



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΣΤΕΡΕΑΣ ΕΛΛΑΔΑΣ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΒΙΟΪΑΤΡΙΚΗ**

**ΑΠΟΚΛΙΣΕΙΣ ΑΠΟ ΤΗΝ ΙΣΟΡΡΟΠΙΑ
HARDY-WEINBERG ΚΑΙ ΟΙ ΕΠΠΤΩΣΕΙΣ ΤΟΥΣ
ΣΤΙΣ ΜΕΛΕΤΕΣ ΓΕΝΕΤΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ**

Θεοδοσία Χαρίτου

**ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ
Υπεύθυνος
Μπάγκος Παντελής
Επίκουρος Καθηγητής**

Λαμία, 2009

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΡΟΛΟΓΟΣ.....	5
ΠΕΡΙΛΗΨΗ.....	6
ABSTRACT.....	7
ΚΕΦΑΛΑΙΟ 1: ΘΕΩΡΗΤΙΚΟ ΜΕΡΟΣ	9
1.ΕΙΣΑΓΩΓΗ.....	10
1.1.ΠΛΗΘΥΣΜΙΑΚΗ ΓΕΝΕΤΙΚΗ.....	10
1.2. ΝΟΜΟΣ HARDY-WEINBERG	13
1.2.3. ΠΑΡΑΓΟΝΤΕΣ ΠΟΥ ΕΠΗΡΕΑΖΟΥΝ ΤΗΝ ΙΣΟΡΡΟΠΙΑ Η-W	17
1.3. ΤΡΟΠΟΙ ΕΛΕΓΧΟΥ ΤΩΝ ΑΠΟΚΛΙΣΕΩΝ ΑΠΟ ΤΗΝ ΗWE	19
1.3.1. ΕΛΕΓΧΟΣ ΚΑΛΗΣ ΠΡΟΣΑΡΜΟΓΗΣ χ^2	20
1.3.2. ΕΛΕΓΧΟΣ ΠΗΛΙΚΟΥ ΠΙΘΑΝΟΦΑΝΕΙΑΣ	22
1.3.3. ΑΚΡΙΒΗΣ ΕΛΕΓΧΟΣ (EXACT TEST FISHER).....	24
1.4. ΜΕΤΡΑ ΥΠΟΛΟΓΙΣΜΟΥ ΤΩΝ ΑΠΟΚΛΙΣΕΩΝ ΑΠΟ ΤΗΝ ΗWE	26
1.5. ΕΓΚΥΡΟΤΗΤΑ ΤΩΝ ΜΕΛΕΤΩΝ ΓΕΝΕΤΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ.....	28
ΚΕΦΑΛΑΙΟ 2: ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ.....	29
2. ΜΕΤΑ-ΑΝΑΛΥΣΗ.....	30
2.1. ΜΕΘΟΔΟΣ ΤΗΣ ΜΕΤΑ-ΑΝΑΛΥΣΗΣ	30
2.2. ΣΤΑΤΙΣΤΙΚΗ ΑΝΑΛΥΣΗ	30
2.2.1 ΜΟΝΤΕΛΟ ΣΤΑΘΕΡΩΝ ΕΠΙΔΡΑΣΕΩΝ (FIXED EFFECT MODEL)....	31
2.2.2. ΜΟΝΤΕΛΟ ΤΥΧΑΙΩΝ ΕΠΙΔΡΑΣΕΩΝ (RANDOM EFFECT MODEL) 32	
2.2.3. ΣΧΕΤΙΚΟΣ ΛΟΓΟΣ ΠΙΘΑΝΟΤΗΤΩΝ ODDS-RATIO (OR)	33
2.4 ΜΕΤΑ-ΑΝΑΛΥΣΗ ΣΤΙΣ ΜΕΛΕΤΕΣ ΓΕΝΕΤΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ.....	37

2.4.2. ΤΡΟΠΟΠΟΙΗΣΗ ΓΙΑ ΤΗΝ ΕΝΣΩΜΑΤΩΣΗ ΤΩΝ ΑΠΟΚΛΙΣΕΩΝ ΑΠΟ ΤΗΝ ΙΣΟΡΡΟΠΙΑ ΗW	39
2.5. ΣΤΑΤΙΣΤΙΚΟ ΠΑΚΕΤΟ STATA (VERSION 10) SOFTWARE	43
2.5.1. ΓΕΝΙΚΗ ΠΕΡΙΓΡΑΦΗ ΤΟΥ STATA.....	44
2.5.2 ΕΙΣΑΓΩΓΗ ΔΕΔΟΜΕΝΩΝ.....	45
ΚΕΦΑΛΑΙΟ 3: ΑΠΟΤΕΛΕΣΜΑΤΑ.....	48
3. ΑΠΟΤΕΛΕΣΜΑΤΑ	49
3.1. ΠΙΝΑΚΑΣ ΜΕ ΤΙΣ 71 ΜΕΤΑ-ΑΝΑΛΥΣΕΙΣ	50
3.2. ΜΕΘΟΔΟΛΟΓΙΑ ΠΡΟΓΡΑΜΜΑΤΟΣ.....	53
3.3. ΠΡΟΓΡΑΜΜΑ STATA	54
3.5. ΕΝΤΟΛΕΣ ΣΤΟ STATA	60
3.6. ΣΥΝΟΠΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ.....	70
ΚΕΦΑΛΑΙΟ 4: ΣΥΜΠΕΡΑΣΜΑΤΑ.....	115
4.1. ΣΥΜΠΕΡΑΣΜΑΤΑ	116
ΚΕΦΑΛΑΙΟ 5: ΒΙΒΛΙΟΓΡΑΦΙΑ	120
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	121
6. ΠΑΡΑΡΤΗΜΑ.....	124

*Αφιερωμένο στους αγαπημένους μου γονείς,
στην αγαπημένη μου αδερφή.*

ΠΡΟΛΟΓΟΣ

Η παρούσα διπλωματική εργασία με θέμα “Αποκλίσεις από την Ισορροπία Hardy-Weinberg και οι επιπτώσεις τους στις μελέτες Γενετικής Συσχέτισης”, πραγματοποιήθηκε εξολοκλήρου στο Τμήμα Πληροφορικής με εφαρμογές στην Βιοϊατρική του Πανεπιστημίου Στερεάς Ελλάδας, κατά τη διάρκεια του ακαδημαϊκού έτους, 2008-2009, υπό την επίβλεψη του Επίκουρου Καθηγητή, κ. Παντελή Μπάγκου.

Σε αυτό το σημείο θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου για την εμπιστοσύνη που μου έδειξε κατά την ανάθεση της διπλωματικής εργασίας και για την πολύτιμη βοήθειά που μου προσέφερε, για τις υποδείξεις και τα ενθαρρυντικά του λόγια, σε όλα τα βήματα, για την διεκπεραίωση της.

Θερμά ευχαριστώ τον επιβλέποντα καθηγητή κ. Ιωαννίδη, για τις γόνιμες συζητήσεις μας, την συμπαράσταση του και την επιμονή του, για την βέλτιστη τελειοποίηση της διπλωματικής μου.

Επίσης, οφείλω να ευχαριστήσω και τα υπόλοιπα δύο μέλη της τριμελούς εξεταστικής επιτροπής, τον Επίκουρο Καθηγητή κ. Πλαγιανάκο Βασίλειο και τον Επίκουρο Καθηγητή κ. Μαγκλογιάννη Ηλία, για το χρόνο που μου αφιέρωσαν.

