdot)$$

(Chung et al., 2013).

Έχοντας ορίσει το μπεϋζιανό μοντέλο τυχαίων επιδράσεων στη μετα-ανάλυση μπορούμε να εξάγουμε εκτιμητές ετερογένειας με τη στατιστική κατά Bayes με τις πιο συνηθισμένες εκ των προτέρων κατανομές για το μ, τ . Η πιο συνηθισμένη κατανομή για το μ αποτελεί η κανονική κατανομή $\mu \sim N(0, 10^6)$, ενώ η εκ των προτέρων κατανομή για το τ (ή το τ^2) μπορεί να έχει επιδράσεις όταν ο αριθμός των μελετών είναι μικρός (Chung, et

al., 2013). Η πιο ευρέως χρησιμοποιημένη εκ των προτέρων κατανομή αποτελεί η αντίστροφη Γάμμα κατανομή *inverse – gamma*($10^{-3}, 10^{-3}$) για το τ^2 και η ομοιόμορφη *uniform*(0,100) για το τ (Chung et al., 2013). Μεταβαίνοντας από τις εκ των προτέρων κατανομές στις εκ των υστέρων κατανομές, οι εκτιμητές για το μ, τ θα είναι οι εκ των υστέρων διάμεσοι (posterior median) για το μ, τ αντίστοιχα, δημιουργώντας έτσι τον εκτιμητή της συνολικής επίδρασης και τον εκτιμητή ετερογένειας που ονομάζεται Full Bayesian (FB). Επομένως, γίνεται αντιληπτό ότι ο εκτιμητής FB είναι ένας μη αρνητικός και επαναληπτικός εκτιμητής.

4.9.4 Εκτιμητής Bayes Modal (BM)

Η μέθοδος Bayes Modal πρόκειται για μια εκ των υστέρων μέθοδο εκτίμησης κατά την οποία έχοντας καθορίσει τις εκ των προτέρων κατανομές μεγιστοποιούμε τις εκ των υστέρων κατανομές για να αποφύγουμε εκτιμήσεις στα όρια σ' ένα λογαριθμικό γραμμικό μοντέλο (Chung et al., 2013). Σε αντίθεση με την εκ των υστέρων διάμεσο η μέθοδος Bayes Modal δεν απαιτεί προσομοίωση για την εύρεση του εκτιμητή αλλά υπολογίζεται έπειτα από μεγιστοποίηση της εκ των υστέρων λογαριθμικής συνάρτησης πιθανοφάνειας του μ, τ (Chung et al., 2013). Θεωρώντας την $\text{Gamma}(\beta, \lambda)$ ως την εκ των προτέρων κατανομή του τ και την ομοιόμορφη κατανομή ως την εκ των προτέρων κατανομή για το μ , η εκ των υστέρων λογαριθμική συνάρτηση πιθανοφάνειας ορίζεται

$$p(\mu, \tau/y) = l(\mu, \tau^2) + (\alpha - 1)\log\tau - \lambda\tau + c_1$$

όπου $l(\mu, \tau^2)$ η συνάρτηση πιθανοφάνειας του μ, τ^2 που δίνεται από την εξίσωση (*). Μεγιστοποιώντας με χρήση της μεθόδου μέγιστης πιθανοφάνειας (ML) και κάνοντας χρήση του αναπτύγματος Taylor προκύπτει ο Bayes Modal εκτιμητής (BM) ως

$$\hat{\tau}_{BM}^2 = \begin{cases} (\beta - 1)\text{Var}(\hat{\tau}_{ML}) & , \text{αν } \hat{\tau}_{ML} = 0 \\ \left(\frac{\hat{\tau}_{ML}}{2} + \frac{\hat{\tau}_{ML}}{2} \sqrt{1 + \frac{4(\beta - 1)\text{Var}(\hat{\tau}_{ML})}{\hat{\tau}_{ML}^2}} \right)^2 & , \text{αν } \hat{\tau}_{ML} > 0 \end{cases}$$

(Chung et al., 2013).

4.10 Εκτιμητής Malzahn, Böhning, and Holling (MBH)

Οι Malzahn, Böhning, & Holling (2000) προτείνουν το μη παραμετρικό εκτιμητή της ετερογένειας Malzahn, Böhning, and Holling (MBH) ο οποίος μπορεί να χρησιμοποιηθεί μόνο για τα δεδομένα με συνεχή αποτελέσματα καθώς έχει δημιουργηθεί για να χρησιμοποιείται μόνο με το μέγεθος επίδρασης της σταθμισμένης μέσης διαφοράς (standardized mean difference). Ο εκτιμητής Malzahn, Böhning, and Holling δίνεται από τη σχέση

$$\hat{\tau}_{MBH}^2 = \frac{\sum_{i=1}^k (1 - \varphi_i)(y_i - \hat{\mu}_{FE})^2}{k - 1} - \frac{1}{k} \sum_{i=1}^k \left(\frac{n_i}{n_{it}n_{ic}} \right) - \frac{1}{k} \sum_{i=1}^k \varphi_i y_i^2$$

όπου $n_i = n_{ic} + n_{it}$ αποτελεί το συνολικό μέγεθος για την i -ιοστή μελέτη, $\hat{\mu}_{FE}$ η εκτιμώμενη συνολική επίδραση για το μοντέλο σταθερών επιδράσεων, φ_i δίνεται από τη σχέση

$$\varphi_i = 1 - \frac{n_i - 4}{J^2(n_i - 2)}$$

και J να αποτελεί τη διόρθωση για τη δημιουργία αμερόληπτων εκτιμητών με τον τύπο

$$J = 1 - \frac{3}{4(n_{ic} + n_{it}) - 9}$$

Από τον ορισμό του συμπεραίνουμε ότι πρόκειται για ένα μη επαναληπτικό και μη αρνητικό εκτιμητή.

4.11 Διαστήματα εμπιστοσύνης για την συνολική επίδραση μ

4.11.1 Wald-type (Wt) διαστήματα εμπιστοσύνης

Η πιο ευρέως χρησιμοποιημένη τεχνική για τη δημιουργία διαστημάτων εμπιστοσύνης για τη συνολική επίδραση αποτελούν τα Wald-type διαστήματα (Wt CI) (DerSimonian & Laird, 1986). Έχοντας το μοντέλο τυχαίων επιδράσεων, τα βάρη τυχαίων επιδράσεων είναι $w_{i,RE} = \frac{1}{\hat{\sigma}_i^2 + \hat{\tau}^2}$

για τη συνολική επίδραση $\hat{\mu}_{RE} = \frac{\sum_{i=1}^k w_{i,RE} y_i}{\sum_{i=1}^k w_{i,RE}}$ και η διακύμανση της

συνολικής επίδρασης όπως αναφέραμε στην παράγραφο 3.4 δίνεται ως $V_M = \frac{1}{\sum_{i=1}^k w_{i,RE}}$. Υποθέτοντας ότι η συνολική επίδραση ασυμπτωτικά κατανέμεται κανονικά μπορούμε να εξάγουμε το $(1 - \alpha)\%$ διάστημα εμπιστοσύνης με βάση την κανονική κατανομή

$$\hat{\mu}_{RE} \pm z_{\alpha/2} \sqrt{V_M}$$

(Sánchez-Meca & Marín-Martínez, 2008).

4.11.2 Knapp and Hartung (KH) διαστήματα εμπιστοσύνης

Σύμφωνα με τον Hartung (1999) η δημιουργία διαστημάτων εμπιστοσύνης για τη συνολική επίδραση μπορεί να δημιουργηθεί από τη t κατανομή με $k - 1$ βαθμούς ελευθερίας. Η προσεγγιστική κατανομή του S στατιστικού είναι

$$S = \frac{\hat{\mu}_{RE} - \mu}{\sqrt{Var(\hat{\mu}_{RE})}} \sim t_{k-1}$$

με $Var(\hat{\mu}_{RE}) = Q_{gen} \frac{1}{(k-1) \sum_{i=1}^k w_{i,RE}}$ όπου Q_{gen} το Generalised Q -στατιστικό (Hartung, 1999). Με βάση τα παραπάνω το $(1 - \alpha)\%$ διάστημα εμπιστοσύνης για τη συνολική επίδραση είναι (Hartung, 1999):

$$\hat{\mu}_{RE} \pm t_{k-1, \alpha/2} \sqrt{Var(\hat{\mu}_{RE})}$$

4.12 Διαστήματα εμπιστοσύνης για την ετερογένεια με τη μέθοδο non-parametric bootstrap

Στην παράγραφο 4.7 παρουσιάστηκε ο εκτιμητής non-parametric bootstrap DLb ο οποίος προκύπτει ως η μέση τιμή των B εκτιμητών $\hat{t}_b^2, b = 1, 2, \dots, B$ που δημιουργούνται από τη μη παραμετρική μέθοδο bootstrap (Kontopantelis et al., 2013). Έχοντας τους παραπάνω B εκτιμητές το $(1 - \alpha)\%$ διάστημα εμπιστοσύνης προκύπτει από τα $\left(\frac{\alpha}{2}\right)\%$ και $\left(1 - \frac{\alpha}{2}\right)\%$ ποσοστιαία σημεία των εκτιμητών $\hat{t}_b^2, b = 1, 2, \dots, B$ (Switzer, Paese, & Drasgow, 1992). Το πλεονέκτημα για τη δημιουργία ενός non-parametric bootstrap διαστήματος εμπιστοσύνης είναι ότι δεν χρειάζεται να στηριχθούμε στην υπόθεση της ύπαρξης κάποιας κατανομής

και μπορούμε να δημιουργήσουμε διαστήματα εμπιστοσύνης για οποιονδήποτε μη αρνητικό εκτιμητή (Switzer et al., 1992). Σύμφωνα με τον Viechtbauer (2007b) τα μη παραμετρικά διαστήματα εμπιστοσύνης δημιουργούν χαμηλές πιθανότητες κάλυψης.

4.13 Συγκεντρωτικοί πίνακες εκτιμητών ετερογένειας

Εκτιμητής	Συμβολισμός	Επαναληπτικός/Μη επαναληπτικός	Θετικός/Μη αρνητικός
DerSimonian and Laird	DL	Μη επαναληπτικός	Μη αρνητικός
General Hedges-Olkin	GHO	Μη επαναληπτικός	Μη αρνητικός
Paule-Mandel	PM	Επαναληπτικός	Μη αρνητικός
Positive DerSimonian and Laird	DLp	Μη επαναληπτικός	Θετικός
Two-step DerSimonian and Laird	DL2	Μη επαναληπτικός	Μη αρνητικός
Two-step General Hedges-Olkin	GHO2	Μη επαναληπτικός	Μη αρνητικός
Hartung-Makambi	HM	Μη επαναληπτικός	Θετικός
Hunter-Schmidt	HS	Μη επαναληπτικός	Μη αρνητικός

Πίνακας 4.1: Εκτιμητές από τη γενική μέθοδο των ροπών

Εκτιμητής	Συμβολισμός	Επαναληπτικός/μη επαναληπτικός	Θετικός/Μη αρνητικός
Maximum likelihood	ML	Επαναληπτικός	Μη αρνητικός
Restricted maximum likelihood	REML	Επαναληπτικός	Μη αρνητικός
Approximate restricted maximum likelihood	AREML	Επαναληπτικός	Μη αρνητικός

Πίνακας 4.2: Εκτιμητές από τη μέθοδο μέγιστης πιθανοφάνειας

Εκτιμητής	Συμβολισμός	Επαναληπτικός/μη επαναληπτικός	Θετικός/Μη αρνητικός
Sidik-Jonkman	SJ	Μη επαναληπτικός	Θετικός
Sidik-Jonkman με αρχικό τον General Hedges-Olkin	SJgho	Μη επαναληπτικός	Θετικός

Πίνακας 4.3: Εκτιμητές από τη μέθοδο σταθμισμένων ελαχίστων τετραγώνων

Εκτιμητής	Συμβολισμός	Επαναληπτικός/μη επαναληπτικός	Θετικός/Μη αρνητικός
Non-Parametric bootstrap DerSimonian and Laird	DLb	Μη επαναληπτικός	Μη αρνητικός

Πίνακας 4.4: Εκτιμητές από τη μέθοδο Non-Parametric bootstrap

Εκτιμητής	Συμβολισμός	Επαναληπτικός/μη επαναληπτικός	Θετικός/Μη αρνητικός
Empirical Bayes	EB	Επαναληπτικός	Μη αρνητικός

Πίνακας 4.5: Εκτιμητής Empirical Bayes

Εκτιμητής	Συμβολισμός	Επαναληπτικός/μη επαναληπτικός	Θετικός/Μη αρνητικός
Rukhin Bayes	RB	Μη επαναληπτικός	Μη αρνητικός
Positive Rukhin Bayes	RBp	Μη επαναληπτικός	Θετικός
Full Bayes	FB	Επαναληπτικός	Μη αρνητικός
Bayes Modal	BM	Επαναληπτικός	Θετικός

Πίνακας 4.6: Εκτιμητές Bayes

Εκτιμητής	Συμβολισμός	Επαναληπτικός/μη επαναληπτικός	Θετικός/Μη αρνητικός
Malzahn, Böhning and Holling	MBH	Μη επαναληπτικός	Μη αρνητικός

Πίνακας 4.7: Μη παραμετρικός εκτιμητής

4.14 Συμπεράσματα από μελέτες προσομοίωσης

Για την αξιολόγηση της καταλληλότητας ενός εκτιμητή έχουν πραγματοποιηθεί μελέτες προσομοίωσης με στόχο την εξέταση της ικανοποίησης των ιδιοτήτων ενός καλού εκτιμητή. Η συμπεριφορά του εκτιμητή ως προς τις ιδιότητες του καθορίζει την ποιότητα του. Έχοντας παρουσιάσει από μαθηματική σκοπιά τους εκτιμητές θα ολοκληρώσουμε με τα συμπεράσματα των μελετών προσομοίωσης για κάθε εκτιμητή ετερογένειας.

4.14.1 Εκτιμητής DL

Ο εκτιμητής DL είναι μη αρνητικός, δηλαδή απαιτεί την αποκοπή στο

μηδέν για αρνητικές τιμές. Η αποκοπή αυτή προκαλεί θετική μεροληψία οδηγώντας επομένως σε αυξημένη εκτίμηση της ετερογένειας για τον εκτιμητή DL (Viechtbauer, 2005). Έπειτα από μελέτες προσομοίωσης έχει παρατηρηθεί ότι ο εκτιμητής έχει καλή συμπεριφορά όταν η ετερογένεια είναι μηδέν ή μικρή και ο αριθμός των μελετών μεγάλος, ενώ μειωμένη μεροληψία παρουσιάζεται στον εκτιμητή DL κατά την ύπαρξη μεγάλης ετερογένειας (Viechtbauer, 2005). Η συμπεριφορά αυτή της υποεκτίμησης μιας αυξημένης ετερογένειας οδηγεί σε χαμηλό τυπικό σφάλμα I και σε χαμηλή πιθανότητα κάλυψης των διαστημάτων εμπιστοσύνης (Sánchez-Meca & Marín-Martínez, 2008). Ως προς την αποδοτικότητα, όταν η ετερογένεια δεν είναι πάρα πολύ μεγάλη ο εκτιμητής DL έχει χαμηλότερο μέσο τετραγωνικό σφάλμα (MSE), είναι δηλαδή πιο αποδοτικός σε σύγκριση με τους εκτιμητές GHO, SJ και PM (Sidik & Jonkman, 2007). Σύμφωνα με τους Jackson et al. (2010) ο εκτιμητής DL είναι μη αποδοτικός όταν το μέγεθος των μελετών σε μία μετα-ανάλυση είναι διαφορετικού μεγέθους και όταν υπάρχει μεγάλη ετερογένεια, ενώ είναι αποδοτικός όταν το συνολικό μέγεθος του δείγματος είναι μεγάλο.

4.14.2 Εκτιμητής GHO

Γνωρίζοντας ότι ο εκτιμητής GHO είναι μη αρνητικός και υποθέτοντας ότι οι πληθυσμιακές διακυμάνσεις είναι αμερόληπτοι εκτιμητές, προκύπτει ότι ο εκτιμητής GHO είναι αμερόληπτος πριν την περικοπή (Viechtbauer, 2007a). Οι DerSimonian Laird (1986) συμπέραναν ότι ο εκτιμητής GHO αποδίδει υψηλότερες τιμές κατά την εκτίμηση της ετερογένειας από τους εκτιμητές DL και REML. Έπειτα ακολουθεί ο εκτιμητής ML με λίγο χαμηλότερες τιμές από τους δύο παραπάνω (DerSimonian & Laird, 1986). Ο εκτιμητής GHO έχει καλή συμπεριφορά στην παρουσία μεγάλης ετερογένειας και μεγάλου αριθμού μελετών ενώ παράλληλα είναι μη αποδοτικός παρουσιάζοντας μεγάλο MSE (Sidik & Jonkman, 2007; Viechtbauer, 2005). Κατά τους Sidik Jonkman (2007) για την ύπαρξη μεγάλης ετερογένειας και για μεσαίο προς μεγάλο αριθμό μελετών ο εκτιμητής GHO έχει μεγαλύτερο MSE συγκριτικά με τους DL, ML, REML, PM και SJ αλλά λιγότερη μεροληψία από τους DL, ML, REML για μεγάλο αριθμό μελετών. Ακόμα, κατά τον Friedman (2000) ο

εκτιμητής GHO είναι πιο αποδοτικός από τον DL στην παρουσία μεγάλης ετερογένειας ενώ το αντίθετο συμβαίνει στην απουσία ετερογένειας. Οι Sidik Jonkman (2007) υποστήριξαν τα αντίθετα συμπεράσματα από αυτά του Friedman (2000) αποδίδοντας ότι η διαφορά αυτή οφείλεται στο γεγονός ότι ο Friedman δεν συμπεριέλαβε την αποκοπή λόγω της μη αρνητικότητας των εκτιμητών και ότι συμπέρανε αποτελέσματα από τις διακυμάνσεις των εκτιμητών.

4.14.3 Εκτιμητής PM

Έχει αποδειχθεί ότι ο εκτιμητής PM έχει θετική μεροληψία για μικρές τιμές ετερογένειας και μικρό αριθμό μελετών. Ενώ, για μεγάλες τιμές ετερογένειας και αριθμό μελετών έχει αρνητική μεροληψία (Sidik & Jonkman, 2007). Σύμφωνα με μία μελέτη προσομοίωσης, ο εκτιμητής PM συμπεριφέρεται καλύτερα ως προς τη μεροληψία από τους εκτιμητές DL και REML (Panityakul, Bumrungsup, & Knapp, 2013). Οι Novianti et al. (2014) συγκρίνοντας περισσότερους εκτιμητές ως προς τη μεροληψία εξήγαγαν ότι ο εκτιμητής PM είναι καλύτερος από άποψη μεροληψίας από τους εκτιμητές DL, DL2, GHO, REML και SJ. Από τις παραπάνω μελέτες γίνεται αντιληπτό ότι ο εκτιμητής PM έχει καλή συμπεριφορά ως προς τη μεροληψία. Ως προς τη αποδοτικότητα, σύμφωνα με τους Knapp Hartung (2003) ο εκτιμητής PM είναι λιγότερο αποδοτικός από τους εκτιμητές DL και REML, ενώ έχει παρόμοιο MSE με τον εκτιμητή SJ κατά τους Sidik & Jonkman (2007). Ακόμα συγκρίνοντας τον εκτιμητή PM με τον DL, οι Bowden et al. (2011) παρατήρησαν ότι όσο αυξάνεται η ετερογένεια τόσο ο εκτιμητής PM γίνεται καλύτερος από τον DL προτείνοντας τη χρήση του εκτιμητή PM καθώς έχει καλή συμπεριφορά ως προς τις ιδιότητες ενός καλού εκτιμητή.

4.14.4 Εκτιμητής DLp

Σύμφωνα με τους Kontopantelis et al. (2013) ο εκτιμητής DLp, ο οποίος είναι θετικός εκτιμητής, έχει χαμηλότερη μεροληψία από τους άλλους θετικούς εκτιμητές SJ, GHO και RBp.

4.14.5 Εκτιμητής DL2

Ο εκτιμητής DL2 προσεγγίζει καλύτερα τον εκτιμητή PM, ο οποίος έχει καλές ιδιότητες σε σύγκριση με τους εκτιμητές DL, GHO και GHO2 (DerSimonian & Kacker, 2007). Ακόμα κατά τους Bhaumik et al (2012) ο εκτιμητής DL2 παρουσιάζει χαμηλή μεροληψία για σπάνια γεγονότα.

4.14.6 Εκτιμητής GHO2

Οι DerSimonian Kacker (2007) επισημαίνουν ότι οι εκτιμητές δύο βημάτων DL2 και GHO2 προσεγγίζουν καλύτερα σε σύγκριση με τους εκτιμητές ενός βήματος DL και GHO τον εκτιμητή PM.

4.14.7 Εκτιμητής HM

Οι Thorlund et al. (2011) συγκρίναν τον εκτιμητή HM με τους εκτιμητές REML, GHO, DL και SJ. Σύμφωνα με την εμπειρική μελέτη τους οι εκτιμητές HM και SJ υπερεκτιμούν την ετερογένεια όταν αυτή κυμαίνεται από μέτρια προς μεγάλη.

4.14.8 Εκτιμητής HS

Οι συγγραφείς δεν προτείνουν τον εκτιμητή HS γιατί έχει σημαντική αρνητική μεροληψία (Viechtbauer, 2005). Ως προς την αποδοτικότητα, ο εκτιμητής HS παρουσιάζει παρόμοιο MSE με τον εκτιμητή ML και χαμηλότερο MSE από τους DL, REML και GHO (Viechtbauer, 2005).

4.14.9 Εκτιμητής ML

Σύμφωνα με τους Panityakul et al (2013) ο εκτιμητής ML έχει σημαντική αρνητική μεροληψία. Σε σύγκριση του εκτιμητή ML με τους εκτιμητές REML, SJ, GHO και PM, ο ML έχει αποδειχθεί ότι παρουσιάζει μεγαλύτερο ποσοστό μεροληψίας ενώ παράλληλα είναι πιο αποδοτικός (Chung, Rabe-Hesketh, & Choi, 2013; Sidik & Jonkman, 2007; Thompson & Sharp, 1999). Το χαμηλό MSE του εκτιμητή ML επιβεβαιώθηκε και από μια άλλη μελέτη συγκρίνοντας τον εκτιμητή με τους DL και REML, ενώ σε σύγκριση με τον εκτιμητή HS παρουσιάζει τις ίδιες τιμές MSE (Viechtbauer, 2005). Κατά την ίδια μελέτη προτείνεται ότι η χρήση του εκτιμητή πρέπει να αποφεύγεται λόγω του μεγάλου ποσοστού της μερολη-

ψίας παρόλο του χαμηλού MSE (Viechtbauer, 2005).

4.14.10 Εκτιμητής REML

Σε μελέτες προσομοίωσης έχει παρατηρηθεί ότι ο εκτιμητής REML υποεκτιμά την ετερογένεια για κάποιες περιπτώσεις (Sidik & Jonkman, 2005, 2007; Viechtbauer, 2005). Από μελέτη προσομοίωσης των Sidik Jonkman (2005) έχουμε το συμπέρασμα ότι ο εκτιμητής REML έχει χαμηλότερη μεροληψία από τον εκτιμητή SJ για μεσαίες προς μεγάλες τιμές της ετερογένειας και παρόμοιο MSE. Κατά τη μελέτη του Viechtbauer (2005) συγκρίνοντας τον εκτιμητή REML με τους εκτιμητές DL, GHO, ML και HS, απέδειξε ότι ο REML μαζί με τον DL έχουν μικρότερο MSE από τον GHO ενώ μεγαλύτερο από τους ML και HS. Σύμφωνα με τον Viechtbauer (2005) ο REML είναι ο καλύτερος εκτιμητής για μεγάλες μελέτες.

4.14.11 Εκτιμητής AREML

Όπως αναφέρθηκε και στην παράγραφο 4.5.3 ο εκτιμητής AREML αποδίδει ακριβώς τις ίδιες τιμές με τον εκτιμητή REML όταν οι πληθυσμιακές διακυμάνσεις είναι ίσες (Viechtbauer, 2005). Κατά την πράξη οι πληθυσμιακές διακυμάνσεις είναι σπάνιο να είναι ίσες και για το λόγο αυτό ο εκτιμητής AREML δεν χρησιμοποιείται (Viechtbauer, 2005). Γίνεται λοιπόν αντιληπτό ότι δεν υπάρχει νόημα να χρησιμοποιηθεί ο εκτιμητής AREML σε κάποια μελέτη προσομοίωσης.

4.14.12 Εκτιμητής SJ

Μελέτες προσομοίωσης προτείνουν ότι ο εκτιμητής SJ έχει χαμηλότερο MSE και σημαντική μικρότερη μεροληψία από τον εκτιμητή DL για μεγάλες τιμές ετερογένειας και μεγάλο αριθμό μελετών, ενώ όσο ο αριθμός μελετών και η ετερογένεια μειώνονται η συμπεριφορά των εκτιμητών αντιστρέφεται (Sidik & Jonkman, 2005). Το παραπάνω συμπέρασμα επιβεβαιώθηκε και από άλλη μελέτη σημειώνοντας ότι την ίδια συμπεριφορά με τον DL σε σύγκριση με τον SJ έχουν και οι εκτιμητές ML, REML, GHO και PM (Panityakul et al., 2013; Sidik & Jonkman, 2007). Οι Sidik Jonkman (2007) συνιστούν τη χρήση του εκτιμητή PM ή

του εκτιμητή SJ καθώς για μεγάλη ετερογένεια είναι οι καλύτεροι ως προς τη μεροληψία. Σε αντίθεση με τους Sidik Jonkman (2007) έρχεται μια εμπειρική μελέτη κατά την οποία ο εκτιμητής SJ αποδείχθηκε να έχει μεγαλύτερη μεροληψία από τη μέθοδο DL (Thorlund et al., 2011).

4.14.13 Εκτιμητής SJgho

Κατά τους Sidik Jonkman (2005) ο εκτιμητής SJgho μαζί με τους εκτιμητές PM και SJ έχουν παρόμοια MSE και χαμηλότερη μεροληψία από τους εκτιμητές DL, GHO, ML και REML.

4.14.14 Εκτιμητής DLb

Σε μελέτη που πραγματοποιήθηκε από τους Kontopantelis et al. (2013) αποδείχθηκε ότι ο εκτιμητής DLb έχει την υψηλότερη μεροληψία για $k = 2$ αλλά αποδεκτή για αριθμό μελετών μεγαλύτερο ή ίσο του πέντε συγκριτικά με υπόλοιπους εκτιμητές όπως DL, RBo, RBp, DLp, GHO, GHO2, REML και SJ αναφέροντας τον από τους καλύτερους εκτιμητές.

4.14.15 Εκτιμητής EB

Όπως αναφέρθηκε και στην παράγραφο 4.8 ο εκτιμητής EB έχει ακριβώς τις ίδιες τιμές με τον εκτιμητή PM (Paule & Mandel, 1982). Για το λόγο αυτό η συμπεριφορά του εκτιμητή EB θα είναι ακριβώς η ίδια με τη συμπεριφορά και τα αποτελέσματα των μελετών προσομοίωσης που εξάγονται για τον PM. Συνεπώς τα συμπεράσματα που διατυπώνονται στην παράγραφο 4.14.3 ισχύουν και για τον εκτιμητή EB.

4.14.16 Εκτιμητής RBo

Σε μελέτη προσομοίωσης που συμπεριλήφθηκε ο εκτιμητής RBo αποδείχθηκε ότι για αριθμό μελετών μικρότερο από πέντε ο RBo έχει μικρότερη μεροληψία από τους εκτιμητές DL, DLp, DLb, DL2, GHO, RBp, GHO2, REML και SJ (Kontopantelis et al., 2013).

4.14.17 Εκτιμητής RBp

Σε μελέτη προσομοίωσης από τους Kontopantelis et al. (2013) η μεροληψία του εκτιμητή RBp ήταν η υψηλότερη σε σύγκριση με τους

εκτιμητές DL, DLp, DLb, DL2, GHO, RBp, RBo, GHO2, REML και SJ για μεγάλο αριθμό μελετών.

4.14.18 Εκτιμητής FB

Σε μελέτη προσομοίωσης που πραγματοποίησαν οι Lambert et al (2005) σύγκριναν 13 εκ των προτέρων κατανομές για το τ (ή το τ^2) για τον εκτιμητή FB, η *Inverse – Gamma*($10^{-3}, 10^{-3}$) για το τ και η *Uniform*(0,100) για το τ^2 είχαν την καλύτερη συμπεριφορά ως προς τη μεροληψία, πιθανότητες κάλυψης και σύγκλισης από τους. Κατά τους Chung et al. (2013) συγκρίνοντας τις δύο παραπάνω κατανομές για τον εκτιμητή FB με τον εκτιμητή EB συμπεράναν ότι η εκ των προτέρων *Uniform*(0,100) για το τ δημιουργεί τη μεγαλύτερη μεροληψία και το μεγαλύτερο ποσό αποδοτικότητας, παράλληλα η εκ των προτέρων για το τ^2 , *Inverse – Gamma*($10^{-3}, 10^{-3}$) έχει μικρότερη μεροληψία από τον εκτιμητή BM για μικρή ετερογένεια αλλά μεγαλύτερη αρνητική μεροληψία από τον BM για μεγάλη ετερογένεια.

4.14.19 Εκτιμητής BM

Ως προς τη μεροληψία ο εκτιμητής BM έχει παρόμοια συμπεριφορά με τους DL και REML ενώ παρουσιάζει καλύτερη συμπεριφορά από τον ML (Chung et al., 2013). Σύμφωνα με τους Chung et al. (2013) όταν η ετερογένεια είναι μηδέν και ειδικά για μικρό αριθμό μελετών και μικρό μέγεθος μελετών ο εκτιμητής BM έχει μεγαλύτερη μεροληψία από τους DL, ML, REML και GHO. Ακόμα κατά την απουσία ετερογένειας, ο BM έχει μεγαλύτερο MSE από τους DL, ML και REML και μικρότερο από τον GHO (Chung et al., 2013).

4.14.20 Εκτιμητής MBH

Έχοντας τον ίδιο αριθμό στην πειραματική ομάδα και στην ομάδα ελέγχου οι Malzahn et al. (2000) σύγκριναν τον εκτιμητή MBH με τον εκτιμητή GHO και απέδειξαν ότι το MSE του GHO είναι πάντα μεγαλύτερο από το MSE του MBH για όλα τα k , ενώ στις περισσότερες περιπτώσεις ο εκτιμητής GHO έχει μικρότερη απόλυτη μεροληψία από τον εκτιμητή MBH. Ακόμα κατά τους Malzahn et al. (2000) ο εκτιμητής DL έχει

μικρότερη μεροληψία συγκριτικά με τους εκτιμητές GHO και MBH, ενώ με την αύξηση της ετερογένειας, η μεροληψία αυξάνεται δραματικά.

ΚΕΦΑΛΑΙΟ 5

ΠΡΟΣΟΜΟΙΩΣΗ

5.1 Σκοπός της έρευνας προσομοίωσης

Στο Κεφάλαιο 3 παρουσιάστηκε ότι η ετερογένεια μπορεί να είναι κλινική, μεθοδολογική ή στατιστική και εκφράζει σε κάθε περίπτωση ένα μέτρο διαφορετικότητας μεταξύ των μελετών. Καθώς κάθε μελέτη μπορεί να έχει διαφορετικά χαρακτηριστικά όπως μεθόδους διεξαγωγής ή χαρακτηριστικά συμμετεχόντων, η ετερογένεια είναι αναπόφευκτη. Κατά την εφαρμογή της μετα-ανάλυσης δεν μπορούμε να έχουμε επίγνωση για την ποσοτικοποίηση αυτής της διαφορετικότητας-ετερογένειας των μελετών και συνεπώς καθίσταται η ανάγκη για την εκτίμηση της. Όσο πιο ακριβή εκτίμηση κάνουμε τόσο πιο αξιόπιστα και έγκυρα συμπεράσματα θα εξάγουμε. Επομένως στη μετα-ανάλυση σημαντικό ρόλο παρουσιάζει η χρήση ενός καλού εκτιμητή ετερογένειας. Ως καλό εκτιμητή ετερογένειας θεωρούμε έναν εκτιμητή που να συμπεριφέρεται όσο το δυνατόν περισσότερο με τις ιδιότητες ενός καλού εκτιμητή όπως αυτές παρουσιάστηκαν στην παράγραφο 4.3. Η ακριβής εκτίμηση θα συμβάλει στον υπολογισμό της συνολικής επίδρασης και κατά αποδοτικότητα στη δημιουργία πιο σωστών αποτελεσμάτων. Για το λόγο αυτό καθίσταται επιτακτική η ανάγκη για τη σύγκριση των εκτιμητών ετερογένειας ως προς τις ιδιότητες τους με στόχο την εύρεση των εκτιμητών ετερογένειας που έχουν τις καλύτερες ιδιότητες. Επομένως οι εκτιμητές ετερογένειας με τις καλύτερες ιδιότητες θα αποτελούν τους πιο κατάλληλους εκτιμητές ετερογένειας για τη σωστή εφαρμογή του μοντέλου της μετα-ανάλυσης. Στη βιβλιογραφία υπάρχουν κάποιες εμπειρικές μελέτες οι οποίες έχουν συγκρίνει μερικούς από τους εκτιμητές ετερογένειας διεξάγοντας μεμονωμένα συμπεράσματα από τις συγκρίσεις αυτές. Επείγει όμως η σύγκριση όλων των εκτιμητών ετερογένειας που υπάρχουν στη μετα-ανάλυση για την εύρεση μέσω μιας μελέτης προσομοίωσης του καλύτερου ή των καλύτερων από τους εκτιμητές ετερογένειας με στόχο την παράθεση

μιας ολοκληρωμένης γνώσης για τη συμπεριφορά και την καταλληλότητα όλων των εκτιμητών ετερογένειας. Με κίνητρο αυτή την ιδέα η παρούσα προσομοίωση αποτελεί μια καινοτομία για την ήδη υπάρχουσα βιβλιογραφία που στοχεύει στην εύρεση του καλύτερου εκτιμητή ή των πρώτων καλύτερων με στόχο την αύξηση της ποιότητας των μελλοντικών μετα-αναλύσεων με τη χρήση των εκτιμητών αυτών.

5.2 Σενάρια προσομοίωσης

Η προσομοίωση πραγματοποιήθηκε στο στατιστικό πακέτο R (R Development Core Team, 2008) <http://www.r-project.org/> (r-project.org) το οποίο αποτελεί μια γλώσσα προγραμματισμού για την επίλυση κυρίως στατιστικών προβλημάτων. Συμπεριλάβαμε και τα δύο είδη αποτελεσμάτων που δημιουργούνται πιο συχνά σε μια κλινική μελέτη, τα διχότομα και τα συνεχή. Για τα διχότομα υπολογίστηκαν όλοι οι εκτιμητές ετερογένειας και για τα συνεχή υπολογίστηκαν οι ίδιοι εκτιμητές μαζί με τον εκτιμητή MBH, που χρησιμοποιείται μόνο για το μέγεθος επίδρασης της τυποποιημένης μέσης διαφοράς. Με στόχο τη λεπτομερή σύγκριση των ιδιοτήτων των εκτιμητών ετερογένειας συμπεριλάβαμε διάφορες τιμές για τον αριθμό των μελετών $k = 10, 20, 30, 50$ τιμές που αντικατοπτρίζουν μικρό, μεσαίο και μεγάλο αριθμό μελετών, διάφορες τιμές για την πραγματική συνολική επίδραση $\theta = 0, 0.3, 0.5, 0.8$, τα επίπεδα σημαντικότητας $\alpha = 0.01, 0.05, 0.10$. Τις τιμές για την ετερογένεια για τα διχότομα αποτελέσματα $\tau^2 = 0, 0.025, 0.07, 0.3$ καθώς και τις τιμές $\tau^2 = 0, 0.01, 0.05, 0.5$ για τα συνεχή αποτελέσματα. Για τον υπολογισμό της ισχύς χρησιμοποιήσαμε τις μηδενικές υποθέσεις $H_0: \theta_0 = 0.3, 0.5, 0.8$ και παρήγαμε τα δεδομένα από την εναλλακτική $H_a: \theta_a = 0$. Το κάθε σενάριο πραγματοποιήθηκε για $B = 1000$ αριθμό επαναλήψεων.

5.3 Μετρήσεις προσομοίωσης

Όπως αναφέρθηκε και στην παράγραφο 4.3 ένας εκτιμητής είναι καλός όταν είναι αμερόληπτος, έχει μικρό μέσο τετραγωνικό σφάλμα, ελαχιστοποιεί το σφάλμα τύπου I, δημιουργεί υψηλή ισχύ, στενά διαστήματα εμπιστοσύνης της συνολικής επίδρασης ή της ετερογένειας,

ενώ παράλληλα μεγάλη πιθανότητα κάλυψής τους. Με στόχο τη σύγκριση των εκτιμητών ετερογένειας χρησιμοποιήσαμε στην προσομοίωση ως μετρήσεις τα μέτρα των ιδιοτήτων των εκτιμητών ετερογένειας. Επομένως οι μετρήσεις που χρησιμοποιήθηκαν στην προσομοίωση είναι:

1) Μεροληψία (Bias)

Ένας καλός εκτιμητής πρέπει να είναι αμερόληπτος δηλαδή η αναμενόμενη τιμή του να είναι ίση με την πραγματική. Στην προσομοίωση χρησιμοποιήσαμε την απόλυτη μέση μεροληψία (average absolute bias) που αποτελεί ένα εμπειρικό μέτρο μεροληψίας και δίνεται ως $bias = \frac{1}{B} \sum_{i=1}^B |\hat{\tau}_i^2 - \tau^2|$ όπου $\hat{\tau}_i^2$ ο παρατηρούμενος εκτιμητής ετερογένειας την i φορά και τ^2 η πραγματική τιμή της ετερογένειας. Όσο πιο κοντά στο μηδέν είναι η απόλυτη μέση μεροληψία τόσο πιο καλός είναι ο εκτιμητής ετερογένειας.

2) Αποδοτικότητα (Efficiency)

Ένας καλός εκτιμητής πρέπει να είναι αποδοτικός. Έχοντας το μέσο τετραγωνικό σφάλμα ως μέτρο αποδοτικότητας, όσο πιο κοντά είναι το μέσο τετραγωνικό σφάλμα στο μηδέν τόσο πιο αποδοτικός είναι ο εκτιμητής. Έχοντας τους εκτιμητές ετερογένειας ως ένα διάνυσμα B διάστασης $\hat{\tau}^2 = (\hat{\tau}_1^2, \dots, \hat{\tau}_B^2)$ και την πραγματική τιμή τ^2 το μέσο τετραγωνικό σφάλμα ορίζεται ως $MSE(\hat{\tau}^2) = \frac{1}{B} \sum_{i=1}^B (\hat{\tau}_i^2 - \tau^2)^2$ (“Mean squared error”, n.d.).

3) Πιθανότητα κάλυψης για το μ (Coverage probability for μ)

Έχοντας $H_0: \mu = \mu_0$ με εναλλακτική την $H_a: \mu = \mu_a$ με μ_0 την πραγματική τιμή της συνολικής επίδρασης, παράγουμε για $i = 1, \dots, B$ την εκτιμώμενη τιμή της συνολικής επίδρασης $\hat{\mu}_i, i = 1, \dots, B$ και το κλειστό $(1 - \alpha)\%$ διάστημα εμπιστοσύνης της εκτιμώμενης συνολικής επίδρασης $[\hat{\mu}_{i,L}, \hat{\mu}_{i,U}]$, $i = 1, \dots, B$. Η πιθανότητα κάλυψης του μ είναι ο αριθμός των φορών που η πραγματική τιμή της συνολικής επίδρασης ανήκει στο $(1 - \alpha)\%$ διάστημα εμπιστοσύνης $[\hat{\mu}_{i,L}, \hat{\mu}_{i,U}]$ της i -οστής επανάληψης, προς τον αριθμό των επαναλήψεων.

$$coverage = \frac{\#(\hat{\mu}_{i,L} \leq \mu_0 \text{ και } \mu_0 \leq \hat{\mu}_{i,U})}{B}$$

4) Πιθανότητα κάλυψης για το τ^2 (Coverage probability for τ^2)

Έχοντας την πραγματική τιμή ετερογένειας τ^2 , παράγουμε για $i = 1, \dots, B$ τον εκτιμητή ετερογένειας $\hat{\tau}_i^2, i = 1, \dots, B$ και το $(1 - \alpha)\%$ διάστημα εμπιστοσύνης του εκτιμητή της ετερογένειας $[\hat{\tau}_{i,L}^2, \hat{\tau}_{i,U}^2], i = 1, \dots, B$. Η πιθανότητα κάλυψης του τ^2 είναι ο αριθμός των φορών που η πραγματική τιμή της ετερογένειας ανήκει στο $(1 - \alpha)\%$ διάστημα εμπιστοσύνης της i -ιστής επανάληψης $[\hat{\tau}_{i,L}^2, \hat{\tau}_{i,U}^2]$, προς τον αριθμό των επαναλήψεων.

$$coverage = \frac{\#(\hat{\tau}_{i,L}^2 \leq \tau^2 \text{ και } \tau^2 \leq \hat{\tau}_{i,U}^2)}{B}$$

5) Εμπειρικό σφάλμα τύπου I (type error I)

Έχοντας $H_0: \mu = \mu_0$ με εναλλακτική την $H_\alpha: \mu = \mu_1$, παράγουμε B φορές το $(1 - \alpha)\%$ διάστημα εμπιστοσύνης από το μ_0 . Θεωρώντας ως $[\hat{\mu}_{i,L}, \hat{\mu}_{i,U}], i = 1, \dots, B$ το κλειστό διάστημα εμπιστοσύνης του, ορίζουμε ως εμπειρικό σφάλμα τύπου I (empirical type error I) τον αριθμό των φορών που το μ_0 δεν ανήκει στο κλειστό διάστημα ως προς τον αριθμό των φορών B .

$$empirical \text{ type error } I = \frac{\#(\hat{\mu}_{i,L} > \mu_0 \text{ ή } \hat{\mu}_{i,U} < \mu_0)}{B}$$

6) Ισχύς (power)

Ισχύς είναι η πιθανότητα να απορρίπτουμε τη μηδενική υπόθεση όταν παράγουμε τα δεδομένα από την εναλλακτική. Θεωρούμε τη μηδενική υπόθεση $H_0: \mu = \mu_0$ με εναλλακτική την $H_\alpha: \mu = \mu_1$. Παράγοντας B φορές το $(1 - \alpha)\%$ διάστημα εμπιστοσύνης από το μ_1 και θεωρώντας ως $[\hat{\mu}_{i,L}, \hat{\mu}_{i,U}], i = 1, \dots, B$ το κλειστό διάστημα εμπιστοσύνης του, ορίζουμε την ισχύ ως τον αριθμό των φορών που το μ_0 δεν ανήκει στο κλειστό

³ Όπου # ο αριθμός των φορών

διάστημα ως προς τον αριθμό των φορών B , έχοντας παράγει τα δεδομένα με μ_1 .

$$power = \frac{\#(\hat{\mu}_{i,L} \geq \mu_0 \parallel \hat{\mu}_{i,U} \leq \mu_0)}{B}$$

5.4 Βήματα προσομοίωσης

Αρχικό βήμα για την προσομοίωση αποτελεί η δημιουργία των δεδομένων. Η δημιουργία των δεδομένων πρέπει να είναι έτσι ώστε να αναπαριστάνει τα αποτελέσματα μιας πραγματικής κλινικής μελέτης. Τα αποτελέσματα σε μια μελέτη μπορεί να είναι είτε διχότομα είτε συνεχή. Η δημιουργία των δεδομένων μας βασίστηκε στη μέθοδο που περιγράφουν οι Novianti et al. (2014) όπως παρατίθεται στη συνέχεια.