Τέλος, ευχαριστώ την οικογένειά μου για την υποστήριξη και την αγάπη της.

ΠΕΡΙΛΗΨΗ

Ένα από τα σημαντικότερα επιτεύγματα στον τομέα της Πληθυσμιακής Γενετικής είναι η διατύπωση του νόμου των Hardy-Weinberg. Ο νόμος αυτός μας επιτρέπει να συγκρίνουμε στοιχεία που προκύπτουν από γενετικές αναλύσεις πληθυσμών. Ας υποθέσουμε έναν πληθυσμό με δύο αλληλόμορφα (A, B) όπου η πιθανότητα για το αλληλόμορφο A είναι π^A και για το B είναι π^B , οι αναμενόμενες συχνότητες των γονοτύπων AA, AB και BB θα είναι $(\pi^A)^2$, $2\pi^A\pi^B$ και $(\pi^B)^2$ αντίστοιχα. Στην περίπτωση όπου οι αναμενόμενες συχνότητες ταιριάζουν με τις παρατηρούμενες, λέμε ότι ο πληθυσμός είναι σε ισορροπία Hardy-Weinberg (HWE). Στις μελέτες γενετικής συσχέτισης ασθενών-μαρτύρων (cases-controls) βασισμένες σε πληθυσμούς, η ομάδα των μαρτύρων αναμένεται να είναι σε ισορροπία HW και με αυτό τον τρόπο ελέγχονται για την ανίχνευση τυχών αποκλίσεων από αυτήν. Η απόκλιση από την ισορροπία HW, θα μπορούσε να είναι ένας δείκτης της αξιοπιστίας της μελέτης δεδομένου ότι η έλλειψη της ευθύνεται για εσφαλμένες γονοτυπικές συσχετίσεις. Η μέθοδος της μετά-ανάλυσης που μελετά γενετικές συσχετίσεις σε πληθυσμούς ασθενών-μαρτύρων χαρακτηρίζεται ως στατιστικό εργαλείο που συνδυάζει τα αποτελέσματα των μελετών. Διάφορες μεθοδολογίες έχουν προταθεί για να προσαρμόζονται τα αποτελέσματα των αναλύσεων για τις αποκλίσεις από την ισορροπία HW. Μια απλή προσέγγιση είναι να εκτελούνται όλες οι απαραίτητες δοκιμές ώστε οι μελέτες που εκτρέπονται σημαντικά να αποκλειστούν. Μία άλλη προσέγγιση είναι να γίνει τροποποίηση των εκτιμήσεων των σχετικών λόγων των πιθανοτήτων (OR) μαζί με τις διακυμάνσεις τους, για την ύπαρξη αποκλίσεων από την ισορροπία HW και να γίνει η ανάλυση χρησιμοποιώντας τις τροποποιημένες εκτιμήσεις. Η μετά-ανάλυση εφαρμόζεται με τρεις διαφορετικές επιλογές με σκοπό να γίνει έλεγχος και να διαπιστωθεί η καλύτερη μεθοδολογία σε κάθε περίπτωση. Αρχικά στην πρώτη επιλογή, δεν εφαρμόζεται καμία τροποποίηση, στην δεύτερη επιλογή γίνονται τροποποιήσεις στους εκτιμητές με μεθόδους που έχουν προταθεί από ερευνητές και στην τρίτη επιλογή αφαιρούνται οι μελέτες εκείνες που δεν είναι σε ισορροπία HW, κρατώντας μόνο εκείνες που είναι σε ισορροπία. Σε αυτή την

πτυχιακή εργασία στο πλαίσιο της μετά-ανάλυσης ελέγξαμε τον βαθμό απόκλισης από την ισορροπία HW σε μελέτες γενετικών συσχετίσεων με όλες τις μεθόδους, με τη βοήθεια του στατιστικού πακέτου STATA 10. Χρησιμοποιήσαμε 71 δημοσιευμένες μετά-αναλύσεις οι οποίες δημοσιεύτηκαν μεταξύ των ετών 2005-2006 και περιέχουν αναλυτικά όλους τους γονοτύπους. Τα συμπεράσματα της εργασίας καταδεικνύουν ότι οι αποκλίσεις από την ισορροπία HW πρέπει να λαμβάνονται υπόψη στη διεξαγωγή μετα-αναλύσεων, ενώ μελλοντική έρευνα πρέπει να γίνει στην κατεύθυνση του να γίνονται παρόμοιες αναλύσεις χωρίς ταυτόχρονα να υποτίθεται ένα συγκεκριμένο μοντέλο κληρονομικότητας.

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Ισορροπία Hardy-Weinberg, μετά-ανάλυση, μελέτες γενετικών συσχετίσεων, γενετική επιδημιολογία, ασθενείς-μάρτυρες, σχετικός λόγος πιθανοτήτων, διακύμανση.

ABSTRACT

The Hardy-Weinberg law states that if some conditions, such as random mating or new no mutations are met, then the genotypes are completely specified by the population allele frequencies. For a simple bi-allelic locus (A, B) where the probability for allele A is π^A and that of B is π^B , the expected frequencies of the three genotypes AA, AB and BB are $(\pi^A)^2$, $2\pi^A\pi^B$ and $(\pi^B)^2$. In case the expected frequencies match the observed ones, the population is in Hardy-Weinberg equilibrium. As it is apparent, deviations from HWE should play an important role in genetic association studies. However, evaluation studies have shown that most genetic-association studies do not appropriately report deviations from HWE. More importantly, the deviations from HWE are an important factor influencing the heterogeneity of the results in a meta-analysis. For dealing with deviations from HWE in meta-analysis, several approaches have been proposed. One approach is to perform the necessary test for detecting deviations and exclude from the analysis those studies

that deviate. Another approach is to adjust the estimates of log-odds-ratio along with their variances for the degree of deviation and perform the analysis using the adjusted estimates. We review all methods that are proposed for adjusting the estimates of genetic association in the context of meta-analysis. The methods are compared on a set of 71 already published meta-analyses. These meta-analyses were performed during the years 2005-2006 and contain analytical information for all genotypes. We conclude that deviations from HWE should be taken into account in a meta-analysis setting, by properly adjusting the variance of the estimates. Future studies should be performed in order to incorporate such adjustments in a genetic-model free approach for meta-analysis, that is, in a methodology that does not assume the genetic model of inheritance beforehand.

KEY WORDS: Hardy-Weinberg equilibrium, meta-analysis, genetic association studies, odds ratio, variances, adjustment, cases-controls, genetic epidemiology.

ΚΕΦΑΛΑΙΟ 1: ΘΕΩΡΗΤΙΚΟ ΜΕΡΟΣ

1.ΕΙΣΑΓΩΓΗ

Τις τελευταίες δύο δεκαετίες, οι μελέτες γενετικής συσχέτισης (genetic association studies) έχουν χρησιμοποιηθεί, για να βρεθούν πιθανές συσχετίσεις μεταξύ πολυμορφισμών σε συγκεκριμένα γονίδια, με τον κίνδυνο νόσησης από διάφορες πολυπαραγοντικές ασθένειες (π.χ. καρκίνος). Οι συγκεκριμένες μελέτες είναι τυπικές cases-controls (ασθενών-μαρτύρων), όπου εξετάζεται η συχνότητα εμφάνισης της εν λόγω μετάλλαξης στους ασθενείς και στους μάρτυρες. Αν προκύψει στατιστικά σημαντική διαφορά μεταξύ της συχνότητας της μετάλλαξης στους ασθενείς και στους μάρτυρες, τότε ενδεχομένως ο πολυμορφισμός (η μετάλλαξη δηλαδή) σχετίζεται με την ασθένεια.