5.4.1 Δημιουργία δεδομένων

5.4.1.1 Διχότομα αποτελέσματα

Σε μια μελέτη ως διχότομο αποτέλεσμα θεωρούμε το γεγονός και το μη γεγονός. Χωρίζουμε το δείγμα στην πειραματική ομάδα (T) και στην ομάδα ελέγχου (C). Για την πειραματική ομάδα (T) θεωρούμε τον αριθμό των γεγονότων a_i , τον αριθμό των μη γεγονότων b_i και n_{it} το μέγεθος της ομάδας. Για την ομάδα ελέγχου (C) θεωρούμε τον αριθμό των γεγονότων c_i , τον αριθμό των μη γεγονότων d_i και n_{ic} το μέγεθος της ομάδας. Για το διχότομο αποτέλεσμα παράγουμε τα δεδομένα έτσι ώστε το μέγεθος επίδρασης να είναι ο λόγος αναλογιών OR. Θεωρώντας k τον αριθμό των μελετών που συμπεριλαμβάνονται σε μία μετα-ανάλυση ακολουθούμε την προσομοίωση των Novianti et al. (2014) με τα εξής βήματα:

- Παράγουμε μέγεθος επίδρασης $\theta_i, i = 1, 2, \dots, k$ για κάθε μελέτη γνωρίζοντας ότι $\theta_i \sim N(\theta, \tau^2)$.

Για τις k μελέτες:

- Προσομοιώνουμε τα μεγέθη επίδρασης $n_i = n_{ic} = n_{it}$ για την πειραματική ομάδα και την ομάδα ελέγχου θεωρώντας το μέγεθος τους ίσο με έναν τυχαίο ακέραιο αριθμό από το 20 ως το 200.
- Παράγουμε τον αριθμό των γεγονότων c_i για την ομάδα ελέγχου με από την διωνυμική κατανομή $Binomial(n_{ic}, p_{ic})$ με

$p_{ic} \sim Uniform(0.05, 0.65)$.

- Παράγουμε τον αριθμό των γεγονότων α_i για την πειραματική ομάδα από την διωνυμική κατανομή $Binomial(n_{it}, p_{it})$ με $p_{it} = \frac{p_{ic}e^{\theta_i}}{1-p_{ic}+p_{ic}e^{\theta_i}}$.
- Υπολογίζουμε τον αριθμό των μη γεγονότων $b_i = n_{it} - \alpha_i$, $d_i = n_{ic} - c_i$.
- Θέτουμε την τιμή 0.5 για οποιοδήποτε μηδενική τιμή του αριθμού των γεγονότων ή των μη γεγονότων.
- Υπολογίζουμε το μέγεθος επίδρασης $y_i = \log(OR)_i = \log\left\{\frac{\alpha_i*d_i}{b_i*c_i}\right\}$ και
- τον εκτιμητή της διακύμανσης $\hat{\sigma}_i^2 = \frac{1}{\alpha_i} + \frac{1}{b_i} + \frac{1}{c_i} + \frac{1}{d_i}$.

5.4.1.2 Συνεχή αποτελέσματα

Έχοντας την πειραματική ομάδα και την ομάδα ελέγχου για το συνεχές αποτέλεσμα προσομοιώνουμε τα δεδομένα έτσι ώστε το μέγεθος επίδρασης να είναι η τυποποιημένη μέση διαφορά Hedges' g.

Θεωρώντας k τον αριθμό των μελετών που συμπεριλαμβάνονται σε μία μετα-ανάλυση ακολουθούμε την προσομοίωση των Novianti et al. (2014) με τα εξής βήματα:

- Παράγουμε το πραγματικό μέγεθος επίδρασης $\theta_i, i = 1, 2, \dots, k$ για κάθε μελέτη γνωρίζοντας ότι $\theta_i \sim N(\theta, \tau^2)$.

Για τις k μελέτες:

- Παράγουμε τα μεγέθη επίδρασης $n_i = n_{ic} = n_{it}$ για την πειραματική ομάδα και την ομάδα ελέγχου θεωρώντας το μέγεθος τους ίσο με έναν τυχαίο ακέραιο αριθμό από το 20 ως το 200.
- Προσομοιώνουμε Z_{ic} παρατηρήσεις σε πλήθος n_{ic} από την κανονική $Z_{ic} \sim N(0, 1)$ για την ομάδα ελέγχου (control group).
- Προσομοιώνουμε Z_{it} παρατηρήσεις σε πλήθος n_{it} από την κανονική $Z_{it} \sim N(\theta_i, 1)$ για την πειραματική ομάδα (treatment group).
- Υπολογίζουμε τη διακύμανση που ορίζεται να είναι η τετραγωνική ρίζα της σταθμισμένης διακύμανσης (pooled

variance) $S_{ip} = \sqrt{\frac{(n_c-1)S_{ic}^2 + (n_t-1)S_{it}^2}{n_c+n_t-2}}$ με S_{ic}^2, S_{it}^2 όπως περιγράφονται παρακάτω.

- Υπολογίζουμε το μέγεθος επίδρασης $y_i = \frac{\bar{Z}_{it} - \bar{Z}_{ic}}{S_{ip}}, \bar{Z}_{it} = \sum_{j=1}^{n_{it}} \frac{Z_{itj}}{n_{it}}, \bar{Z}_{ic} = \sum_{j=1}^{n_{ic}} \frac{Z_{icj}}{n_{ic}}$.
- Υπολογίζουμε τη διακύμανση $s_i^2 = \frac{8+y_i^2}{4n_i}$ (όταν $n_{ic} = n_{it}$).
- Χρησιμοποιούμε τη διόρθωση $J_i = 1 - \frac{3}{8n_i-9}$ επειδή κατά την παραγωγή της εκτίμησης της συνολικής επίδρασης δημιουργείται μεροληψία.
- Υπολογίζουμε την αμερόληπτη εκτίμηση του μεγέθους επίδρασης είναι $J_i * y_i$ και
- τον εκτιμητή της διακύμανσης $\hat{\sigma}_i^2 = J_i^2 * s_i^2$.

Περιγραφή των διακυμάνσεων της πειραματικής ομάδας και της ομάδας ελέγχου

Θεωρούμε σε κάθε μελέτη $i = 1, 2, \dots, k$, n_{ic} άτομα και μια τιμή Z_{ic} για κάθε άτομο στην ομάδα ελέγχου. Επομένως έχουμε τις παρατηρήσεις Z_{icj} όπου $i = 1, 2, \dots, k$ η κάθε μελέτη και $j = 1, 2, \dots, n_{ic}$ τα άτομα της κάθε μελέτης για την ομάδα ελέγχου C . Ακόμα, θεωρούμε σε κάθε μελέτη $i = 1, 2, \dots, k$, n_{it} άτομα και μια τιμή Z_{it} για κάθε άτομο στην πειραματική ομάδα και τις παρατηρήσεις Z_{itj} οι παρατηρήσεις για την πειραματική ομάδα T με $Z_{itj} \sim N(0,1)$. Στην περίπτωση της i μελέτης για την πειραματική ομάδα θα έχουμε παρατηρήσεις $Z_{it1}, \dots, Z_{itn_{ic}}, n_{ic}$ σε αριθμό με $Z_{it1} \sim N(0,1), \dots, Z_{itn_{ic}} \sim N(0,1)$ και δειγματική διακύμανση

$$S_{it}^2 = \frac{1}{n_{it} - 1} \sum_{j=1}^{n_{it}} (Z_{itj} - \bar{Z}_{it})^2$$

με $\bar{Z}_{it} = \sum_{j=1}^{n_{it}} \frac{Z_{itj}}{n_{it}}$.

Ανάλογα στην περίπτωση της i -οστής μελέτης για την ομάδα ελέγχου θα έχουμε

$Z_{ic1} \sim N(\theta_i, 1), \dots, Z_{icn_{ic}} \sim N(\theta_i, 1)$ και δειγματική διακύμανση

$$S_{ic}^2 = \frac{1}{n_{ic} - 1} \sum_{j=1}^{n_{ic}} (Z_{icj} - \bar{Z}_{ic})^2$$

$$\text{με } \bar{Z}_{ic} = \sum_{j=1}^{n_{ic}} \frac{Z_{icj}}{n_{ic}}.$$

Έχοντας τις δειγματικές διακυμάνσεις για την i -ιστή μελέτη σε κάθε περίπτωση η διακύμανση για τις k μελέτες θα είναι η τετραγωνική ρίζα της σταθμισμένης διακύμανσης (pooled variance)

$$S_{ip} = \sqrt{\frac{(n_{ic} - 1)S_{ic}^2 + (n_{it} - 1)S_{it}^2}{n_{ic} + n_{it} - 2}}$$

Η παραγωγή των διχότομων και συνεχών αποτελεσμάτων πραγματοποιήθηκε σε ξεχωριστούς κώδικες που παρατίθεται στο παράρτημα. Για το κάθε σενάριο εκτελώντας $B = 1000$ αριθμό επαναλήψεων παράγουμε τα δεδομένα και μέσω της βιβλιοθήκης MASS (<https://cran.r-project.org/web/packages/MASS/MASS.pdf>) της R αποθηκεύονται αυτόματα σε αρχείο. Έπειτα τα δεδομένα είναι αποθηκευμένα στον υπολογιστή για την περεταίρω ανάλυση και για τη διευκόλυνση της διαδικασίας της προσομοίωσης.

5.4.2 Υπολογισμός Μετρήσεων

Δεύτερο βήμα αποτελεί ο υπολογισμός των μετρήσεων προσομοίωσης. Βασικό ρόλο για τον υπολογισμό των μετρήσεων αποτελεί ο υπολογισμός των εκτιμητών ετερογένειας. Στην προσομοίωση συμπεριλάβαμε όλους τους εκτιμητές ετερογένειας που παρουσιάστηκαν στο Κεφάλαιο 4 έπειτα από ανασκόπηση της βιβλιογραφίας. Επειδή είχαμε δύο είδη αποτελεσμάτων, τα διχότομα και τα συνεχή, κατά την παραγωγή δεδομένων ο υπολογισμός των μετρήσεων πραγματοποιήθηκε σε διαφορετικούς κώδικες για τα δύο είδη δεδομένων. Επομένως δημιουργήθηκαν δύο κώδικες για τον υπολογισμό των εκτιμητών ετερογένειας και κατά αποδοτικότητα των μετρήσεων για τα διχότομα δεδομένα και για τα συνεχή αντίστοιχα.

5.4.2.1 Οι εκτιμητές της προσομοίωσης

Κατά τη βιβλιογραφία οι εκτιμητές ετερογένειας όπως παρουσιάστηκαν

στο Κεφάλαιο 4 αποτελούν σε συνολικό αριθμό 23 για τα διχότομα και 24 για τα συνεχή, καθώς ο εκτιμητής MBH χρησιμοποιείται μόνο για την τυποποιημένη μέση διαφορά. Επειδή οι εκτιμητές PM και EB παίρνουν ακριβώς τις ίδιες τιμές Paule & Mandel (1982) συμπεριλάβαμε στην ανάλυση μόνο τον EB. Το γεγονός βέβαια αυτό υποδηλώνει ότι τα συμπεράσματα για τον εκτιμητή PM θα ισχύουν ακριβώς τα ίδια και για τον εκτιμητή EB. Οι δύο εκτιμητές μπορούν να προγραμματιστούν μέσω της βιβλιοθήκης metafor Viechtbauer (2013) (<http://www.metafor-project.org/doku.php>) και ο υπολογισμός του EB πραγματοποιήθηκε μέσω αυτής της βιβλιοθήκης. Ακόμα όπως αναφέραμε ο εκτιμητής AREML αποτελεί μια προσέγγιση του REML. Η προσέγγιση είναι εφικτή όταν οι πληθυσμιακές διακυμάνσεις είναι ίσες (Viechtbauer, 2005). Επειδή στην πράξη οι πληθυσμιακές διακυμάνσεις δεν είναι ίσες ο εκτιμητής AREML αποκλείστηκε από την περαιτέρω ανάλυση. Πέρα από τη σύγκριση των εκτιμητών που αναφέραμε προσομοιώσαμε και εκτιμητές με την μέθοδο Knapp and Hartung για τη σύγκριση της μεθόδου δημιουργίας διαστημάτων εμπιστοσύνης για το μ με τη μέθοδο δημιουργίας Wald type. Η σύγκριση αυτή πραγματοποιήθηκε για τις μετρήσεις των διαστημάτων εμπιστοσύνης, του σφάλματος τύπου I και της ισχύς καθώς η μεροληψία και η αποδοτικότητα αποδίδουν τις ίδιες τιμές. Για τη σύγκριση αυτή συμπεριλάβαμε ενδεικτικά κάποιους εκτιμητές τους DL, GHO, ML και REML συμβολίζοντας τους με DLknh, GHOknh, MLknh, REMLknh αντίστοιχα. Οι 14 εκτιμητές DL, GHO, HS, SJ, ML, REML, EB, DL2, GHO2, SJgho, DLknh, GHOknh, MLknh, REMLknh υπολογίστηκαν με τη χρήση της βιβλιοθήκης metafor, ενώ 7 εκτιμητές DLp, DLb, HM, RBp, RBo, MBH, BM υπολογίστηκαν μέσω του προγραμματισμού συναρτήσεων. Σύμφωνα με τους Kontopantelis et al (2013) για να γίνει εφικτός ο υπολογισμός του εκτιμητή RBo χρειαζόμασταν τα μεγέθη των μελετών n_i τα οποία προσομοιώσαμε να παράγονται από την ομοιόμορφη κατανομή *Uniform*(50,500). Ο εκτιμητής FB ως αποτέλεσμα υπολογισμού μιας μαρκοβιανής αλυσίδας υπολογίστηκε στο μπεϋζιανό περιβάλλον του WinBUGS (Lunn et al., 2000) www.mrc-bsu.cam.ac.uk/bugs/ και μέσω της βιβλιοθήκης της R, R2WinBUGS (<http://cran.r-project.org/web/packages/R2WinBUGS/R2WinBUGS.pdf>)

στέλλονταν αυτόματα τα δεδομένα στο WinBUGS υπολογίζονταν ο εκτιμητής FB και έπειτα παίρναμε τα αποτελέσματα πίσω στην R. Για τον υπολογισμό του εκτιμητή FB παίρναμε κάθε προσομοιωμένο σετ y_i , $\hat{\sigma}_i^2$ και τις εκ των προτέρων κατανομές, την $uniform(0,100)$ για το τ και την $N(0, 10^6)$ για το μ . Έπειτα μέσω της μεθόδου MCMC εκτιμούσαμε τις παραμέτρους μ και τ εκτελώντας 1000 επαναλήψεις μέσα στη μαρκοβιανή αλυσίδα (iteration burn-in). Ο εκτιμητής FB και η συνολική επίδραση παράγονταν από τη διάμεσο των εκτιμώμενων παραμέτρων τ και μ . Για τη δημιουργία της παραπάνω διαδικασίας δημιουργήθηκε ένα ξεχωριστό αρχείο που χρησιμοποιήθηκε στο WinBUGS και παρατίθεται στο παράρτημα.

5.4.2.2 Μετρήσεις

Πέρα από τον υπολογισμό των εκτιμητών ετερογένειας υπολογίστηκε η συνολική επίδραση και το $(1 - \alpha)\%$ διάστημα εμπιστοσύνης της είτε μέσω των βιβλιοθηκών *metafor*, *R2WinBUGS* είτε μέσω του προγραμματισμού ανάλογα με τον τρόπο που είχε υπολογιστεί ο κάθε εκτιμητής ετερογένειας. Ακόμα υπολογίστηκε το $(1 - \alpha)\%$ bootstrap διάστημα εμπιστοσύνης των εκτιμητών ετερογένειας μέσω της βιβλιοθήκης *boot* (<https://cran.r-project.org/web/packages/boot/boot.pdf>). Μέσα στις B επαναλήψεις υπολογίζονταν οι εκτιμητές ετερογένειας, τα διαστήματα εμπιστοσύνης για τη συνολική επίδραση και αθροιστές για τις μετρήσεις προσομοίωσης. Μετά το πέρας των B επαναλήψεων υπολογίστηκαν οι μετρήσεις

- 1) μεροληψία (bias)
- 2) αποτελεσματικότητα (efficiency)
- 3) πιθανότητα κάλυψης για το μ (coverage probability for μ)
- 4) πιθανότητα κάλυψης για το τ^2 (coverage probability for τ^2)
- 5) εμπειρικό σφάλμα τύπου I (type error I)
- 6) ισχύς (power)

Ως αποτέλεσμα για κάθε σενάριο λαμβάναμε τις μετρήσεις προσομοίωσης για όλους τους εκτιμητές ετερογένειας για συγκεκριμένη τιμή του αριθμού των μελετών k , της πραγματικής τιμής της συνολικής επίδρασης μ_0 , της εναλλακτικής πραγματικής τιμής της συνολικής επίδρασης μ_1 , της

πραγματικής τιμής της ετερογένειας τ^2 και του επιπέδου σημαντικότητας α . Τα αποτελέσματα της προσομοίωσης αποθηκεύονταν αυτόματα σε αρχείο στον υπολογιστή για πιο γρήγορη και εύκολη εξαγωγή των αποτελεσμάτων. Έπειτα δημιουργήθηκαν κατάλληλοι πίνακες και διαγράμματα με στόχο την εξαγωγή των συμπερασμάτων από τη σύγκριση των ιδιοτήτων των εκτιμητών ετερογένειας για τα διάφορα σενάρια.

5.5 Εξαγωγή συμπερασμάτων για τα δεδομένα με διχότομο αποτέλεσμα

Η εξαγωγή συμπερασμάτων πραγματοποιήθηκε έπειτα από τη σύγκριση όλων των μετρήσεων που χρησιμοποιήθηκαν στην προσομοίωση για τα διάφορα σενάρια. Συγκεκριμένα παραθέτουμε τα διαγράμματα στην ενότητα 5.9 για τα σενάρια $\theta = 0, 0.3, 0.5, 0.8$ και $k = 10, 20, 30, 50$ για μια οπτική και πιο εύκολη αντιληπτή εικόνα της συμπεριφοράς των εκτιμητών.

5.5.1 Σύγκριση των εκτιμητών ετερογένειας ως προς τη μεροληψία

Από τους Πίνακες 5.1(α), 5.1(β) παρατηρούμε τα διαγράμματα για τη συμπεριφορά των εκτιμητών ως προς τη μεροληψία για τις διάφορες τιμές της ετερογένειας και για τα σενάρια $\theta = 0, 0.3, 0.5, 0.8$ και $k = 10, 20, 30, 50$.

Απ' όλα τα διαγράμματα παρατηρούμε ότι η μεροληψία όλων των εκτιμητών έχει ανοδική πορεία καθώς η ετερογένεια αυξάνεται.

Ακόμα η μεροληψία όλων των εκτιμητών σε όλες τις περιπτώσεις μειώνεται καθώς ο αριθμός των μελετών αυξάνεται.

Ο εκτιμητής RBp έχει τη μεγαλύτερη μεροληψία με διαφορά από όλους τους υπόλοιπους εκτιμητές ετερογένειας για όλες τις τιμές της ετερογένειας και της συνολικής επίδρασης θ , όταν ο αριθμός των μελετών στη μετα-ανάλυση είναι $k = 20, 30, 50$.

Ενώ όταν έχουμε μικρό αριθμό μελετών, $k = 10$, ο εκτιμητής RBp έχει τη μεγαλύτερη μεροληψία για τιμές ετερογένειας μέχρι 0.07, αλλά για μεγάλη ετερογένεια τη μεγαλύτερη μεροληψία έχει ο FB ακολουθώντας οι BM, HE και RBp.

Σε όλα τα σενάρια, όταν ο αριθμός των μελετών στη μετα-ανάλυση είναι $k = 20, 30, 50$, για τιμές ετερογένειας μέχρι 0.07 , έπειτα από τον εκτιμητή RBr, μεγαλύτερη μεροληψία έχει ο εκτιμητής SJ.

Ακόμα, για όλες τις τιμές της συνολικής επίδρασης θ και για μικρό αριθμό μελετών, $k = 10$, ο εκτιμητής RBr έχει με διαφορά μεγαλύτερη μεροληψία για τιμές ετερογένειας μέχρι 0.07 από τον εκτιμητή FB, ενώ για μεγαλύτερη ετερογένεια ο εκτιμητής FB έχει μεγαλύτερη μεροληψία από τον εκτιμητή RBr με διαφορά. Όμως κατά την απουσία ετερογένειας, έπειτα από τον εκτιμητή RBr, μεγαλύτερη μεροληψία έχουν οι εκτιμητές SJ και FB με παρόμοια συμπεριφορά.

Ο εκτιμητής HM έχει τη μεγαλύτερη μεροληψία έπειτα από τους RBr και SJ για όλα τα σενάρια με αριθμό μελετών $k = 20, 30, 50$ κατά την απουσία ετερογένειας, ενώ για αριθμό μελετών $k = 10$ έχει τη μεγαλύτερη μεροληψία από τους υπόλοιπους εκτιμητές εκτός από τους RBr, SJ και HE, ο οποίος έχει λίγο μεγαλύτερη μεροληψία. Ο εκτιμητής HM κατά την απουσία ετερογένειας έχει μεγάλη μεροληψία συγκριτικά με τους περισσότερους εκτιμητές. Όσο αυξάνεται η ετερογένεια, η μεροληψία του μειώνεται, έχοντας τη μικρότερη μεροληψία απ' όλους τους εκτιμητές για τιμές ετερογένειας κοντά στο 0.07 . Κατά την παρουσία μεγάλης ετερογένειας, για μικρό αριθμό μελετών έχει μικρή μεροληψία, ενώ με την αύξηση του αριθμού των μελετών αυξάνεται και η μεροληψία. Ο εκτιμητής DLb έχει την πιο μικρή μεροληψία μετά τον εκτιμητή HM κατά την παρουσία ετερογένειας όταν ο αριθμός των μελετών είναι μικρός. Εάν έχουμε μεγάλη ετερογένεια και ο αριθμός των μελετών είναι μεγάλος ο DLb όπως και ο HM έχει τιμές μεροληψίας στις τιμές που κυμαίνονται οι περισσότεροι εκτιμητές.

Για $k = 20, 30, 50$ και $\tau^2 \geq 0.07$, οι εκτιμητές HE και RBo έχουν μεγαλύτερη μεροληψία συγκριτικά με τους περισσότερους εκτιμητές.

Κατά την απουσία ετερογένειας μικρότερη μεροληψία παρουσιάζει ο εκτιμητής ML και έπειτα ακολουθεί ο εκτιμητής HS με λίγο μεγαλύτερη μεροληψία.

Για μικρό αριθμό μελετών, $k = 10$, ο εκτιμητής FB έχει τη μεγαλύτερη μεροληψία για μεγάλη ετερογένεια, και τη δεύτερη μεγαλύτερη έπειτα από

τον εκτιμητή RBr για μικρή ετερογένεια. Όμως για αριθμό μελετών $k = 20, 30, 50$ η μεροληψία του εκτιμητή FB κυμαίνεται στις τιμές μεροληψίας των περισσότερων εκτιμητών.

Έπειτα από τους εκτιμητές RBr, SJ, HE, RBo τη μεγαλύτερη μεροληψία έχει ο εκτιμητής SJgho, ενώ οι υπόλοιποι έχουν παρόμοια συμπεριφορά για όλες τις περιπτώσεις όταν έχουμε $k = 20, 30, 50$, ενώ για μικρό αριθμό μελετών $k = 10$ εκτός από τους παραπάνω εκτιμητές, μεροληψία με διαφορά τιμών από τους υπόλοιπους εκτιμητές έχει ο εκτιμητής BM και FB.

5.5.2 Σύγκριση των εκτιμητών ετερογένειας ως προς το μέσο τετραγωνικό σφάλμα

Στους Πίνακες 5.2(α), 5.2(β) παρουσιάζεται η συμπεριφορά των εκτιμητών ως προς το μέσο τετραγωνικό σφάλμα. Παρατηρούμε ότι η συμπεριφορά των εκτιμητών για τη αποδοτικότητα είναι παρόμοια με τη συμπεριφορά ως προς τη μεροληψία. Σε όλες τις περιπτώσεις καθώς η ετερογένεια αυξάνεται, το MSE αυξάνεται για κάθε εκτιμητή.

Ακόμα καθώς ο αριθμός των μελετών αυξάνεται το MSE για κάθε εκτιμητή μειώνεται για τις διάφορες τιμές της συνολικής επίδρασης θ .

Γενικά ο εκτιμητής RBr έχει το μεγαλύτερο MSE με διαφορά από τους υπόλοιπους εκτιμητές.

Οι SJ και HE έχουν λίγο μεγαλύτερο MSE από τους περισσότερους εκτιμητές.

Ο εκτιμητής HM έχει το ίδιο MSE με τους περισσότερους εκτιμητές για μικρή και μεγάλη ετερογένεια, ενώ για την τιμή της ετερογένειας κοντά στο 0.07 παρουσιάζει το μικρότερο MSE από όλους τους εκτιμητές.

Ο εκτιμητής DLb έχει μικρό MSE γενικά για όλες τις περιπτώσεις.

Ο εκτιμητής RBo έχει μικρότερο MSE από τους SJ και HE για μικρό αριθμό μελετών, ενώ για μεγάλο αριθμό μελετών έχει μεγαλύτερο MSE.

Ακόμα, ο εκτιμητής SJgho έχει γενικά μεσαίο MSE συγκριτικά με τους υπόλοιπους εκτιμητές, παρουσιάζοντας μικρότερο MSE από τους SJ, HE και RBo αλλά μεγαλύτερο στις περισσότερες περιπτώσεις από τους υπόλοιπους.

Ο ΒΜ έχει μεγάλο MSE σε σύγκριση με τους περισσότερους εκτιμητές όταν ο αριθμός των μελετών είναι μικρός, $k = 10$.

5.5.3 Σύγκριση των εκτιμητών ετερογένειας ως προς την πιθανότητα κάλυψης μ

Κατά την προσομοίωση υπολογίσαμε για κάθε εκτιμητή ετερογένειας την πιθανότητα κάλυψης του μ . Στους Πίνακες 5.3 (α), 5.3 (β) παρουσιάζονται τα γραφήματα της πιθανότητας κάλυψης του μ ως προς την ετερογένεια για όλα τα σενάρια. Έχοντας επίπεδο σημαντικότητας α , ορίζουμε την πιθανότητα κάλυψης για την ετερογένεια στην ιδεατή περίπτωση να είναι $\hat{p} = 1 - \alpha$. Θεωρούμε ότι ένας εκτιμητής με πιθανότητα κάλυψης του μ ίση με p έχει καλή συμπεριφορά όταν η τιμή της πιθανότητας κάλυψης του βρίσκεται μέσα στο διάστημα $\hat{p} \pm 1.96 \sqrt{\frac{\hat{p}(1-\hat{p})}{B}}$, όπου B ο αριθμός των επαναλήψεων. Με επίπεδο σημαντικότητας $\alpha_2 = 0.05$, η πιθανότητα κάλυψης του μ στην ιδεατή περίπτωση είναι ίση με $\hat{p} = 0.95$. Συνεπώς, η πιθανότητα κάλυψης ενός εκτιμητή πρέπει να βρίσκεται μέσα στο διάστημα $(l_2, u_2) = (0.93, 0.96)$. Εκτιμητές που ξεπερνάνε αυτό το διάστημα δεν έχουν καλή συμπεριφορά ως προς την πιθανότητα κάλυψης του μ .

Για μικρό αριθμό μελετών $k = 10$, κατά την ύπαρξη ετερογένειας η πιθανότητα κάλυψης του μ για τους περισσότερους εκτιμητές ξεπερνάει το διάστημα (l_2, u_2) αφού η πιθανότητα κάλυψης για τους περισσότερους εκτιμητές βρίσκεται κάτω από το όριο l_2 .

Για μεγαλύτερο αριθμό μελετών η πιθανότητα κάλυψης του μ για τους περισσότερους εκτιμητές βρίσκονται μέσα στο επιτρεπτό διάστημα (l_2, u_2) . Παρατηρούμε ότι με την αύξηση του αριθμού των μελετών η πιθανότητα κάλυψης για το μ από τους περισσότερους εκτιμητές βρίσκεται μέσα στο επιτρεπτό διάστημα. Επομένως οι εκτιμητές έχουν καλύτερη συμπεριφορά καθώς ο αριθμός των μελετών αυξάνεται.

Ακόμα παρατηρούμε ότι καθώς αυξάνεται η ετερογένεια οι πιθανότητες κάλυψης της συνολικής επίδρασης βρίσκονται πιο κοντά στο επιτρεπτό διάστημα.

Η πιθανότητα κάλυψης του μ για τον εκτιμητή RBp είναι η πιο υψηλή από όλους τους υπόλοιπους εκτιμητές και έχει τιμές πολύ πιο υψηλές από το άνω όριο u_1 για όλες τις περιπτώσεις, ενώ μόνο για μικρό αριθμό μελετών $k = 10$ βρίσκεται μέσα στο επιτρεπτό διάστημα (l_2, u_2) μόνο κατά την ύπαρξη μεγάλης ετερογένειας με τιμές κοντά στο 0.30. Ο εκτιμητής SJ παρουσιάζει πολύ πιο υψηλή από το επιτρεπτό όριο πιθανότητα κάλυψης για το μ , οι τιμές είναι μικρότερες από αυτή του εκτιμητή RBp αλλά μεγαλύτερες από τους περισσότερους εκτιμητές. Η πιθανότητα κάλυψης του μ είναι πιο υψηλή από το επιτρεπτό όριο για μικρές και μεσαίες τιμές ετερογένειας, ενώ για μεγάλη ετερογένεια κοντά στο 0.30 βρίσκεται μέσα στο επιτρεπτό όριο.

Ο εκτιμητής HM έχει πιθανότητα κάλυψης για το μ μεγαλύτερη από το άνω όριο του επιτρεπτού διαστήματος, για μεσαία ετερογένεια βρίσκεται μέσα στο επιτρεπτό όριο και για μεγάλη ετερογένεια μπορεί να μειώνεται με τιμές χαμηλότερες από το κάτω όριο του επιτρεπτού διαστήματος.

Ο εκτιμητής FB έχει υψηλότερη πιθανότητα κάλυψης του μ από το επιτρεπτό διάστημα για μικρό αριθμό μελετών $k = 10$ κατά την απουσία μικρής ή μεσαίας ετερογένειας, ενώ για μεγάλη ετερογένεια βρίσκεται μέσα στο διάστημα (l_2, u_2) . Όσο ο αριθμός των μελετών αυξάνεται τόσο οι τιμές της πιθανότητας κάλυψης του μ βρίσκονται μέσα στο επιτρεπτό διάστημα, έτσι ώστε για $k = 20$ να βρίσκονται μέσα στο επιτρεπτό διάστημα για μεσαία ετερογένεια, για $k = 30$ να βρίσκονται για μικρή ετερογένεια μέσα στο διάστημα και για $k = 50$ να έχει τιμές πιθανότητας κάλυψης μέσα στο διάστημα για όλες τις τιμές ετερογένειας. Ο εκτιμητής RBo παρουσιάζει την πιο χαμηλή πιθανότητα κάλυψης, πολύ μικρότερη από το κάτω άκρο l_2 του διαστήματος σε όλες τις περιπτώσεις κατά τις οποίες η ετερογένεια είναι μεσαία και μεγάλη, ενώ κατά την απουσία ετερογένειας ή για μικρή τιμή ετερογένειας μέχρι 0.025, βρίσκεται μέσα στο επιτρεπτό διάστημα σε όλες τις περιπτώσεις. Για μικρό αριθμό μελετών $k = 10$, η πιθανότητα κάλυψης του μ για τον εκτιμητή RBo μειώνεται συνεχώς καθώς αυξάνεται η ετερογένεια. Για μεσαία και μεγάλο αριθμό μελετών, η πιθανότητα κάλυψης του μειώνεται για μεσαία ετερογένεια ενώ για μεγάλη ετερογένεια πλησιάζει το κάτω όριο του επι-

τρεπτού διαστήματος.

Σε όλες τις περιπτώσεις παρατηρούμε ότι η πιθανότητα κάλυψης του μ των εκτιμητών που προσαρμόζουν πιο καλά από όλους τους υπόλοιπους είναι οι εκτιμητές Knapp Hartung (DLknh, GHOknh, MLknh, REMLknh) καθώς βρίσκονται μέσα στο επιτρεπτό διάστημα για όλες τις τιμές της ετερογένειας. Το αποτέλεσμά μας αυτό επιβεβαιώνεται από τα αποτελέσματα των (IntHout, Ioannidis, & Borm, 2014), οι οποίοι συμπέραναν ότι ο εκτιμητής DLknh συμπεριφέρεται καλύτερα από τον εκτιμητή DL.

5.5.4 Σύγκριση των εκτιμητών ετερογένειας ως προς την πιθανότητα κάλυψης τ^2

Κατά την προσομοίωση υπολογίσαμε για κάθε εκτιμητή ετερογένειας την πιθανότητα κάλυψης της ετερογένειας τ^2 . Στους Πίνακες 5.4(α), 5.4(β) παρουσιάζονται τα γραφήματα της πιθανότητας κάλυψης του τ^2 ως προς την ετερογένεια για όλα τα σενάρια. Θα συγκρίνουμε τους εκτιμητές έχοντας επίπεδο σημαντικότητας $\alpha_2 = 0.05$ και το επιτρεπτό διάστημα της πιθανότητας κάλυψης ενός εκτιμητή $(l_2, u_2) = (0.93, 0.96)$.

Αρχικά παρατηρούμε ότι οι πιθανότητες κάλυψης για την ετερογένεια τ^2 είναι πολύ μικρότερες για κάποιους εκτιμητές από το διάστημα (l_2, u_2) ενώ για κάποιους άλλους είναι λίγο υψηλότερες. Το γεγονός ότι δεν παίρνουμε τιμές κοντά στο 0.95 έχει διαπιστωθεί από τον (Viechtbauer, 2007b) καθώς η μέθοδος που χρησιμοποιούμε για τη δημιουργία της πιθανότητας κάλυψης του τ^2 είναι η μη-παραμετρική μέθοδος bootstrap (non parametric bootstrap).

Παρατηρούμε ακόμα ότι καθώς αυξάνεται η ετερογένεια, οι πιθανότητες κάλυψης της ετερογένειας απομακρύνονται από το επιτρεπτό διάστημα, ενώ με την αύξηση του αριθμού των μελετών οι πιθανότητες κάλυψης προσαρμόζουν καλύτερα.

Για τους περισσότερους θετικούς εκτιμητές όπως οι BM, RBp, SJ και HM η πιθανότητα κάλυψης της ετερογένειας είναι μηδενική κατά την απουσία ετερογένειας, ενώ με την παρουσία ετερογένειας έχουμε δραματική αύξηση της πιθανότητας κάλυψης. Επομένως αυτοί οι εκτιμητές συμπερι-

φέρονται καλύτερα κατά την παρουσία ετερογένειας.

Ο εκτιμητής BM έχει την χαμηλότερη πιθανότητα κάλυψης της ετερογένειας από όλους τους εκτιμητές για όλες τις περιπτώσεις. Ο SJgho, ο οποίος είναι και αυτός θετικός έχει πολύ χαμηλή πιθανότητα κάλυψης από τους περισσότερους εκτιμητές εκτός από τον BM.

Πέρα από τους παραπάνω θετικούς εκτιμητές, οι υπόλοιποι εκτιμητές DL, HE, HS, ML, REML, EB, DL2 και GHO2 έχουν κοινή συμπεριφορά ως προς τη πιθανότητα κάλυψης της ετερογένειας με παρόμοιες τιμές. Κατά την απουσία της ετερογένειας έχουν μεγαλύτερη πιθανότητα κάλυψης από το επιτρεπτό διάστημα, για μικρή ετερογένεια μέχρι 0.025 η πιθανότητα κάλυψης μειώνεται διαπερνώντας το επιτρεπτό διάστημα, ενώ για μεγαλύτερη τιμή της ετερογένειας, η πιθανότητα κάλυψης κυμαίνεται στις ίδιες τιμές οι οποίες είναι πολύ χαμηλότερες από το επιτρεπτό διάστημα. Ακόμα, για τους εκτιμητές αυτούς παρατηρούμε ότι έχουν πιθανότητα κάλυψης πιο κοντά στο επιτρεπτό διάστημα όταν ο αριθμός των μελετών αυξάνεται. Επομένως, οι εκτιμητές αυτοί έχουν την καλύτερη συμπεριφορά για τιμές ετερογένειας μέχρι 0.025 και προσαρμόζουν καλύτερα όσο αυξάνεται ο αριθμός των μελετών.

Από τους εκτιμητές DL, HE, HS, ML, REML, EB, DL2 και GHO2 που έχουν παρόμοια συμπεριφορά ως προς την πιθανότητα κάλυψης της ετερογένειας, την πιο χαμηλή πιθανότητα κάλυψης έχει ο εκτιμητής RBo, ακολουθώντας ο ML και ο HS. Πιο κοντά στο επιτρεπτό διάστημα για όλες τις τιμές ετερογένειας βρίσκεται ο εκτιμητής EB για μεσαίο και μεγάλο αριθμό μελετών, ενώ για μικρό αριθμό μελετών $k = 10$ η πιθανότητα κάλυψης του κυμαίνεται με τις τιμές της πιθανότητας κάλυψης των υπόλοιπων εκτιμητών.

Ο εκτιμητής HM έχει την πιο κοντινή πιθανότητα κάλυψης κοντά στο επιτρεπτό διάστημα από όλους τους εκτιμητές για μεσαία ετερογένεια κοντά στο 0.025 για μικρό αριθμό μελετών $k = 10$, για $k = 20$ τη μεγαλύτερη πιθανότητα κάλυψης έχει για ετερογένεια μεταξύ του 0.025 και 0.07, ενώ για μεγαλύτερο αριθμό μελετών $k = 30, 50$ τη μεγαλύτερη πιθανότητα κάλυψης έχει για ετερογένεια κοντά στο 0.07.

Για μεγάλη ετερογένεια και μικρό αριθμό μελετών $k = 10$ ο εκτιμητής

RBr έχει την πιο κοντινή πιθανότητα κάλυψης στο επιτρεπτό διάστημα, για αριθμό μελετών $k = 20$ καλύτερη συμπεριφορά παρουσιάζουν οι εκτιμητές RBr και SJ χωρίς μεγάλες διαφορές τιμών ενώ για μεγάλο αριθμό μελετών $k = 30, 50$ ο SJ προσαρμόζει καλύτερα. Ο εκτιμητής SJ προσαρμόζει καλύτερα ως προς τη πιθανότητα κάλυψης για μεσαία προς μεγάλη ετερογένεια μόνο όταν ο αριθμός των μελετών είναι μικρός $k = 10$.

5.5.5 Σύγκριση των εκτιμητών ετερογένειας ως προς το εμπειρικό type error I

Κατά την προσομοίωση υπολογίσαμε για κάθε εκτιμητή ετερογένειας το εμπειρικό type error I. Στους Πίνακες 5.5(α), 5.5(β) παρουσιάζονται τα γραφήματα του εμπειρικού type error I ως προς την ετερογένεια για όλα τα σενάρια όταν το επίπεδο σημαντικότητας είναι $\alpha = 0.05$. Ένας εκτιμητής λέμε ότι έχει καλή συμπεριφορά όταν το εμπειρικό type error I και βρίσκεται μέσα στο διάστημα $\alpha \pm 1.96 \sqrt{\frac{\alpha(1-\alpha)}{B}}$, όπου B ο αριθμός των επαναλήψεων. Για επίπεδο σημαντικότητας $\alpha_2 = 0.05$, για να έχει καλή συμπεριφορά ο εκτιμητής πρέπει να βρίσκεται επομένως στο διάστημα $(L_2, U_2) = (0.03, 0.06)$. Εκτιμητές που ξεπερνάνε αυτό το διάστημα δεν έχουν καλή συμπεριφορά ως προς το εμπειρικό type error I. Το εμπειρικό type error I έχει άμεση σχέση με τις πιθανότητες κάλυψης της συνολικής επίδρασης καθώς όταν έχουμε αυξημένο εμπειρικό type error I σημαίνει ότι απορρίπτεται πιο συχνά η μηδενική μέθοδος από όσο θα έπρεπε και επομένως η πιθανότητα κάλυψης για τη συνολική επίδραση θα είναι μικρότερη. Συνεπώς τη συμπεριφορά των εκτιμητών ως προς τις πιθανότητες κάλυψης της συνολικής επίδρασης θα διατηρείται στα ίδια πλαίσια και για το I.

Για μικρό αριθμό μελετών, $k = 10$, οι περισσότεροι εκτιμητές έχουν εμπειρικό type error I με μεγάλη διαφορά τιμών από τις τιμές του διαστήματος (L_2, U_2) , είτε πολύ μεγαλύτερες είτε μικρότερες. Στην πλειοψηφία των εκτιμητών για μικρό αριθμό μελετών, $k = 10$, οι τιμές του εμπειρικού σφάλματος τύπου I βρίσκονται πάνω από το άνω όριο U_2

για μεσαία και μεγάλη ετερογένεια.