Κάνοντας μια τυπική μελέτη ασθενών-μαρτύρων μπορούμε να ελέγξουμε, αν ο πολυμορφισμός αυτός επηρεάζει την λειτουργία ή την παραγωγή ενός μορίου, που αποτελεί βασικό παράγοντα σε μία ασθένεια. Παρ' όλα αυτά η διαδικασία επιλογής των μαρτύρων, απαραίτητων για την παραπάνω μελέτη, αποτελεί θα λέγαμε ένα από τα βασικά προβλήματα της έρευνας.

Βασική προϋπόθεση στην έρευνα αυτή, είναι η χρήση μεγάλου δείγματος για την ύπαρξη μίας ενδεχόμενης συσχέτισης. Ωστόσο, σε πολλές περιπτώσεις έχουμε αντιφατικά αποτελέσματα μεταξύ διάφορων ερευνητικών ομάδων και σε πολλές περιπτώσεις είναι ιδιαίτερα δύσκολο να αποσαφηνιστεί για ποιο λόγο προκύπτουν αντικρουόμενα συμπεράσματα (Trikalinos, Salanti et al. 2006). Ένα ιδιαίτερο πρόβλημα που πρέπει να αποσαφηνιστεί σε αυτές τις μελέτες είναι αν η κατανομή των γονοτύπων στους μάρτυρες αντιστοιχεί σε αυτά που προβλέπει η αρχή των Hardy-Weinberg ή αλλιώς αν βρίσκονται σε Hardy-Weinberg (H-W) ισορροπία (Hardy 1908).

1.1.ΠΛΗΘΥΣΜΙΑΚΗ ΓΕΝΕΤΙΚΗ

Κύριος τομέας της επιστημονικής μελέτης, είναι η Πληθυσμιακή Γενετική, η οποία αναπτύχθηκε στην πρώτη δεκαετία του 20ου αιώνα και από τότε μέχρι σήμερα έχει προχωρήσει σημαντικά, χάρη στην συντονισμένη προσπάθεια ερευνητών από

όλο τον κόσμο. Η Πληθυσμιακή Γενετική ασχολείται με τη μελέτη της κατανομής των γονιδίων στους πληθυσμούς και του τρόπου με τον οποίο οι συχνότητες των γονιδίων και των γονοτύπων εμφανίζονται. Επίσης ασχολείται με την γενετική δομή των πληθυσμών, δηλαδή με τους τύπους και τις συχνότητες των γονιδίων και των γονοτύπων που τους αντιπροσωπεύουν καθώς και με τους μηχανισμούς (φυσική επιλογή, γενετική μετάλλαξη, γενετική παρέκκλιση, γονιδιακή ροή) οι οποίοι μεταβάλλουν την γενετική δομή των πληθυσμών (Khoury and Little 2000). Αποτελεί βασική επιστήμη για την γενετική βελτίωση οργανισμών, ενώ συμμετέχει στην σύγχρονη εξελικτική σύνθεση, μια ενοποιημένη θεωρία για την προέλευση και συμπεριφορά όλων των μορφών ζωής στον πλανήτη. Όπως αναφέρει και το όνομά της, έχει μια προοπτική με κέντρο τον πληθυσμό, και όχι το άτομο.

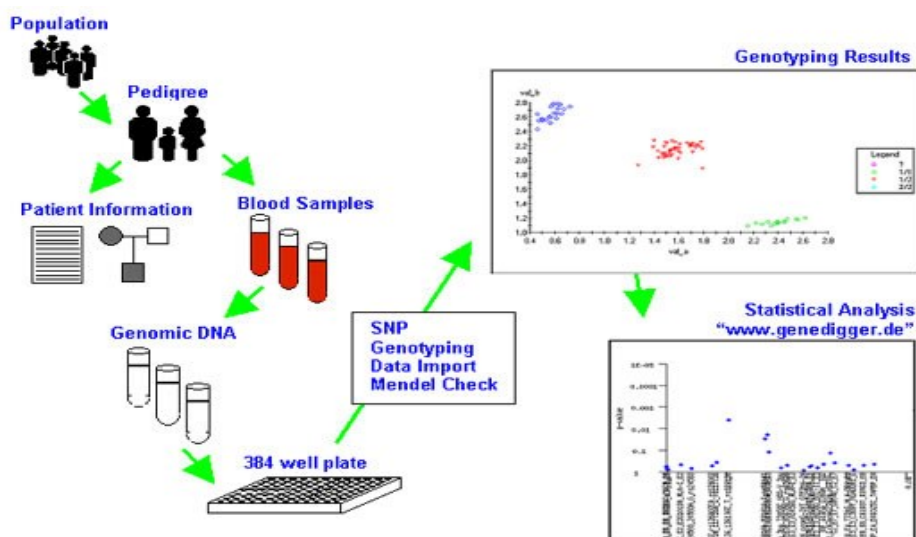
Συγκεκριμένα, οι πληθυσμιακές γενετικές μελέτες συμβάλλουν στην κατανόηση της εξελικτικής ιστορίας και δυναμικής ενός είδους, καθώς και στην πορεία διασποράς του από το αρχικό σημείο εμφάνισης μέχρι την εγκαθίδρυσή του σε άλλα γεωγραφικά πλάτη. Το γεγονός αυτό θεωρείται πολύ σημαντικό σε σχέση με την εφαρμογή μεθόδων βιολογικού ελέγχου. Σχεδόν όλα τα είδη που υπάρχουν στη φύση ζουν οργανωμένα σε πληθυσμούς.

Με τον όρο «πληθυσμός» προσδιορίζεται μια ομάδα ατόμων του ίδιου είδους που συνυπάρχουν στο χώρο και στο χρόνο και μπορούν να αναπαράγονται μεταξύ τους με φυλετική αναπαραγωγή. Στους φυσικούς πληθυσμούς εμφανίζονται από την γενετικοί πολυμορφισμοί, οι οποίοι τους δίνουν την δυνατότητα να προσαρμόζονται στις περιβαλλοντικές αλλαγές. Η ύπαρξη γενετικών πολυμορφισμών είναι απαραίτητη για την επιβίωση του πληθυσμού και κατ' επέκταση και του είδους, ακόμα και αν αυτοί έχουν αρνητικό αντίκτυπο σε κάποια άτομα του πληθυσμού που τους φέρουν. Αυτό δεν σημαίνει ότι όλοι οι φυσικοί πληθυσμοί εμφανίζουν γενετικούς πολυμορφισμούς με την ίδια συχνότητα. Η γενετική ποικιλότητα διαφέρει από πληθυσμό σε πληθυσμό (M.W. Thomson and R.RMcInnes).

Αν παραδείγματος χάριν, ο πληθυσμός έχει μέγεθος N ατόμων και είναι διπλοειδής, τότε κάθε γενετικός τόπος θα αντιπροσωπεύεται $2N$ φορές στον πληθυσμό. Αν υποθέσουμε ότι για ένα γενετικό τόπο υπάρχουν στον πληθυσμό μόνο δύο αλληλόμορφα A και B , ένα άτομο μπορεί να έχει γονότυπο AA , AB ή BB . Η γονοτυπική συχνότητα δείχνει πόσες φορές υπάρχει ένας συγκεκριμένος γονότυπος

προς το σύνολο των γονοτύπων. Αντίστοιχα, η γονιδιακή συχνότητα δείχνει πόσες φορές υπάρχει ένα αλληλόμορφο προς το σύνολο των αλληλομόρφων.