Όσο αυξάνεται ο αριθμός των μελετών, τόσο πιο κοντά είναι οι τιμές του εμπειρικού σφάλματος τύπου I των εκτιμητών στο επιτρεπτό διάστημα (L_2, U_2) .

Ο εκτιμητής RBp έχει εμπειρικό type error I πολύ μικρότερο από το κάτω όριο του διαστήματος L_2 για όλα τα σενάρια, με την αύξηση της ετερογένειας αυξάνεται και το type error I παραμένοντας πολύ μικρότερο από το κάτω όριο L_2 , ενώ μόνο για μικρό αριθμό μελετών $k = 10$ και ετερογένεια κοντά στο 0,3 βρίσκεται μέσα στο επιτρεπτό διάστημα.

Ο εκτιμητής FB έχει χαμηλότερο εμπειρικό type error I από το κάτω όριο L_2 μόνο για μικρό αριθμό μελετών $k = 10$, ενώ για μεγάλη ετερογένεια ($\tau^2 = 0.30$) βρίσκεται μέσα στο επιτρεπτό διάστημα. Για $k = 20$, το εμπειρικό type error I του εκτιμητή FB βρίσκεται κάτω από το επιτρεπτό διάστημα κατά την απουσία ετερογένειας ή για μικρή ετερογένεια, ενώ για μικρή προς μεσαία ετερογένεια βρίσκεται μέσα στο επιτρεπτό διάστημα. Για $k = 30$, υπάρχουν περιπτώσεις κατά τις οποίες το εμπειρικό type error I μπορεί να βρίσκεται κάτω από το επιτρεπτό διάστημα για πολύ μικρή ετερογένεια, ενώ για μεγάλο αριθμό μελετών $k = 50$, το εμπειρικό type error I του εκτιμητή βρίσκεται μέσα στο επιτρεπτό διάστημα για όλες τις τιμές ετερογένειας.

Ο εκτιμητής SJ έχει πολύ χαμηλότερο εμπειρικό type error I από το κάτω όριο του διαστήματος L_2 για όλες τις περιπτώσεις κατά την απουσία ετερογένειας ή κατά την ύπαρξη μικρής και μεσαίας ετερογένειας. Το εμπειρικό type error I αυξάνεται με την αύξηση της ετερογένειας και για μεγάλη ετερογένεια, βρίσκεται μέσα στο επιτρεπτό όριο εκτός από την περίπτωση $\theta = 0, k = 50$.

Το εμπειρικό type error I του εκτιμητή HM είναι χαμηλότερο από το κάτω όριο του L_2 , κατά την απουσία ετερογένειας ή για μικρή ετερογένεια, για μεσαία ετερογένεια βρίσκεται μέσα στο διάστημα (L_2, U_2) , ενώ για μεγάλη ετερογένεια βρίσκεται πάνω από το άνω όριο U_2 .

Το εμπειρικό type error I του εκτιμητή RBo βρίσκεται μέσα στο επιτρεπτό διάστημα κατά την απουσία ετερογένειας ή για μικρή ετερογένεια μέχρι την τιμή 0,025 για όλα τα σενάρια, ενώ για μεγαλύτερη ετερογένεια το

εμπειρικό type error I του εκτιμητή είναι μεγαλύτερο από το επιτρεπτό διάστημα. Για μικρό αριθμό μελετών, $k = 10$, με την αύξηση της ετερογένειας αυξάνεται και το εμπειρικό type error I. Για μεσαίο και μεγάλο αριθμό μελετών, $k = 20, 30, 50$ πολύ μεγάλο εμπειρικό type error I παρουσιάζεται για μεσαία ετερογένεια, ενώ για μεγάλη ετερογένεια το εμπειρικό type error I είναι λίγο μεγαλύτερο από το άνω όριο u_3 .

Εκτός από τους εκτιμητές RBo, RBp, FB, HM, SJ που αναφέρθηκαν παραπάνω όλοι οι υπόλοιποι εκτιμητές έχουν παρόμοια συμπεριφορά ως προς το εμπειρικό type error I.

Την καλύτερη συμπεριφορά για όλες τις τιμές ετερογένειας παρουσιάζουν οι εκτιμητές Knapp and Hartung (DLknh, GHOknh, MLknh, REMLknh).

Ακόμα για μικρό αριθμό μελετών, $k = 10$ και κατά την παρουσία ετερογένειας το εμπειρικό type error I για τους εκτιμητές ML, HS είναι αρκετά πιο υψηλό από το άνω όριο καθώς και μεγαλύτερο από τους υπόλοιπους εκτιμητές εκτός από τον RBo.

5.5.6 Σύγκριση των εκτιμητών ετερογένειας ως προς την ισχύ

Ένας εκτιμητής είναι καλός όταν μεγιστοποιεί την ισχύ. Θέλουμε λοιπόν εκτιμητές οι οποίοι θα έχουν ισχύ πιο κοντά στη μονάδα. Στους Πίνακες 5.6(α), 5.6(β) παρατηρούμε τα διαγράμματα της ισχύς όλων των εκτιμητών για τις διάφορες τιμές της ετερογένειας, για $k = 10, 20, 30, 50$ και $\theta = 0.3, 0.5, 0.8$.

Αρχικά παρατηρούμε ότι κατά την αύξηση της ετερογένειας η ισχύς μειώνεται για όλους τους εκτιμητές ετερογένειας. Το φαινόμενο αυτό ισχύει κατά την ύπαρξη μικρής συνολικής επίδρασης ($\theta = 0.3$) σ' όλες τις περιπτώσεις. Όσο μεγαλώνει το θ και ο αριθμός των μελετών έχουμε την ιδεατή περίπτωση να παρουσιάζουν όλοι οι εκτιμητές ισχύ σχεδόν ίση με τη μονάδα για όλες τις τιμές ετερογένειας.

Όταν δεν έχουν όλοι οι εκτιμητές ισχύ κοντά στη μονάδα, ο εκτιμητής RBp έχει τη μικρότερη ισχύ εκτός από $k = 10$ και μεγάλη ετερογένεια κοντά στο 0.30. Έπειτα μικρότερη ισχύ έχει για τις περιπτώσεις αυτές ο FB και ο SJ για $\theta = 0.3$, ενώ όσο το θ αυξάνεται, μεγαλώνει και η ισχύς

του εκτιμητή. Κατά την απουσία ετερογένειας και για μικρό αριθμό μελετών $k = 10, 20$ οι εκτιμητές HS και ML έχουν τη μεγαλύτερη ισχύ για όλες τις τιμές ετερογένειας, ενώ κατά την ύπαρξη ετερογένειας και για $k = 10$ υψηλή ισχύ έχουμε και από τον εκτιμητή RBo.

5.5.7 Σύγκριση των εκτιμητών ως προς τα επίπεδα σημαντικότητας

Κατά την προσομοίωση υπολογίσαμε τις μετρήσεις της πιθανότητας κάλυψης για το μ , της πιθανότητας κάλυψης για την ετερογένεια, του σφάλματος τύπου I και της ισχύς για τα επίπεδα σημαντικότητας $\alpha = 0.01, 0.05, 0.10$. Στις παραγράφους 5.5.3, 5.5.4, 5.5.5 και 5.5.6 εξάγαμε συμπεράσματα για τους εκτιμητές ως προς τις μετρήσεις με επίπεδο σημαντικότητας $\alpha = 0.05$. Σκοπός μας είναι η σύγκριση των εκτιμητών για τις παραπάνω μετρήσεις ως προς τα διάφορα επίπεδα σημαντικότητας με στόχο την εξαγωγή συμπερασμάτων στο ευρύ φάσμα των παραμέτρων που μπορεί να πραγματοποιηθεί μια μετα-ανάλυση.

Ως προς τις πιθανότητες κάλυψης, είτε την πιθανότητα κάλυψης του μ είτε την πιθανότητα κάλυψης της ετερογένειας, έχοντας επίπεδο σημαντικότητας α , ορίζουμε την πιθανότητα κάλυψης για την ετερογένεια στην ιδεατή περίπτωση να είναι $\hat{p} = 1 - \alpha$. Θεωρούμε ότι ένας εκτιμητής με πιθανότητα κάλυψης της ετερογένειας ίση με p έχει καλή συμπεριφορά όταν η τιμή της πιθανότητας κάλυψης βρίσκεται μέσα

στο διάστημα $\hat{p} \pm 1.96 \sqrt{\frac{\hat{p}(1-\hat{p})}{B}}$, όπου B ο αριθμός των επαναλήψεων.

Επομένως για επίπεδο σημαντικότητας $\alpha_1 = 0.01$, ένας εκτιμητής έχει καλή συμπεριφορά όταν οι πιθανότητες κάλυψης του βρίσκονται μέσα στο διάστημα $(l_1, u_1) = (0.98, 0.99)$. Με επίπεδο σημαντικότητας $\alpha_2 = 0.05$, οι πιθανότητες κάλυψης πρέπει να βρίσκονται στο διάστημα $(l_2, u_2) = (0.93, 0.96)$. Ενώ με επίπεδο σημαντικότητας $\alpha_3 = 0.10$, πρέπει να βρίσκονται στο διάστημα $(l_3, u_3) = (0.88, 0.91)$.

Ως προς το type error I, με επίπεδο σημαντικότητας α , ένας εκτιμητής ορίζουμε να έχει καλή συμπεριφορά όταν βρίσκεται μέσα στο διάστημα $\alpha \pm 1.96 \sqrt{\frac{\alpha(1-\alpha)}{B}}$, όπου B ο αριθμός των επαναλήψεων. Επομένως για

επίπεδο σημαντικότητας $\alpha_1 = 0.01$, ένας εκτιμητής έχει καλή συμπεριφορά όταν το type error I βρίσκεται μέσα στο διάστημα $(L_1, U_1) = (0.0038, 0.0161)$. Με επίπεδο σημαντικότητας $\alpha_2 = 0.05$, το type error I πρέπει να βρίσκεται στο διάστημα $(L_2, U_2) = (0.04, 0.06)$. Ενώ με επίπεδο σημαντικότητας $\alpha_3 = 0.10$, πρέπει να βρίσκονται στο διάστημα $(L_3, U_3) = (0.08, 0.19)$. Από τους Πίνακες 5.13, 5.14, 5.15 και 5.16 παρατηρούμε τα διαγράμματα της πιθανότητας κάλυψης μ , της πιθανότητας κάλυψης της ετερογένειας, του σφάλματος τύπου I και της ισχύς αντίστοιχα, ως προς τις τιμές ετερογένειας για τα σενάρια με συνολική επίδραση $\theta = 0, 0.5$, αριθμό μελετών $k = 10, 30$ και επίπεδα σημαντικότητας $\alpha = 0.01, 0.05, 0.10$ όταν έχουμε διχότομα αποτελέσματα.

5.5.7.1 Σύγκριση ως προς την πιθανότητα κάλυψης του μ

Από τους Πίνακες 5.13 (α), 5.13 (β) παρατηρούμε τα διαγράμματα της πιθανότητας κάλυψης του μ ως προς τις τιμές της ετερογένειας. Έχοντας επίπεδο σημαντικότητας $\alpha_1 = 0.01$, η πιθανότητα κάλυψης είναι $\hat{p} = 0.99$, για επίπεδο σημαντικότητας $\alpha_2 = 0.05$, η πιθανότητα κάλυψης είναι $\hat{p} = 0.95$ και για επίπεδο σημαντικότητας $\alpha_3 = 0.10$, η πιθανότητα κάλυψης είναι $\hat{p} = 0.90$. Επομένως για διαφορετικά επίπεδα σημαντικότητας έχουμε και διαφορετικές τιμές των πιθανοτήτων κάλυψης. Συνεπώς οι πιθανότητες κάλυψης των εκτιμητών θα κυμαίνονται σε διαφορετικές τιμές ανάλογα με το επίπεδο σημαντικότητας που έχουμε. Παρατηρούμε ότι οι πιθανότητες κάλυψης έχουν παρόμοια συμπεριφορά απ' όλους τους εκτιμητές. Διαφορετική συμπεριφορά παρατηρούμε από τον εκτιμητή SJ, ο οποίος έχει πιθανότητες κάλυψης μέσα ή κοντά στο διάστημα που επιθυμούμε για α_1 επίπεδο σημαντικότητας. Ενώ όσο αυξάνεται το επίπεδο σημαντικότητας, η πιθανότητα κάλυψης του SJ αυξάνεται και κυμαίνεται αρκετά πιο μακριά από το επιθυμητό διάστημα.

5.5.7.2 Σύγκριση ως προς την πιθανότητα κάλυψης της ετερογένειας

Από τους Πίνακες 5.14 (α), 5.14 (β) παρατηρούμε τα διαγράμματα της πι-

θανότητας κάλυψης της ετερογένειας ως προς τις τιμές της ετερογένειας. Όπως είναι φυσικό όσο μικρότερο είναι το επίπεδο σημαντικότητας τόσο μεγαλύτερες πιθανότητες κάλυψης μπορούμε να πάρουμε. Επομένως για επίπεδο σημαντικότητας $\alpha_1 = 0.01$ έχουμε λίγο πιο υψηλές πιθανότητες κάλυψης, με επίπεδο σημαντικότητας $\alpha_2 = 0.05$ λίγο πιο χαμηλές και με $\alpha_3 = 0.10$ ακόμα λιγότερο χαμηλές. Πέρα από τη φυσιολογική συμπεριφορά των τιμών μεταξύ των επιπέδων σημαντικότητας, η συμπεριφορά των εκτιμητών είναι παρόμοια στα διάφορα επίπεδα σημαντικότητας. Αλλαγή συμπεριφοράς παρατηρείται από τον εκτιμητή HM ως προς τα διάφορα επίπεδα σημαντικότητας. Συγκεκριμένα, ο εκτιμητής HM έχει την πιο κοντινή πιθανότητα κάλυψης για ετερογένεια $0.025 \leq \tau^2 \leq 0.07$ απ' όλους τους εκτιμητές όταν ο αριθμός των μελετών είναι μικρός, $k = 10$. Όταν το επίπεδο σημαντικότητας είναι $\alpha_1 = 0.01$ ο εκτιμητής κυμαίνεται στις ίδιες τιμές. Με μεγαλύτερο επίπεδο σημαντικότητας $\alpha = 0.05, 0.10$ ο εκτιμητής HM έχει την πιο κατάλληλη πιθανότητα κάλυψης απ' όλους τους εκτιμητές μόνο για ετερογένεια κοντά στο 0.07, ενώ για τιμή ετερογένειας κοντά στο 0.025 παρουσιάζει χαμηλότερη πιθανότητα κάλυψης συγκριτικά με τους περισσότερους εκτιμητές. Ο εκτιμητής RBp βρίσκεται πολύ κοντά στο επιθυμητό διάστημα για επίπεδο σημαντικότητας $\alpha_1 = 0.01$ μαζί με τον SJ όταν έχουμε μικρό αριθμό μελετών, $k = 10$. Για επίπεδο σημαντικότητας $\alpha_2 = 0.05$ έχει λίγο χαμηλότερη πιθανότητα κάλυψης αλλά υψηλή από τους περισσότερους, ενώ για επίπεδο σημαντικότητας $\alpha_3 = 0.10$ κυμαίνεται με τις τιμές των πιθανοτήτων κάλυψης των περισσότερων εκτιμητών.

5.5.7.3 Σύγκριση ως προς το εμπειρικό type error I

Από τους Πίνακες 5.15 (α), 5.15 (β) παρατηρούμε τα διαγράμματα του εμπειρικού σφάλματος τύπου I ως προς τις τιμές της ετερογένειας, με τα οποία συμπεραίνουμε ότι οι εκτιμητές ετερογένειας έχουν παρόμοια συμπεριφορά για τα διάφορα επίπεδα σημαντικότητας.

5.5.7.4 Σύγκριση ως προς την ισχύ

Από τον Πίνακα 5.16 παρατηρούμε τα διαγράμματα της ισχύς ως προς τις

τιμές της ετερογένειας. Όπως είναι φυσικό, οι εκτιμητές έχουν μεγαλύτερη ισχύ καθώς αυξάνεται το επίπεδο σημαντικότητας. Ακόμα συμπεραίνουμε ότι όλοι οι εκτιμητές παρουσιάζουν παρόμοια συμπεριφορά.

5.6 Εξαγωγή συμπερασμάτων για τα δεδομένα με συνεχή αποτέλεσμα

5.6.1 Σύγκριση των εκτιμητών ετερογένειας ως προς τη μεροληψία

Από τους Πίνακες 5.7(α), 5.7(β) παρατηρούμε τα διαγράμματα για τη συμπεριφορά των εκτιμητών ως προς την μεροληψία για τις διάφορες τιμές της ετερογένειας και για τα σενάρια $\theta = 0, 0.3, 0.5, 0.8$ και $k = 10, 20, 30, 50$.

Απ' όλα τα διαγράμματα παρατηρούμε ότι η μεροληψία όλων των εκτιμητών έχει ανοδική πορεία καθώς η ετερογένεια αυξάνεται. Μερικοί εκτιμητές όπως RBr, SJ, DLp, FB και HE η μεροληψία τους παρουσιάζει καθοδική πορεία για ετερογένεια μέχρι 0.01 και ανοδική πορεία για μεγαλύτερη ετερογένεια.

Η μεροληψία των εκτιμητών αυξάνεται δραματικά για τιμές ετερογένειας από 0.05 ως 0.50 παρουσιάζοντας παρόμοια συμπεριφορά.

Αξίζει να σημειωθεί ότι η μεροληψία του εκτιμητή RBr είναι μεγαλύτερη από τους υπόλοιπους εκτιμητές κατά την απουσία ετερογένειας ή για μικρή ετερογένεια μέχρι 0.01, αντίθετα για μεγαλύτερη ετερογένεια ο RBr έχει τη μικρότερη μεροληψία από όλους τους εκτιμητές.

Παρόμοια συμπεριφορά με τον RBr παρουσιάζει ο εκτιμητής SJ, με μεροληψία μεγαλύτερη από τους περισσότερους εκτιμητές, αλλά μικρότερη από τον εκτιμητή RBr για ετερογένεια μέχρι 0.01. Παράλληλα κατά την ύπαρξη μεγάλης ετερογένειας παρουσιάζει μικρότερη μεροληψία από τους περισσότερους εκτιμητές.

Έχοντας ετερογένεια μικρότερη από 0.05 παρατηρούμε διαφορετικές συμπεριφορές ως προς τη μεροληψία των εκτιμητών. Πιο συγκεκριμένα, για ετερογένεια κοντά στο 0.01, τη μικρότερη μεροληψία στις περισσότερες περιπτώσεις έχουν οι εκτιμητές DLp και HM και έπειτα ακολουθούν οι εκτιμητές FB και DLb.

Κατά την απουσία ετερογένειας, έπειτα από τον RBr, μεγαλύτερη μερο-

ληγία από τους υπόλοιπους εκτιμητές έχει ο SJ για όλες τις περιπτώσεις, ενώ για μικρό αριθμό μελετών $k = 10$, μεγάλη μεροληψία έχουν εκτός από τον εκτιμητή SJ οι εκτιμητές FB και DLp.

Συμπεραίνουμε λοιπόν ότι κατά την παρουσία ετερογένειας και για τιμές μέχρι 0.05 καλή συμπεριφορά ως προς την ετερογένεια έχουν οι εκτιμητές DLp, HM, FB και DLb, με τον DLp να παρουσιάζει λίγο καλύτερη συμπεριφορά από τους παραπάνω, ενώ για ετερογένεια από 0.05 και μεγαλύτερη καλή συμπεριφορά παρουσιάζει ο εκτιμητής RBp και έπειτα ο SJ.

5.6.2 Σύγκριση των εκτιμητών ετερογένειας ως προς το μέσο τετραγωνικό σφάλμα

Από τους Πίνακες 5.8(α), 5.8(β) παρατηρούμε τα διαγράμματα για τη συμπεριφορά των εκτιμητών ως προς το μέσο τετραγωνικό σφάλμα.

Για την ετερογένεια μεταξύ των τιμών $0 \leq \tau^2 \leq 0.05$, το μέσο τετραγωνικό σφάλμα όλων των εκτιμητών βρίσκεται κοντά στο μηδέν, γεγονός που υποδεικνύει ότι οι εκτιμητές ετερογένειας συμπεριφέρονται καλά για τις παραπάνω τιμές ετερογένειας.

Αντίθετα, για ετερογένεια μεταξύ των τιμών $0.05 \leq \tau^2 \leq 0.5$, το μέσο τετραγωνικό σφάλμα αυξάνεται σε μεγάλο βαθμό διατηρώντας την ίδια συμπεριφορά για όλους τους εκτιμητές.

Παρατηρούμε ότι η συμπεριφορά του μέσου τετραγωνικού σφάλματος του εκτιμητή RBp είναι παρόμοια με τη συμπεριφορά της μεροληψίας του.

Για μικρές τιμές ετερογένειας $\tau^2 = 0, 0.01$, το μέσο τετραγωνικό σφάλμα εκτιμητή RBp έχει τις μεγαλύτερες τιμές σε σύγκριση με τους άλλους εκτιμητές, ενώ για μεγάλες τιμές ετερογένειας $\tau^2 = 0.05, 0.5$, ο εκτιμητής RBp έχει το μικρότερο μέσο τετραγωνικό σφάλμα.

Για μικρές τιμές ετερογένειας, ο RBp έχει το μικρότερο μέσο τετραγωνικό σφάλμα. Κατά την παρουσία ετερογένειας και για τιμές κοντά στο 0.01 καλή συμπεριφορά ως προς την ετερογένεια έχουν οι εκτιμητές DLp, HM, ενώ για τιμές μεταξύ $0.01 < \tau^2 < 0.05$, οι εκτιμητές HM, DLp, FB και SJ. Για ετερογένεια από 0.05 και μεγαλύτερη, καλή

συμπεριφορά παρουσιάζει ο εκτιμητής RBp.

5.6.3 Σύγκριση των εκτιμητών ετερογένειας ως προς την πιθανότητα κάλυψης μ

Στους Πίνακες 5.9(α), 5.9(β) παρουσιάζονται τα διαγράμματα της πιθανότητας κάλυψης του μ για τα διάφορα σενάρια. Παρατηρούμε ότι η πιθανότητα κάλυψης μειώνεται καθώς η ετερογένεια αυξάνεται με δραματική μείωση για τιμές ετερογένειας μεταξύ $0.05 < \tau^2 < 0.5$. Επομένως συμπεραίνουμε ότι οι εκτιμητές για αυτές τις τιμές ετερογένειας δεν προσαρμόζουν καλά ως προς την πιθανότητα κάλυψης του μ .

Κατά την απουσία ετερογένειας οι εκτιμητές HM, DLp, SJ, FB και RBp στις περισσότερες περιπτώσεις βρίσκονται έξω από το επιτρεπτό διάστημα, ενώ οι υπόλοιποι εκτιμητές βρίσκονται μέσα στο επιτρεπτό διάστημα. Όταν οι τιμές ετερογένειας κυμαίνονται κοντά στο 0.01 κυρίως οι εκτιμητές HM, DLp, SJ, FB βρίσκονται μέσα ή πιο κοντά στο επιτρεπτό διάστημα και έπειτα ακολουθεί ο εκτιμητής RBp, ενώ για τιμές ετερογένειας κοντά στο 0.05 πιο κοντά στο επιτρεπτό διάστημα βρίσκεται ο εκτιμητής RBp. Όλοι οι υπόλοιποι εκτιμητές, πέρα από τους HM, DLp, SJ, FB και RBp, παρουσιάζουν κοινή συμπεριφορά και προσαρμόζουν χειρότερα από τους παραπάνω κατά την παρουσία ετερογένειας.

5.6.4 Σύγκριση των εκτιμητών ετερογένειας ως προς την πιθανότητα κάλυψης τ^2

Στους Πίνακες 5.10(α), 5.10(β) παρουσιάζονται τα διαγράμματα της πιθανότητας κάλυψης του τ^2 για τα διάφορα σενάρια. Όπως και στα διαγράμματα κάλυψης του μ έτσι και στα διαγράμματα κάλυψης της ετερογένειας παρατηρούμε ότι η πιθανότητα κάλυψης μειώνεται δραματικά για τιμές ετερογένειας μεταξύ $0.05 < \tau^2 < 0.5$. Επομένως συμπεραίνουμε ότι οι εκτιμητές για αυτές τις τιμές ετερογένειας δεν προσαρμόζουν καλά ως προς την πιθανότητα κάλυψης της ετερογένειας. Ακόμα γίνεται αντιληπτό ότι η συμπεριφορά των εκτιμητών είναι παρόμοια για τα διάφορα σενάρια ως προς τη συνολική επίδραση θ και το μέγεθος των μελετών k .

Ο εκτιμητής BM έχει πολύ χαμηλή πιθανότητα κάλυψης της ετερογένειας για κάθε σενάριο και για όλες τις τιμές της ετερογένειας.

Οι εκτιμητές SJ και HM έχουν τη μεγαλύτερη πιθανότητα κάλυψης για τιμές ετερογένειας $0.01 \leq \tau^2 \leq 0.05$.

Ο εκτιμητής RBp έχει τη μεγαλύτερη πιθανότητα κάλυψης της ετερογένειας κατά την παρουσία ετερογένειας όταν έχουμε μικρό αριθμό μελετών, $k = 10$. Για μεγαλύτερο αριθμό μελετών, όταν έχουμε μικρή ετερογένεια έχει από τις πιο χαμηλές πιθανότητες κάλυψης, ενώ για ετερογένεια μεγαλύτερη από 0.5 έχει τη μεγαλύτερη πιθανότητα κάλυψης. Οι εκτιμητές DL, HE, HS, ML, REML, EB, DL2, GHO2, RBo και MBH έχουν παρόμοια συμπεριφορά. Οι παραπάνω εκτιμητές προσαρμόζουν καλά κατά την απουσία ετερογένειας και καθώς η ετερογένεια αυξάνεται, η πιθανότητα κάλυψής τους μειώνεται.

Παρόμοια συμπεριφορά με τους παραπάνω εκτιμητές αλλά με πολύ χαμηλότερη πιθανότητα κάλυψης παρουσιάζει ο εκτιμητής SJgho.

5.6.5 Σύγκριση των εκτιμητών ετερογένειας ως προς το εμπειρικό type error I

Από τους Πίνακες 5.11(α), 5.11(β) παρατηρούμε τα διαγράμματα του εμπειρικού σφάλματος τύπου I για τα διάφορα σενάρια. Για τις τιμές ετερογένειας $0.05 \leq \tau^2 \leq 0.5$, το εμπειρικό type error I αυξάνεται δραματικά για όλα τα σενάρια. Συμπεραίνουμε λοιπόν ότι οι εκτιμητές γι' αυτές τις τιμές ετερογένειας δεν προσαρμόζουν καλά ως προς το εμπειρικό type error I.

Παρατηρούμε ότι η συμπεριφορά των εκτιμητών είναι παρόμοια για τα διάφορα σενάρια ως προς τη συνολική επίδραση θ και το μέγεθος των μελετών k .

Ο εκτιμητής RBp έχει εμπειρικό σφάλμα τύπου I πιο κοντά απ' όλους τους υπόλοιπους εκτιμητές για τιμές ετερογένειας $0.01 \leq \tau^2 \leq 0.5$.

Έπειτα από τον RBp και για τις ίδιες τιμές ετερογένειας ακολουθούν κατά σειρά οι εκτιμητές SJ, DLp, HM αποτελώντας τους εκτιμητές με τιμές πιο κοντά στο επιθυμητό διάστημα. Πέρα από τους εκτιμητές RBp, SJ, DLp, HM, ο εκτιμητής FB βρίσκεται στις ίδιες τιμές όταν ο αριθμός των

μελετών είναι μικρός, $k = 10$, ενώ με την αύξηση του αριθμού των μελετών ή την αύξηση του θ , ο εκτιμητής FB έχει τιμές που κυμαίνονται με όλους τους υπόλοιπους εκτιμητές. Οι υπόλοιποι εκτιμητές έχουν καλή συμπεριφορά κατά την απουσία ετερογένειας, παρουσιάζουν παρόμοια συμπεριφορά κατά την παρουσία ετερογένειας, με τον RBo να παρουσιάζει το πιο απομακρυσμένο εμπειρικό σφάλμα τύπου I για όλες τις τιμές ετερογένειας.

5.6.6 Σύγκριση των εκτιμητών ετερογένειας ως προς την ισχύ

Από τους Πίνακες 5.12(α), 5.12(β) παρατηρούμε τα διαγράμματα για την ισχύ ως προς τις διάφορες τιμές της ετερογένειας. Παρατηρούμε ότι η συμπεριφορά των εκτιμητών είναι παρόμοια για όλα τα σενάρια. Ακόμη, η ισχύς κυμαίνεται κοντά στη μονάδα για μικρές τιμές ετερογένειας ενώ όσο η ετερογένεια αυξάνεται η ισχύς μειώνεται ελαφρώς με την αύξηση της ετερογένειας όταν ο αριθμός των μελετών είναι μικρός. Με την αύξηση του αριθμού των μελετών η ισχύς κυμαίνεται κοντά στη μονάδα για όλες τις τιμές ετερογένειας. Ακόμα παρατηρούμε ότι καθώς αυξάνεται το θ , τόσο η ισχύς κυμαίνεται όλο και πιο κοντά στη μονάδα για όλες τις τιμές της ετερογένειας.

Παρατηρούμε ότι όλοι οι εκτιμητές έχουν παρόμοιες τιμές με μια διαφοροποίηση του εκτιμητή RBp, ο οποίος για όλες τις τιμές ετερογένειας έχει τη χαμηλότερη ισχύ. Ακόμα χαμηλή ισχύ παρουσιάζουν οι εκτιμητές SJ, DLp και HM για μεγάλες τιμές ετερογένειας.

5.6.7 Σύγκριση των εκτιμητών ως προς τα επίπεδα σημαντικότητας

Όπως αναφέραμε και στην παράγραφο 5.5.7 σκοπός μας αποτελεί η σύγκριση των εκτιμητών για τις μετρήσεις της πιθανότητας κάλυψης για το μ , της πιθανότητας κάλυψης της ετερογένειας, του σφάλματος τύπου I και της ισχύς για τα επίπεδα σημαντικότητας $\alpha = 0.01, 0.05, 0.10$. Έχοντας ορίσει τα διαστήματα (l_1, u_1) , (l_2, u_2) και (l_3, u_3) για τις πιθανότητες κάλυψης και τα διαστήματα (L_1, U_1) , (L_2, U_2) , (L_3, U_3) για το type error I ως προς τα επίπεδα σημαντικότητας $\alpha_1 = 0.01$, $\alpha_2 = 0.05$ και $\alpha_3 = 0.10$ αντίστοιχα, θα εξετάσουμε εάν οι εκτιμητές προσαρμόζουν

καλά ως προς τις μετρήσεις στα διαστήματα αυτά. Από τους Πίνακες 5.17, 5.18, 5.19 και 5.20 παρατηρούμε τα διαγράμματα της πιθανότητας κάλυψης μ , της πιθανότητας κάλυψης της ετερογένειας, του σφάλματος τύπου I και της ισχύς αντίστοιχα, ως προς τις τιμές ετερογένειας για τα σενάρια με συνολική επίδραση $\theta = 0, 0.5$, αριθμό μελετών $k = 10, 30$ και επίπεδα σημαντικότητας $\alpha = 0.01, 0.05, 0.10$ όταν έχουμε συνεχή αποτελέσματα. Όπως αναφέραμε και παραπάνω είναι φυσικό οι πιθανότητες κάλυψης, το εμπειρικό σφάλμα τύπου I και η ισχύς να έχουν διαφορετικές τιμές για τα διάφορα επίπεδα σημαντικότητας και σκοπός μας αποτελεί η εξέταση της ύπαρξης διαφορετικής συμπεριφοράς από τους εκτιμητές. Παρατηρούμε ότι οι εκτιμητές έχουν παρόμοια συμπεριφορά για την πιθανότητα κάλυψης του μ για τα διάφορα επίπεδα σημαντικότητας.

5.6.7.1 Σύγκριση ως προς την πιθανότητα κάλυψης του μ

Από τους Πίνακες 5.17(α), 5.17(β) παρατηρούμε τα διαγράμματα της πιθανότητας κάλυψης του μ ως προς τις τιμές της ετερογένειας και συμπεραίνουμε ότι οι εκτιμητές έχουν παρόμοια συμπεριφορά για την πιθανότητα κάλυψης του μ για τα διάφορα επίπεδα σημαντικότητας.

5.6.7.2 Σύγκριση ως προς την πιθανότητα κάλυψης της ετερογένειας

Από τους Πίνακες 5.18(α), 5.18(β) παρατηρούμε τα διαγράμματα της πιθανότητας κάλυψης της ετερογένειας ως προς τις τιμές της ετερογένειας και συμπεραίνουμε ότι οι εκτιμητές έχουν παρόμοια συμπεριφορά.

5.6.7.3 Σύγκριση ως προς το εμπειρικό type error I

Από τους Πίνακες 5.19(α), 5.19(β) παρατηρούμε τα διαγράμματα του εμπειρικού type error I ως προς τις τιμές της ετερογένειας και συμπεραίνουμε ότι οι εκτιμητές έχουν παρόμοια συμπεριφορά.

5.6.7.4 Σύγκριση ως προς την ισχύ

Από τον Πίνακα 5.20 παρατηρούμε τα διαγράμματα της ισχύς ως προς τις τιμές της ετερογένειας και συμπεραίνουμε ότι οι εκτιμητές έχουν παρό-

μοια συμπεριφορά.

5.7 Συμπεράσματα

Έπειτα από τις συγκρίσεις που κάναμε για τους εκτιμητές θα εξάγουμε ένα γενικό συμπέρασμα για τη συμπεριφορά των εκτιμητών ως προς όλες τις μετρήσεις.

5.7.1 Συμπεράσματα για τα διχότομα δεδομένα

Όσον αφορά τα δεδομένα με διχότομο αποτέλεσμα, συμπεραίνουμε ότι όλοι οι εκτιμητές συμπεριφέρονται χειρότερα καθώς αυξάνεται η ετερογένεια, ενώ συμπεριφέρονται καλύτερα καθώς αυξάνεται ο αριθμός των μελετών. Παρατηρούμε ότι με τη μείωση του αριθμού των μελετών ή την αύξηση της ετερογένειας αυξάνεται η μεροληψία, το μέσο τετραγωνικό σφάλμα, οι πιθανότητες κάλυψης και το εμπειρικό *type error* I τείνουν να κυμαίνονται πιο μακριά από το επιτρεπτό διάστημα και η ισχύς μειώνεται. Ακόμα, μείωση της ισχύς παρατηρούμε και από την αύξηση της συνολικής επίδρασης. Ο εκτιμητής RBp έχει καλή συμπεριφορά μόνο για μικρό αριθμό μελετών $k = 10$ και για μεγάλη ετερογένεια κοντά στο 0.03, ενώ στις υπόλοιπες περιπτώσεις παρουσιάζει τη χειρότερη συμπεριφορά. Ο εκτιμητής FB έχει μεγάλη μεροληψία, χαμηλή ισχύ ενώ παράλληλα βρίσκεται μέσα στο επιτρεπτό διάστημα για την πιθανότητα κάλυψης και το εμπειρικό σφάλμα τύπου I για μεγάλη ετερογένεια και όταν ο αριθμός των μελετών είναι μικρός, $k = 10$. Όσο αυξάνει ο αριθμός των μελετών τόσο πιο καλή συμπεριφορά παρουσιάζει ο εκτιμητής και για τις μικρότερες τιμές ετερογένειας. Έτσι λοιπόν για μεγάλο αριθμό μελετών παρουσιάζει καλή συμπεριφορά. Ο εκτιμητής HM έχει την καλύτερη συμπεριφορά από τους υπόλοιπους εκτιμητές για μεσαία ετερογένεια κοντά στην τιμή 0.07. Ο εκτιμητής SJ έχει καλή συμπεριφορά μόνο για μεγάλη ετερογένεια. Η συμπεριφορά του εκτιμητή EB κυμαίνεται με τη συμπεριφορά των περισσότερων εκτιμητών ως προς όλες τις μετρήσεις. Όσο αφορά την πιθανότητα κάλυψης της ετερογένειας παρατηρούμε ότι συμπεριφέρεται καλύτερα από τους άλλους εκτιμητές καθώς προσαρμόζει καλύτερα στο επιτρεπτό διάστημα για μεσαίο και μεγάλο αριθμό μελετών. Οι εκτιμητές Knapp and Hartung έχουν την

καλύτερη συμπεριφορά από τους υπόλοιπους εκτιμητές ως προς τις πιθανότητες κάλυψης και το εμπειρικό σφάλμα τύπου I. Οι θετικοί εκτιμητές BM, RBr, SJ, HM έχουν καλύτερη συμπεριφορά όπως είναι φυσικό κατά την παρουσία ετερογένειας το οποίο συμπεραίνουμε και αν παρατηρήσουμε την πιθανότητα κάλυψης της ετερογένειας. Ο εκτιμητής DLb έχει γενικά από την πιο χαμηλή μεροληψία και μέσο τετραγωνικό σφάλμα κατά την παρουσία ετερογένειας. Βρίσκεται μέσα στο επιτρεπτό διάστημα της πιθανότητας κάλυψης της ετερογένειας και του εμπειρικού σφάλματος τύπου I εκτός από κάποιες περιπτώσεις μεγάλης ετερογένειας και κατά την απουσία ετερογένειας και έχει πολύ υψηλή ισχύ. Επομένως ο εκτιμητής DLb έχει πολύ καλή συμπεριφορά εκτός από κάποιες περιπτώσεις της ύπαρξης μεγάλης ετερογένειας. Ο εκτιμητής SJgho παρουσιάζει μέτρια συμπεριφορά καθώς έχει μεσαίες τιμές ως προς τους υπόλοιπους εκτιμητές ως προς τη μεροληψία, το μέσο τετραγωνικό σφάλμα και την ισχύ. Παρουσιάζει πολύ χαμηλή πιθανότητα κάλυψης της ετερογένειας και βρίσκεται μέσα στο διάστημα κατά πλειοψηφία όσον αφορά την πιθανότητα κάλυψης της συνολικής επίδρασης και του εμπειρικού σφάλματος τύπου I. Τέλος, παρατηρούμε ότι σε γενικά πλαίσια όλοι οι εκτιμητές έχουν παρόμοια συμπεριφορά ως προς τα διάφορα επίπεδα σημαντικότητα.

5.7.2 Συμπεράσματα για τα συνεχή δεδομένα

Όσον αφορά τα δεδομένα με συνεχή αποτέλεσμα, συμπεραίνουμε ότι κατά την αύξηση της ετερογένειας, αυξάνεται η μεροληψία, το μέσο τετραγωνικό σφάλμα, μειώνεται η ισχύς και παίρνουμε τιμές έξω από το επιτρεπτό διάστημα για τις πιθανότητες κάλυψης και για το εμπειρικό σφάλμα τύπου I. Παρατηρούμε ότι για ετερογένεια κοντά στο 0.05 και μεγαλύτερη το γεγονός αυτό αυξάνεται δραματικά και επομένως για μεγάλη ετερογένεια δεν έχουμε καλή προσαρμογή των εκτιμητών. Κατά την παρουσία ετερογένειας και για τιμές μέχρι 0.05 καλή συμπεριφορά ως προς την ετερογένεια έχουν οι εκτιμητές DLp, HM, FB και SJ, με τον SJ να προσαρμόζει λίγο καλύτερα από τους παραπάνω. Για μεγάλη ετερογένεια με τιμές μεγαλύτερες από 0.05, ο εκτιμητής RBr παρουσιάζει τη μικρότερη μεροληψία, το μικρότερο μέσο τετραγωνικό σφάλμα,

βρίσκεται πιο κοντά στις πιθανότητες κάλυψης αλλά έχει τη μικρότερη ισχύ. Επομένως ο εκτιμητής RBp έχει καλή συμπεριφορά για μεγάλη ετερογένεια αλλά δεν είναι ισχυρός. Οι εκτιμητές DL, HE, HS, ML, REML, EB, DL2, GHO2, RBo, BM, SJgho και MBH προσαρμόζουν καλά κατά την απουσία ετερογένειας. Οι παραπάνω εκτιμητές έχουν παρόμοια συμπεριφορά για όλες τις τιμές ετερογένειας με τους BM και SJgho να παρουσιάζουν πιο μέτρια συμπεριφορά κατά την πιθανότητα κάλυψης της ετερογένειας. Οι εκτιμητές Knapp and Hartung κυμαίνονται στις τιμές που έχουν οι παραπάνω εκτιμητές, επομένως για τα συνεχή αποτελέσματα δεν κάνουν την καλύτερη προσαρμογή καθώς πιο κοντά στα επιθυμητά διαστήματα βρίσκονται οι εκτιμητές DLp, HM, FB και SJ. Τέλος παρατηρούμε ότι όλοι οι εκτιμητές έχουν παρόμοια συμπεριφορά ως προς τα διάφορα επίπεδα σημαντικότητας.

5.8 Γενικό συμπέρασμα

Σκοπός της προσομοίωσης αποτέλεσε η σύγκριση όλων των εκτιμητών ετερογένειας με τις μετρήσεις εκτίμησης της μεροληψίας, του μέσου τετραγωνικού σφάλματος, τις πιθανότητες κάλυψης για την συνολική επίδραση και την ετερογένεια, του εμπειρικού σφάλματος τύπου I και της ισχύς. Έπειτα από την πραγματοποίηση της μελέτης εξαγάγαμε σημαντικά αποτελέσματα. Βασικό συμπέρασμα αποτελεί ότι οι εκτιμητές έχουν διαφορετική συμπεριφορά ως προς τα δύο είδη αποτελεσμάτων, δηλαδή ως προς τα διχότομα και τα συνεχή. Ως προς τα δύο είδη αποτελεσμάτων η συμπεριφορά των εκτιμητών χειροτερεύει με την αύξηση της ετερογένειας. Με την αύξηση του αριθμού των μελετών, όλοι οι εκτιμητές συμπεριφέρονται καλύτερα στα διχότομα αποτελέσματα, ενώ κάτι τέτοιο δεν είναι εμφανές στα συνεχή αποτελέσματα. Ακόμα, από τη συμπεριφορά των εκτιμητών για τα συνεχή αποτελέσματα παρατηρούμε ότι δεν συμπεριφέρονται καθόλου καλά για ετερογένεια από 0.05 και μεγαλύτερη, καθώς έχουμε δραματική αλλαγή της συμπεριφοράς ως προς όλες τις μετρήσεις (μεγάλη μεροληψία, μέσο τετραγωνικό σφάλμα καθώς και πιθανότητες κάλυψης και εμπειρικού σφάλματος τύπου I πολύ πιο μακριά από τα επιτρεπτά διαστήματα). Επιπρόσθετα, με την αύξηση του θ έχουμε πιο ισχυρούς εκτιμητές και για τα δύο είδη αποτελεσμάτων.