Αυτές οι δυο παράμετροι αποτελούν σημαντικά στοιχεία της Πληθυσμιακής ανάλυσης. Συχνά υφίστανται τεράστιες διαφορές στις συχνότητες εμφάνισης κάποιων αλληλομόρφων ανάμεσα σε πληθυσμιακές ομάδες. Οι διαφορές αυτές εντοπίζονται τόσο σε αλληλόμορφα που προκαλούν γενετικές ασθένειες όσο και σε γενετικά ουδέτερα αλληλόμορφα (τα αποκαλούμε γενετικούς δείκτες), όπως είναι ορισμένες ομάδες αίματος, πολυμορφισμοί πρωτεϊνών ή DNA κλπ. Αν και τα πρώτα έχουν ιδιαίτερη σημασία για τον προσδιορισμό του κινδύνου επανεμφάνισης γενετικών ασθενειών σε συγκεκριμένες πληθυσμιακές ομάδες, οι δεύτεροι είναι σημαντικοί ως δείκτες της πρόσφατης εξέλιξης.



Εικόνα 1: Αναπαράσταση της συλλογής ενός πληθυσμού, καταγραφή των πληροφοριών για τους ασθενείς και με βάση του DNA γίνεται στατιστική ανάλυση των γονοτυπικών αποτελεσμάτων.

Η γνώση της ύπαρξης πολυμορφικών γονιδίων που σχετίζονται με ασθένειες και τα οποία εμφανίζονται είναι συνήθη σε διαφορετικούς πληθυσμούς, καθώς και της ενδεχόμενης επίδρασης τους στην αναπαραγωγική αρμοστικότητα, μπορεί να αποδειχθεί πολύτιμη στην κλινική διάγνωση και στη γενετική καθοδήγηση (εικόνα 1).

Το γνωστικό επίπεδο της Πληθυσμιακής Γενετικής συγγενεύει με αυτό της Γενετικής Επιδημιολογίας, που ασχολείται κατά κύριο λόγο με ασθένειες που έχουν πολύπλοκα πρότυπα κληρονομικότητας ή εξαρτώνται από το συνδυασμό

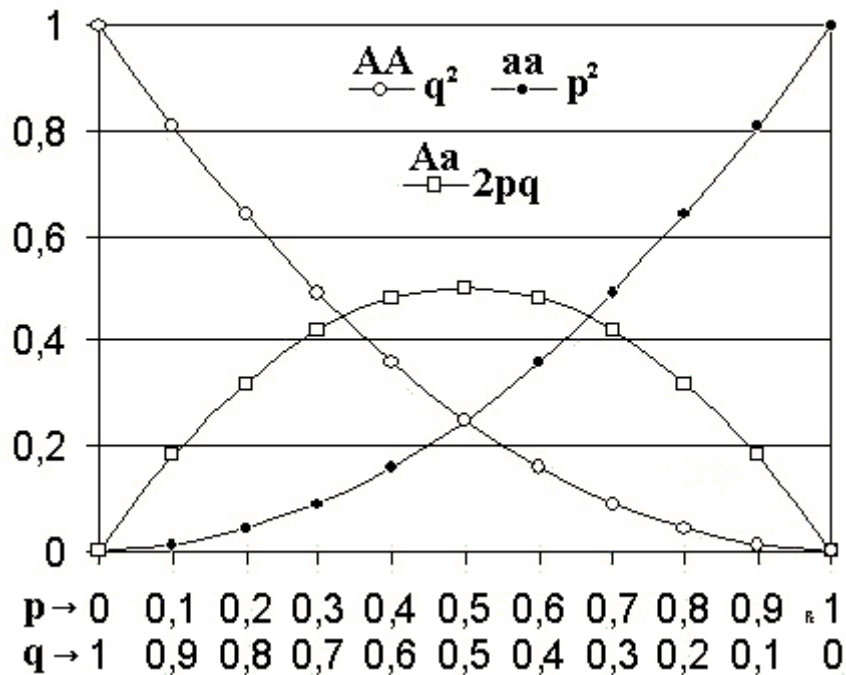
κληρονομήσιμων και περιβαλλοντικών παραγόντων (Khoury 1997). Η θεμελιώδης αρχή της Πληθυσμιακής Γενετικής, είναι η ισορροπία Hardy-Weinberg, η οποία μελετά τις προϋποθέσεις και τους παράγοντες που μπορεί να προκαλούν πραγματική ή φαινομενική απόκλιση από την ισορροπία σε πραγματικούς, και όχι σε ιδεατούς πληθυσμούς. Επίσης προσδιορίζονται οι συχνότητες και οι ρυθμοί μετάλλαξης συγκεκριμένων γονιδίων κλινικού ενδιαφέροντος.

1.2. ΝΟΜΟΣ HARDY-WEINBERG

Ένα από τα σημαντικότερα επιτεύγματα στον τομέα της Πληθυσμιακής Γενετικής είναι η διατύπωση του νόμου των Hardy-Weinberg, η οποία μπορεί να θεωρηθεί ως η λογική επέκταση των νόμων του Mendel πάνω σε πληθυσμούς. Οι βάσεις της τίθενται το 1908 όταν ο Βρετανός μαθηματικός George Hardy και ο Γερμανός γιατρός, Wilhelm Weinberg διατυπώνουν, ανεξάρτητα ο ένας απ' τον άλλον, την αρχή της σταθερότητας των γονοτυπικών συχνοτήτων σε μεγάλους πληθυσμούς (Hardy 1908).

Για ένα ζεύγος αλληλόμορφων (A, B) όπου η πιθανότητα για το αλληλόμορφο A είναι p και για το B είναι q , τότε οι συχνότητες των γονοτύπων AA, AB και BB θα είναι p^2 , $2pq$ και q^2 αντίστοιχα (εικόνα 2). Μια σημαντική συνέπεια που προκύπτει από τις σχέσεις ανάμεσα σε γονότυπο και συχνότητες αλληλομόρφων είναι ότι οι αναλογίες των γονοτύπων δεν θα μεταβάλλονται από γενεά σε γενεά. Η γενεά που θα προέλθει από τυχαίες διασταυρώσεις ατόμων ενός πληθυσμού στον οποίο οι γονότυποι A/A, A/B, B/B, βρίσκονται σε αναλογίες $p^2: 2pq: q^2$ (εικόνα 3), θα έχει γονοτύπους με τις ίδιες σχετικές αναλογίες.

Το γεγονός ότι οι γονότυποι κατανέμονται ανάλογα με τις συχνότητες των ατομικών αλληλομόρφων στον πληθυσμό και παραμένουν σταθερές από γενεά σε γενεά αποτελεί τη θεμελιώδη αρχή του νόμου Hardy-Weinberg. Ο νόμος ερμηνεύει τα αίτια της σταθερότητας των γονιδιακών συχνοτήτων, εφαρμόζοντας το διωνυμικό θεώρημα (Emigh 1980).



Εικόνα 2: Ο νόμος Hardy-Weinberg για δύο αλληλόμορφα: Ο οριζόντιος άξονας δείχνει τις συχνότητες των αλληλόμορφων p και q και ο κάθετος άξονας δείχνει τις γονοτυπικές συχνότητες. Κάθε γραμμή δείχνει ένα από τους τρεις πιθανούς γονότυπους.

Για να ισχύει ο νόμος πρέπει, να ισχύουν κάποιες προϋποθέσεις, όπως το μέγεθος ενός πληθυσμού να είναι αρκετά μεγάλο. Οι διασταυρώσεις μεταξύ των ατόμων να γίνονται τυχαία και ανεξάρτητα. Οι γονείς να περνούν στους απογόνους τους όλα τα αλληλόμορφα με την ίδια πιθανότητα. Να μην προκύπτουν νέα γονίδια από μεταλλάξεις. Να μην υπάρχει μετανάστευση, το γονίδιο να μην υφίσταται επιλεκτική πίεση, δηλαδή να μην επηρεάζει τον αριθμό των απογόνων του ατόμου (Hardy 1908). Συχνά όμως, δεν εμφανίζονται αποκλίσεις από την ισορροπία αυτή, στους φυσικούς πληθυσμούς.

Ένας πληθυσμός που διέπεται από τα βασικά χαρακτηριστικά του νόμου Hardy-Weinberg θεωρείται ότι βρίσκεται σε ισορροπία Hardy-Weinberg. Ο Hardy και ο Weinberg ανέπτυξαν μία απλή εξίσωση που μπορεί να χρησιμοποιηθεί για να ανακαλύψει τις πιθανές γονοτυπικές συχνότητες σε έναν πληθυσμό και η οποία μπορεί να ακολουθήσει τις αλλαγές που εμφανίζονται από μία γενιά σε μια άλλη. Αυτή η εξίσωση είναι γνωστή ως εξίσωση ισορροπίας Hardy-Weinberg.

Ας θεωρήσουμε ένα ζεύγος συνεπικρατών αλληλομόρφων A και B (Zhou, Lange et al. 2009), που δημιουργούν δύο γονότυπους A/A, A/B, B/B που αντιστοιχούν στους φαινοτύπους A, AB, B. Αν στον πληθυσμό που μελετάμε τηρούνται ορισμένες βασικές προϋποθέσεις τότε η συχνότητα των γονοτύπων που αναμένουμε είναι για τους AA: p^2 , τους AB: $2pq$ και τους BB: q^2 .

➤ Εξίσωση ισορροπίας:

$$p^2 + 2pq + q^2 = 1 \quad (1)$$

➤ Συχνότητες γονοτύπων :

$$P = \frac{N_{AA}}{N} \quad H = \frac{N_{AB}}{N} \quad Q = \frac{N_{BB}}{N} \quad (2)$$

➤ Συχνότητες αλληλομόρφων

$$p = \frac{N_{AA} + \frac{1}{2}N_{AB}}{N} \quad q = \frac{N_{BB} + \frac{1}{2}N_{AB}}{N} \quad (3)$$

Επειδή υπάρχουν μόνο δύο αλληλόμορφα γονίδια στην περίπτωση αυτή, το άθροισμα των συχνοτήτων είναι ίσο με 100%, το οποίο προκύπτει από την εξίσωση $p+q=1$. Δεδομένου ότι αυτό ισχύει τότε ισχύει και το ότι $p=1-q$.

Το γεγονός ότι όλες οι πιθανότητες των πιθανών συνδυασμών αλληλόμορφων γονιδίων εμφανίζονται τυχαία ήταν ένα βήμα να γίνει κατανοητό ότι η γνώση για Hardy-Weinberg είναι η εξίσωση $(p+q)^2=1$. Εάν p είναι η συχνότητα του ενός αλληλομόρφου και q η συχνότητα του άλλου, το p^2 είναι η συχνότητα των ομόζυγων κυρίαρχων γονιδίων (AA) σε έναν πληθυσμό, το $2pq$ των ετεροζυγωτικών (AB) και το q^2 η συχνότητα για των ομόζυγων γονιδίων (BB).

Ο νόμος των Hardy Weinberg ισχύει και στην περίπτωση που έχουμε περισσότερα από δύο αλληλόμορφα (A, B)

- Για 3 αλληλόμορφα (A,B και C) με συχνότητες p, q, και r:

$$(p + q + r)^2 = p^2 + 2pq + q^2 + 2pr + 2qr + r^2 \quad (4)$$

- Για 4 αλληλόμορφα (A, B, C, και D) με συχνότητες p, q, r, και s αντίστοιχα είναι:

$$(p + q + r + s)^2 = p^2 + 2pq + q^2 + 2pr + 2qr + r^2 + 2ps + 2qs + 2rs + s^2 \quad (5)$$

- Για τον υπολογισμό των συχνοτήτων των αλληλομόρφων σε γονιδιακούς τόπους με περισσότερα από δύο αλληλόμορφα δίνεται ο μαθηματικός τύπος:

$$p_i = \frac{N_{ii} + \frac{1}{2} \sum_{j=1}^n N_{ij}}{N} \quad (6)$$

Ένας πληθυσμός επομένως, βρίσκεται σε ισορροπία κατά Hardy-Weinberg όταν οι γονοτυπικές και οι γονιδιακές του συχνότητες παραμένουν σταθερές από γενιά σε γενιά. Η γονοτυπική αυτή ισορροπία που βασίζεται πάνω σε σταθερές γονιδιακές συχνότητες και τυχαίες διασταυρώσεις είναι γνωστή ως νόμος Hardy-Weinberg. Πριν από την έρευνα του Hardy και του Weinberg, θεωρήθηκε ότι τα κυρίαρχα αλληλόμορφα γονίδια πρέπει, με την πάροδο του χρόνου, να είναι αναπόφευκτα περισσότερα σε πλήθος από τα υπολειπόμενα αλληλόμορφα γονίδια. Σύμφωνα με αυτήν την λανθασμένη ιδέα, τα κυρίαρχα αλληλόμορφα γονίδια αυξάνονται πάντα στη συχνότητα από γενεά σε γενεά.

Ο Hardy και ο Weinberg ήταν σε θέση να καταδείξουν με την εξίσωσή τους ότι τα κυρίαρχα αλληλόμορφα γονίδια μπορούν εύκολα να μειωθούν στην συχνότητα. Ωστόσο οποιαδήποτε στατιστικά σημαντική απόκλιση από τη συγκεκριμένη αρχή μπορεί να επηρεάσει σημαντικά τη στατιστική ανάλυση και να οδηγήσει σε ψευδή συμπεράσματα. Το γεγονός αυτό αποτελεί πρόβλημα καθώς παραβιάζεται η ισορροπία H-W. Στις περισσότερες περιπτώσεις όπου η ισορροπία H-W παραβιάζεται, αυτό δεν αναφέρεται. Μπορεί πολύ εύκολα κάποιος να κάνει ένα απλό test (Εικόνα 4), μέσω διαδικτύου για να ελέγξει τη συμβατότητα με το HWE σε

κλάσματα του δευτερολέπτου και να διαπιστώσει εάν οι κατανομές των γονοτύπων διαφέρουν σημαντικά από τις προβλεπόμενες της ισορροπίας.

Chang Bioscience

\$99 Mini Centrifuge

\$69 Pipette

Online | Download | Store | Contact | About

Hardy-Weinberg Equilibrium
Check Hardy-Weinberg equilibrium. Chi-square, HWE Correlation R.
www.goldenhelix.com/SNP_Variation/

Ads by Google

Hardy-Weinberg Equilibrium Calculator [Click here for more biology tools.](#)

Allele Frequencies in Population:

	A (p)	a (q)	Total
Frequency:			1
Number:			

Clear

Genotype Frequencies in Population:

	AA (p ²)	Aa (2pq)	aa (q ²)	Total
Frequency:				1
Number:				

Εικόνα 3: Δυνατότητα υπολογισμού για έλεγχο της ισορροπίας Hardy-Weinberg μέσω διαδικτύου με το HWE calculator. Με απλή πληκτρολόγηση στα κελιά των συχνοτήτων των αλληλομόρφων. Πρέπει το σύνολο να δίνει άθροισμα 1 (<http://www.changbioscience.com/genetics/hardy.html>).