Επιπλέον, όλοι οι εκτιμητές έχουν παρόμοια συμπεριφορά για τα διάφορα επίπεδα σημαντικότητας για τα δύο είδη αποτελεσμάτων.

Για την συνολική εικόνα ως προς τα διχότομα και τα συνεχή αποτελέσματα παραθέτουμε τους Πίνακες 5.21 και 5.22 αντίστοιχα. Από τον Πίνακα 5.22 παρατηρούμε ότι στα συνεχή αποτελέσματα οι εκτιμητές DLp, HM, SJ, RBr και FB έχουν καλύτερη συμπεριφορά από τους υπόλοιπους εκτιμητές οι οποίοι παρουσιάζουν παρόμοια συμπεριφορά. Το γεγονός βέβαια αυτό δεν σημαίνει ότι οι υπόλοιποι εκτιμητές δεν προσαρμόζουν καλά, αλλά ότι έχουν χειρότερη προσαρμογή.

Οι εκτιμητές Knapp and Hartung έχουν την καλύτερη συμπεριφορά όπως έχει αποδειχθεί ως προς τα διαστήματα εμπιστοσύνης, της πιθανότητας κάλυψης και του εμπειρικού σφάλματος τύπου I για τα διχότομα αποτελέσματα, ενώ για τα συνεχή αποτελέσματα οι εκτιμητές Knapp and Hartung κυμαίνονται με τους περισσότερους εκτιμητές.

Ο εκτιμητής FB έχει καλή συμπεριφορά καθώς αυξάνεται ο αριθμός των μελετών και για τα δύο είδη αποτελεσμάτων.

Ο εκτιμητής SJ δεν έχει καθόλου καλή συμπεριφορά για ετερογένεια μέχρι 0.07, ενώ για μεγαλύτερη ετερογένεια έχει καλή συμπεριφορά για τα διχότομα αποτελέσματα. Ως προς τα συνεχή αποτελέσματα ο εκτιμητής SJ έχει την καλύτερη συμπεριφορά.

Ο HM έχει την καλύτερη συμπεριφορά για μέτρια ετερογένεια κοντά στο 0.07 για τα διχότομα και από τις καλύτερες συμπεριφορές για τα συνεχή αποτελέσματα.

Ο εκτιμητής EB κυμαίνεται με τους περισσότερους εκτιμητές παρουσιάζοντας γενικά καλή προσαρμογή για τα διχότομα, ενώ για τα συνεχή απλά κυμαίνεται με τους περισσότερους εκτιμητές.

Ο RBr δεν έχει καθόλου καλή συμπεριφορά για τα διχότομα, ενώ για τα συνεχή παρουσιάζει μέτρια συμπεριφορά καθώς έχει θετικές αλλά και αρνητικές συμπεριφορές.

Οι εκτιμητές DL, GHO, DL2, GHO2 και REML κυμαίνονται με τους περισσότερους παρουσιάζοντας θετική συμπεριφορά σ' όλες τις μετρήσεις για τα διχότομα και για τα συνεχή αποτελέσματα.

Ο DLp κυμαίνεται με τους περισσότερους εκτιμητές για τα διχότομα και παρουσιάζει καλή συμπεριφορά για τα συνεχή μαζί με τον εκτιμητή HM

έπειτα από τον SJ.

Ο εκτιμητής BM δεν έχει καλή συμπεριφορά για τα διχότομα καθώς έχει πολύ χαμηλή πιθανότητα κάλυψης για την ετερογένεια και μέτρια συμπεριφορά ως προς τις άλλες μετρήσεις παρουσιάζοντας θετική αλλά και αρνητική συμπεριφορά όπως υψηλή μεροληψία, μέσο τετραγωνικό σφάλμα και πιθανότητα κάλυψης της ετερογένειας, ενώ για τα συνεχή κυμαίνεται με τους περισσότερους εκτιμητές.

Ο εκτιμητής HE κυμαίνεται με τους περισσότερους εκτιμητές και για τα δύο αποτελέσματα, παρουσιάζοντας αρνητική συμπεριφορά ως προς τη μεροληψία και το μέσο τετραγωνικό σφάλμα για τα διχότομα αποτελέσματα.

Οι εκτιμητές HS και ML κυμαίνονται με τους περισσότερους εκτιμητές και για τα δύο αποτελέσματα, παρουσιάζοντας αρνητική συμπεριφορά ως προς την πιθανότητα κάλυψης της ετερογένειας για τα διχότομα αποτελέσματα.

Ο εκτιμητής SJgho παρουσιάζει μέτρια συμπεριφορά καθώς κυμαίνεται με τους περισσότερους εκτιμητές και για τα δύο αποτελέσματα, παρουσιάζοντας αρνητική συμπεριφορά ως προς τη μεροληψία και το μέσο τετραγωνικό σφάλμα και την πιθανότητα κάλυψης της ετερογένειας για τα διχότομα αποτελέσματα.

Ο εκτιμητής RBo κυμαίνεται με τους περισσότερους εκτιμητές και για τα δύο αποτελέσματα, παρουσιάζοντας αρνητική συμπεριφορά ως προς τις πιθανότητες κάλυψης και του εμπειρικού σφάλματος τύπου I για τα διχότομα αποτελέσματα.

Ο εκτιμητής DLb έχει την πιο μικρή μεροληψία και καλή συμπεριφορά για τις υπόλοιπες συγκρίσεις για τα διχότομα αποτελέσματα ενώ για τα συνεχή κυμαίνεται με την πλειοψηφία.

Γενικά κατά την απουσία ετερογένειας, οι περισσότεροι εκτιμητές προσαρμόζουν καλύτερα, με τον ML και έπειτα τον HS να συμπεριφέρονται καλύτερα για όλες τις μετρήσεις από όλους τους εκτιμητές όταν τα αποτελέσματα είναι διχότομα. Κατά την απουσία ετερογένειας για τα συνεχή αποτελέσματα οι εκτιμητές DLp, HM, FB και SJ που αποτελούν τους καλύτερους για την ύπαρξη της ετερογένειας μαζί με τον εκτιμητή RBp παρουσιάζουν τη χειρότερη συμπεριφορά. Όπως και στα

συνεχή αποτελέσματα οι περισσότεροι εκτιμητές προσαρμόζουν καλύτερα, με τον ML και έπειτα τον HS να συμπεριφέρονται καλύτερα για όλες τις μετρήσεις απ' όλους τους εκτιμητές.

Από τα παραπάνω όταν τα αποτελέσματα είναι διχότομα προτείνουμε την χρήση του εκτιμητή DLb για όλες τις τιμές κατά την ύπαρξη ετερογένειας, και τον HM για μέτρια ετερογένεια κοντά στο 0.07.

Όταν τα αποτελέσματα είναι συνεχή προτείνουμε τους εκτιμητές SJ και έπειτα τους DLr και HM για όλες τις τιμές κατά την ύπαρξη ετερογένειας. Κατά την απουσία της ετερογένειας προτείνουμε τη χρήση του εκτιμητή ML ή του HS και για τα δύο είδη αποτελεσμάτων.

	Bias	MSE	Coverage probability for μ	Coverage probability for τ^2	Type error I	power
DL	⊕	⊕	⊕	⊕	⊕	⊕
EB	⊕	⊕	⊕	⊕	⊕	⊕
DLp	⊕	⊕	⊕	–	⊕	⊕
GHO	⊕	⊕	⊕	⊕	⊕	⊕
DL2	⊕	⊕	⊕	⊕	⊕	⊕
GHO2	⊕	⊕	⊕	⊕	⊕	⊕
HM	⊕	⊕	⊕	⊕	⊕	⊕
HE	×	×	⊕	⊕	⊕	⊕
HS	⊕	⊕	⊕	×	⊕	⊕
ML	⊕	⊕	⊕	×	⊕	⊕
REML	⊕	⊕	⊕	⊕	⊕	⊕
SJ	×	×	×	×	×	×
SJgho	×	×	⊕	×	⊕	⊕
DLb	⊕	⊕	⊕	–	⊕	⊕
RBo	⊕	⊕	×	×	×	⊕
RBp	×	×	×	×	×	×
FB	×	×	⊕	–	⊕	⊕
BM	×	×	⊕	×	⊕	⊕
DLkha	⊕	⊕	⊕	⊕	⊕	×
GHOkha	⊕	⊕	⊕	⊕	⊕	×
MLkha	⊕	⊕	⊕	⊕	⊕	×
REMLkha	⊕	⊕	⊕	⊕	⊕	×

Πίνακας 5.21: Συνολική συμπεριφορά των εκτιμητών για διχότομο αποτέλεσμα

Υπόμνημα πίνακα 5.21	
⊕	καλή συμπεριφορά μόνο για μέτρια ετερογένεια κοντά στο 0.07
⊕	καλή συμπεριφορά για όλες τις τιμές ετερογένειας
⊕	καλή συμπεριφορά καθώς αυξάνεται ο αριθμός των μελετών
–	δεν πραγματοποιήθηκε η μέτρηση (χρονικά ανέφικτο ή σφάλμα αποκοπής για τον DLp)
×	αρνητική συμπεριφορά για όλες τις τιμές ετερογένειας
×	αρνητική συμπεριφορά για ετερογένεια μέχρι 0.07 (για μεγαλύτερη ετερογένεια έχει καλή συμπεριφορά)

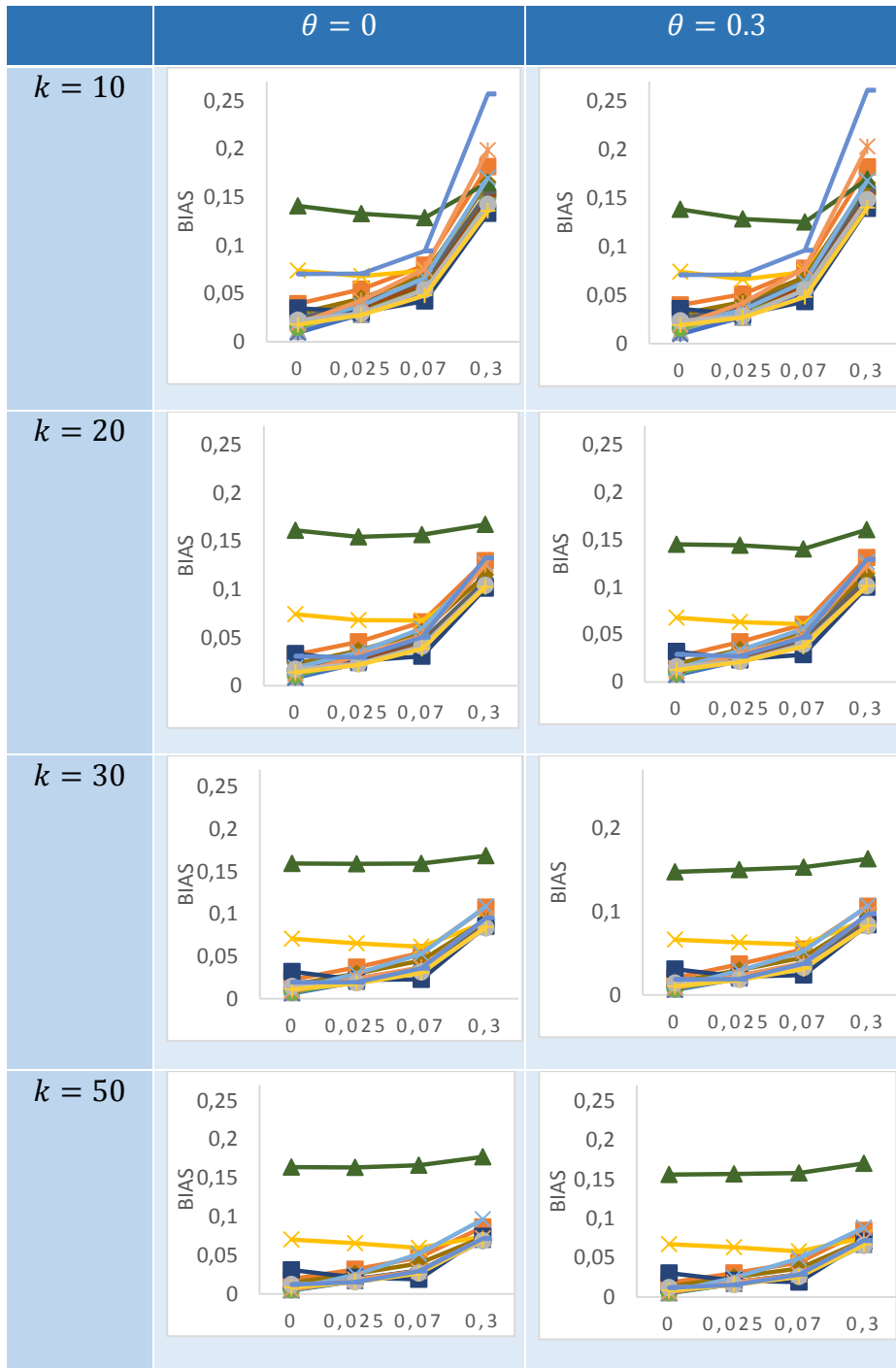
	Bias	MSE	Coverage probability for μ	Coverage probability for τ^2	Type error I	power
DL	~	~	~	~	~	⊕
EB	~	~	~	~	~	⊕
DLp	⊕	⊕	⊕	—	⊕	⊕
GHO	~	~	~	~	~	⊕
DL2	~	~	~	~	~	⊕
GHO2	~	~	~	~	~	⊕
HM	⊕	⊕	⊕	⊕	⊕	⊕
HE	~	~	~	~	~	⊕
HS	~	~	~	~	~	⊕
ML	~	~	~	~	~	⊕
REML	~	~	~	~	~	⊕
SJ	⊕	⊕	⊕	⊕	⊕	⊕
SJgho	~	~	~	~	~	⊕
DLb	~	~	~	—		⊕
RBo	~	~	~	~	~	⊕
RBp	×	×	⊕	⊕	⊕	⊕
FB	⊕	⊕	⊕	—	⊕	⊕
BM	~	~	~	~	~	⊕
MBH	~	~	~	~	~	⊕
DLknh	~	~	~	~	~	⊕
GHOknh	~	~	~	~	~	⊕
MLknh	~	~	~	~	~	⊕
REMLknh	~	~	~	~	~	⊕

Πίνακας 5.22: Συνολική συμπεριφορά των εκτιμητών για συνεχή αποτέλεσμα

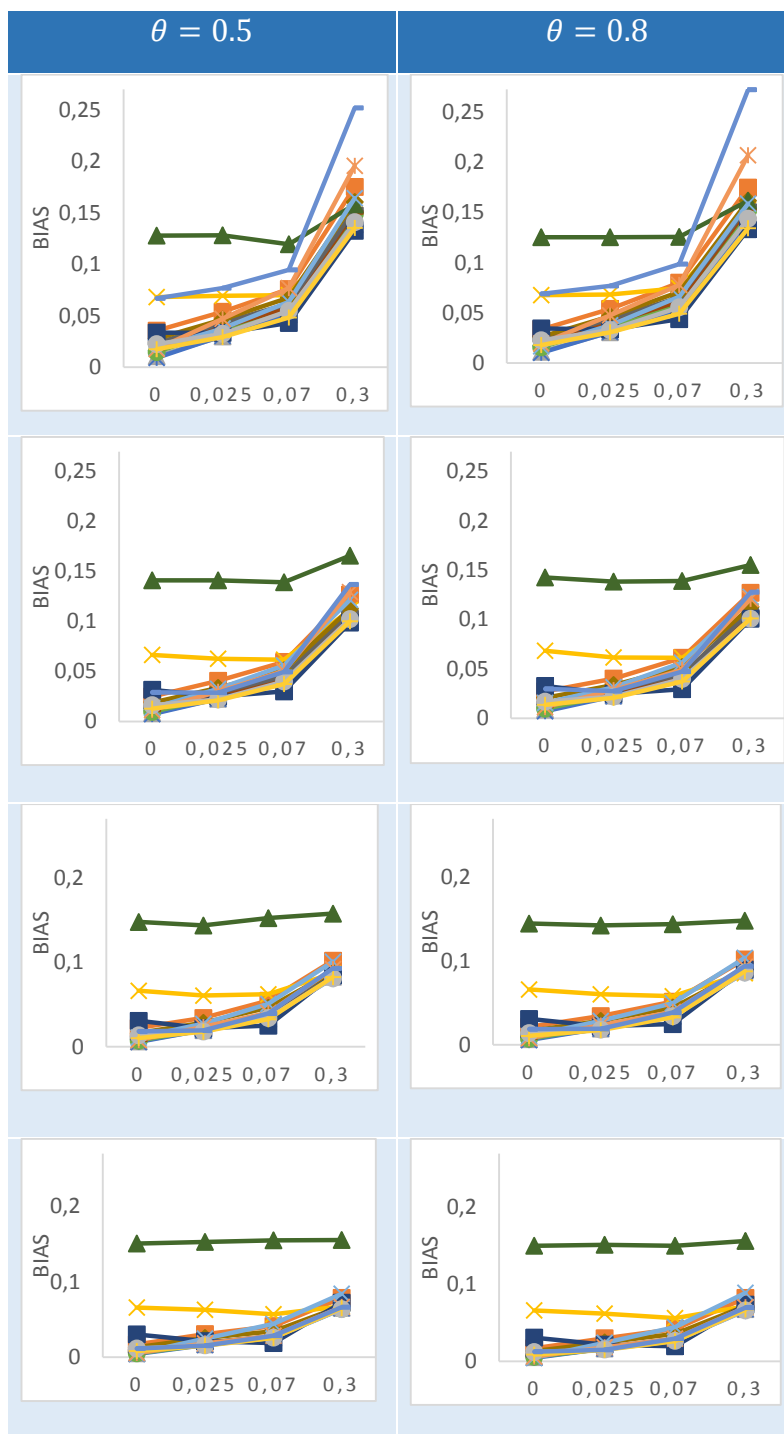
Υπόμνημα πίνακα 5.22	
⊕	καλή συμπεριφορά για όλες τις τιμές ετερογένειας
⊕	καλή συμπεριφορά για τιμές της ετερογένειας μέχρι 0.05
⊕	καλή συμπεριφορά καθώς μειώνεται ο αριθμός των μελετών
—	δεν πραγματοποιήθηκε η μέτρηση (χρονικά ανέφικτο ή σφάλμα αποκοπής για τον DLp)
~	παρόμοια συμπεριφορά με τους περισσότερους εκτιμητές για όλες τις τιμές ετερογένειας
×	την πιο αρνητική συμπεριφορά για ετερογένεια μέχρι και 0.01 (για μεγαλύτερη ετερογένεια έχει καλή συμπεριφορά)

5.9 Διαγράμματα

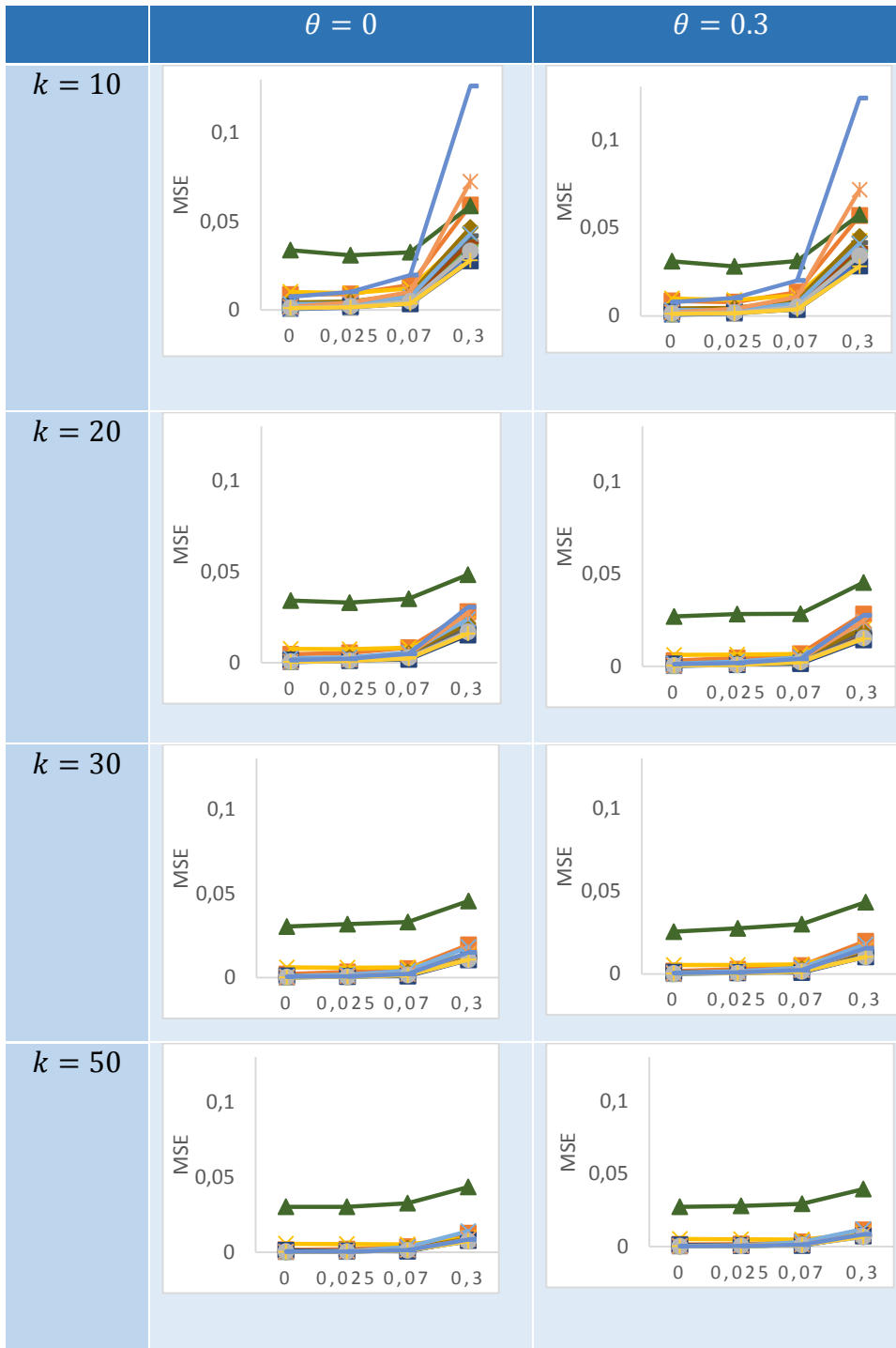
Στην ενότητα αυτή παρατίθενται όλα τα διαγράμματα για τα διχότομα και συνεχή αποτελέσματα για τα διάφορα σενάρια κατά τα οποία πραγματοποιήθηκε η προσομοίωση. Στον οριζόντιο άξονα παριστάνονται οι πραγματικές τιμές της ετερογένειας για κάθε περίπτωση, ενώ στον κατακόρυφο οι διάφορες τιμές μέτρησης (μεροληψία, μέσο τετραγωνικό σφάλμα, πιθανότητες κάλυψης, σφάλμα τύπου I και ισχύς). Κάθε γραμμή αντιπροσωπεύει και έναν διαφορετικό εκτιμητή ετερογένειας. Στο τέλος της ενότητας 5.9 παρατίθενται υπομνήματα για την καλύτερη κατανόηση των διαγραμμάτων.



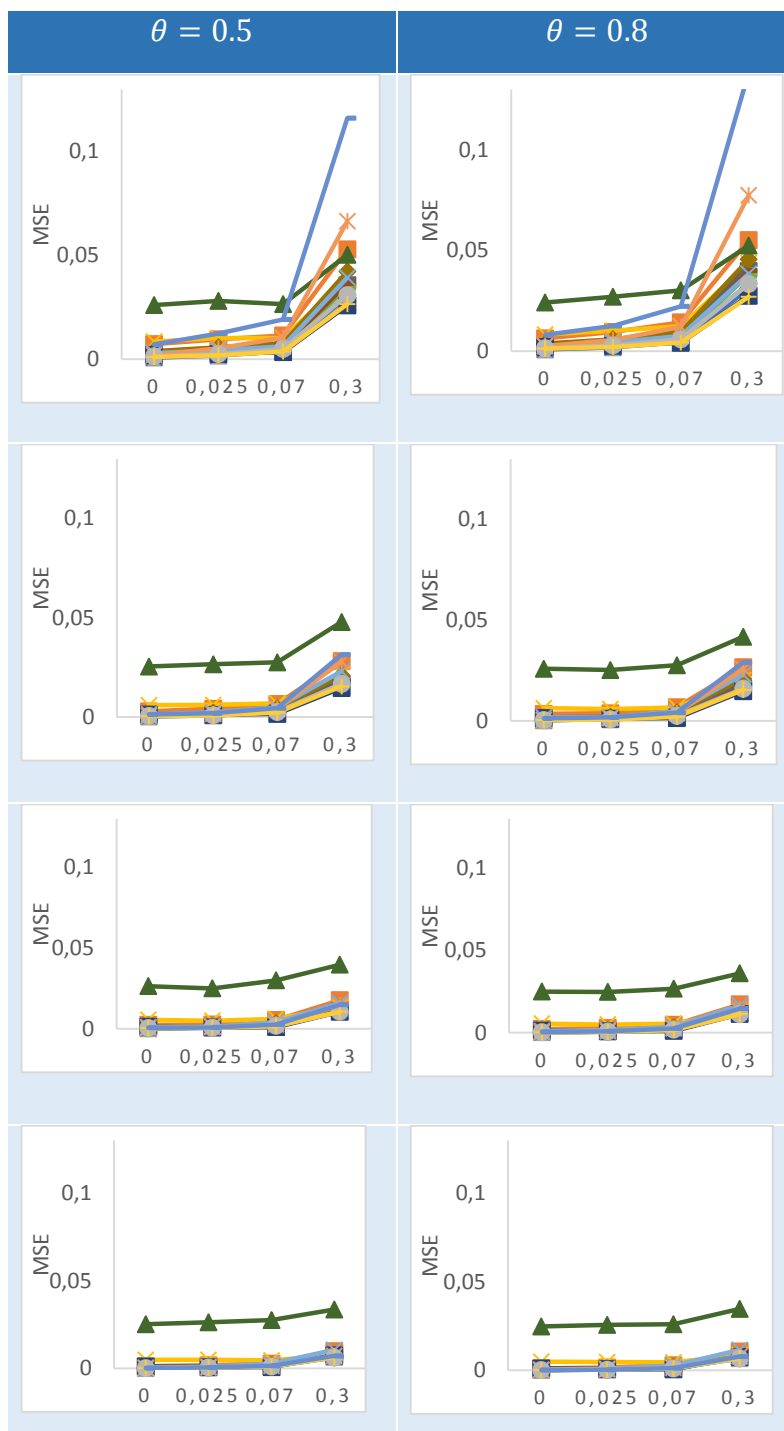
Πίνακας 5.1(α): Μεροληψία (bias) για τα διχότομα αποτελέσματα



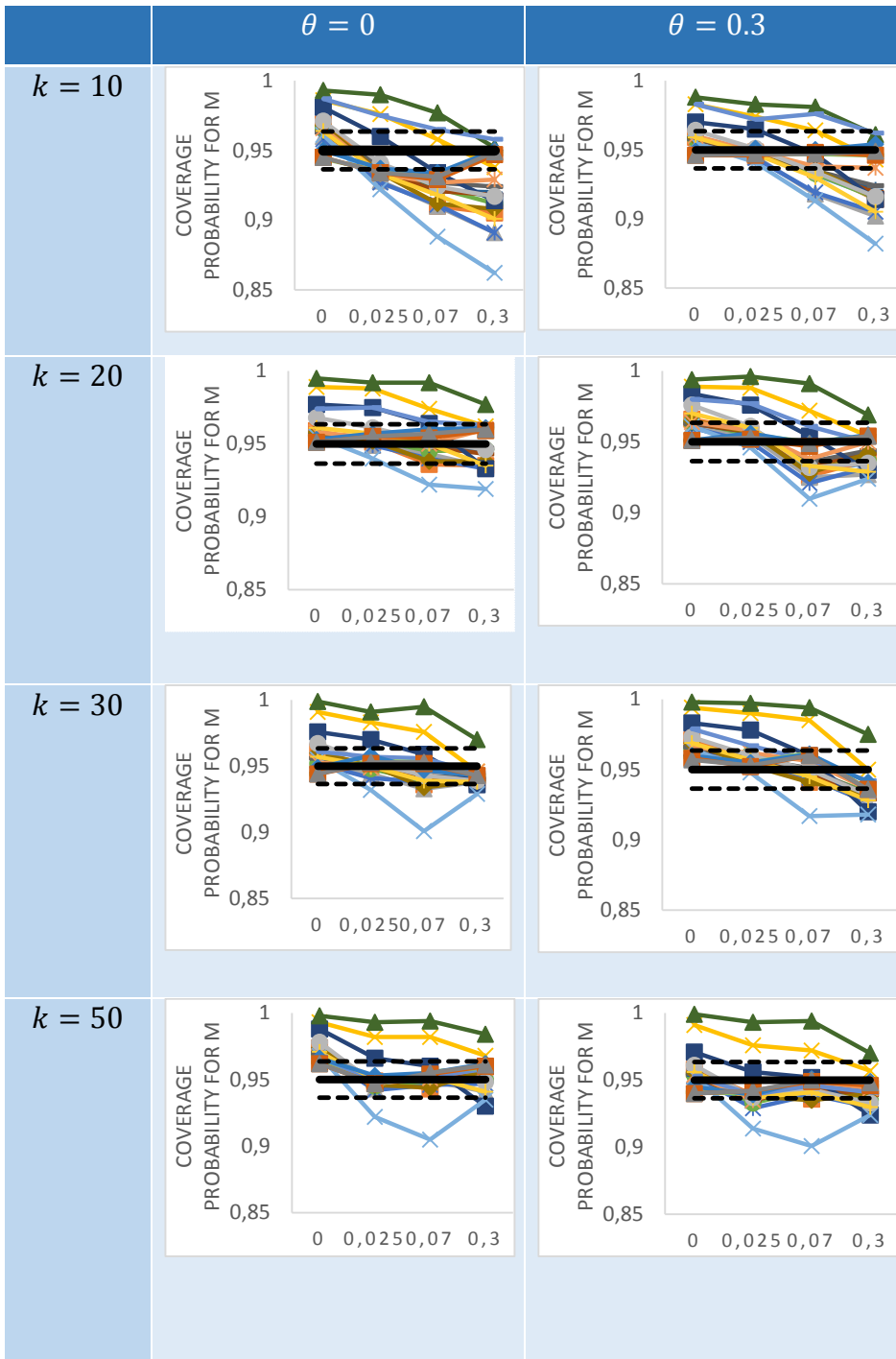
Πίνακας 5.1(β): Μεροληψία(bias) για τα διχότομα αποτελέσματα



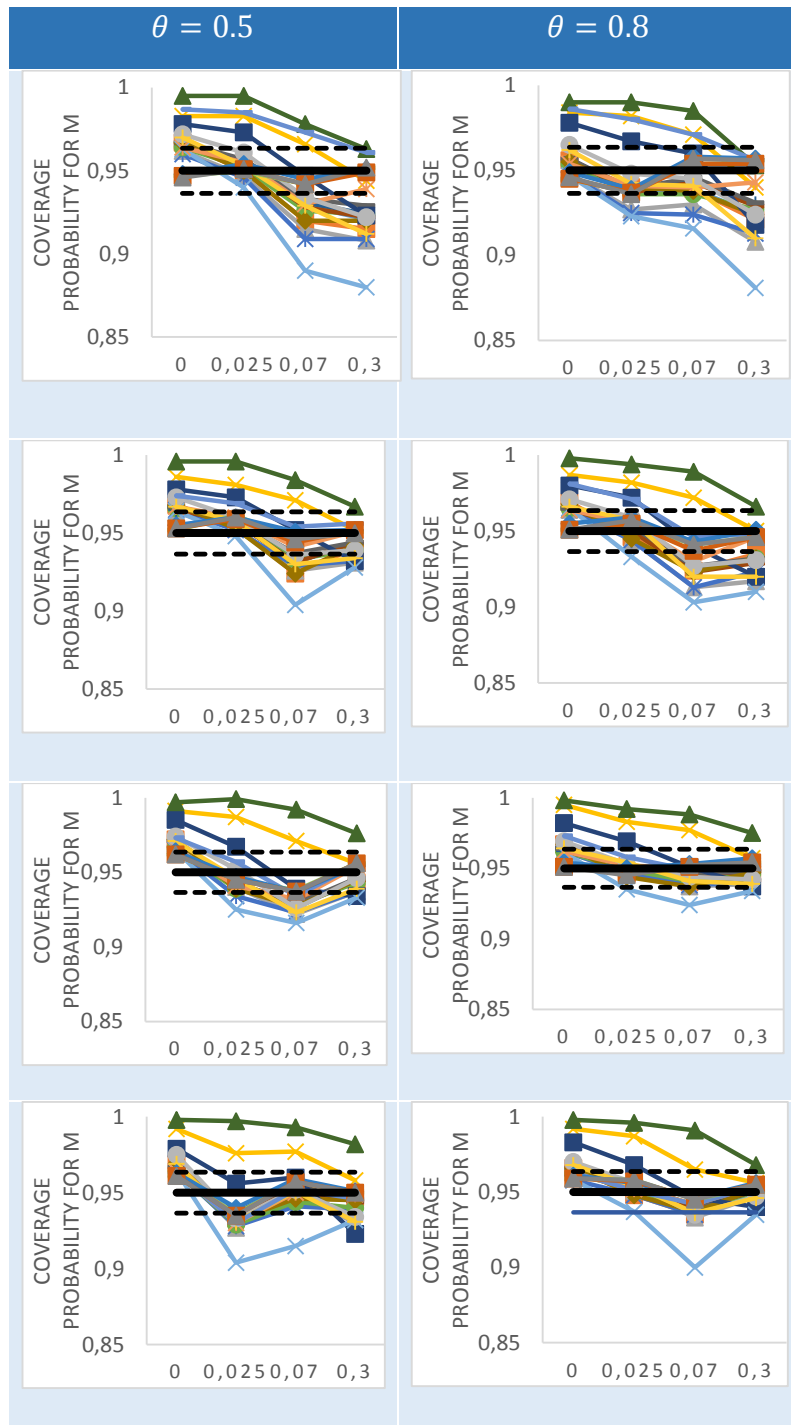
Πίνακας 5.2(α) : MSE για τα διχότομα αποτελέσματα



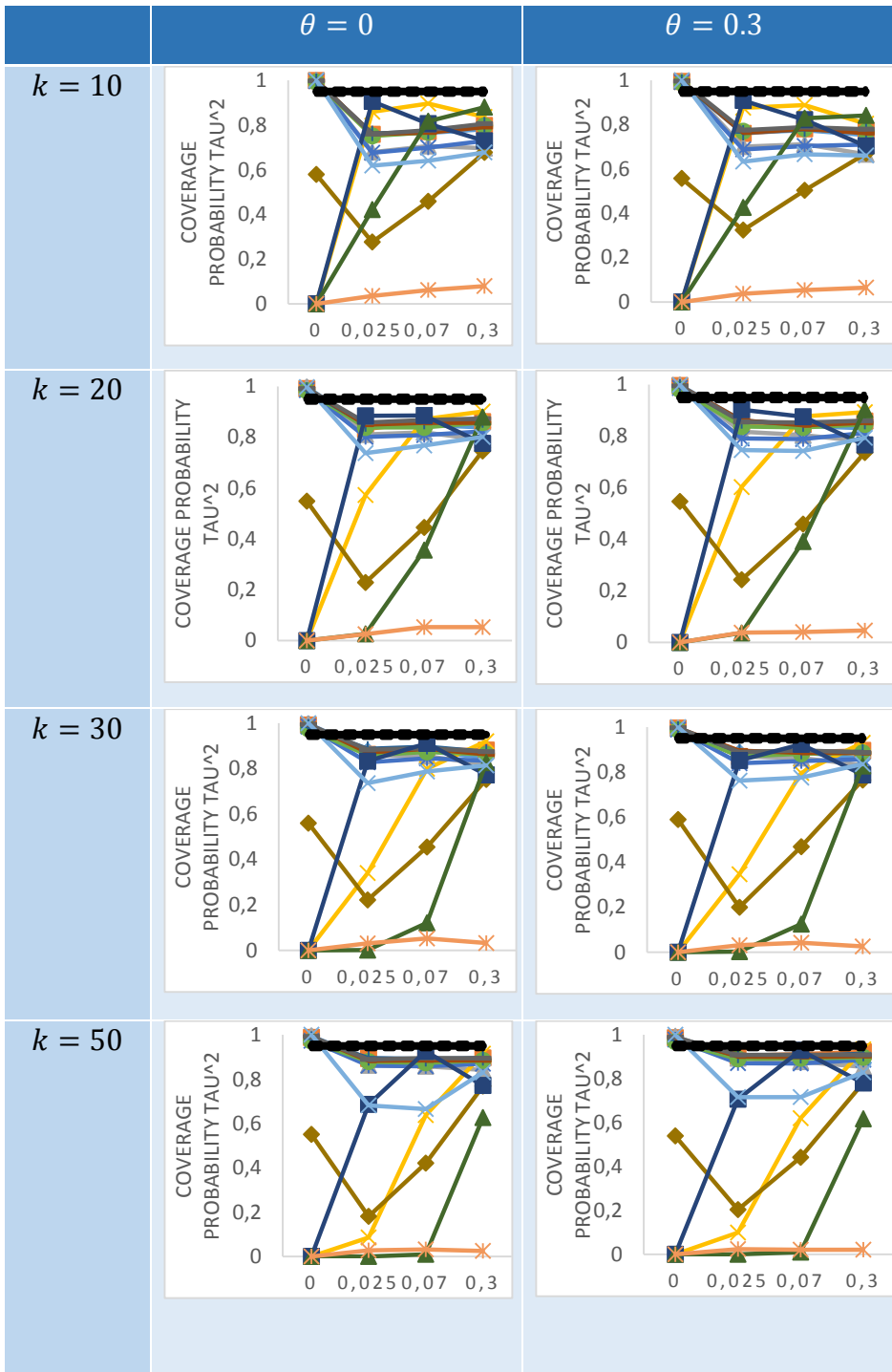
Πίνακας 5.2(β): MSE για τα διχότομα αποτελέσματα



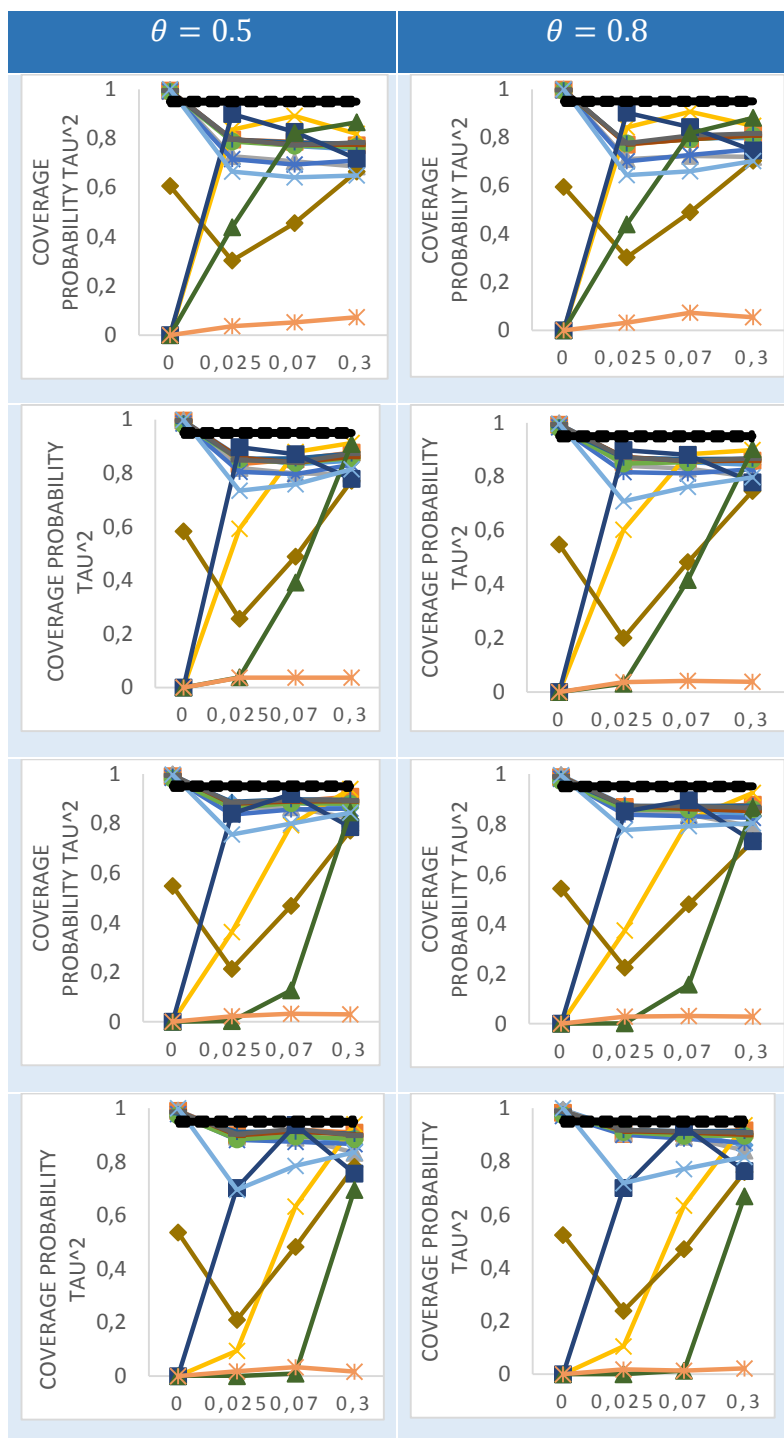
Πίνακας 5.3(α): Πιθανότητα κάλυψης για το μ για τα διχότομα αποτελέσματα