1.2.3. ΠΑΡΑΓΟΝΤΕΣ ΠΟΥ ΕΠΗΡΕΑΖΟΥΝ ΤΗΝ ΙΣΟΡΡΟΠΙΑ H-W

Οι γενετικές συσχετίσεις που προέρχονται από μελέτες ασθενών-μαρτύρων ενδέχεται να είναι εσφαλμένες όταν η γονοτυπική κατανομή των φυσιολογικών ελέγχων αποκλίνει από την ισορροπία Hardy-Weinberg. Τότε έχουμε υπερεπάρκεια ή έλλειψη από ομοζυγώτες. Η απόκλιση από το HWE μπορεί να οφείλεται σε εργαστηριακά και γονοτυπικά σφάλματα, στη διαστρωμάτωση του πληθυσμού, σε λάθη επιλογής ή και σε εργαστηριακά σφάλματα (Weiss, Silverman et al. 2001; Xu, Turner et al. 2002).

Στον νόμο Hardy-Weinberg γίνονται ορισμένες θεμελιώδης παραδοχές που δεν ισχύουν πάντοτε στους πραγματικούς ανθρώπινους πληθυσμούς:

1. Τυχαίο ζευγάρισμα, δηλαδή χαρακτηρίζεται με μικρή ή καθόλου ομοιοτυπική σύζευξη ή ενδογαμία.

2. Σταθερό ρυθμό μετάλλαξης, όσα μεταλλαγμένα αλληλόμορφα χάνονται λόγω θανάτου, αντικαθίστανται από νέες μεταλλάξεις.

3. Δεν υπάρχει επιλογή υπέρ ή κατά ενός συγκεκριμένου φαινοτύπου. Όλοι οι γονότυποι ενός γενετικού τόπου είναι εξίσου βιώσιμοι.

4. Ο πληθυσμός είναι αρκετά μεγάλος, ώστε να μη λαμβάνει χώρα τυχαία μεταβίβαση κάποιου γονοτύπου.

5. Δεν έχει αλλάξει τίποτα στη δομή του πληθυσμού λόγω μετανάστευσης, η οποία μπορεί βαθμιαία να μεταβάλλει τις γονιδιακές συχνότητες, αυξάνοντας ή μειώνοντας τον αριθμό των ατόμων με κάποιον συγκεκριμένο γονότυπο.

Αν σε έναν δεδομένο πληθυσμό μία ή περισσότερες από αυτές τις παραδοχές δεν ισχύουν, οι γονότυποι στον πληθυσμό αυτό είναι δυνατόν να μη βρίσκονται σε ισορροπία Hardy-Weinberg. Όταν οι διασταυρώσεις γίνονται μεταξύ συγγενικών ατόμων έχουμε ομομιξία. Τέτοιες διασταυρώσεις δεν αλλάζουν τις γονιδιακές συχνότητες, αυξάνουν όμως το ποσοστό των ομοζυγωτικών ατόμων.

Με τον όρο μετάλλαξη εννοούμε μια αλλαγή σε επίπεδο γενετικού υλικού ενός ατόμου. Έτσι εισάγονται καινούρια αλληλόμορφα στον πληθυσμό, αλλά από μόνες τους δεν μπορούν να προκαλέσουν αλλαγές των γονιδιακών συχνοτήτων. Απαραίτητη προϋπόθεση είναι η διατήρηση και η διασπορά τους στον πληθυσμό.

Ένας άλλος παράγοντας που μπορεί να προσθέσει καινούρια αλληλόμορφα στους πληθυσμούς ή να αλλάζει τις γονιδιακές συχνότητες είναι η γονιδιακή ροή που είναι το αποτέλεσμα μετανάστευσης ατόμων από τον έναν πληθυσμό στον άλλο. Η συνέπεια της γονιδιακής ροής εξαρτάται από τις διαφορές της γενετικής δομής των πληθυσμών (διαφορές γονιδιακών συχνοτήτων), από το ρυθμό και το μέγεθος της μετανάστευσης και από το μέγεθος του πληθυσμού.

Η τυχαία γενετική εκτροπή είναι η προσέγγιση της ισορροπίας σε συνθήκες τυχαίων διασταυρώσεων, είναι τέλεια όταν το μέγεθος του πληθυσμού είναι άπειρο. Η τυχαία γενετική εκτροπή αφορά ακριβώς τις περιπτώσεις εκείνες που ένας πληθυσμός είναι πολύ μικρός και συνεπώς μια μικρή αλλαγή μπορεί να έχει μεγάλη επίπτωση στις γονιδιακές συχνότητες της επόμενης γενιάς. Η γενετικής εκτροπή αναφέρεται σε περιπτώσεις όπου ένας πληθυσμός δημιουργείται από πολύ λίγα άτομα ενός άλλου, μεγάλου πληθυσμού (π.χ. σε περιπτώσεις μετανάστευσης λίγων ατόμων ενός είδους σε περιοχές που αυτό δεν προϋπάρχει).

Τα φαινόμενα «bottleneck effects», είναι επίσης δείγμα ακραίας γενετικής εκτροπής και αναφέρονται σε περιπτώσεις όπου ένας μεγάλος πληθυσμός αντιμετωπίζει κάποιου είδους περιβαλλοντική πίεση, με αποτέλεσμα μόνο λίγα άτομα από αυτόν να επιβιώνουν ή λίγα μόνο άτομα να συμμετέχουν στην δημιουργία της επόμενης γενιάς, με αποτέλεσμα την δραστική αλλαγή των γονιδιακών συχνοτήτων. Και οι τρεις αυτές περιπτώσεις οδηγούν σε αλλαγές των γονιδιακών συχνοτήτων λόγω του μικρού μεγέθους του πληθυσμού.

Η φυσική επιλογή μπορεί να οριστεί απλά ως η διαφορική αναπαραγωγή γονοτύπων, που καθορίζεται από το γεγονός ότι μερικοί γονότυποι αυξάνουν τις πιθανότητες επιβίωσης και αναπαραγωγής των φορέων τους σε σχέση με άλλους. Η φυσική επιλογή μπορεί να οφείλεται στη διαφορική επιβίωση ή στη γονιμότητα ή και στα δυο. Είναι η κινητήρια δύναμη της εξέλιξης που ωθεί τις όποιες γονιδιακές αλλαγές προς συγκεκριμένη κατεύθυνση. Είναι προφανές ότι για να δράσει η φυσική επιλογή απαραίτητη είναι η ύπαρξη γενετικής ποικιλότητας.

1.3. ΤΡΟΠΟΙ ΕΛΕΓΧΟΥ ΤΩΝ ΑΠΟΚΛΙΣΕΩΝ ΑΠΟ ΤΗΝ HWE

Για να βρούμε αν ο πληθυσμός που μελετάμε είναι σε ισορροπία Hardy-Weinberg, χρησιμοποιούμε κατάλληλες στατιστικές μεθόδους. Οι τρεις διαφορετικοί τρόποι ελέγχου των αποκλίσεων με p -value $>5\%$ είναι οι εξής :

- Ο έλεγχος καλής προσαρμογής χ^2
- Έλεγχος πηλίκου πιθανοφάνειας (likelihood-ratio test)
- Ακριβές έλεγχος του Fisher (Exact test)

Και οι τρεις αυτοί δείκτες υπολογίζουν τις αποκλίσεις, όπως αυτές προκύπτουν από τις διαφορές στις συχνότητες των παρατηρούμενων και των αναμενόμενων γονοτύπων, για κάθε γενετικό τόπο (Emigh 1980).