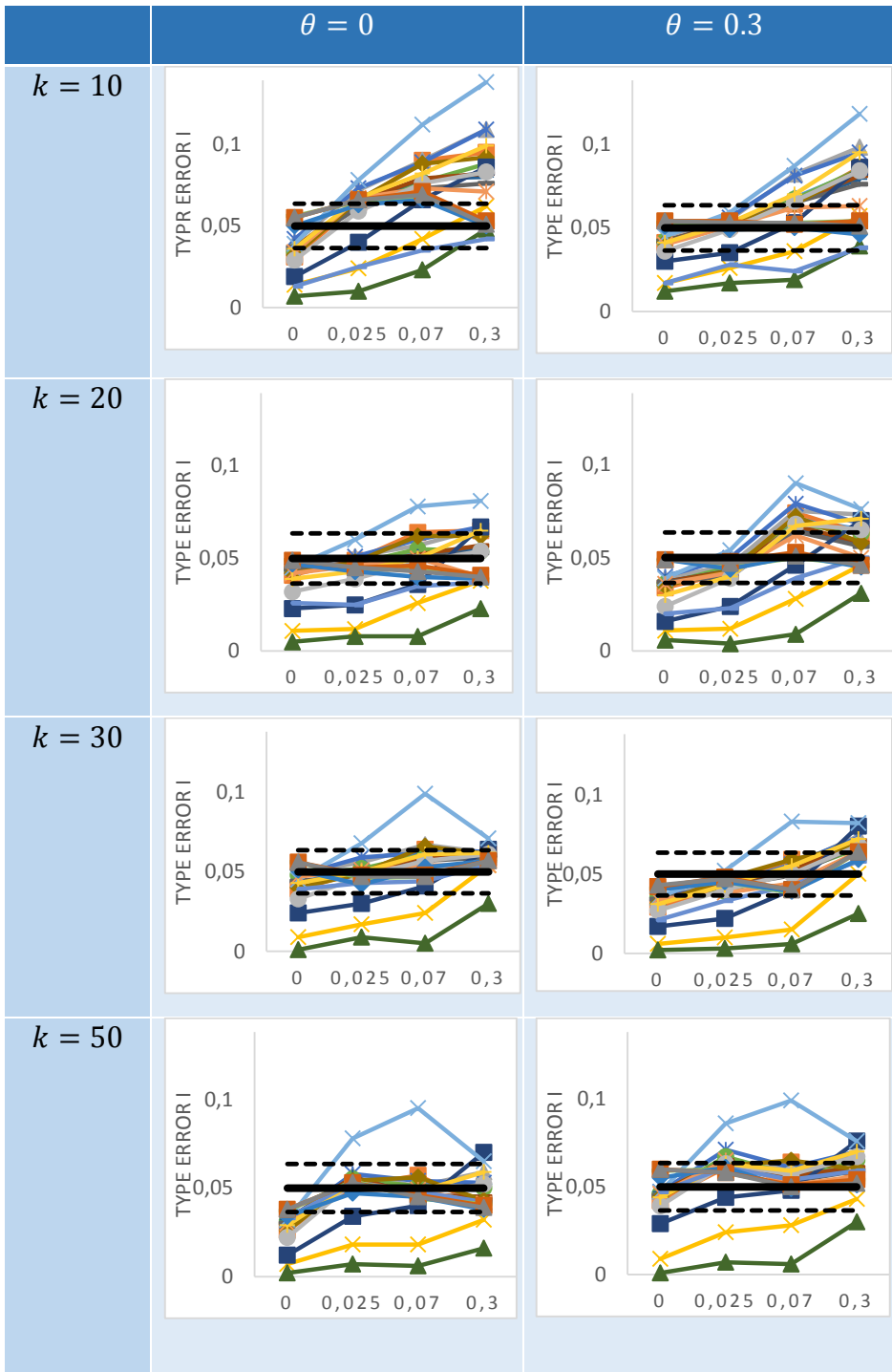
Πίνακας 5.3(β): Πιθανότητα κάλυψης για το μ για τα διχότομα αποτελέσματα



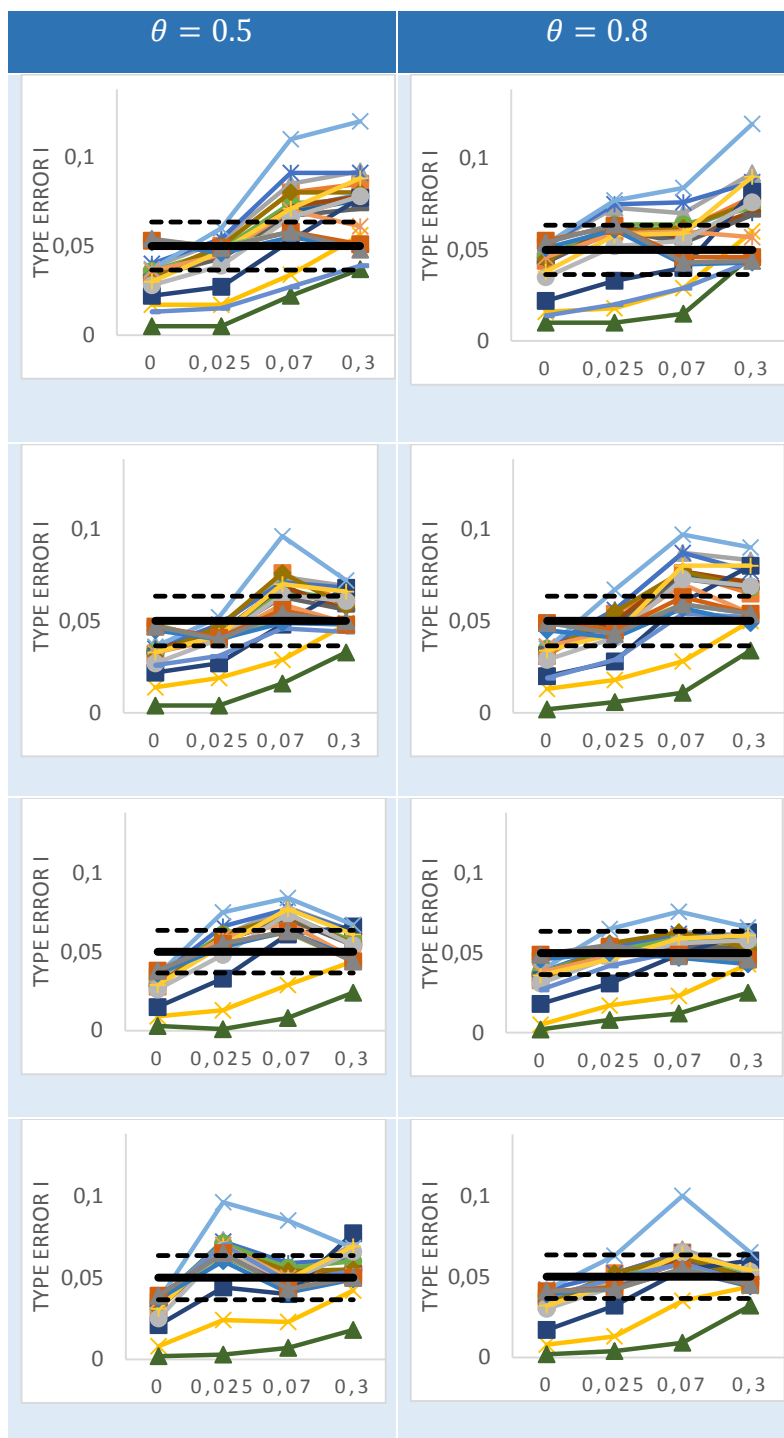
Πίνακας 5.4(α): Πιθανότητα κάλυψης για το τ^2 για τα διχότομα αποτελέσματα



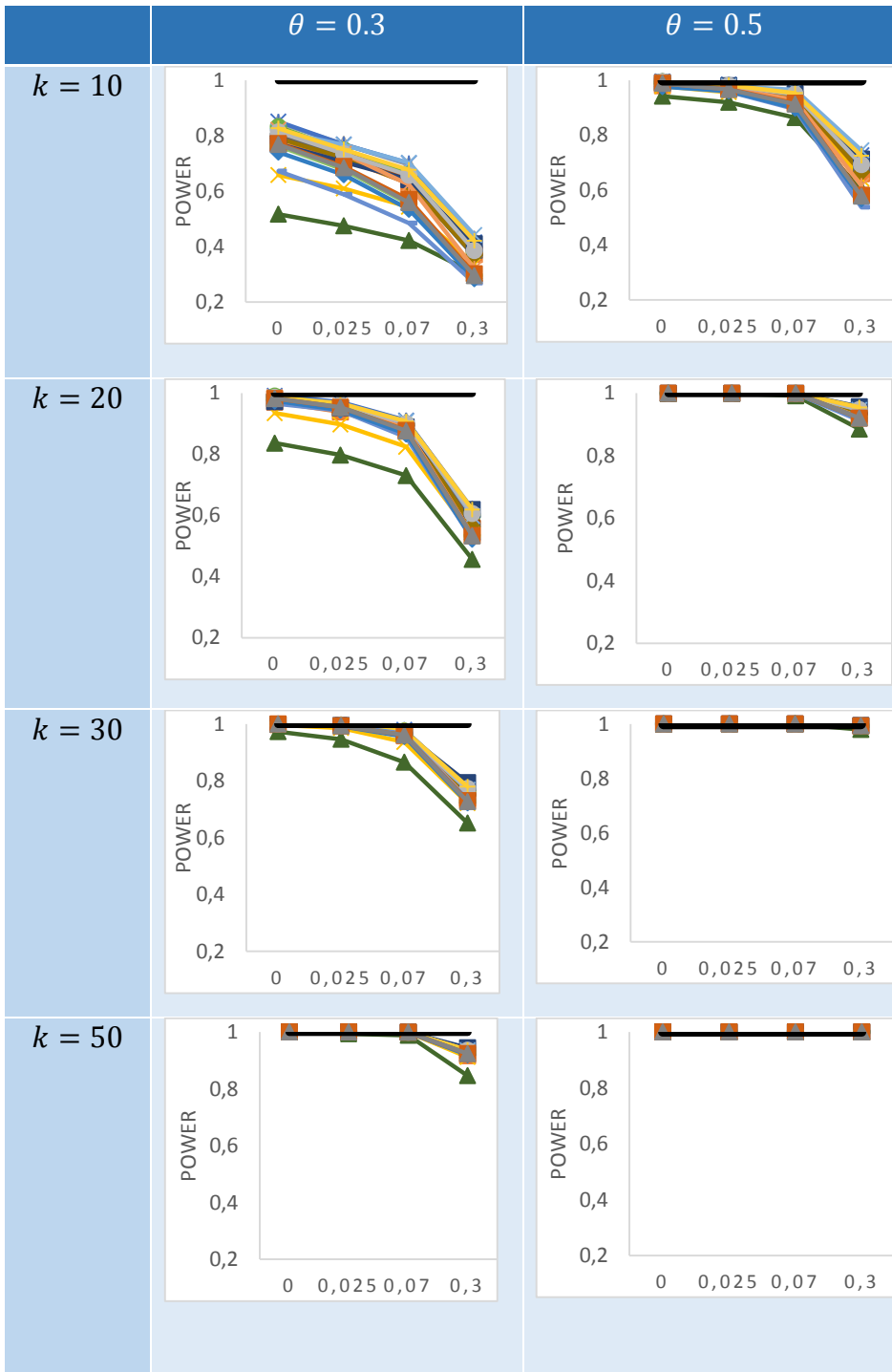
Πίνακας 5.4(β): Πιθανότητα κάλυψης για το τ^2 για τα διχότομα αποτελέσματα



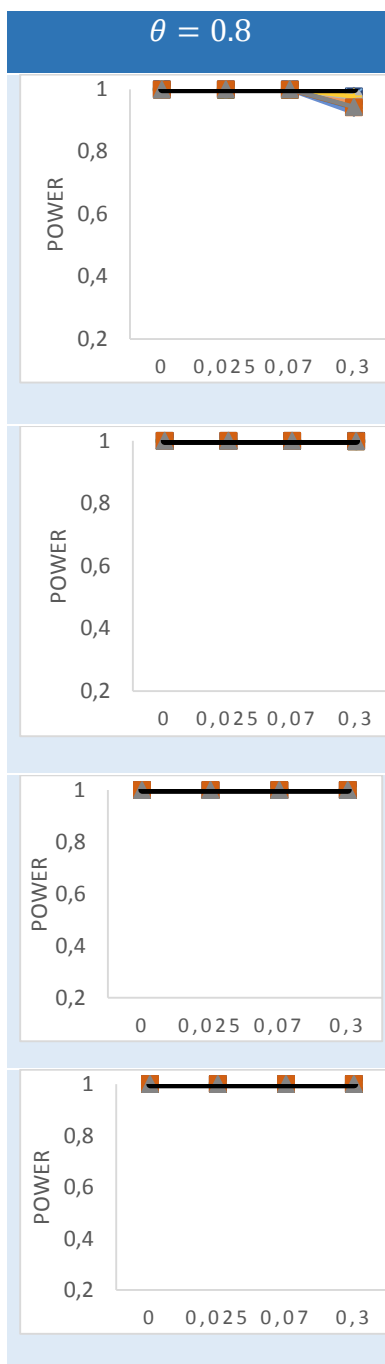
Πίνακας 5.5(α): Σφάλμα τύπου I για τα διχότομα αποτελέσματα



Πίνακας 5.5(β): Σφάλμα τύπου I για τα διχότομα αποτελέσματα

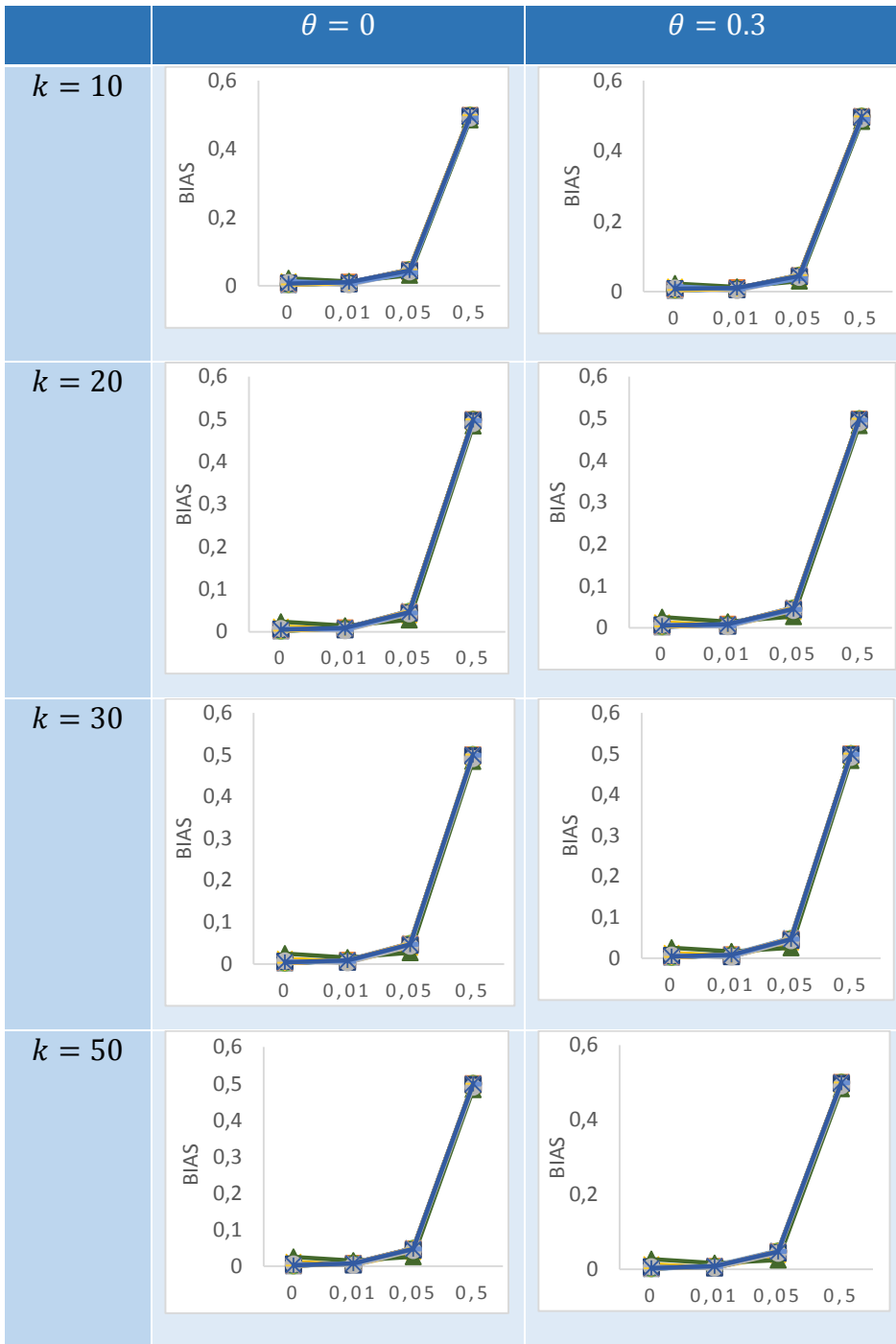


Πίνακας 5.6(α): Ισχύς για τα διχότομα αποτελέσματα

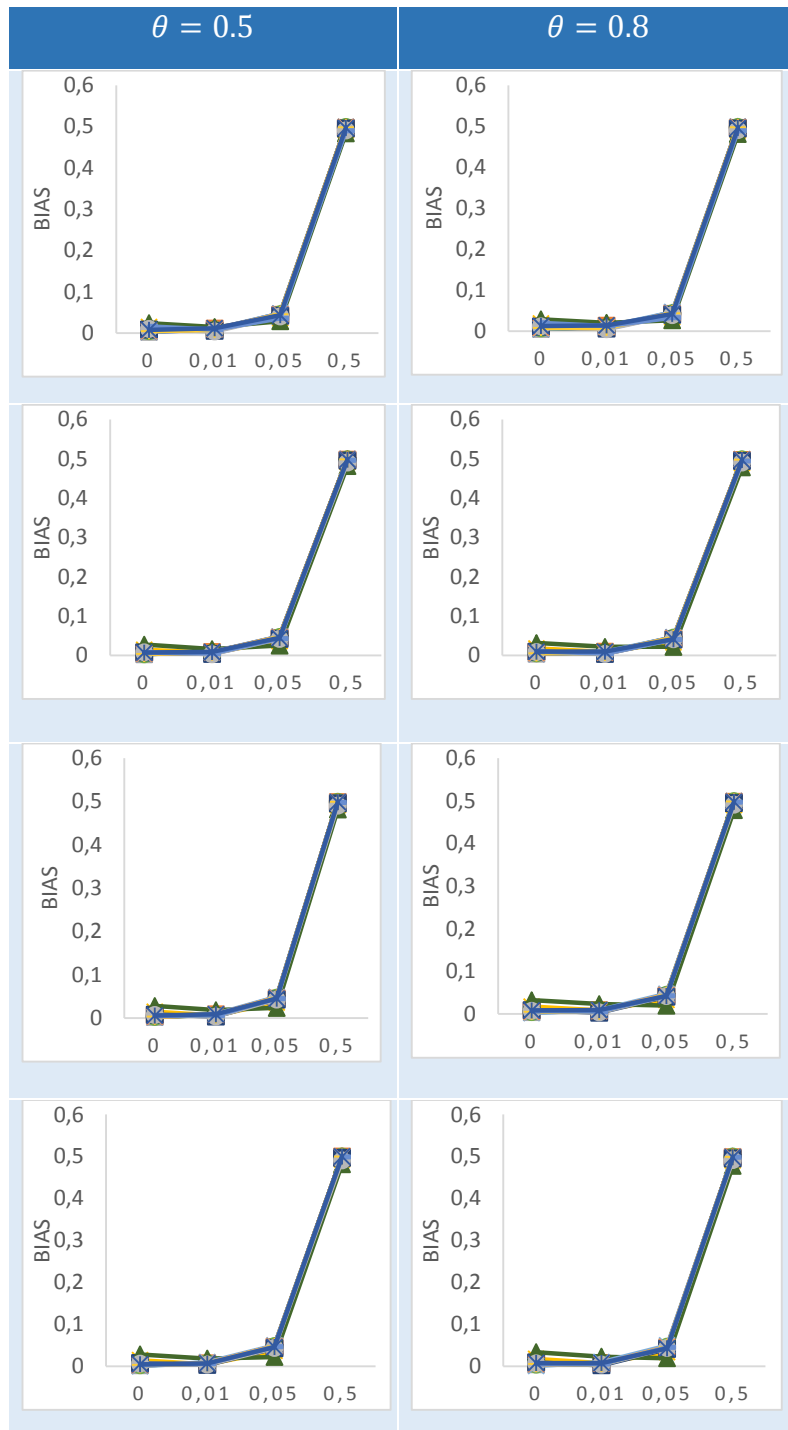


Πίνακας 5.6(β):

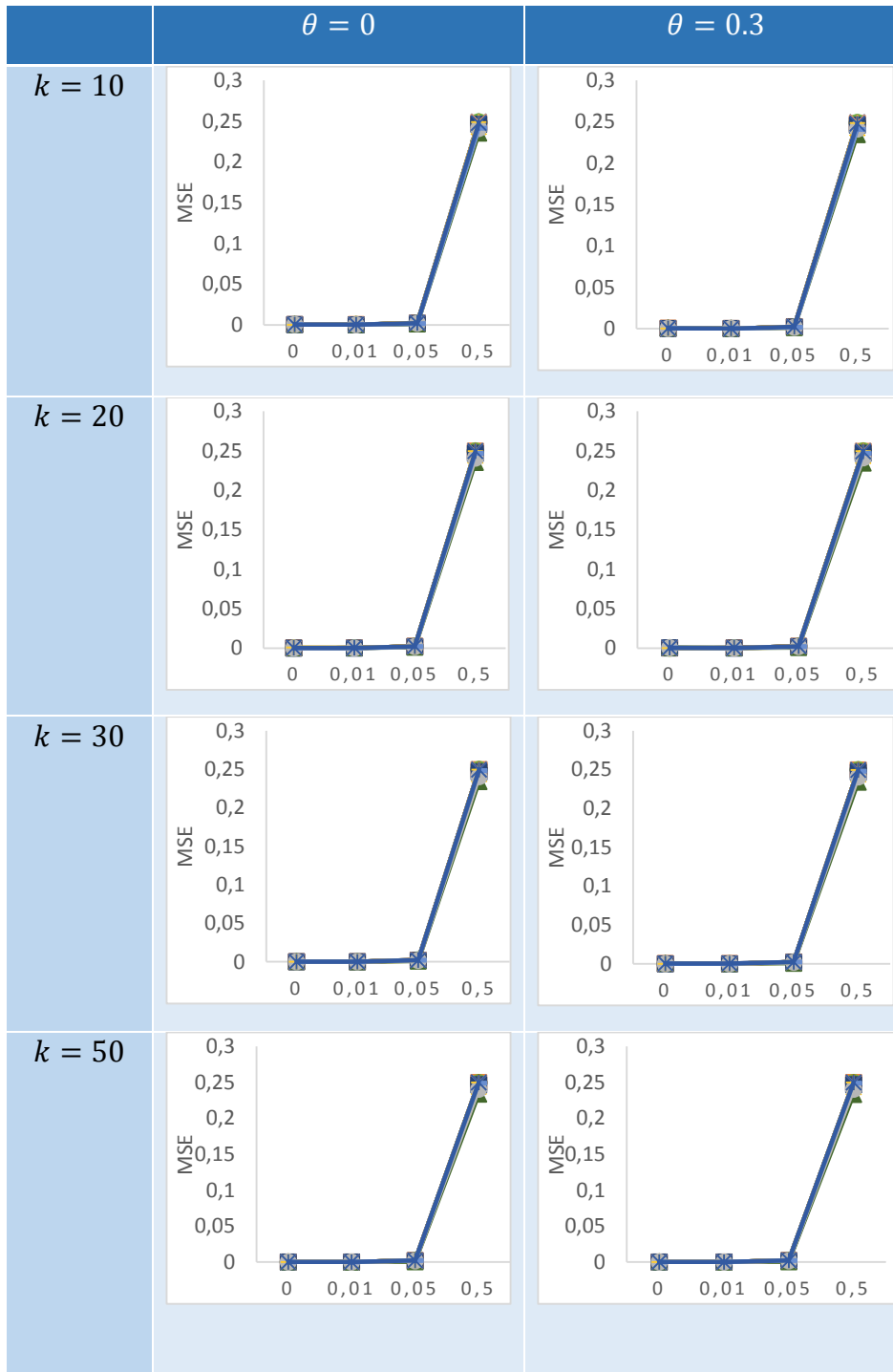
Ισχύς για τα διχότομα αποτελέσματα



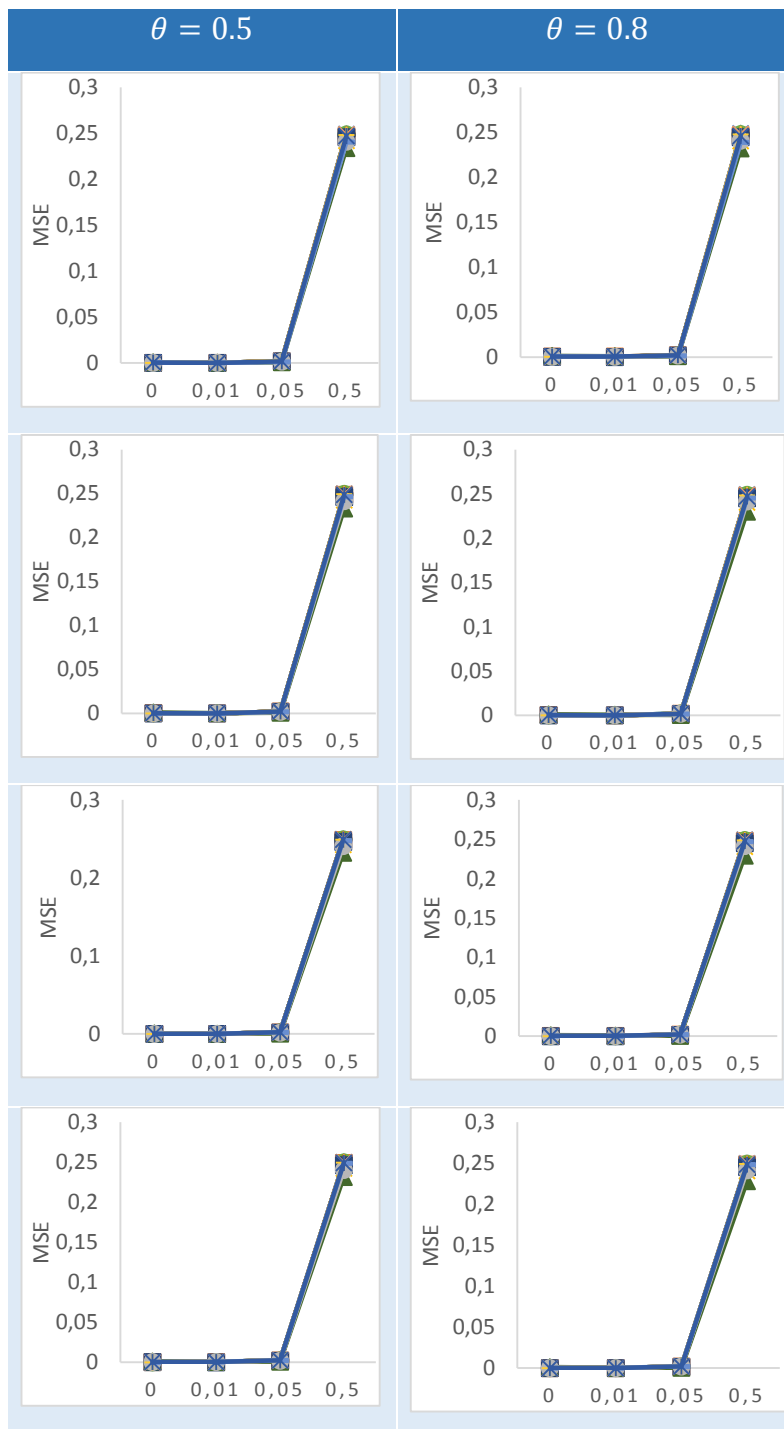
Πίνακας 5.7 (α): Μεροληψία (bias) για τα συνεχή αποτελέσματα



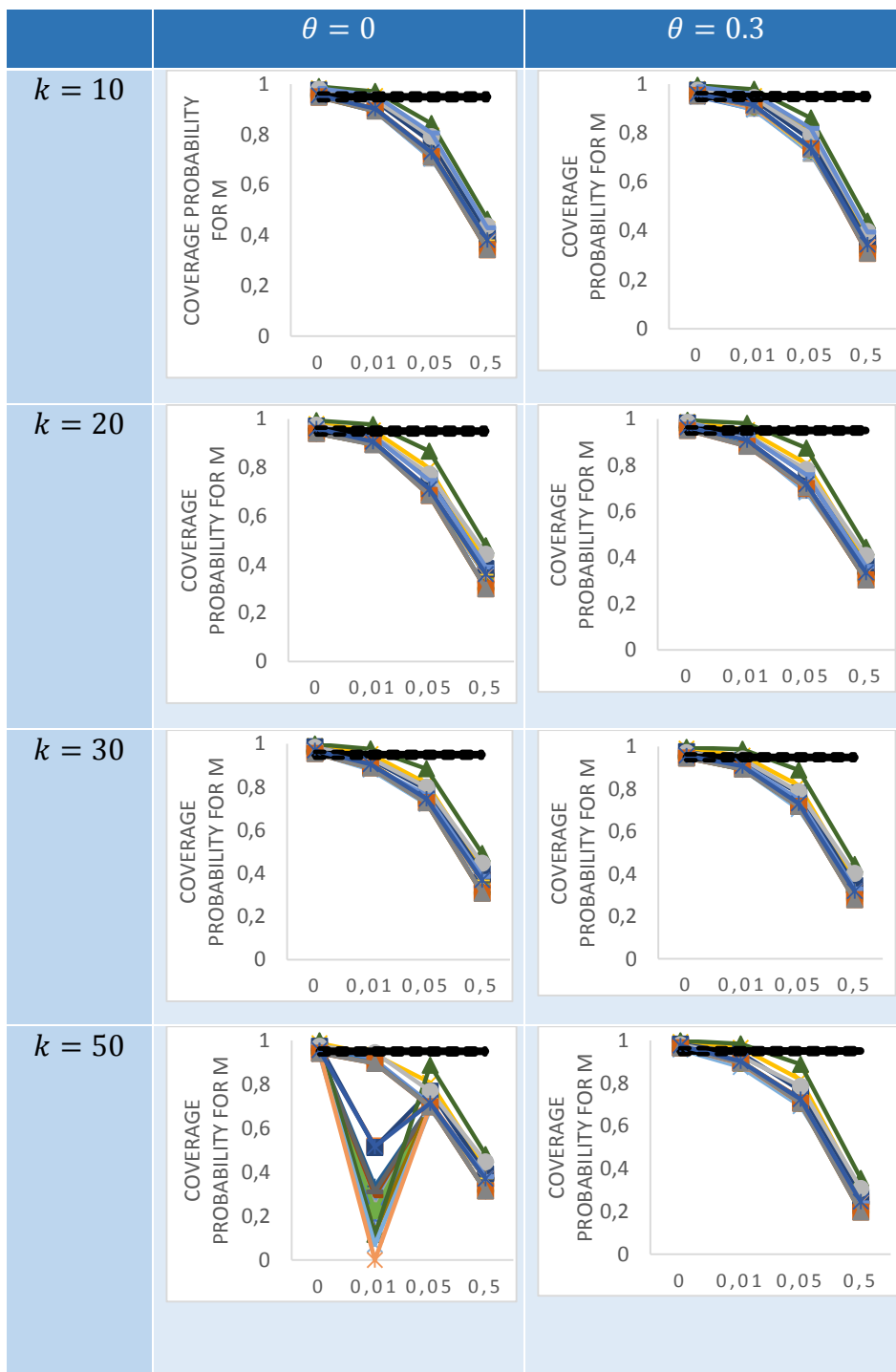
Πίνακας 5.7 (β): Μεροληγία (bias) για τα συνεχή αποτελέσματα



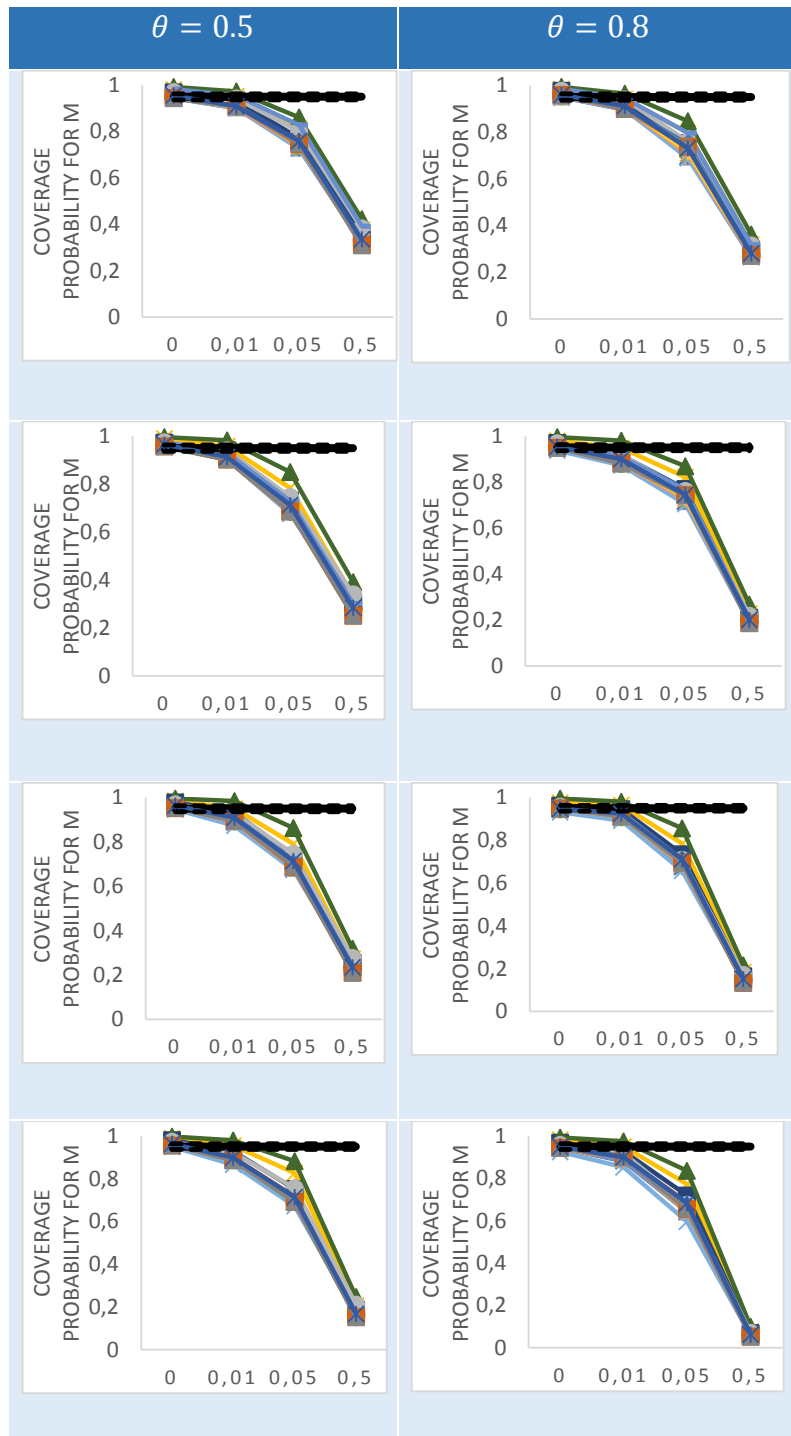
Πίνακας 5.8 (α): MSE για τα συνεχή αποτελέσματα



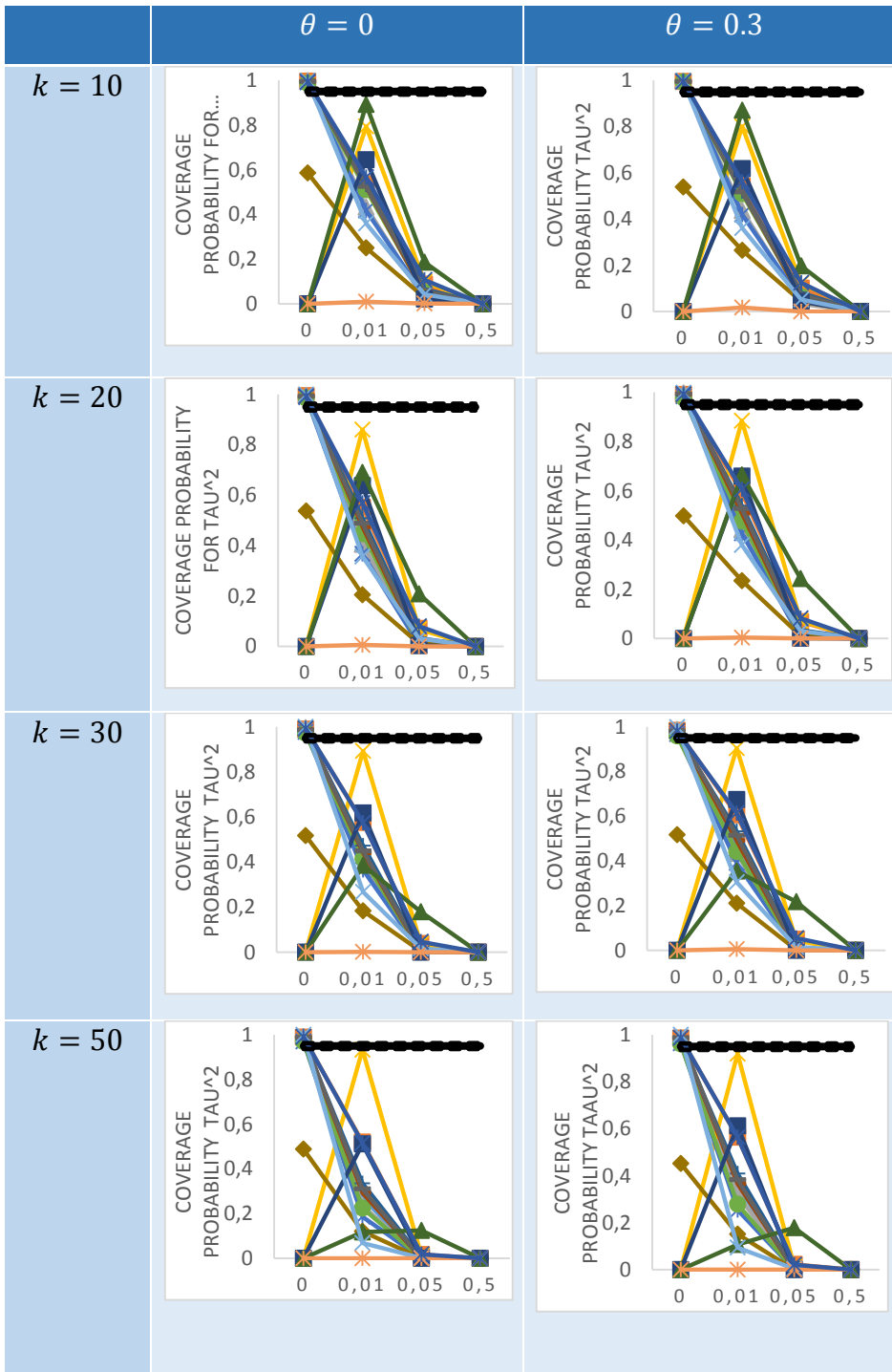
Πίνακας 5.8 (β): MSE για τα συνεχή αποτελέσματα



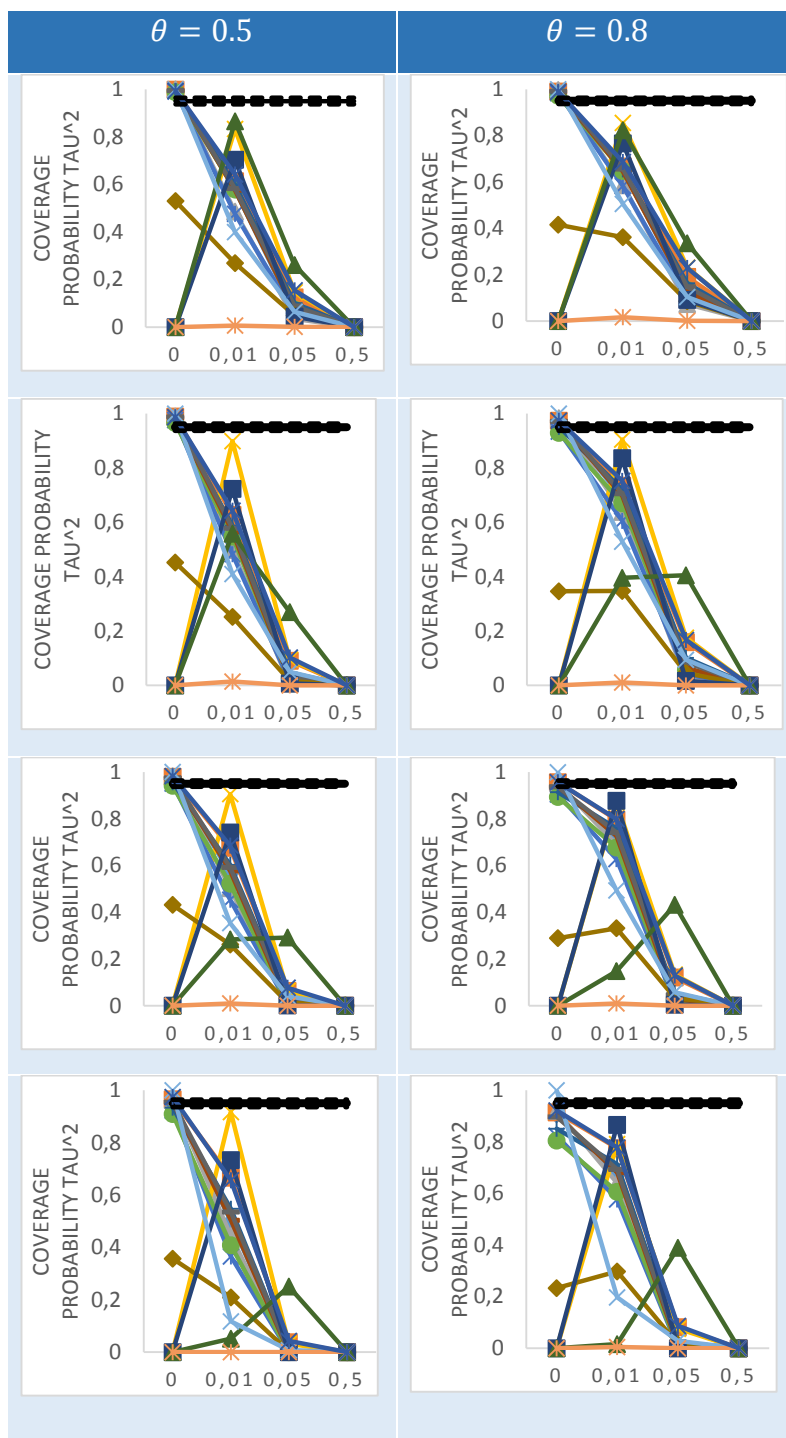
Πίνακας 5.9(α): Πιθανότητα κάλυψης για το μ για τα συνεχή αποτελέσματα



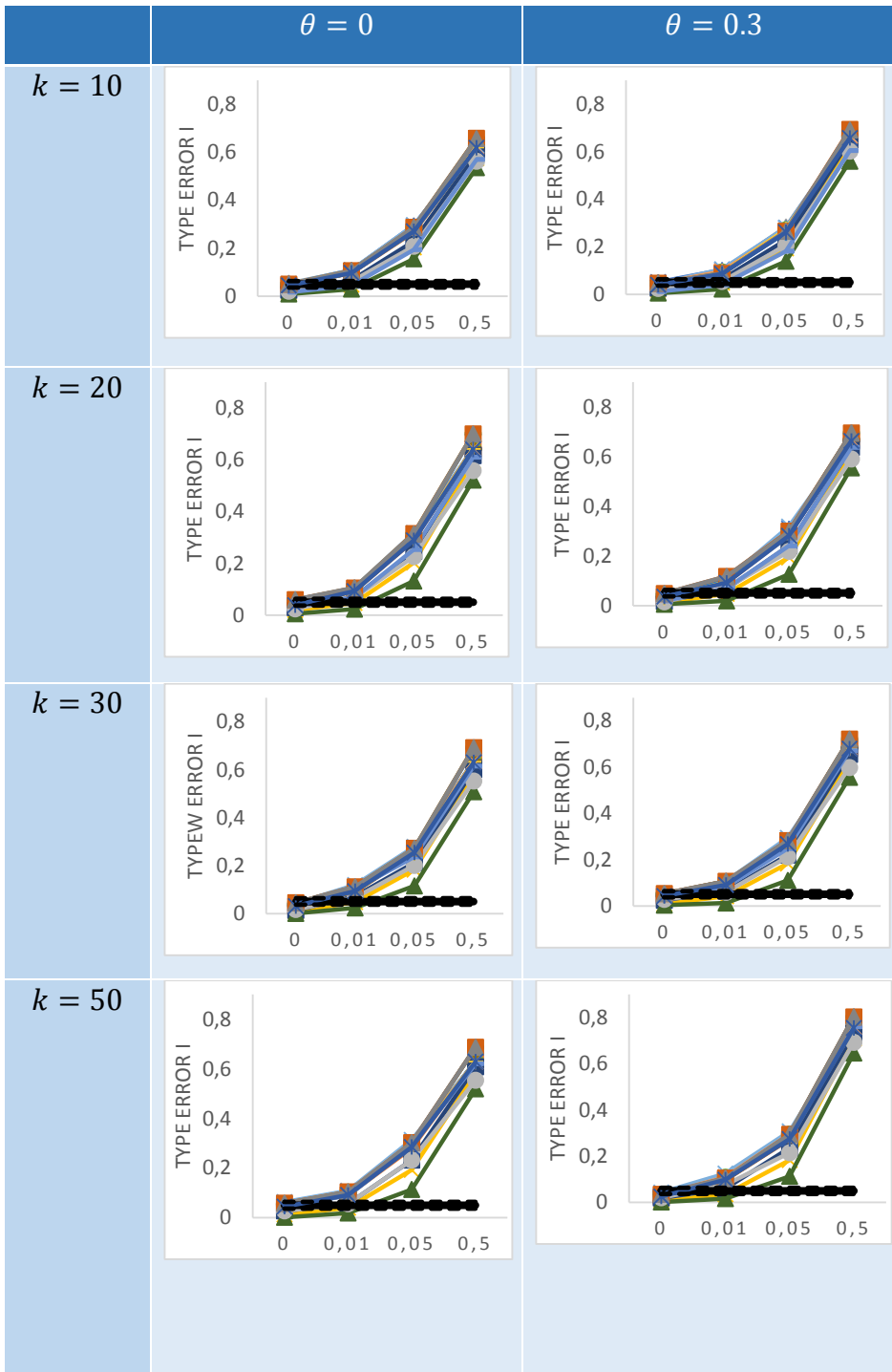
Πίνακας 5.9(β): Πιθανότητα κάλυψης για το μ για τα συνεχή αποτελέσματα



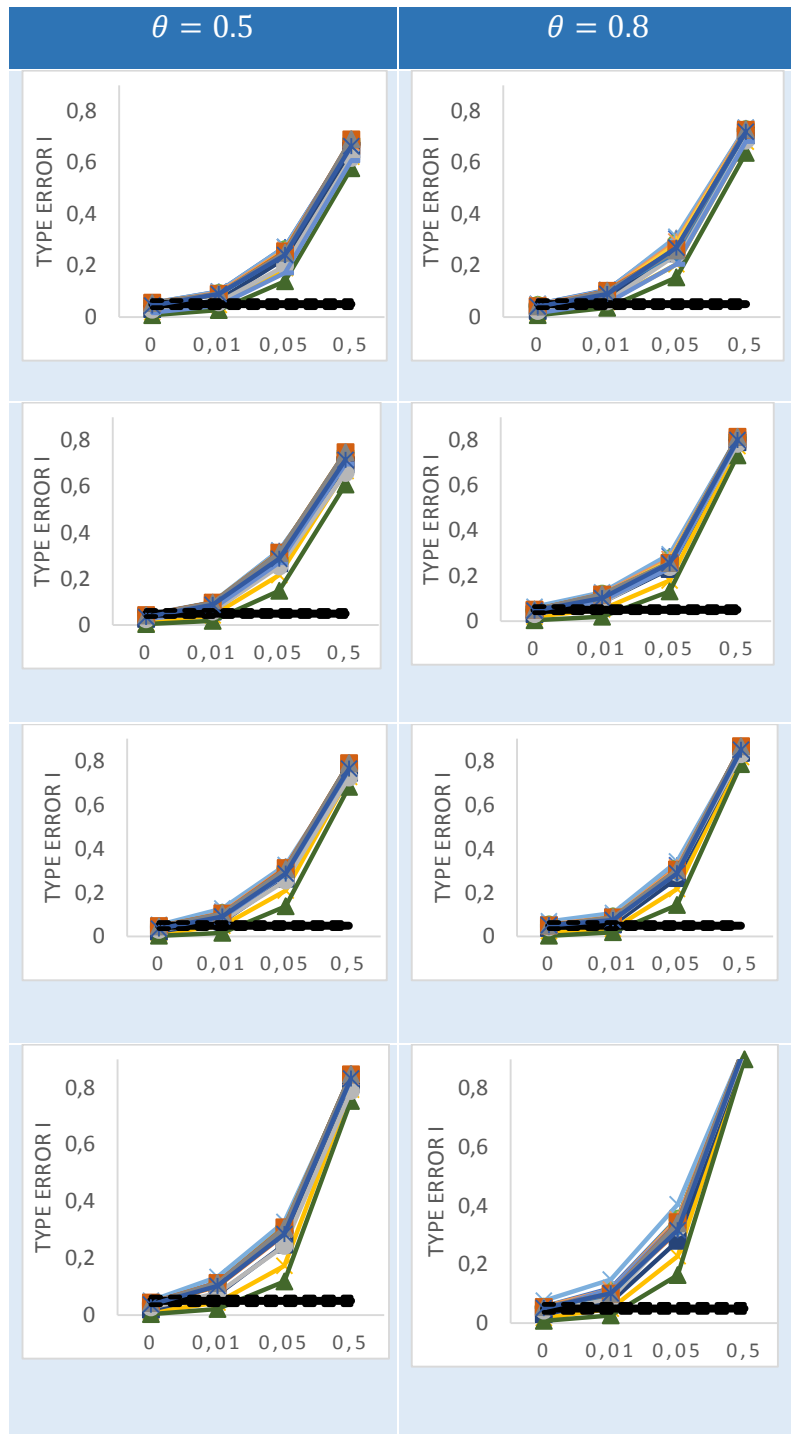
Πίνακας 5.10(α): Πιθανότητα κάλυψης για το τ^2 για τα συνεχή αποτελέσματα



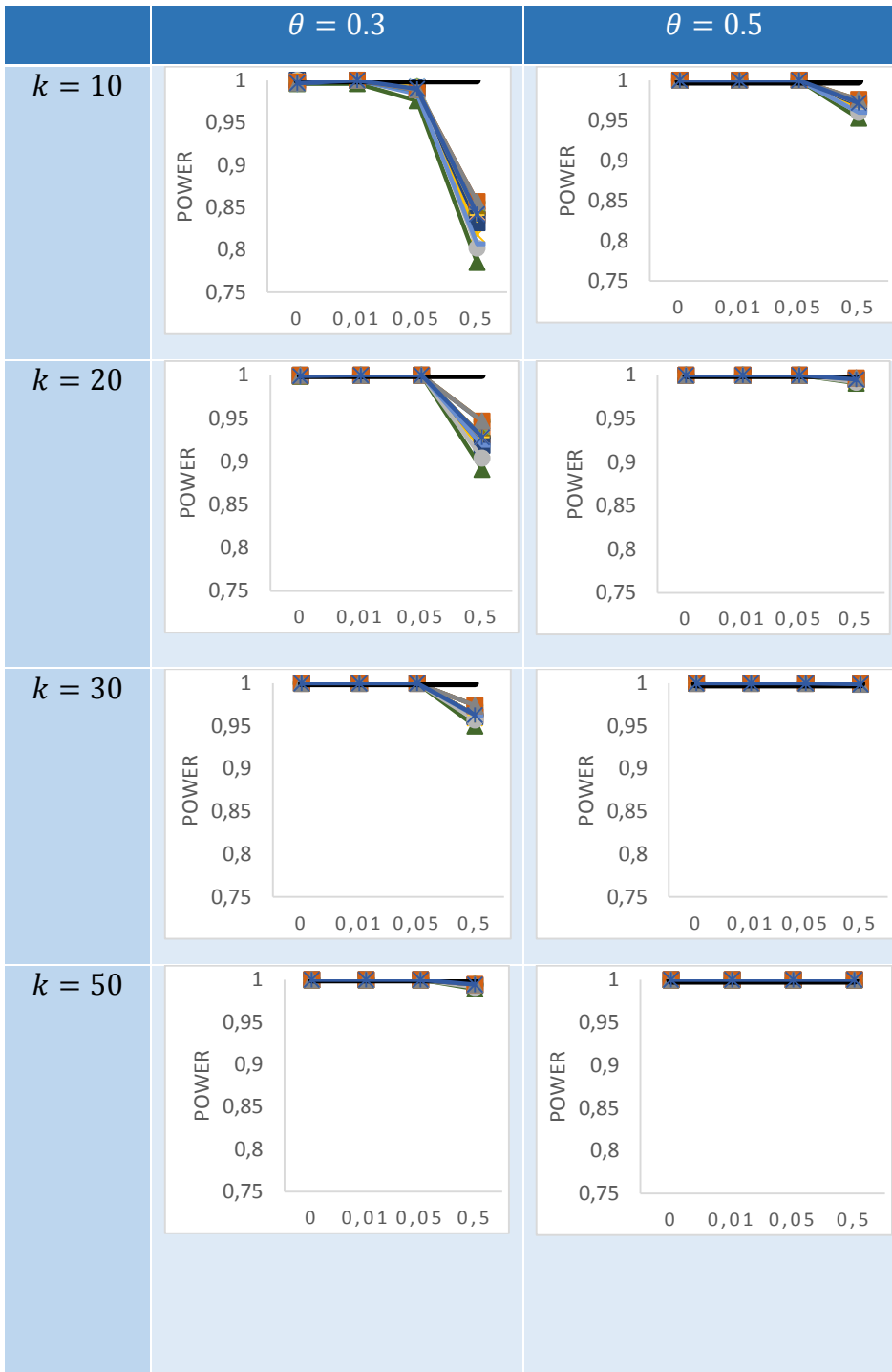
Πίνακας 5.10(β): Πιθανότητα κάλυψης για το τ^2 για τα συνεχή αποτελέσματα



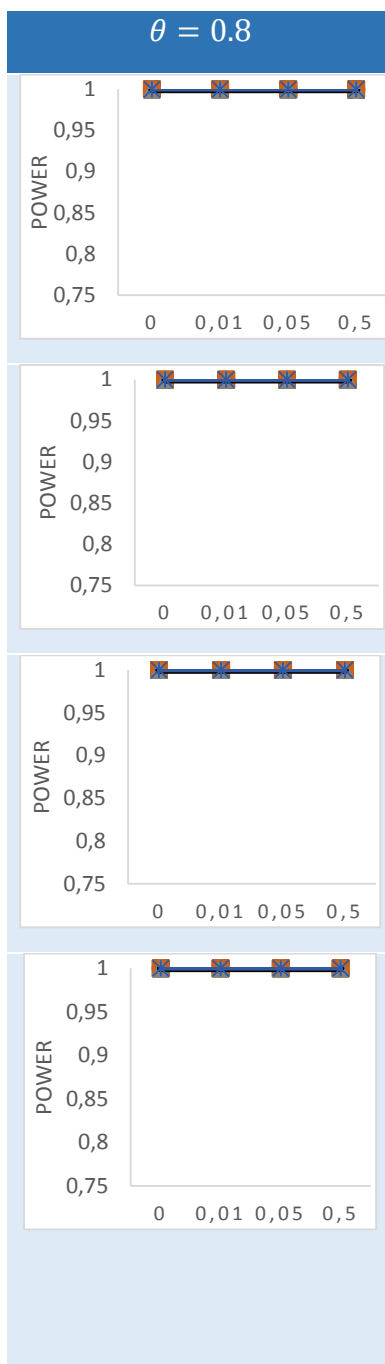
Πίνακας 5.11(α): Σφάλμα τύπου I για τα συνεχή αποτελέσματα



Πίνακας 5.11(β): Σφάλμα τύπου I για τα συνεχή αποτελέσματα

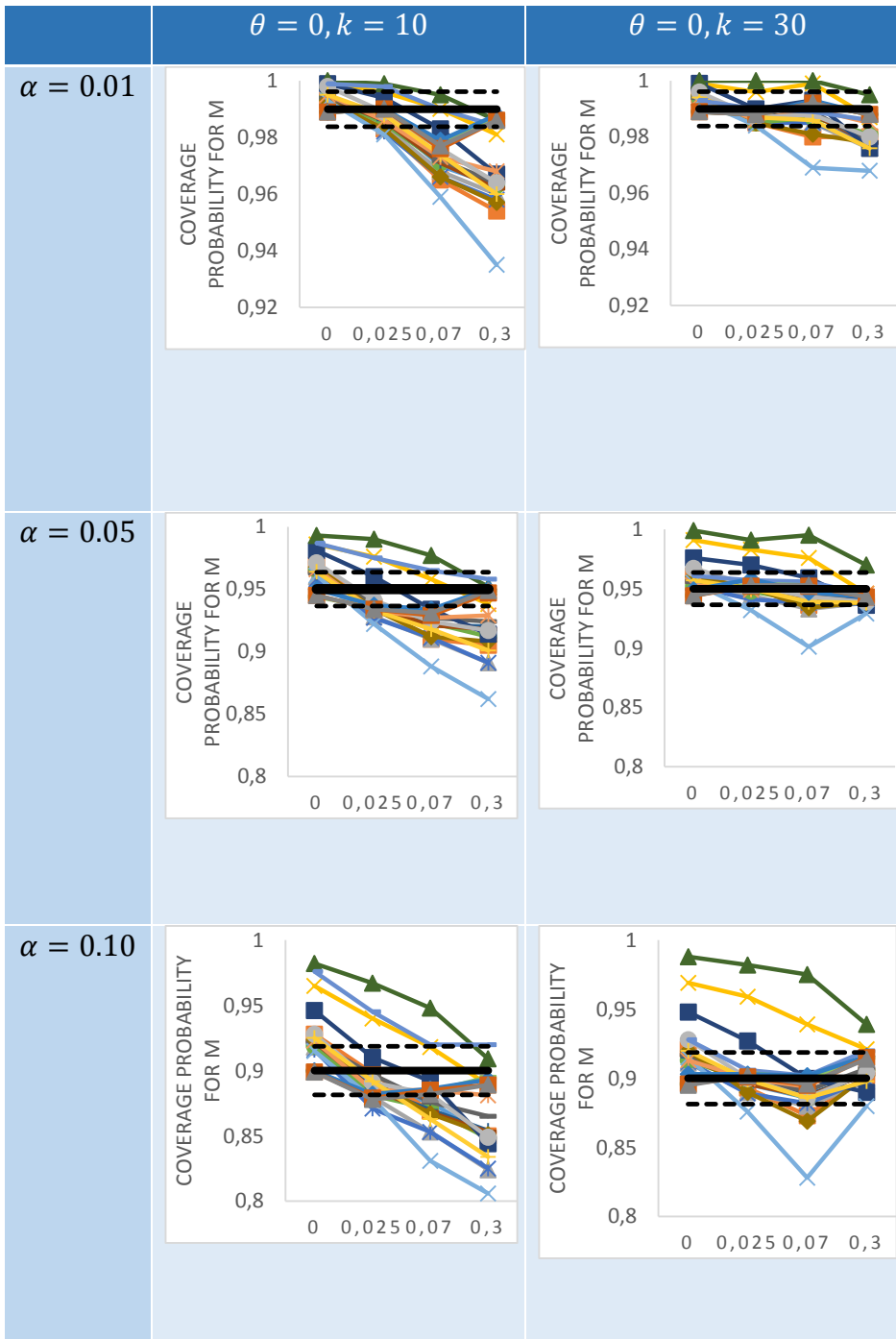


Πίνακας 5.12(α): Ισχύς για τα συνεχή αποτελέσματα

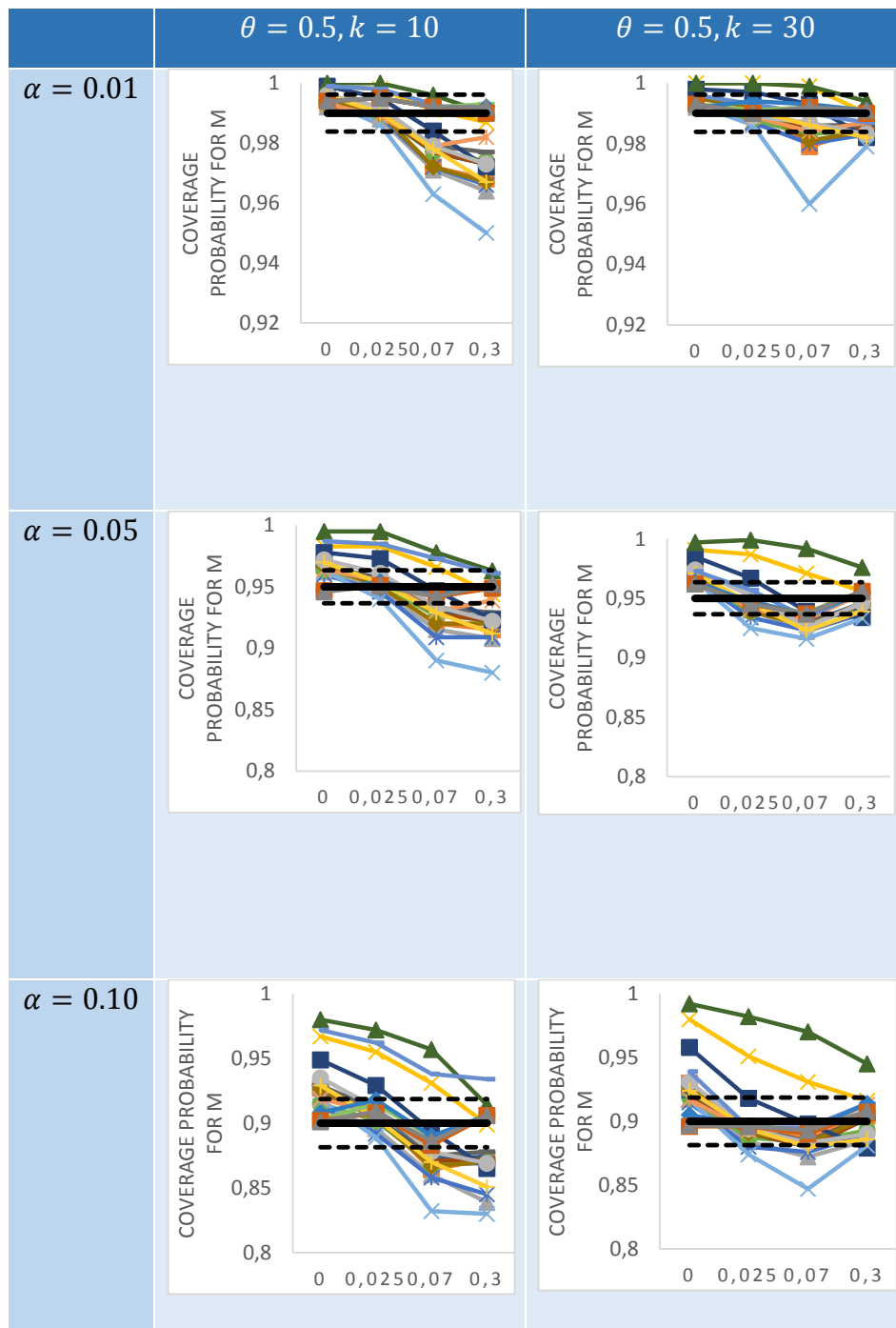


Πίνακας 5.12(β):

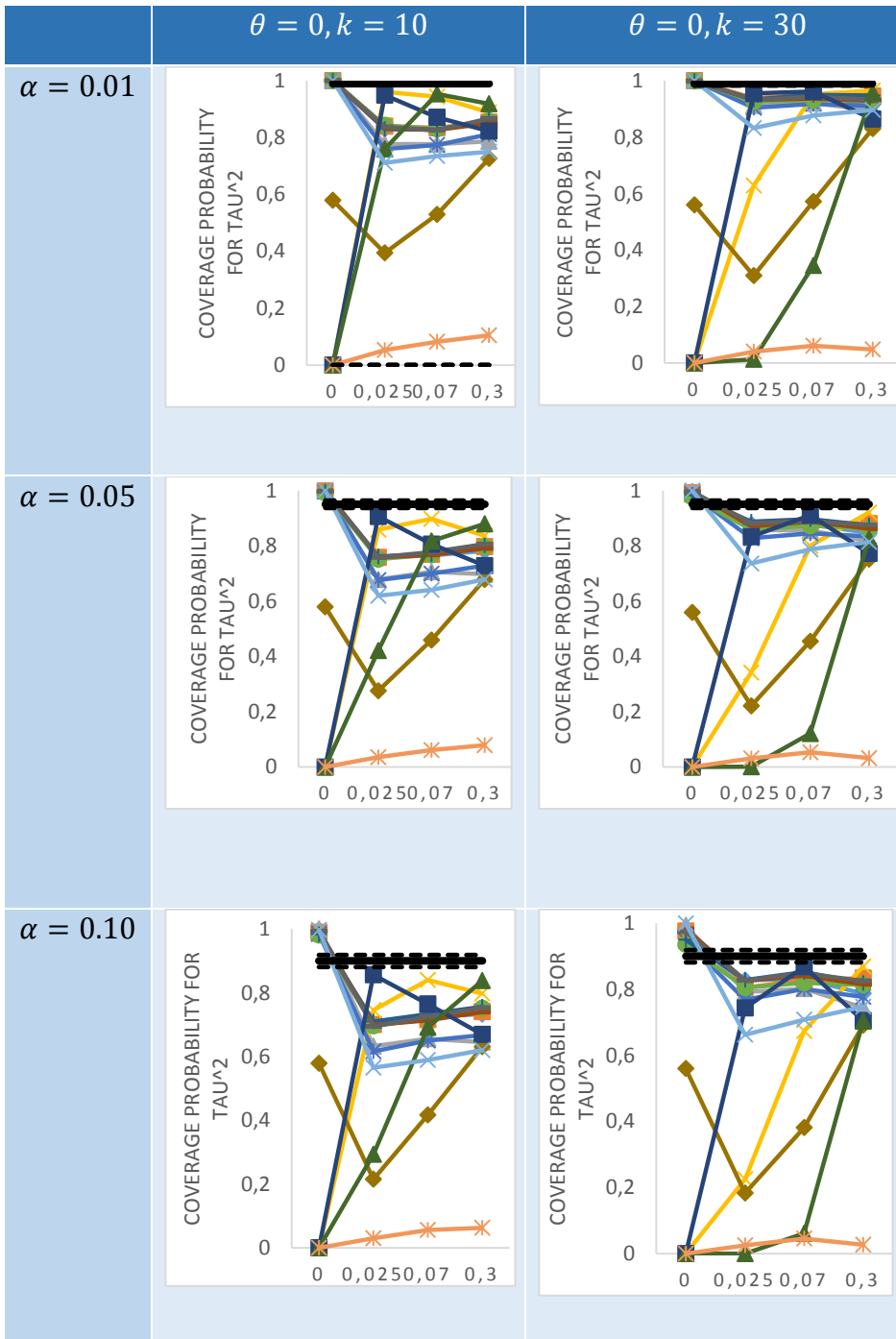
Ισχύς για τα συνεχή αποτελέσματα



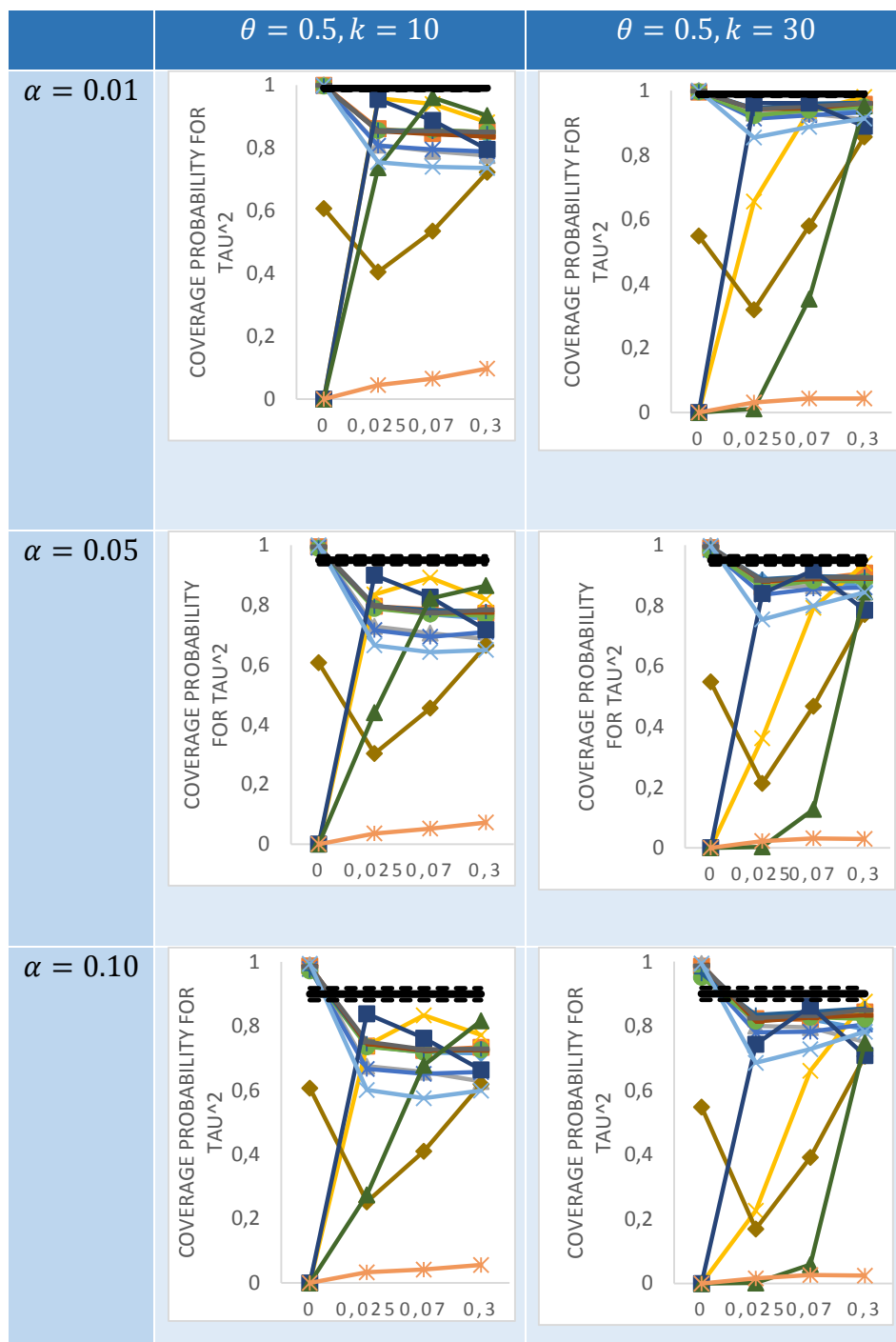
Πίνακας 5.13 (α): Σύγκριση της πιθανότητας κάλυψης του μ για τα διάφορα επίπεδα σημαντικότητας στα διχότομα αποτελέσματα



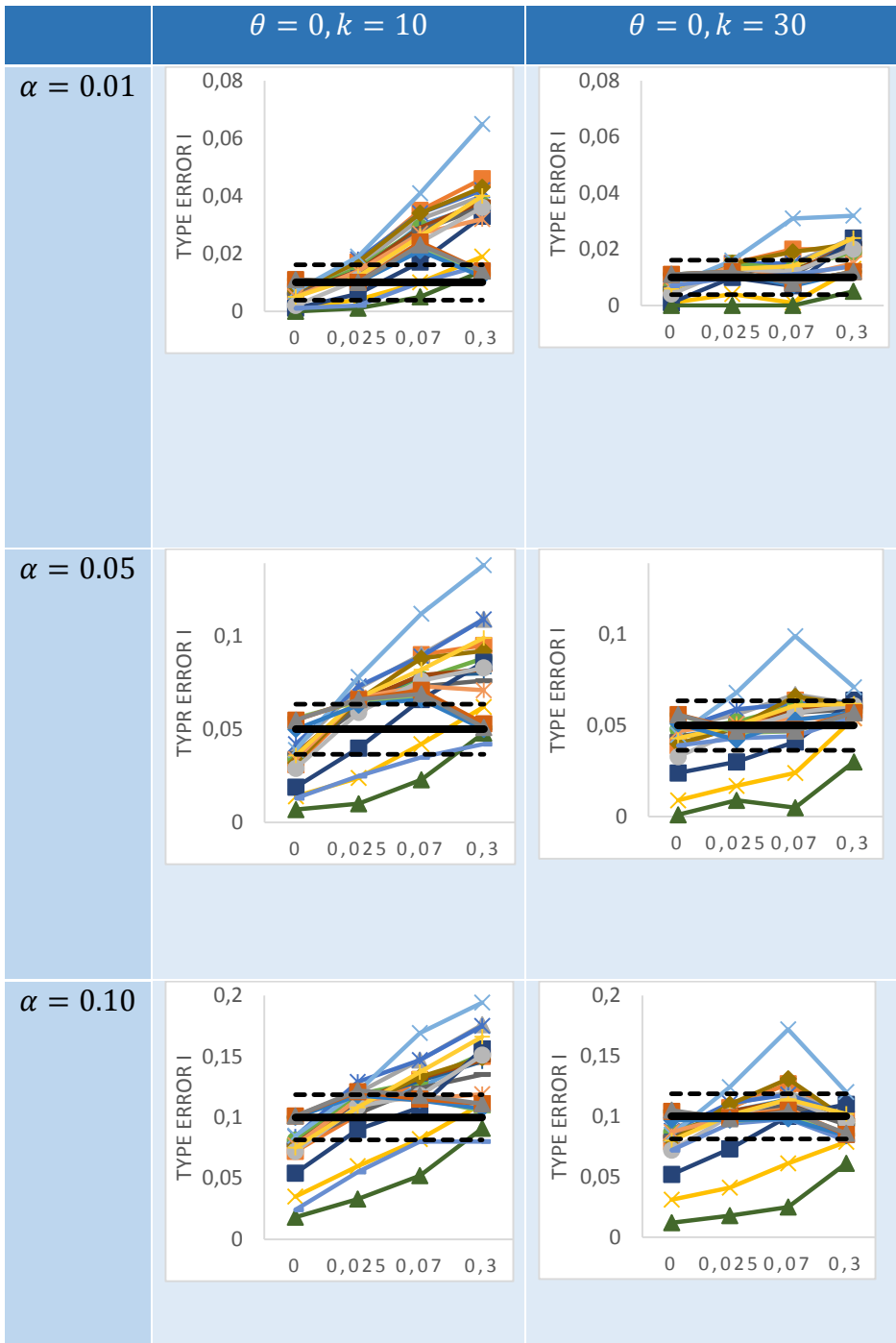
Πίνακας 5.13 (β): Σύγκριση της πιθανότητας κάλυψης του μ για τα διάφορα επίπεδα σημαντικότητας στα διχότομα αποτελέσματα



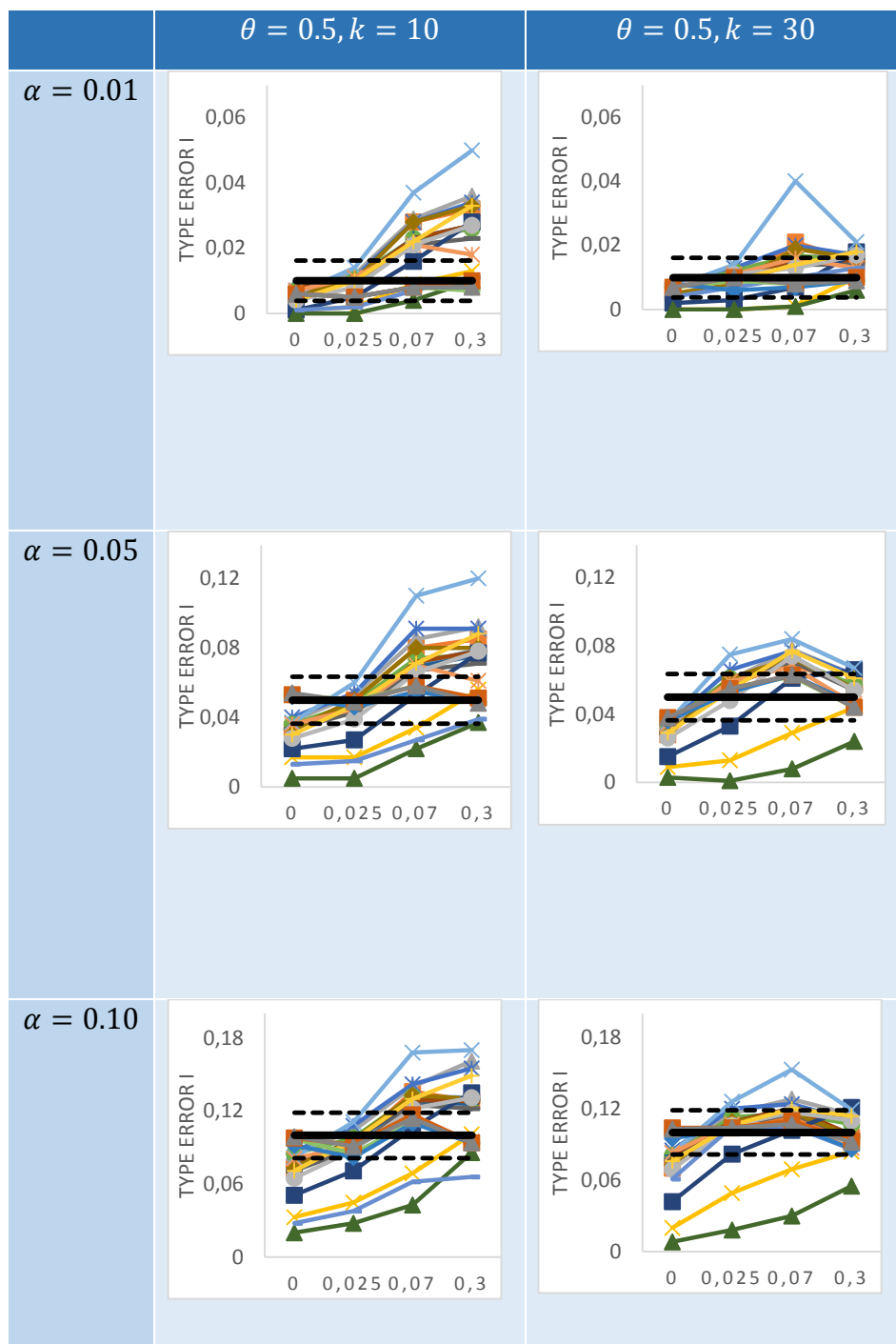
Πίνακας 5.14 (α): Σύγκριση της πιθανότητας κάλυψης του τ^2 για τα διάφορα επίπεδα σημαντικότητας στα διχότομα αποτελέσματα



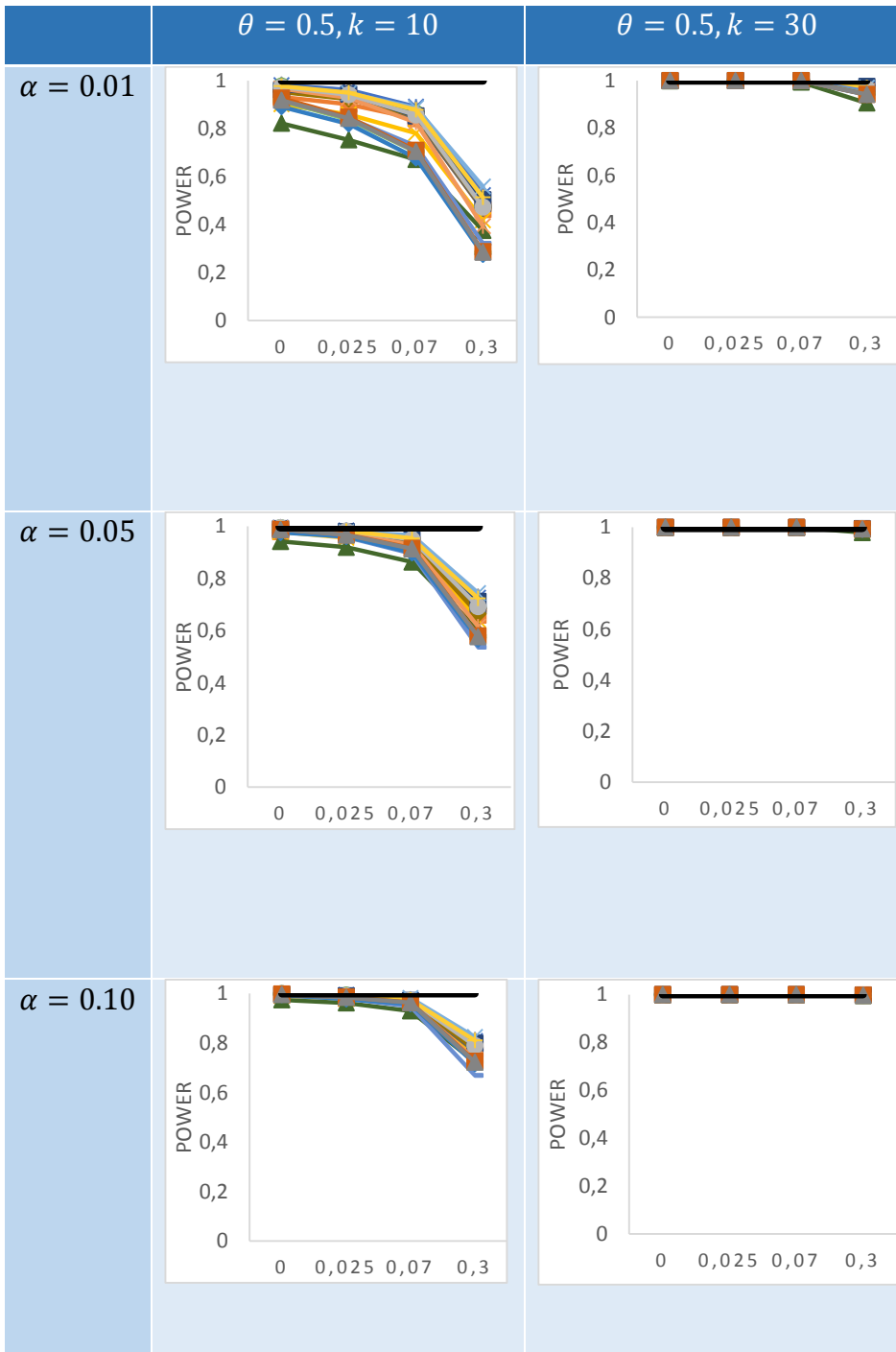
Πίνακας 5.14 (β): Σύγκριση της πιθανότητας κάλυψης του τ^2 για τα διάφορα επίπεδα σημαντικότητας στα διχότομα αποτελέσματα