Genotype	Observed	Expected	
AA	120	111.11	
Aa	160	177.78	
aa	80	71.11	
total	360	360.00	
Allele	Observed	Frequency	Std. Err.
A	400	0.5556	0.0194
a	320	0.4444	0.0194
total	720	1.0000	
Estimated disequilibrium coefficient (D) =			0.0247
Hardy-Weinberg Equilibrium Test:			
Pearson chi2 (1) =	3.600	Pr=	0.0578
likelihood-ratio chi2 (1) =	3.601	Pr=	0.0578
Exact significance prob =			0.0557

Πίνακας 1: Παράδειγμα ελέγχου των αποκλίσεων. Για το test ισορροπίας HW με τους τρεις τρόπους ελέγχου, με τη χρήση του προγράμματος Stata. Βαθμοί ελευθερίας για τον χ^2 του Fisher (1).

1.3.1. ΕΛΕΓΧΟΣ ΚΑΛΗΣ ΠΡΟΣΑΡΜΟΓΗΣ χ^2

Σε κάθε στατιστική έρευνα ελέγχουμε για HWE χρησιμοποιώντας τον έλεγχο χ^2 ή τεστ ακρίβειας στους μάρτυρες (Emigh 1980). Ο έλεγχος καλής προσαρμογής συγκρίνει τις παρατηρούμενες από τις αναμενόμενες συχνότητες των διαφόρων γονοτύπων. Για λόγους συμφωνίας χρησιμοποιούμε στις μελέτες τους όρους «παραβιάσεις» για να περιγράψουμε αποκλίσεις από το HWE που πετυχαίνουν τυπική στατιστική σημασία, και «αποκλίσεις», για να περιγράψουμε οποιαδήποτε παρέκκλιση από τις συχνότητες ακριβείας που αναμένονται υπό το HWE, ανεξάρτητα από το αν η παρέκκλιση είναι στατιστικά σημαντική.

Έστω ότι έχουμε τυχαίες παρατηρήσεις χ_1, \dots, χ_n από έναν πληθυσμό X με κατανομή $f(x)$ και έστω $f_0(x)$ είναι κάποια γνωστή κατανομή. Μας ενδιαφέρει να ελέγξουμε τη μηδενική υπόθεση.

$$H_0: f(x) = f_0(x)$$

Έναντι της εναλλακτικής:

$$H_a: f(x) \neq f_0(x)$$

Για έναν τέτοιο έλεγχο εργαζόμαστε ως εξής:

Διαιρούμε το δειγματικό χώρο σε k κατηγορίες (τμήματα) T_1, T_2, \dots, T_k και υπολογίζουμε τις πιθανότητες p_i σε κάθε κατηγορία με τη βοήθεια της κατανομής που ορίζει η αρχική υπόθεση. Αν η H_0 είναι σωστή ο αριθμός O_i των παρατηρούμενων μετρήσεων (observed) που ανήκουν στην i κατηγορία δε θα πρέπει να διαφέρει πολύ από τον αντίστοιχο αναμενόμενο αριθμό (expected), $E_i = np_i$ που βασίζεται στην $f_0(x)$.

Ο έλεγχος αυτός μπορεί να εφαρμοστεί σε πολλά προβλήματα, καθώς το μόνο που απαιτεί είναι η γνώση του πλήθους που αντιστοιχεί σε κάθε κατηγορία, ενώ είναι ανεξάρτητος από το βαθμό κανονικότητας της κατανομής των δεδομένων.

Το κριτήριο που μπορεί να μετρήσει τη συνολική διαφορά μεταξύ των παρατηρούμενων και αναμενόμενων τιμών οφείλεται στον K. Pearson (1900), ονομάζεται χ^2 κριτήριο, και η στατιστική του συνάρτηση δίνεται από την σχέση:

$$\chi^2 = \sum_{i=1}^k \frac{(O-E)^2}{E} \sim \chi^2_{k-1} \quad \text{όταν ισχύει η } H_0 \quad (7)$$

με κρίσιμη περιοχή

$$\chi^2 \geq \chi^2_{k-1, \alpha}$$

Αν s παράμετροι της γνωστής κατανομής $f_0(x)$ είναι γνωστές τότε τις εκτιμούμε, με αποτέλεσμα οι βαθμοί ελευθερίας της χ^2 να μειωθούν κατά s και τελικά να έχουμε $k-1-s$ βαθμούς ελευθερίας. Επίσης η προσέγγιση της χ^2

κατανομής είναι καλύτερη όταν τα p_i είναι σχεδόν ίσα, η επιλογή των κατηγοριών γίνεται ανεξάρτητα από τη μορφή των δεδομένων και τα $E_i \geq 5$ (ή τουλάχιστον 80% των $E_i \geq 5$ και κανένα $E_i < 1$).

➤ Επομένως για δύο αλληλόμορφα είναι:

$$\chi^2 = \frac{(N_{AA} - p^2 N)^2}{p^2 N} + \frac{(N_{AB} - 2pqN)^2}{2pqN} + \frac{(N_{aa} - q^2 N)^2}{q^2 N} \quad (8)$$

Για κάθε γενετικό τόπο υπολογίζει το άθροισμα αυτών των κλασμάτων και ακολούθως δίνει την πιθανότητα να είναι σε ισορροπία. Το πρόβλημα με αυτό το κριτήριο είναι η πολύ μεγάλη βαρύτητα που δίνει σε περιπτώσεις που ενώ η αναμενόμενη συχνότητα ενός γενοτύπου είναι πολύ μικρή, για κάποιο λόγο εμφανίζονται ένα ή περισσότερα άτομα αυτού του γενοτύπου. Επειδή οι αναμενόμενοι γονότυποι είναι στον παρανομαστή του κλάσματος, το χ^2 γίνεται πολύ μεγάλο. Είναι δυνατόν δηλαδή ένας πληθυσμός που βρίσκεται σε ισορροπία, να εμφανίζεται εκτός ισορροπίας λόγω της ύπαρξης ενός μόνο ατόμου.

Συνεπώς αυτό που υπολογίζει ο έλεγχος καλής προσαρμογής είναι η πιθανότητα, βάσει των γονοτύπων, που έχει ένας πληθυσμός να είναι σε ισορροπία. Το σύνηθες επίπεδο σημαντικότητας που χρησιμοποιείται είναι 5%, που σημαίνει ότι θεωρούμε ότι ένας πληθυσμός είναι σε ισορροπία για ένα γενετικό τόπο αν το κριτήριο χ^2 δίνει πιθανότητα πάνω από 5%.