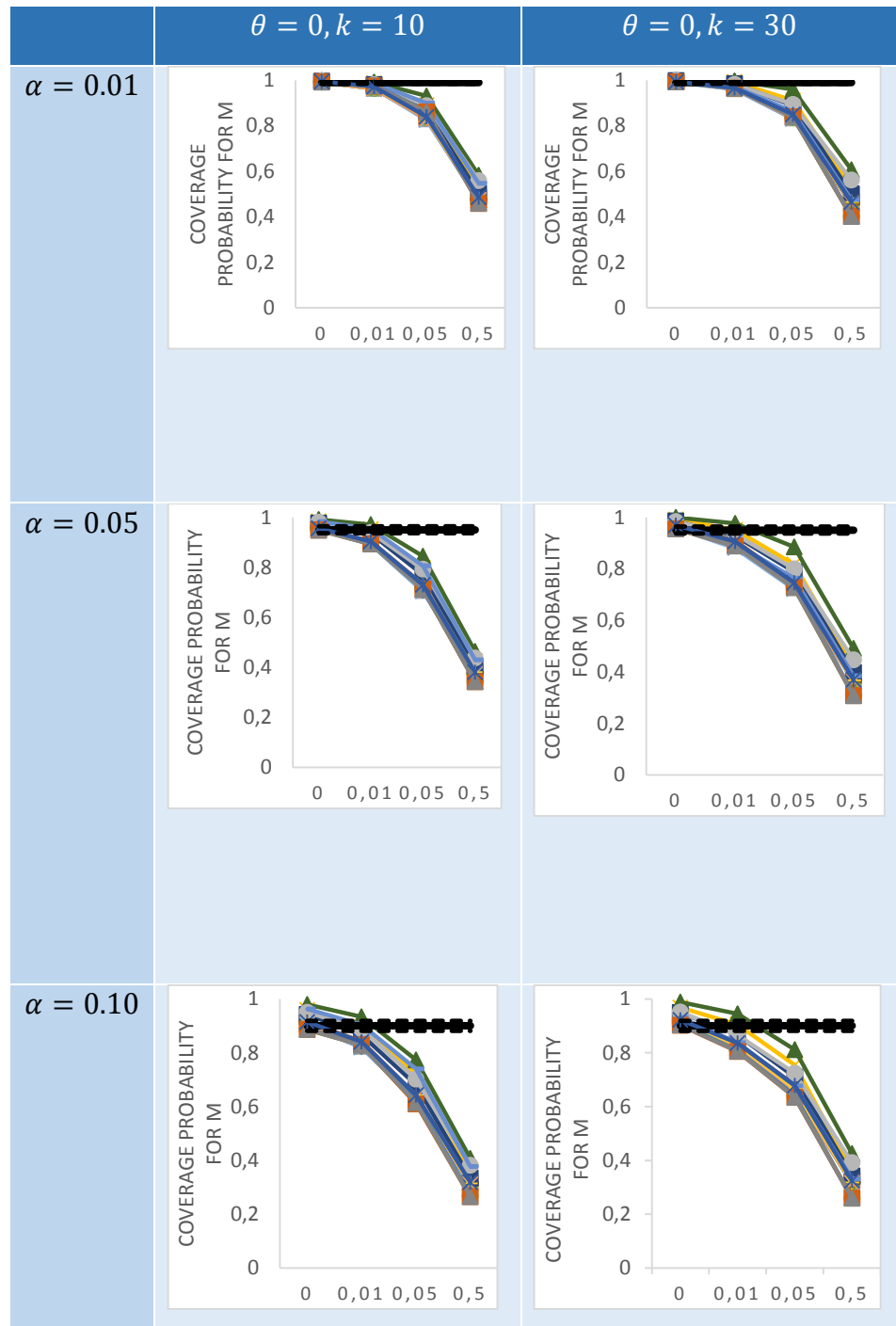
Πίνακας 5.15 (α): Σύγκριση του σφάλματος τύπου I για τα διάφορα επίπεδα σημαντικότητας στα διχότομα αποτελέσματα



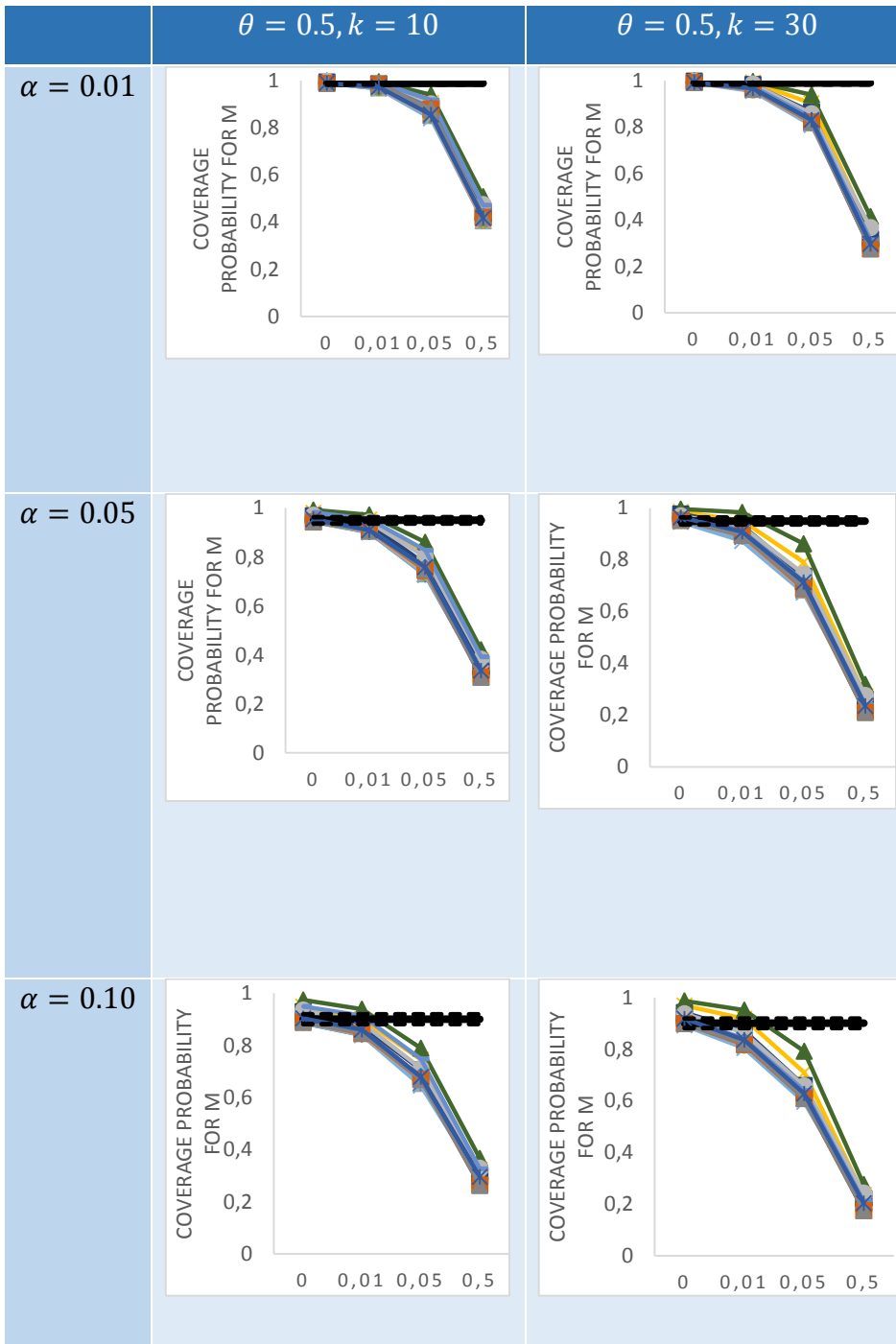
Πίνακας 5.15 (β): Σύγκριση του σφάλματος τύπου I για τα διάφορα επίπεδα σημαντικότητας στα διχότομα αποτελέσματα



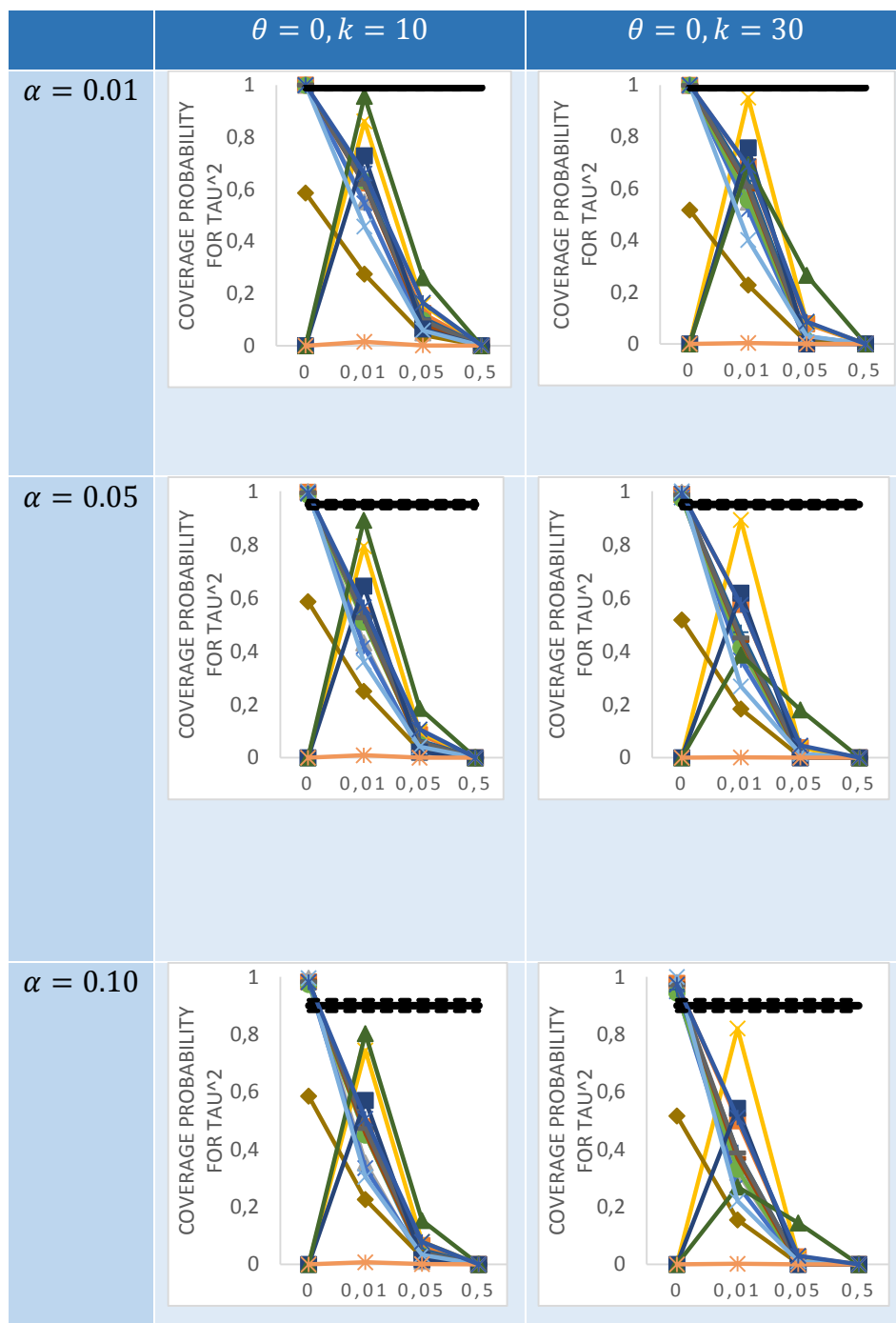
Πίνακας 5.16: Σύγκριση της ισχύς για τα διάφορα επίπεδα σημαντικότητας στα διχότομα αποτελέσματα



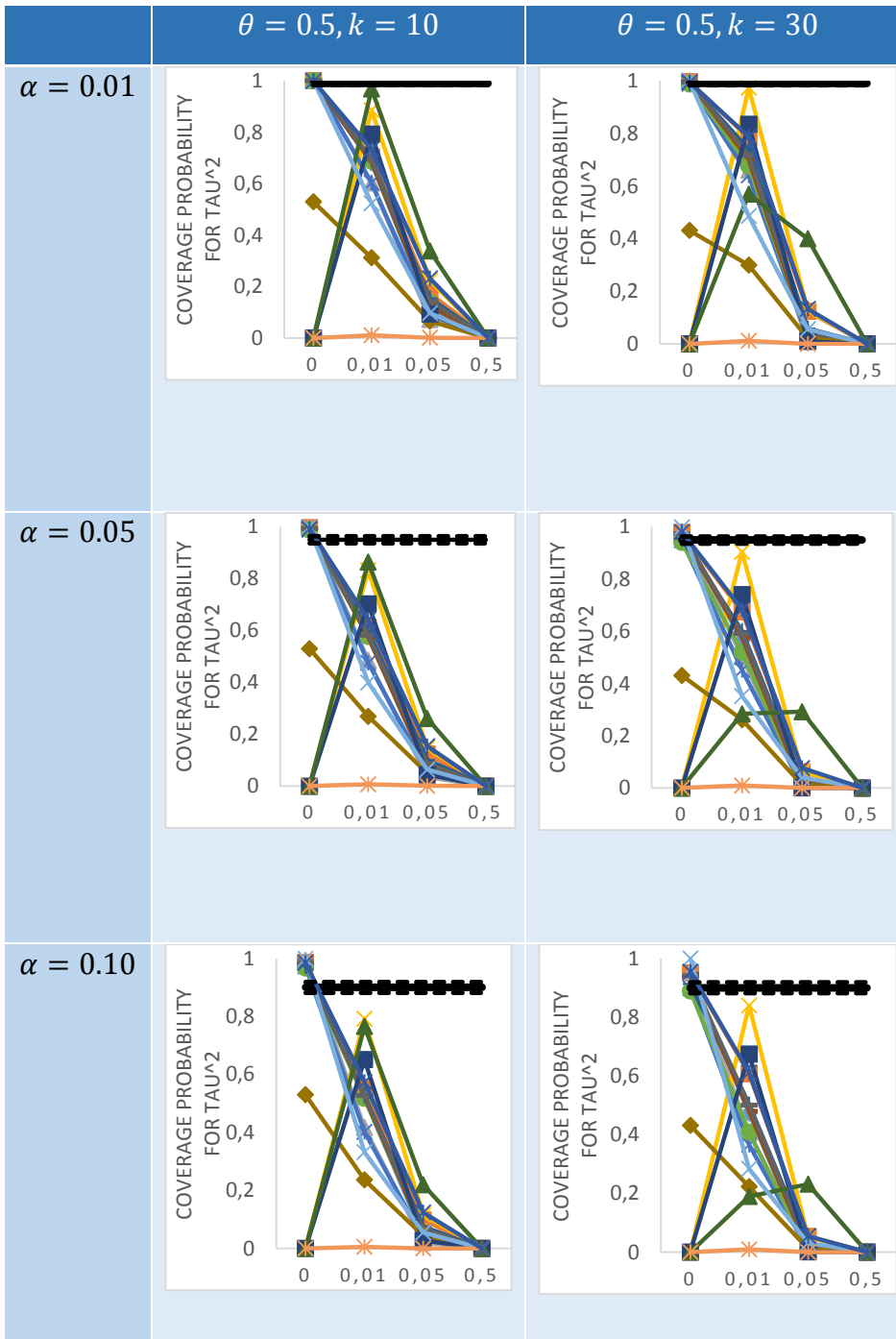
Πίνακας 5.17 (α): Σύγκριση της πιθανότητας κάλυψης του μ για τα διάφορα επίπεδα σημαντικότητας στα συνεχή αποτελέσματα



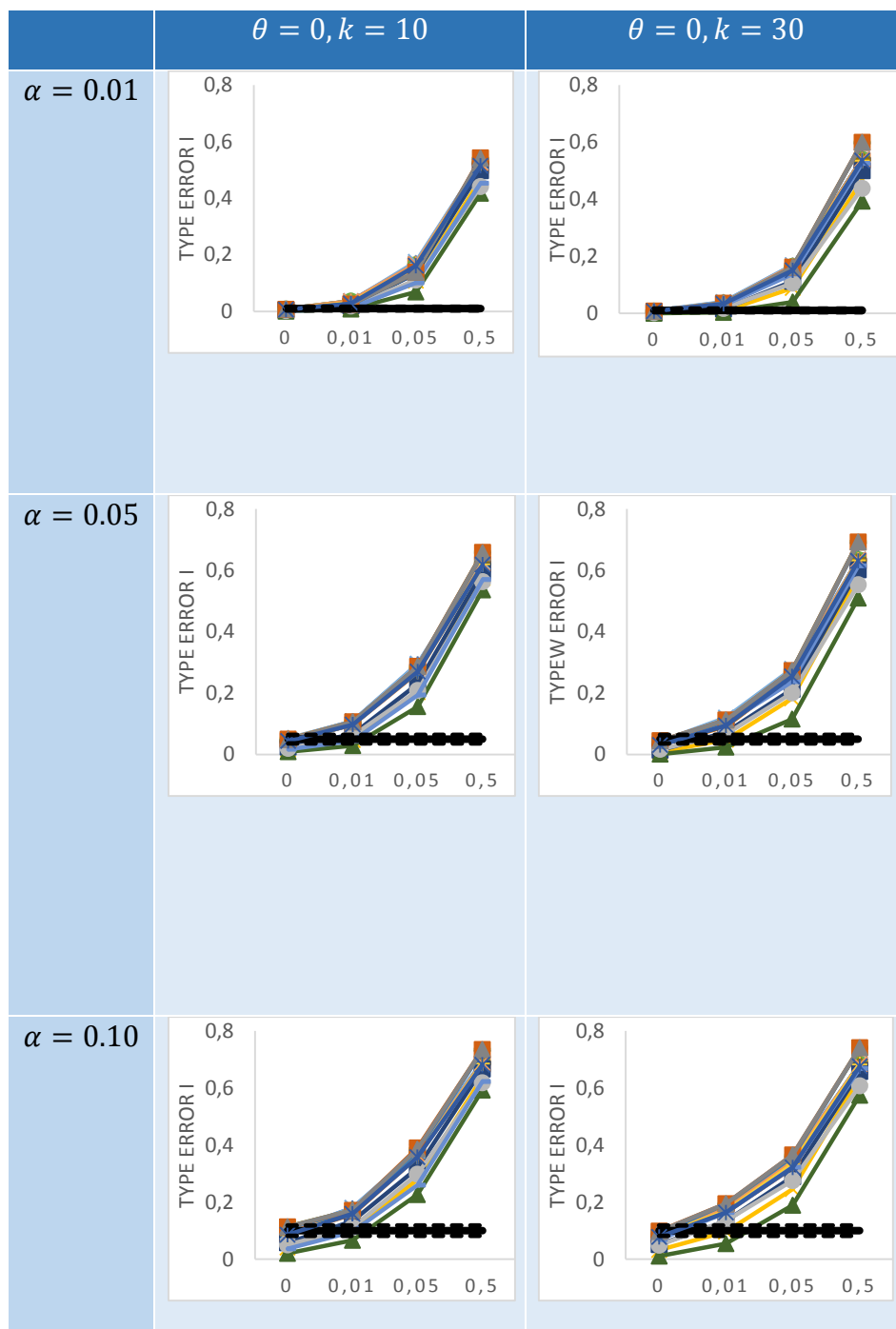
Πίνακας 5.17 (β): Σύγκριση της πιθανότητας κάλυψης του μ για τα διάφορα επίπεδα σημαντικότητας στα συνεχή αποτελέσματα.



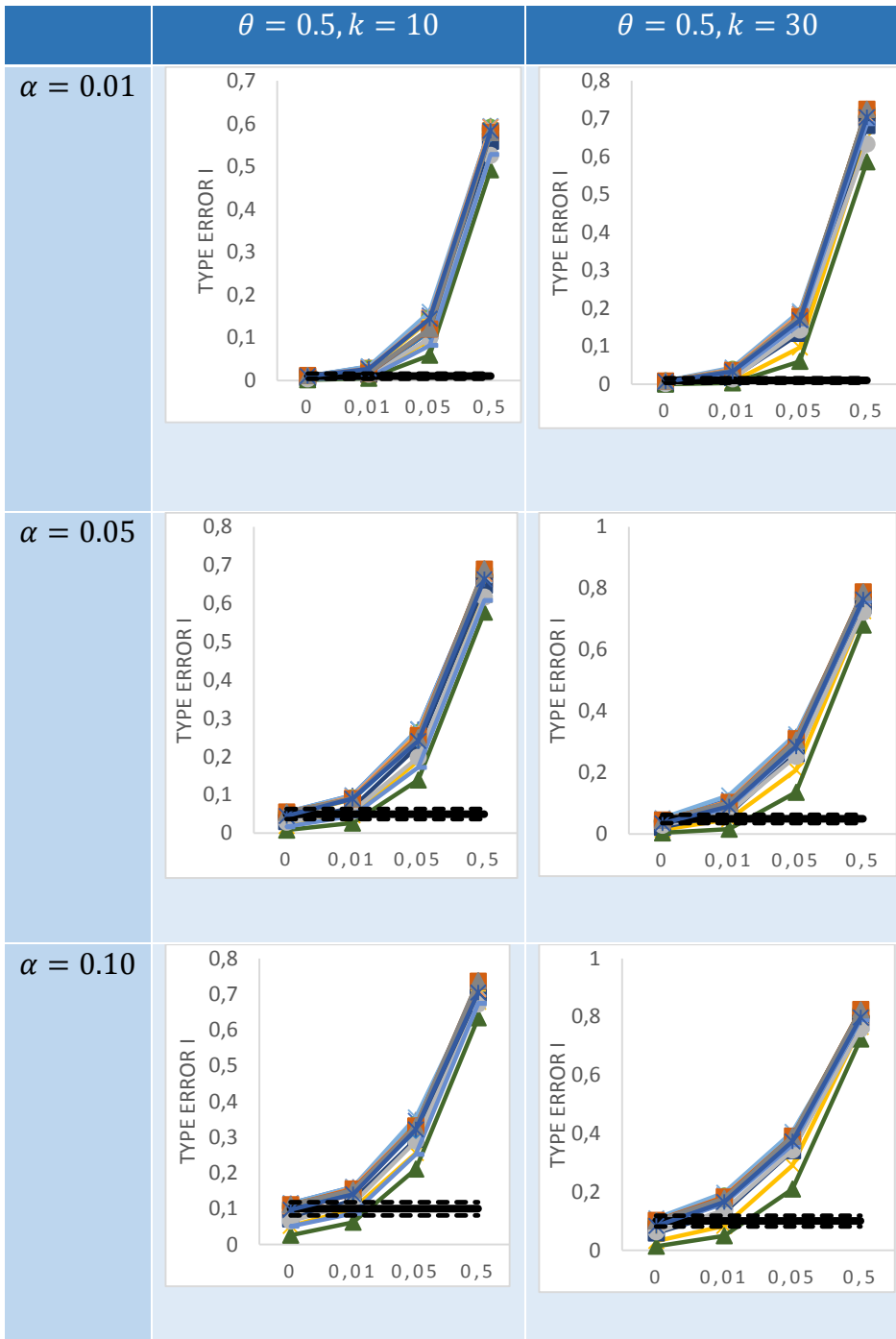
Πίνακας 5.18 (α): Σύγκριση της πιθανότητας κάλυψης του τ^2 για τα διάφορα επίπεδα σημαντικότητας στα συνεχή αποτελέσματα



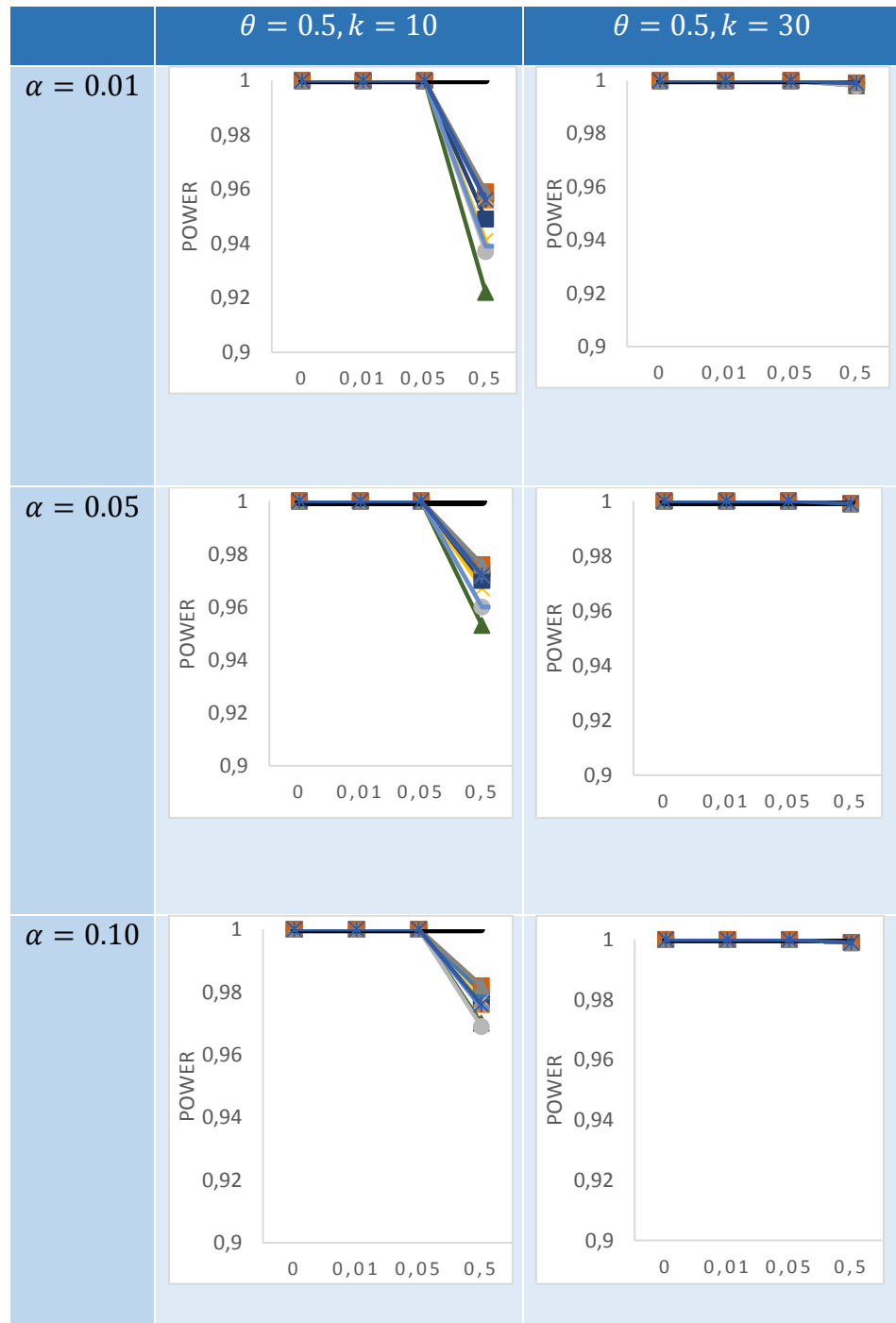
Πίνακας 5.18 (β): Σύγκριση της πιθανότητας κάλυψης του τ^2 για τα διάφορα επίπεδα σημαντικότητας στα συνεχή αποτελέσματα



Πίνακας 5.19 (α): Σύγκριση του σφάλματος τύπου I για τα διάφορα επίπεδα σημαντικότητας στα συνεχή αποτελέσματα

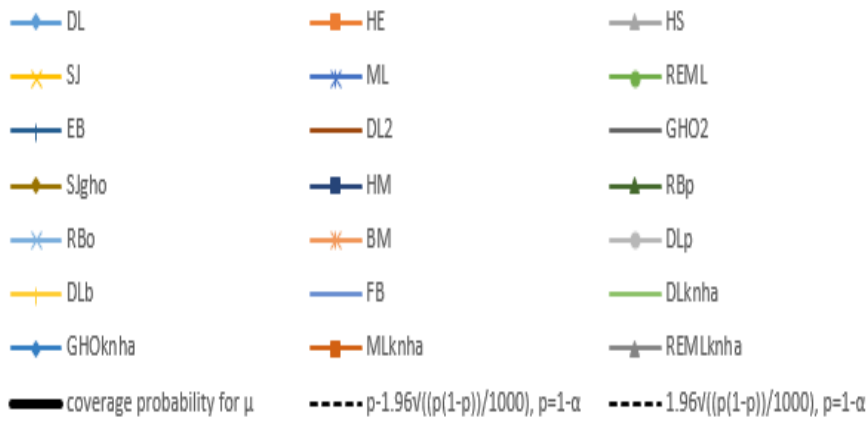


Πίνακας 5.19 (β): Σύγκριση του σφάλματος τύπου I για τα διάφορα επίπεδα σημαντικότητας στα συνεχή αποτελέσματα

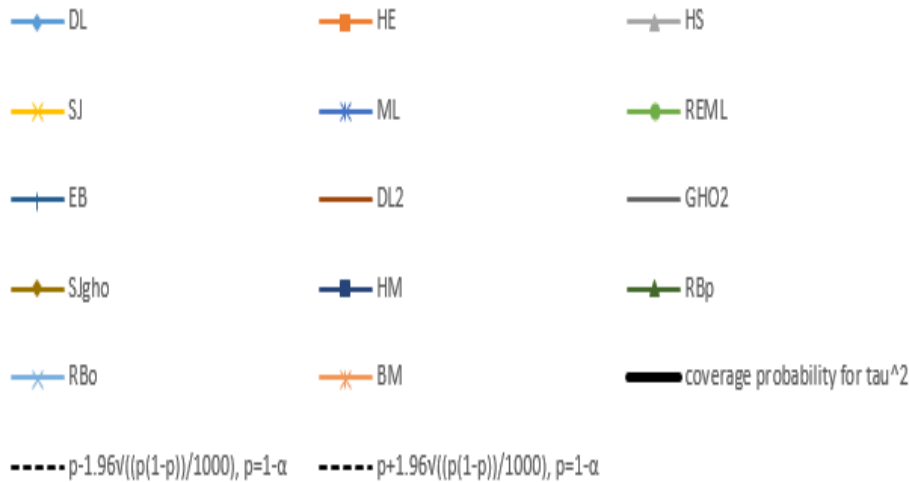


Πίνακας 5.20: Σύγκριση της ισχύς για τα διάφορα επίπεδα σημαντικότητας στα συνεχή αποτελέσματα.

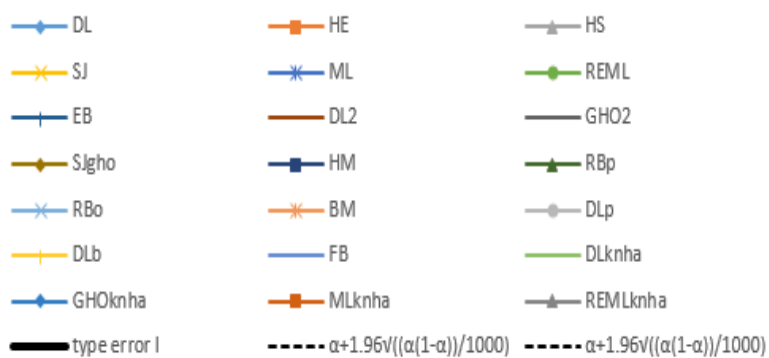
Έπειτα από την παρουσίαση των διαγραμμάτων παρατίθενται τα υπομνήματα για τα παραπάνω διαγράμματα στους Πίνακες 5.23, 5.24, 5.25, 5.26 όταν τα αποτελέσματα είναι διχότομα:



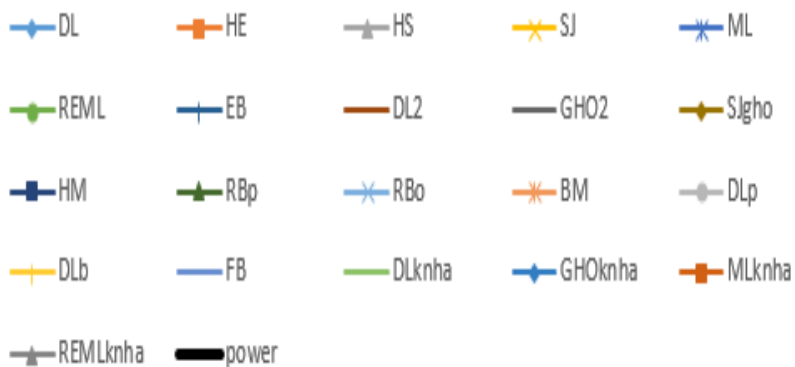
Πίνακας 5.23: Υπόμνημα για τις πιθανότητες κάλυψης του μ με επίπεδο σημαντικότητας α



Πίνακας 5.24: Υπόμνημα για τις πιθανότητες κάλυψης της ετερογένειας με επίπεδο σημαντικότητας α



Πίνακας 5.25: Υπόμνημα για το σφάλμα τύπου I με επίπεδο σημαντικότητας α



Πίνακας 5.26: Υπόμνημα για το σφάλμα τύπου I με επίπεδο σημαντικότητας α

Όταν τα αποτελέσματα είναι συνεχή έχουμε έναν επιπλέον εκτιμητή τον MBH, ο οποίος συμβολίζεται όπως φαίνεται στο Πίνακα 5.27:

—*— MBH

Πίνακας 5.27:

Συμβολισμός MBH

ΠΑΡΑΡΤΗΜΑ

Στο παράρτημα παρατίθενται οι κώδικες που χρησιμοποιήθηκαν για τη διεξαγωγή της μελέτης προσομοίωσης στο στατιστικό πακέτο της R και στο WinBUGS. Όπως έχουμε αναφέρει πραγματοποιήθηκαν δύο κώδικες για την παραγωγή των δύο ειδών δεδομένων που χρησιμοποιήθηκαν στη μελέτη προσομοίωσης, των διχότομων και συνεχών δεδομένων. Μετά από τη δημιουργία των δεδομένων σε κάθε σενάριο τα δεδομένα αποθηκεύονταν μέσω της βιβλιοθήκης MASS σε αρχείο στον υπολογιστή. Έπειτα πραγματοποιήθηκαν δύο κώδικες για τον υπολογισμό των μετρήσεων στα δύο είδη δεδομένων. Για τον υπολογισμό του εκτιμητή FB ως αποτέλεσμα μιας μαρκοβιανής αλυσίδας προγραμματίσαμε στο μεϋζιανό περιβάλλον του WinBUGS. Παραθέτουμε τους κώδικες για την πραγματοποίηση ενός σεναρίου $\tau^2 = 0, \theta = 0, k = 10, a = 0.05$.

Κατά τον υπολογισμό των διαστημάτων εμπιστοσύνης για τους εκτιμητές από την metafor με την μη παραμετρική μέθοδο bootstrap χρησιμοποιήθηκε η διαδικασία από (Viechtbauer, 2015).

Προεργασία για την πραγματοποίηση της μελέτης προσομοίωσης αποτελεί η εγκατάσταση των απαραίτητων βιβλιοθηκών στο πακέτο λογισμικού της R:

```
install.packages ("MASS")  
install.packages ("metafor")  
install.packages ("boot")  
install.packages ("R2WinBUGS")
```

1. Δημιουργία δεδομένων για τα διχότομα αποτελέσματα

```
# generator  
library (MASS)  
generate <-function (tau, theta, k, B) {
```

```
#initialize matrix
pinakasyi <-matrix (1: k, ncol=1)
pinakasvi <-matrix (1: k, ncol=1)

for (i in 1: B) {

  #generate data

  #generate the treatment effect  $\theta_i$  for each trial
  thi <-rnorm (k, theta, sqrt (tau))
  #generate within-study sample sizes for the treatment (T)
  #and control (C) groups, nit and nic, respectively
  ni <- sample (20:200, k, replace=T)
  nit =ni
  nic =ni
  #obtain the outcome ci for the control group
  pic <-runif (k, 0.05, 0.65)
  #ci, response
  ci <-rbinom (k, nic, pic)
  #obtain the outcome ai for the treatment group
  pit =pic*exp (thi) / (1- pic + pic*exp (thi))
  #ai, response
  ai <-rbinom (k, nit, pit)
  #calculate no response bi, di
  bi = nit-ai
  di = nic-ci

  #if any of ai, bi, ci, di is zero put the value 0.5
  for (j in 1: k) {
    if (ai[j] == 0) {
      ai[j] = ai[j] + 0.5
    }
    if (bi[j] == 0) {
```

```

    bi[j] = bi[j] + 0.5
  }
  if (ci[j] == 0) {
    ci[j] = ci[j] + 0.5
  }
  if (di[j] == 0) {
    di[j] = di[j] + 0.5
  }
}
#calculate the treatment effect
yi = log ((ai*di) / (bi*ci))
#estimate the within-study variance
vi = 1/ai + 1/bi + 1/ci + 1/di
pinakasyi <- cbind (pinakasyi, yi)
pinakasvi<- cbind (pinakasvi, vi)
#end of generate data
}
write.matrix (pinakasvi, file = "pinakas_vi1.csv", sep = ",")
write.matrix (pinakasyi, file = "pinakas_yi1.csv", sep = ",")

}
generate (tau=0, theta=0, k=10, B=1000)
#end of generate

```

2. Δημιουργία δεδομένων για τα συνεχή αποτελέσματα

```

# generator
library (MASS)

generate <- function (tau, theta, k, B) {
  #initialize
  ni = c (NA, k)

```

```

nit = c (NA, k)
nic = c (NA, k)
Azic = c (NA, k)
Azit = c (NA, k)
sip = c (NA, k)
y = c (NA, k)
s = c (NA, k)
J = c (NA, k)
yi = c (NA, k)
vi = c (NA, k)
sit = c (NA, k)
sic = c (NA, k)

```

#initialize matrix

```

pinakasyi <- matrix (1: k, ncol=1)
pinakasvi<-matrix (1: k, ncol=1)
pinakasnit <- matrix (1: k, ncol=1)
pinakasJ <- matrix (1: k, ncol=1)

for (i in 1: B) {
  #generate the treatment effect  $\theta_i$  for each trial
  thi <- rnorm (k, theta, sqrt (tau))
  #generate within-study sample sizes for the treatment (T)
  #and control (C) groups, nit and nic, respectively
  ni <- sample (20:200, k, replace=T)

  for (j in 1: k) {

    nit[j] = ni[j]
    nic[j] = ni[j]
    #generate nic observations zic for the control group
    zic <- rnorm (nic[j], 0, 1)
    #nit observations zit for the treatment group
    zit<- rnorm (nit[j], thi, 1)

```

```

Azic[j] = sum (zic) /nic[j]
Azit[j] = sum (zit) /nit[j]

sit[j] = (1/ (nit[j]-1))*sum ((zit-Azit[j]) ^2)
sic[j] = (1/ (nic[j]-1))*sum ((zic-Azic[j]) ^2)
sip[j] = sqrt (((nit[j]-1)* ((sit[j]) ^2) + (nic[j]-1)*((sic[j]) ^2)) / (2*ni[j]-2))

y[j] = (Azit[j]-Azic[j]) /sip[j]

s[j] = (8+(y[j]) ^2) / (4*ni[j])
J[j] = 1- (3/ (8*ni[j]-9))
#calculate the treatment effect
yi[j] = J[j]*y[j]
#estimate the within-study variance
vi[j] = ((J[j]) ^2)*s[j]
}
pinakasyi <-cbind (pinakasyi, yi)
pinakasvi <-cbind (pinakasvi, vi)
pinakasnit <-cbind (pinakasnit, nit)
pinakasJ <-cbind (pinakasJ, J)

#end of generate data

}
write.matrix (pinakasvi, file = "pinakas_vi.csv", sep = ",")
write.matrix (pinakasyi, file = "pinakas_yi.csv", sep = ",")
write.matrix (pinakasnit, file="pinakas_nit.csv", sep = ",")
write.matrix (pinakasJ, file="pinakas_J.csv", sep = ",")
}

generate (tau=0, theta=0, k=10, B=1000)
#end of generate

```

3. Υπολογισμός μετρήσεων για τα διχότομα αποτελέσματα (συνεχή αποτελέσματα με bold)

Για τον υπολογισμό των μετρήσεων στα συνεχή αποτελέσματα ο κώδικας είναι ίδιος με τον κώδικα στα διχότομα εκτός από κάποιες διαφορές και τον υπολογισμό ενός ακόμα εκτιμητή, του MBH ο οποίος χρησιμοποιείται μόνο για τα συνεχή. Οι διαφορές καθώς και ο υπολογισμός του εκτιμητή MBH παρατίθενται παρακάτω

```
#packages
library (metafor)
library (boot)
library (R2WinBUGS)
#k = number of studies
#tau=heterogeneity
# B =number of repetitions

###do the simulation
simulation<-
function(tau,theta,theta0,k,B,a,fileResult,fileyi,filevi,workdir,winbugsdir)
(simulation<-
function(tau,theta,theta0,k,B,a,fileResult,fileyi,filevi,filenit,fileJ,work
dir,winbugsdir)
{
  #initialize heterogeneity estimators
  t1<-rep (NA, B)
  t2<-rep (NA, B)
  ⋮
  ⋮
  t20<-rep (NA, B)
  t21<-rep (NA, B)
  (t22<-rep (NA, B))
  #initialize lower and upper interval for summary estimate
  l1<-rep (NA, B)
  r1<-rep (NA, B)
  l2<-rep (NA, B)
```

```
r2<-rep (NA, B)
  ⋮
  ⋮
l20<-rep (NA, B)
r20<-rep (NA, B)
l21<-rep (NA, B)
r21<-rep (NA, B)
(l22<-rep (NA, B)
r22<-rep (NA, B))
#initialize lower and upper interval for heterogeneity estimate
lh1<-rep (NA, B)
rh1<-rep (NA, B)
lh2<-rep (NA, B)
rh2<-rep (NA, B)
  ⋮
  ⋮
lh17<-rep (NA, B)
rh17<-rep (NA, B)
lh18<-rep (NA, B)
rh18<-rep (NA, B)
(lh19<-rep (NA, B)
rh19<-rep (NA, B))

#initialize counter to compute observed bias

coo1=0
coo2=0
  ⋮
  ⋮
coo20=0
coo21=0
(coo22=0)

#initialize counter to compute Mean Square Error (MSE)
```

```
cc1=0
cc2=0
⋮
⋮
cc20=0
cc21=0
(cc22=0)
```

```
#initialize counter to compute how many times theta containing
#in confidence interval of summary estimate this is useful
# for computing coverage probability for summary estimate.
```

```
counter1=0
counter2=0
⋮
⋮
counter20=0
counter21=0
(counter22=0)
```

```
#initialize counter to compute how many times theta doesn't
# containing in confidence interval of summary estimate
#this is useful for computing type error I
```

```
count1=0
count2=0
⋮
⋮
count20=0
count21=0
(count22=0)
```

```
#initialize counter to compute how many times theta containing
#in confidence interval of heterogeneity estimate
#this is useful for computing coverage probability
#for heterogeneity estimate.
```

```
co1=0
```



```
co2=0
```

```
⋮  
⋮
```

```
co17=0
```

```
co18=0
```

```
(co19=0)
```

```
#initialize counter to compute how many times theta0 doesn't  
# containing in confidence interval of summary estimate from  
# simulated data with theta, this is useful for computing power.
```

```
c1=0
```

```
c2=0
```

```
⋮  
⋮
```

```
c20=0
```

```
c21=0
```

```
(c22=0)
```

```
#import tables with generated data
```

```
pinakasyi <- as.matrix (read.table (fileyi, header=T, row.names=1, sep =  
",", as.is=F))
```

```
colnames (pinakasyi) = NULL
```

```
rownames (pinakasyi) = NULL
```

```
pinakasvi <- as.matrix (read.table (filevi, header=T, row.names=1, sep =  
",", as.is=F))
```

```
colnames (pinakasvi) = NULL
```

```
rownames (pinakasvi) = NULL
```

```
(pinakasnit <- as.matrix (read.table (filenit, header=T, row.names=1,  
sep = ",", as.is=F))
```

```
colnames (pinakasnit) = NULL
```

```
rownames (pinakasnit) = NULL
```

```

pinakasJ <- as.matrix (read.table (fileJ, header=T, row.names=1, sep
= ",", as.is=F))
colnames (pinakasJ) = NULL
rownames (pinakasJ) = NULL

#end of import tables
ni <- runif (k, 50, 500)
#create function metanalysisi for computing estimators
metanalysisi<- function(y=yi, v=vi)
{

wi =1/v
mektimitis = sum (wi * y) /sum (wi)
Q=sum (wi*((y-mektimitis) ^2))

thm=(Q^2)/((2*(k-1)+Q)*(sum(wi)-(sum(wi^2)/sum(wi))))
#estimator HM
thmteliko = thm

ymeso = sum(y) /k
trb0= sum ((y-ymeso) ^2)/ (k+1) - ((sum (ni)-k)*(k+1)*sum (v))/ (sum
(ni-k+2)*k*(k+1))
#estimator Rukhin Bayes RBo
trb0teliko=max (0, trb0)
#estimator Rukhin Bayes RBp
trbpteliko =sum ((yi-ymeso) ^2)/ (k+1)

result2<<-list (thmteliko=thmteliko, trb0teliko = trb0teliko, trbpteliko =
trbpteliko)
}
(
metanalysisi <- function(y=yi, v=vi, n=ni, s=J)
{
wi = 1/v

```

```

mektimitis = sum (wi*y) / sum (wi)
Q = sum (wi * ((y-mektimitis) ^2))
thm = (Q ^2) / ((2*(k-1) +Q)*(sum (wi)-(sum (wi^2)/sum (wi))))
#estimator HM
thmteliko = thm

ymeso = sum (y) /k
trb0 = sum ((y-ymeso) ^2) / (k+1) - ((sum (ni)-k) * (k+1) * sum(v)) /
(sum (ni-k+2) *k * (k+1))
#estimator Rukhin Bayes RBo
trb0teliko = max (0, trb0)
#estimator Rukhin Bayes RBp
trbpteliko = sum ((yi-ymeso) ^2) / (k+1)

fi = 1-(2*n-4) / ((s ^2)*(2*n-2))
tmbh = sum ((1-fi) * ((y-mektimitis) ^2)) / (k-1) - (2/k) * sum (1/n) -
(1/k) * sum (fi*(y ^2))
#estimator MBH
mbhteliko = max (0, tmbh)

result2 <<- list (thmteliko = thmteliko, trb0teliko = trb0teliko,
trbpteliko
= trbpteliko, mbhteliko = mbhteliko)

}
)
#end of function metanalysisi
#function randomeffect
#compute (1-a) % C.I. for summary estimate
randomeffect <- function (timi, y=yi, v=vi)
{
wir = 1/ (v+ timi)
mre =sum (wir * y) /sum (wir)
sem2 = 1/ sqrt (sum (wir))

```

```

cl2 = mre + qnorm(c (a/2, 1-a/2))*sem2
result3<<-list (cl2 = cl2)
}

#end of function randomeffect
#function to obtain estimator DL from data in non-parametric bootstrap
func1 <- function (dat, indices) {
  res1 <- rma (yi, vi, data=dat, method="DL", subset=indices)
  c (res1$tau2, res1$se.tau2^2)
}
#end of function
#function to obtain estimator HE from data in non-parametric bootstrap
func2 <- function (dat, indices) {
  res2 <- rma (yi, vi, data=dat, method="HE", subset=indices)
  c (res2$tau2, res2$se.tau2^2)
}
#function to obtain estimator HS from data in non-parametric bootstrap
func3 <- function (dat, indices) {
  res3 <- rma (yi, vi, data=dat, method="HS", subset=indices)
  c (res3$tau2, res3$se.tau2^2)
}
#function to obtain estimator SJ from data in non-parametric bootstrap
func4 <- function (dat, indices) {
  res4 <- rma (yi, vi, data=dat, method="SJ", subset=indices)
  c (res4$tau2, res4$se.tau2^2)
}
#function to obtain estimator ML from data in non-parametric bootstrap
func5 <- function (dat, indices) {

  res5 <- try (rma (yi, vi, data=dat, method="ML", subset=indices),
silent=TRUE)
  if (is.element ("try-error", class (res5))) {
    c (NA, NA)
  } else {

```

```
      c (res5$tau2, res5$se.tau2^2)
    }
  }
#function to obtain estimator REML from data in non-parametric bootstrap
func6 <- function (dat, indices) {
  res6<-try (rma (yi, vi, data=dat, method="REML", subset=indices),
  silent=TRUE)
  if (is.element ("try-error", class (res6))) {
    c (NA, NA)
  } else {
    c (res6$tau2, res6$se.tau2^2)
  }
}
#function to obtain estimator EB from data in non-parametric bootstrap
func7 <- function (dat, indices) {
  res7<-try (rma (yi, vi, data =dat, method="EB", subset=indices),
  silent=TRUE)

  if (is.element ("try-error", class (res7))) {
    c (NA, NA)
  } else {
    c (res7$tau2, res7$se.tau2^2)
  }
}
#function to obtain two-step (DL2) estimator from data in non-parametric
bootstrap
func8 <- function (dat, indices) {
  res.DL2 <- rma (yi, vi, data=dat, method="GENQ", weights=1 / (vi +
res1$tau2), subset=indices)
  res8 <- rma (yi, vi, data=dat, tau2=res.DL2$tau2, subset=indices)
  c (res8$tau2, res8$se.tau2^2)
}
#function to obtain two-step (GHO2) estimator from data in non-
parametric bootstrap
```

```

func9 <- function (dat, indices) {
  res.HE2 <- rma (yi, vi, data=dat, method="GENQ", weights=1/ (vi +
res2$tau2), subset=indices)
  res9 <- rma (yi, vi, data=dat, tau2=res.HE2$tau2, subset=indices)
  c (res9$tau2, res9$se.tau2^2)
}
#function to obtain SJgho estimator from data in non-parametric bootstrap
func10 <- function (dat, indices) {
  res10<-rma (yi, vi, data=dat, method="SJ", control = list
(tau2.init=res2$tau2), subset=indices)
  c (res10$tau2, res10$se.tau2^2)
}
#function to obtain DLknha estimator from data in non-parametric
# bootstrap
func11 <- function (dat, indices) {
  res11<-rma (yi, vi, data=dat, method="DL", knha=TRUE,
subset=indices)
  c (res11$tau2, res11$se.tau2^2)
}
#function to obtain HEknha estimator from data in non-parametric
# bootstrap
func12 <- function (dat, indices) {
  res12<- rma (yi, vi, data=dat, method="HE", knha=TRUE,
subset=indices)
  c (res12$tau2, res12$se.tau2^2)
}
#function to obtain MLknha estimator from data in non-parametric
# bootstrap
func13 <- function (dat, indices) {
  res13<-try (rma (yi, vi, data=dat, method="ML", knha=TRUE,
subset=indices), silent=TRUE)
  if (is.element ("try-error", class (res13))) {
    c (NA, NA)
  } else {

```

```

    c (res13$tau2, res13$se.tau2^2)
  }
}
#function to obtain REMLknh estimator from data in non-parametric
# bootstrap
func14 <- function (dat, indices) {
  res14<- try (rma (yi, vi, data=dat, method="REML", knha=TRUE,
subset=indices), silent=TRUE)
  if (is.element ("try-error", class (res14))) {
    c (NA, NA)
  } else {
    c (res14$tau2, res14$se.tau2^2)
  }
}
#function to obtain HM estimator from data in non-parametric bootstrap
func15 <- function (d, j){
  d2 <- d[j,]
  wi = 1/d2$vi
  mektimitis = sum (wi*d2$yi)/sum (wi)
  Q=sum (wi*((d2$yi-mektimitis) ^2))
  thm=(Q^2)/((2*(k-1)+Q)*(sum(wi)-(sum(wi^2)/sum(wi))))
  thmteliko = thm
  c (thmteliko)
}
#function to obtain RBp estimator from data in non-parametric bootstrap
func16<- function (d, j) {
  d2 <- d [j,]
  ymeso = sum (d2$yi)/k
  trbpteliko = sum ((d2$yi-ymeso) ^2) / (k+1)
  c (trbpteliko)
}
#function to obtain RBo estimator from data in non-parametric bootstrap
func17<- function (d, j)
  d2 <- d [j,]

```

```

y meso = sum (d2$yi) /k
trb0 = sum ((d2$yi-y meso) ^2) / (k+1) - ((sum (ni)-k)* (k+1) *sum
(d2$vi)) / (sum (ni-k+2) * k * (k+1))
trb0teliko=max (0, trb0)
c (trb0teliko)
}
#function to obtain BM estimator from data in non-parametric bootstrap
func18<- function (dat, indices) {
resm <- rma (yi, vi, data=dat, method="ML", tau2 = sqrt (res5$tau2),
subset = indices)
seml = resm$se.tau2
tml = sqrt (res5$tau2)
if (tml == 0)
{
t bm = (ab-1)*(seml ^2)
} else {
t bm = (tml/2+ (tml/2)*sqrt (1+ (4*(ab-1)*(seml^2))/res5$tau2)) ^2
}
c (t bm)
}
(
func19 <- function (d, j) {
d2 <- d [j,]
wi = 1 / d2$vi
mektimitis = sum (wi*d2$yi) / sum (wi)
fi = 1- (2*d2$ni-4) / (((d2$J) ^2)*(2 * d2$ni -2))
tmbh = sum ((1-fi) * ((d2$yi - mektimitis) ^2)) / (k-1) - (2/k) * sum
(1/ d2$ni) - (1/k) * sum (fi * (d2$yi ^2))
#estimator MBH
mbhteliko = max (0, tmbh)
c (mbhteliko)
}
)

```



```
for (i in 1: B)
{ #open big loop -----

#Import data.....
yi <- pinakasyi [, i]
vi <- pinakasvi [, i]
(ni <- pinakasnit [, i]
J <- pinakasJ [, i])
dat <- data.frame (yi, vi)
(dat <- data.frame (yi, vi, ni, J))
#End of import data.....
#Computation heterogeneity estimators.....

#rma function from package metafor
#estimators from metafor
#Dersimonian and Laird
res1 <- rma (yi = yi, vi = vi, method="DL")
#Hedges
res2 <- rma (yi = yi, vi = vi, method="HE")
#Hunter-Schmidt
res3 <- rma (yi = yi, vi = vi, method="HS")

#Sidik-Jonkman
res4 <- rma (yi = yi, vi = vi, method="SJ")
#Maximum-likelihood
res5<- rma (yi = yi, vi = vi, method="ML")
#Restricted Maximum Likelihood
res6<- rma (yi = yi, vi = vi, method="REML")
#Paule-Mandel
res7<- rma (yi = yi, vi = vi, method="EB")
#two-step Dersimonian and Laird (DL2)
res.DL2 <- rma (yi = yi, vi = vi, method = "GENQ", weights=1/ (vi +
res1$tau2))
res8 <- rma (yi=yi, vi=vi, tau2=res.DL2$tau2)
```

```

#two-step Hedges-Olkin (GHO2)
res.HE2 <- rma (yi = yi, vi = vi, method="GENQ", weights=1/ (vi +
res2$tau2))
res9 <- rma (yi = yi, vi = vi, tau2 = res.HE2$tau2)
#Sidik-Jonkman with initial GHO estimator (SJgho)
res10<-rma (yi = yi, vi = vi, method="SJ", control = list (tau2.init
=res2$tau2))
# Knapp and Hartung estimators (DLknha, HEknha, MLknha,
# REMLknha)
res11 <- rma (yi = yi, vi = vi, method= "DL", knha=TRUE)
res12 <- rma (yi = yi, vi = vi, method= "HE", knha=TRUE)
res13 <- rma (yi = yi, vi = vi, method= "ML", knha=TRUE)
res14 <- rma (yi = yi, vi = vi, method="REML" , knha=TRUE)

#estimators without metafor
#Positive Dersimonian and Laird
  if (res1$tau2 <= 0)
  {
    tdlp =0.01
  } else {
    tdlp = res1$tau2
  }
#Non-parametric bootstrap DL (DLb) estimator
# bootstrapping with 500 replications with calling the function func1
apot1<- boot (dat, func1, R=500)
#end of non-parametric bootstrap for DL estimator
#non-parametric bootstrap version of the DL (DLb) estimator
dlb =mean (apot1$t [1:500, 1])
data <-list (k = k, y = dat$yi, v=dat$vi)
# Fully Bayesian estimator
FullyBayesian.sim <- bugs (data, inits = NULL, c ("mean" ,"tau"),
model. file="c:/FullyBayesian.odc", debug=F, bugs.directory =
winbugsdir, program="WinBUGS", working.directory = workdir)
#keep estimate FB

```

```
fb <- FullyBayesian.sim$mean [2]
#estimator FB
fullb = fb$tau

fbb <- FullyBayesian.sim$mean [1]
#summary estimate of FB estimator
fulb = fbb$mean
x <- FullyBayesian.sim$sd [1]
clfb = x$mean

#call function metanalysisi to compute estimators

  metanalysisi (yi, vi)
(
  metanalysisi (yi, vi, ni, J)
)
#keep estimators out of function metanalysisi
#estimator HM
hm = result2$thmteliko
#estimator RBp
rbp = result2$trbpteliko
#estimator RBo
rbo = result2$trb0teliko
(
#estimator MBH
mbh = result2$mbhteliko
)
#estimator BM
ab = 2
resm <- rma (yi = yi, vi = vi, method="ML", tau2 = sqrt (res5$tau2))
seml = resm$se.tau2
tml = sqrt (res5$tau2)
if (tml == 0)
```

```

{
  tbm = (ab-1)*(seml ^2)
} else {
  tbm = (tml/2+ (tml/2)*sqrt (1+ (4*(ab-1)*(seml ^2))/res5$tau2)) ^2
}
#end of computation heterogeneity estimators.....
#Computation for summary estimate
# (1 - a) % confidence interval for summary estimate from metafor
d1=res1$b + qnorm(c (a/2, 1-a/2))*res1$se
d2=res2$b + qnorm(c (a/2, 1-a/2))*res2$se
d3=res3$b + qnorm(c (a/2, 1-a/2))*res3$se
d4=res4$b + qnorm(c (a/2, 1-a/2))*res4$se
d5=res5$b + qnorm(c (a/2, 1-a/2))*res5$se
d6=res6$b + qnorm(c (a/2, 1-a/2))*res6$se
d7=res7$b + qnorm (c (a/2, 1-a/2))*res7$se
d8=res8$b + qnorm(c (a/2, 1-a/2))*res8$se
d9=res9$b + qnorm (c (a/2, 1-a/2))*res9$se
d10=res10$b+qnorm(c (a/2, 1-a/2))*res10$se
d11 = res11$b + qt (c (1-a/2, 1-a/2), df = k-1)*res11$se
d12 = res12$b + qt (c (1-a/2, 1-a/2), df = k-1)*res12$se
d13 = res13$b + qt (c (1-a/2, 1-a/2), df = k-1)*res13$se
d14 = res14$b + qt (c (1-a/2, 1-a/2), df = k-1)*res14$se

#call function randomeffect
##keep (1 - a) % confidence interval of summary estimates
randomeffect (hm)
d15 = result3$c12
randomeffect (rbp)
d16 = result3$c12
randomeffect (rbo)
d17 = result3$c12
randomeffect (tbm)
d18=result3$c12

```

```

d19 = fulb + qnorm (c (a/2, 1-a/2))*clfb
randomeffect (tdlp)
d20 = result3$cla1
randomeffect (dlb)
d21=result3$cla1
(
randomeffect (mbh)
d22=result3$cla1
)

#Computation (1 - a) % confidence interval for heterogeneity
#NON PARAMETRIC BOOTSRAP

#non-parametric bootstrap confidence interval for DL estimator
#bootstrapping with 500 replications with calling the function func1
apot1 <- boot (dat, func1, R=500)
# get (1 - a) % percentile confidence interval
interval1 <- boot.ci (apot1, conf = 1-a, type="perc", index=1:2)
#keep (1 - a) % percentile confidence interval
test1 = getElement (interval1, "percent")
x1 = test1 [, 4]
y1 = test1 [, 5]
#end of non-parametric bootstrap confidence interval for DL estimator
apot2 <- boot (dat, func2, R=500)
interval2 <- boot.ci (apot2, conf= 1-a, type="perc", index = 1:2)
test2 = getElement (interval2, "percent")
x2 = test2 [, 4]
y2 = test2 [, 5]
#
apot3 <- boot (dat, func3, R=500)
interval3 <- boot.ci (apot3, conf = 1-a, type="perc", index=1:2)
test3 = getElement (interval3, "percent")
x3 = test3 [, 4]
y3 = test3 [, 5]

```

```
#
apot4 <- boot (dat, func4, R=500)
interval4 <- boot.ci (apot4, conf = 1-a, type = "perc", index=1:2)
test4 = getElement (interval4, "percent")
x4 = test4 [, 4]
y4 = test4 [, 5]
#
apot5 <- boot (dat, func5, R=500)
interval5 <- boot.ci (apot5, conf=1-a, type="perc", index=1:2)
test5 = getElement (interval5, "percent")
x5 = test5 [, 4]
y5 = test5 [, 5]
#
apot6 <- boot (dat, func6, R=500)
interval6 <- boot.ci (apot6, conf = 1-a, type="perc", index=1:2)
test6 = getElement (interval6, "percent")
x6 = test6 [, 4]
y6 = test6 [, 5]
#
apot7 <- boot (dat, func7, R=500)
interval7 <- boot.ci (apot7, conf=1-a, type="perc", index=1:2)
test7 = getElement (interval7, "percent")
x7 = test7 [, 4]
y7 = test7 [, 5]
#
apot8 <- boot (dat, func8, R=500)
interval8 <- boot.ci (apot8, conf = 1-a, type = "perc", index=1:2)
test8 = getElement (interval8, "percent")
x8 = test8 [, 4]
y8 = test8 [, 5]
#
apot9 <- boot (dat, func9, R=500)
interval9 <- boot.ci (apot9, conf = 1-a, type = "perc", index = 1:2)
test9 = getElement (interval9, "percent")
```

```
x9 = test9 [, 4]
y9 = test9 [, 5]
#
apot10 <- boot (dat, func10, R=500)
interval10 <- boot.ci (apot10, conf = 1-a, type = "perc", index = 1:2)
test10 = getElement (interval10, "percent")
x10 = test10 [, 4]
y10 = test10 [, 5]
#
apot11 <- boot (dat, func11, R=500)
interval11 <- boot.ci (apot11, conf = 1-a, type = "perc", index=1:2)
test11 = getElement (interval11, "percent")
x11 = test11 [, 4]
y11 = test11 [, 5]
#
apot12 <- boot (dat, func12, R=500)
interval12 <- boot.ci (apot12, conf = 1-a, type = "perc", index=1:2)
test12 = getElement (interval12, "percent")
x12 = test12 [, 4]
y12 = test12 [, 5]
#
apot13 <- boot (dat, func13, R=500)
interval13 <- boot.ci (apot13, conf = 1-a, type = "perc", index=1:2)
test13 = getElement (interval13, "percent")
x13 = test13 [, 4]
y13 = test13 [, 5]
#
apot14 <- boot (dat, func14, R=500)
interval14 <- boot.ci (apot14, conf = 1-a, type = "perc", index = 1:2)
test14 = getElement (interval14, "percent")
x14 = test14 [, 4]
y14 = test14 [, 5]
#
apot15 <- boot (dat, func15, R=500)
```

```
interval15 <- boot.ci (apot15, conf = 1-a, type = "perc")
test15 = getElement (interval15, "percent")
x15 = test15 [, 4]
y15 = test15 [, 5]
#
apot16 <- boot (dat, func16, R=500)
interval16 <- boot.ci (apot16, conf = 1-a, type = "perc")
test16 = getElement (interval16, "percent")
x16 = test16 [, 4]
y16 = test16 [, 5]
#
apot17 <- boot (dat, func17, R=500)
interval17 <- boot.ci (apot17, conf = 1-a, type = "perc")
test17 = getElement (interval17, "percent")
x17 = test17 [, 4]
y17 = test17 [, 5]
#
apot18 <- boot (dat, func18, R=500)
interval18 <- boot.ci (apot18, conf = 1-a, type = "perc")
test18 = getElement (interval18, "percent")
x18 = test18 [, 4]
y18 = test18 [, 5]
(
  apot19 <- boot (dat, func19, R=500)
  interval19 <- boot.ci (apot19, conf = 1-a, type = "perc")
  test19 = getElement (interval19, "percent")
  x19 = test19 [, 4]
  y19 = test19 [, 5]
)

#END OF NON PARAMETRIC BOOTSRAP

#heterogeneity estimators
t1 [i] <- res1$tau2
```



```
t2 [i] <- res2$tau2
```

```
⋮  
⋮
```

```
t13 [i] <- res13$tau2
```

```
t14 [i] <- res14$tau2
```

```
t15 [i] <- hm
```

```
t16 [i] <- rbp
```

```
t17 [i] <- rbo
```

```
t18 [i] <- tbm
```

```
t19 [i] <- fullb
```

```
t20 [i] <- tdlp
```

```
t21 [i] <- dlb
```

```
(t22 [i] <- mbh)
```

```
#calculations for observed bias
```

```
coo1 = coo1 + abs (t1 [i]-tau)
```

```
coo2 = coo2 + abs (t2 [i]-tau)
```

```
⋮  
⋮
```

```
coo20 = coo20 + abs (t20 [i]-tau)
```

```
coo21 = coo21 + abs (t21 [i]-tau)
```

```
(coo22 = coo22 + abs (t22 [i]-tau))
```

```
#calculations for mean square error (MSE)
```

```
cc1 = cc1 + (t1 [i]-tau) ^2
```

```
cc2 = cc2 + (t2 [i]-tau) ^2
```

```
⋮  
⋮
```

```
cc20 = cc20 + (t20 [i]-tau) ^2
```

```
cc21 = cc21 + (t21 [i]-tau) ^2
```

```
( cc22 = cc22 + (t22 [i]-tau) ^2 )
```

```
# lower and upper interval for summary estimate
```

```
l1 [i] < -d1 [1]
```

```
r1 [i] <- d1 [2]
#
l2 [i] <- d2 [1]
r2 [i] <- d2 [2]
  ⋮
  ⋮
l20 [i] <- d20 [1]
r20 [i] <- d20 [2]
#
l21 [i] <- d21 [1]
r21 [i] <- d21 [2]
(
l22 [i] <- d22 [1]
r22 [i] <- d22 [2]
)
#end of lower and upper interval for summary estimate

#lower and upper interval for heterogeneity estimators
lh1 [i] <- x1
rh1 [i] <- y1
#
lh2 [i] <- x2
rh2 [i] <- y2
  ⋮
  ⋮
lh17 [i] <- x17
rh17 [i] <- y17
#
lh18 [i] <- x18
rh18 [i] <- y18
(
lh19 [i] <- x19
rh19 [i] <- y19
```

```
)
#end of lower and upper interval for heterogeneity estimators

#calculations for coverage probability for summary estimates
if (l1 [i] <= theta && theta <= r1 [i]) {
  counter1 = counter1 + 1
}
#
if (l2 [i] <= theta && theta <= r2 [i]) {
  counter2=counter2+1
}
  ⋮
  ⋮
if (l20 [i] <= theta && theta <= r20 [i]) {
  counter20 = counter20 + 1
}
#
if (l21 [i] <= theta && theta <= r21 [i]) {
  counter21 = counter21 + 1
}

(
if (l22 [i] <= theta && theta <= r22 [i]) {
  counter22 = counter22 + 1
}
)
#end calculations for coverage probability for summary estimates
#calculations for coverage probability for heterogeneity estimators
if (lh1 [i] <= tau && tau <= rh1 [i]) {
  co1 = co1 + 1
}
#
if (lh2 [i] <= tau && tau <= rh2 [i]) {
```

```

    co2 = co2 + 1
  }
    ⋮
    ⋮
  if (lh17 [i] <= tau && tau <= rh17 [i]) {
    co17 = co17 + 1
  }
  #
  if (lh18 [i] <= tau && tau <= rh18 [i]) {
    co18 = co18 + 1
  }
  (
    if (lh19 [i] <= tau && tau <= rh19 [i]) {
    co19 = co19 + 1
    }
  )

  #end calculations for coverage probability
  # for heterogeneity estimators

  #calculations for type error I
  if (l1 [i] > theta || r1 [i] < theta) {
    count1 = count1 + 1
  }
  if (l2 [i] > theta || r2 [i] < theta) {
    count2 = count2 + 1
  }
    ⋮
    ⋮
  if (l20 [i] > theta || r20 [i] < theta) {
    count20 = count20 + 1
  }
  #

```

```
if (l21 [i] > theta || r21 [i] < theta) {
  count21 = count21 + 1
}
(
  if (l22 [i] > theta || r22 [i] < theta) {
    count22 = count22 + 1
  }
)
#end calculations for type error I

#calculations for power
if (theta0 < l1 [i] || r1 [i] < theta0) {
  c1 = c1 + 1
}
#
if (theta0 < l2 [i] || r2 [i] < theta0) {
  c2 = c2 + 1
}
  ⋮
  ⋮
if (theta0 < l20 [i] || r20 [i] < theta0) {
  c20 = c20 + 1
}
#
if (theta0 < l21 [i] || r21 [i] < theta0) {
  c21 = c21 + 1
}
(   if (theta0 < l22 [i] || r22 [i] < theta0) {
    c22 = c22 + 1
  }
)

#end calculations for power
```

```

} #close big loop-----
#observed bias
bias1 = coo1/B
bias2 = coo2/B
  ⋮
  ⋮
bias20 = coo20/B
bias21 = coo21/B
(bias22 = coo22/B)

#Mean Square Error (MSE)
mse1=cc1/B
mse2=cc2/B
  ⋮
  ⋮
mse20 = cc20/B
mse21 = cc21/B
(mse22 = cc22/B)
#coverage probability for summary estimate
# proportion of the confidence intervals which
# containing the population mean
coveragem1 = counter1/B
coveragem2 = counter2/B
  ⋮
  ⋮
coveragem20 = counter20/B
coveragem21 = counter21/B
(coveragem22 = counter22/B)

#coverage probability for heterogeneity estimators
coveragetau1 = co1/B
coveragetau2 = co2/B
  ⋮
  ⋮

```

```

coveragetau17 = co17/B
coveragetau18 = co18/B
(coveragetau19 = co19/B)

```

#type error I

```

typerror1 = count1/B
typerror2 = count2/B
  ⋮
  ⋮
typerror20 = count20/B
typerror21 = count21/B
(typerror22 = count22/B)

```

#power

```

pow1 = c1/B
pow2 = c2/B
  ⋮
  ⋮
pow20 = c20/B
pow21 = c21/B
(pow22 = c22/B)

```

w1 = c (bias1, bias2, bias3, bias4, bias5, bias6, bias7, bias8, bias9, bias10, bias11, bias12, bias13, bias14, bias15, bias16, bias17, bias18, bias19, bias20, bias21 (, **bias22**))

bia = round (w1, digits=5)

w2 = c (mse1, mse2, mse3, mse4, mse5, mse6, mse7, mse8, mse9, mse10, mse11, mse12, mse13, mse14, mse15, mse16, mse17, mse18, mse19, mse20, mse21 (, **mse22**))

mse = round (w2, digits=5)

w3 = c (coveragem1, coveragem2, coveragem3, coveragem4, coveragem5, coveragem6, coveragem7, coveragem8, coveragem9, coveragem10, coveragem11, coveragem12, coveragem13, coveragem14,

```
coveragem15, coveragem16, coveragem17, coveragem18, coveragem19,
coveragem20, coveragem21 (, coveragem22) )
```

```
coverage_probability_for_μ = round (w3, digits=5)
```

```
w4 = c (coveragetau1, coveragetau2, coveragetau3, coveragetau4,
coveragetau5, coveragetau6, coveragetau7, coveragetau8, coveragetau9,
coveragetau10, coveragetau11, coveragetau12, coveragetau13,
coveragetau14, coveragetau15, coveragetau16, coveragetau17,
coveragetau18, 1 (,coveragetau19), 1, 1)
```

```
(w4 = c (coveragetau1, coveragetau2, coveragetau3, coveragetau4,
coveragetau5, coveragetau6, coveragetau7, coveragetau8,
coveragetau9, coveragetau10, coveragetau11, coveragetau12,
coveragetau13, coveragetau14, coveragetau15, coveragetau16,
coveragetau17, coveragetau18 ,coveragetau19, 1, 1))
```

```
coverage_probability_for_tau = round (w4, digits=5)
```

```
w5 = c (typerror1, typerror2, typerror3, typerror4, typerror5, typerror6,
typerror7, typerror8, typerror9, typerror10 ,typerror11, typerror12,
typerror13, typerror14, typerror15, typerror16, typerror17, typerror18,
typerror19, typerror20, typerror21 (, typerror22) )
```

```
type_error_I = round (w5, digits=5)
```

```
w6 = c (pow1, pow2, pow3, pow4, pow5, pow6, pow7, pow8, pow9,
pow10, pow11, pow12, pow13, pow14, pow15, pow16, pow17, pow18,
pow19, pow20, pow21 (, pow22) )
```

```
power = round (w6, digits=5)
```

```
bias <- bia
```

```
bias <- cbind (bias, mse)
```

```
bias <- cbind (bias, coverage_probability_for_μ)
```

```
bias <- cbind (bias, coverage_probability_for_tau)
```

```
bias <- cbind (bias, type_error_I)
```

```
bias <- cbind (bias, power)
```

```
write.csv2 (bias, fileResult, row.names = c ("DL", "HE", "HS", "SJ",
"ML", "REML", "EB", "DL2", "GHO2", "SJgho", "DLknh",
```



```

"GHOknha", "MLknha", "REMLknha", "HM", "RBp", "RBo", "BM",
"FB", "DLp", "DLb" (, "MBH"))
}
myrand <- sample (1:2000, 1)
original <- "C: /tempBugs/"
semifinal <- paste (original, myrand, sep="")
workdir <- paste (semifinal, "/", sep="")
dir.create (workdir, showWarnings = F, recursive=T)

winbugsdire = "C:/Program Files/WinBUGS1"

simulation (tau=0, theta=0, theta0=0, k=10, B=1000, a=0.05, fileResult =
"senario tau=0 theta=0 k=10 a=0, 05.csv", fileyi="pinakas_yi1.csv",
filevi="pinakas_vi1.csv", workdir, winbugsdire)
(simulation (tau=0, theta=0, theta0=0, k=10, B=1000, a=0.01,
fileResult = "senario tau=0 theta=0 k=10 a=0,01.csv", fileyi =
"pinakas_yi1.csv", filevi = "pinakas_vi1.csv", filenit =
"pinakas_nit1.csv", fileJ = "pinakas_J1.csv", workdir, winbugsdire))

```

#end of simulation

4. Υπολογισμός του εκτιμητή FB στο WinBUGS

```

model{

for(i in 1:k){

#for each study calculate  $y_i, \theta_i$ 

w[i]<-1/v[i]
y[i]~dnorm(theta[i],w[i])
theta[i]~dnorm(mean,prec)

```

```
}  
#take initial values for tau and mean from distributions Uniform(0, 100)  
#and  $N(0, 10^6)$  respectively  
mean~dnorm(0,0.000001)  
tau2<-pow(tau,2)  
prec<-1/tau2  
tau~dunif(0,100)  
  
}
```

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Αλυσίδα Μαρκόφ. (2015, May 10). In Βικιπαίδεια. Retrieved May 5, 2015 from https://el.wikipedia.org/wiki/Αλυσίδα_Μαρκόφ.
2. Bhaumik, D. K., Amatya, A., Normand, S.-L., Greenhouse, J., Kaizar, E., Neelon, B., & Gibbons, R. D. (2012). Meta-Analysis of Rare Binary Adverse Event Data. *Journal of the American Statistical Association*, **107**(498), 555–567.
3. Bootstrapping (statistics). (2015, August 18). In Wikipedia. Retrieved August 22, 2015 from [https://en.wikipedia.org/wiki/Bootstrapping_\(statistics\)](https://en.wikipedia.org/wiki/Bootstrapping_(statistics)).
4. Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis* (first). Wiley.
5. Bowden, J., Tierney, J. F., Copas, A. J., & Burdett, S. (2011). Quantifying, displaying and accounting for heterogeneity in the meta-analysis of RCTs using standard and generalised Q statistics. *BMC Medical Research Methodology*, **11**, 41.
6. Γαλάνης, Π. (2009). Συστηματική ανασκόπηση και μετα-ανάλυση. *Αρχαία Ελληνικής Ιατρικής*, **26**(6), 826–841.
7. Cauty A., & Ripley B. (2015). boot: Bootstrap R (S-Plus) Functions. R package version 1.3-17.
8. Chung, Y., Rabe-Hesketh, S., & Choi, I.-H. (2013). Avoiding zero between-study variance estimates in random-effects meta-analysis. *Statistics in Medicine*, **32**(23), 4071–4089.
9. Chung, Y., Rabe-Hesketh, S., Dorie, V., Gelman, A., & Liu, J. (2013). A Nondegenerate Penalized Likelihood Estimator for Variance Parameters in Multilevel Models. *Psychometrika*, **78**(4), 685–709.
10. Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrics*, **10**, 101-129.
11. Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (second). Lawrence Erlbaum Associates.

12. Dan Jackson, J. B. (2010). How does the DerSimonian and Laird procedure for random effects meta-analysis compare with its more efficient but harder to compute counterparts? *Journal of Statistical Planning and Inference*, **140**(4), 961–970.
13. DerSimonian, R., & Kacker, R. (2007). Random-effects model for meta-analysis of clinical trials: An update. *Contemporary Clinical Trials*, **28**(2), 105–114.
14. DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, **7**(3), 177–188.
15. Friedman, L. (2000). Estimators of Random Effects Variance Components in Meta-Analysis. *Journal of Educational and Behavioral Statistics*, **25**(1), 1–12.
16. Hartung, J. (1999). An Alternative Method for Meta-Analysis. *Biometrical Journal*, **41**(8), 901–916.
17. Hartung, J., & Makambi, K. H. (2003). Reducing the Number of Unjustified Significant Results in Meta-analysis. *Communications in Statistics - Simulation and Computation*, **32**(4), 1179–1190.
18. Hedges, L. (1981). "Distribution theory for Glass' estimator of effect size and related estimators". *Journal of Educational Statistics* **6** (2), 107–128.
19. Higgins, J., & Green, S. (2008). *Cochrane Handbook for Systematic Reviews of Interventions*. Wiley Blackwell.
20. Higgins, J., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, **21**(11), 1539–1558.
21. Hunter, J. E., & Schmidt, F. (2004). *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings* (2nd Edition). Newbury Park, CA: Sage.
22. IntHout, J., Ioannidis, J. P., & Borm, G. F. (2014). The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Medical Research Methodology*, **14**(1), 25.
23. Isenburg, M. (2015, August 4). *LibGuides: Introduction to Evidence-Based Practice : Types of Studies*. Retrieved from

- <http://guides.mclibrary.duke.edu/c.php?g=158201&p=1036068>
24. Kacker, R. N. (2004). Combining information from interlaboratory evaluations using a random effects model. *Metrologia*, **41**(3), 132.
 25. Καράσσα, Φ. (2006). Αρχές και μεθοδολογία της συστηματικής ανασκόπησης της βιβλιογραφίας. *Ελληνική Ρευματολογία*, **17**(4), 289–297.
 26. Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine*, **22**(17), 2693–2710.
 27. Kontopantelis, E., Springate, D. A., & Reeves, D. (2013). A Re-Analysis of the Cochrane Library Data: The Dangers of Unobserved Heterogeneity in Meta-Analyses. *PLoS ONE*, **8**(7), e69930. <http://doi.org/10.1371/journal.pone.0069930>.
 28. Light, R.J., Pillemer D.B. (1984). *Summing up: The Science of Reviewing Research*. Cambridge, Massachusetts.: Harvard University Press.
 29. Lunn, D.J., Thomas, A., Best, N., and Spiegelhalter, D. (2000) WinBUGS -- a Bayesian modelling framework: concepts, structure and extensibility. *Statistics and Computing*, 10:325--337.
 30. Malzahn, U., Böhning, D., & Holling, H. (2000). Nonparametric estimation of heterogeneity variance for the standardised difference used in meta-analysis. *Biometrika*, **87**(3), 619–632.
 31. Markov chain Monte Carlo. (2015, August 5). In Wikipedia. Retrieved August 6, 2015 from https://en.wikipedia.org/wiki/Markov_chain_Monte_Carlo.
 32. Mavridis, D., & Salanti, G. (2014). Exploring and accounting for publication bias in mental health: a brief overview of methods. *Evidence Based Mental Health*, **17**(1), 11–15.
 33. Mean squared error. (n.d.). In Wikipedia. Retrieved June 2, 2015 from https://en.wikipedia.org/wiki/Mean_squared_error.
 34. Morris, C. N. (1983). Parametric Empirical Bayes Inference: Theory and Applications. *Journal of the American Statistical Association*, **78**(381), 47–55.
 35. Nikolakopoulou, A., Mavridis, D., & Salanti, G. (2014). Demysti-

- fyng fixed and random effects meta-analysis. *Evidence-Based Mental Health*, **17**(2), 53–57.
36. Novianti, P. W., Roes, K. C. B., & van der Tweel, I. (2014). Estimation of between-trial variance in sequential meta-analyses: a simulation study. *Contemporary Clinical Trials*, **37**(1), 129–138.
37. Panityakul, T., Bumrungsup, C., & Knapp, G. (2013). On Estimating Residual Heterogeneity in Random-Effects Meta-Regression: A Comparative Study. *Journal of Statistical Theory and Applications*, **12**(3), 253.
38. Παπαϊωάννου, Τ., Φερεντίνο, Κ. (2000). *Μαθηματική Στατιστική (Δεύτερη Έκδοση)*. Σταμούλη.
39. Πατελάρου, Ε., & Μπροκαλάκη, Η. (2010). Μεθοδολογία της Συστηματικής Ανασκόπησης και Μετα-ανάλυσης. *ΝΟΣΗΛΕΥΤΙΚΗ*, **49**(2), 122–130.
40. Paule, R., & Mandel, J. (1982). Consensus Values and Weighting Factors. *JOURNAL OF RESEARCH of the National Bureau of Standards*, **87**, 377 – 385.
41. R Development Core Team (2008). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
42. Rukhin, A. L. (2013). Estimating heterogeneity variance in meta-analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **75**(3), 451–469.
43. Sánchez-Meca, J., & Marín-Martínez, F. (2008). Confidence intervals for the overall effect size in random-effects meta-analysis. *Psychological Methods*, **13**(1), 31–48.
44. Searle, S. R. (1979). *Notes on Variance Component Estimation - A Detailed Account of Maximum Likelihood and Kindred Methodology*. Biometrics Unit, New York State College of Agriculture and Life Sciences, Cornell University, Ithaca, New York.
45. Sidik, K., & Jonkman, J. N. (2007). A comparison of heterogeneity variance estimators in combining results of studies. *Statistics in*

- Medicine, **26**(9), 1964–1981.
46. Sturtz, S., Ligges, U., and Gelman, A. (2005). R2WinBUGS: A Package for Running WinBUGS from R. *Journal of Statistical Software*, **12**(3), 1-16.
 47. Switzer, F. S., Paese, P. W., & Drasgow, F. (1992). Bootstrap estimates of standard errors in validity generalization. *Journal of Applied Psychology*, **77**(2), 123–129.
 48. Thompson, S. G., & Sharp, S. J. (1999). Explaining heterogeneity in meta-analysis: a comparison of methods. *Statistics in Medicine*, **18**(20), 2693–2708.
 49. 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.
 50. Venables, W. N. & Ripley, B. D. (2002). *Modern Applied Statistics with S*. Fourth Edition. Springer, New York. ISBN 0-387-95457-0
 51. Veroniki, A. et al., (2015). Methods to estimate the between-study variance and its uncertainty in meta-analysis. John Wiley & Sons, Ltd. Advance online publication. doi: 10.1002/jrsm.1164
 52. Viechtbauer, W. (2005). Bias and Efficiency of Meta-Analytic Variance Estimators in the Random-Effects Model. *Journal of Educational and Behavioral Statistics*, **30**(3), 261–293.
 53. Viechtbauer, W. (2007a). Approximate Confidence Intervals for Standardized Effect Sizes in the Two-Independent and Two-Dependent Samples Design. *Journal of Educational and Behavioral Statistics*, **32**(1), 39–60.
 54. Viechtbauer, W. (2007b). Confidence intervals for the amount of heterogeneity in meta-analysis. *Statistics in Medicine*, **26**(1), 37–52.
 55. Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, **36**(3), 1-48. URL <http://www.jstatsoft.org/v36/i03/>.

56. Viechtbauer, W. (2014, August 17). Bootstrapping with Meta-Analytic Models. The metafor Package: A Meta-Analysis Package for R. Retrieved August 4, 2015 from [http://www.metafor-project.org/doku.php/tips:bootstrapping_with_ma?s\[\]=bootstrap](http://www.metafor-project.org/doku.php/tips:bootstrapping_with_ma?s[]=bootstrap)