1.3.2. ΕΛΕΓΧΟΣ ΠΗΛΙΚΟΥ ΠΙΘΑΝΟΦΑΝΕΙΑΣ

Ο έλεγχος πηλικού πιθανοφάνειας (likelihood-ratio test) είναι ένας στατιστικός έλεγχος καλής προσαρμογής μεταξύ δύο μοντέλων. Καθορίζει ο κριτήριο για την επιλογή ανάμεσα σε δύο πιθανά μοντέλα. Ένα περισσότερο σύνθετο μοντέλο συγκρίνεται με ένα λιγότερο σύνθετο για να διαπιστωθεί αν αυτό ταιριάζει περισσότερο σε ένα συγκεκριμένο σύνολο δεδομένων. Αν ισχύει τότε οι παράμετροι

του μοντέλου χρησιμοποιούνται στις επόμενες αναλύσεις. Ο έλεγχος πηλίκου πιθανοφάνειας είναι έγκυρος όταν χρησιμοποιηθεί για να συγκρίνει 2 σύνθετα μοντέλα. Το σύνθετο μοντέλο διαφέρει από το πιο απλό στην προσθήκη περισσότερων παραμέτρων. Η προσθήκη περισσότερων παραμέτρων οδηγεί σε υψηλότερη βαθμολογία της πιθανοφάνειας.

Ο έλεγχος της υπόθεσης για το αν ένας πληθυσμός είναι σε ισορροπία HW γίνεται με τον έλεγχο του πηλίκου πιθανοφάνειας. Σύμφωνα με την μηδενική υπόθεση της HWE, ο έλεγχος της πιθανοφάνειας κατανέμεται ασυμπτωτικά ως χ^2 τυχαία μεταβλητή με έναν βαθμό ελευθερίας. Εδώ N είναι ο συνολικός αριθμός των ατόμων και N_χ ο αριθμός των αλληλόμορφων ή των γονότυπων ανάλογα με τον δείκτη χ που αντιπροσωπεύει.

$$LRTS = -2 \ln \left[\frac{(N)^N (N_A)^{N_A} (N_B)^{N_B} 2^{N_{AB}}}{(2N)^{2N} (N_{AA})^{N_{AA}} (N_{AB})^{N_{AB}} (N_{BB})^{N_{BB}}} \right]$$

Η πιθανοφάνεια των δεδομένων όταν δεν είναι σε ισορροπία είναι:

$$L(P_{AA}, P_{AB}) = \frac{N!}{N_{AA}! N_{AB}! N_{BB}!} P_{AA}^{N_{AA}} P_{AB}^{N_{AB}} P_{BB}^{N_{BB}}$$

Αν είναι σε ισορροπία τότε υπάρχει σχέση μεταξύ των γονοτυπικών συχνοτήτων και των αλληλόμορφων και η πιθανοφάνεια είναι:

$$L(P_A) = \frac{N!}{N_{AA}! N_{AB}! N_{BB}!} (P_A^2)^{N_{AA}} (2P_A(1-P_A))^{N_{AB}} ((1-P_A)^2)^{N_{BB}}$$

Ο έλεγχος του πηλίκου πιθανοφάνειας είναι ο λογάριθμος του πηλίκου των δύο πιθανοφανειών:

$$LR = 2 \cdot \ln \left(\frac{L(P_{AA}, P_{AB})}{L(P_A)} \right) \quad (9)$$

Μας ενδιαφέρει να ελέγξουμε τη μηδενική υπόθεση, όπου ο πληθυσμός είναι σε ισορροπία HW.

$$H_0: P(AA) = P(A), P(A) \\ P(AB) = 2P(A)P(B) \\ P(BB) = (1 - P(A))^2$$

Έναντι της εναλλακτικής, όταν ο πληθυσμός δεν είναι σε ισορροπία:

$$H_\alpha: P(AA) \neq P(A), P(A) \\ P(AB) \neq 2P(A)P(B) \\ P(BB) \neq (1 - P(A))^2$$

Ο στατιστικός έλεγχος του πληθυσμίου πιθανοφάνειας ακολουθεί έναν έλεγχο καλής προσαρμογής χ^2 . Για να προσδιοριστεί αν η διαφορά μεταξύ των δύο μοντέλων είναι σημαντική, εξετάζονται οι βαθμοί ελευθερίας.

1.3.3. ΑΚΡΙΒΗΣ ΕΛΕΓΧΟΣ (EXACT TEST FISHER)

Ο ακριβής έλεγχος του Fisher, είναι ένας μη παραμετρικός έλεγχος, για ανεξαρτησία και ομοιογένεια. Χρησιμοποιείται για ανάλυση διακριτών δεδομένων όταν τα δύο ανεξάρτητα δείγματα έχουν μικρό μέγεθος (Wigginton, Cutler et al. 2005). Έστω δύο ανεξάρτητα τυχαία δείγματα με :

$$N = \sum_i R_i = \sum_j R_j$$

	Χαρακτηριστικό A	Χαρακτηριστικό B	Σύνολο
Ομάδα A	a	b	$R_1 = a+b$
Ομάδα B	c	d	$R_2 = c+d$
Σύνολο	$C_1 = a+c$	$C_2 = b+d$	$N = R_1 + R_2$ ή $C_1 + C_2$

Η ακριβής πιθανότητα του Fisher για τον συγκεκριμένο πίνακα 2x2 δίνεται από την παρακάτω σχέση:

$$P = \frac{(R_1!R_2!)(C_1!C_2!)}{N! \prod_{i,j} a_{ij}}$$

Αρχικά υπολογίζονται οι συχνότητες των αλληλομόρφων και οι αναμενόμενες τιμές τους. Στην συνέχεια γίνεται ο στατιστικός έλεγχος χ^2 . Ο έλεγχος αυτός είναι κατάλληλος, αλλά μπορεί να είναι ανακριβής στην περίπτωση που ένα αλληλόμορφο είναι σπάνιο. Γι' αυτό μία καλύτερη προσέγγιση είναι ο ακριβής έλεγχος των γονοτυπικών πιθανοτήτων καθορίζεται από την σχέση 10. Αυτή η προσέγγιση είναι ανάλογη με τον ακριβή έλεγχο του Fisher, όπου υπολογίζεται η πιθανότητα της παρατήρησης ενός δείγματος. Για να γίνει έλεγχος της ισορροπίας HW, εξετάζεται ένα δείγμα που αποτελείται από 2N αλληλόμορφα και περιλαμβάνει n_A αλληλόμορφα τύπου A, n_B αλληλόμορφα τύπου B, n_{AA} αριθμό γονοτύπων τύπου A, n_{AB} γονότυπους τύπου AB, n_{BB} γονότυπους τύπου BB.

$$P_{HWE} = \sum_{\dot{n}_{AB}} I[P(N_{AB} = n_{AB} | N, n_A) \geq P(N_{AB} = \dot{n}_{AB} | N, n_A)] \cdot P(N_{AB} = \dot{n}_{AB} | N, n_A) \quad (10)$$

Κατά συνέπεια, στις περιπτώσεις της ισορροπίας HW, η πιθανότητα να παρατηρηθούν ακριβώς n_{AB} ετεροζυγώτες σε ένα δείγμα ατόμων N με αλληλόμορφα γονίδια n_A υπολογίζεται από την παρακάτω σχέση και ισχύει για κάθε πιθανό αριθμό ετεροζυγωτών :

$$P(N_{AB} = n_{AB} | N, n_A) = \frac{2^{n_{AB}} N!}{n_{AA}! n_{AB}! n_{BB}!} \times \frac{n_A! n_B!}{(2N)!}$$

Σημειώνεται ότι οι αριθμοί των ομόζυγων AA και BB γονότυπων είναι :

$$n_{AA} = (n_A - n_{AB})/2 \quad n_{BB} = (n_B - n_{AB})/2$$

Το p-value υπολογίζεται με το άθροισμα των πιθανοτήτων όλων των ακραίων περιπτώσεων και του δοσμένου πίνακα. Χρησιμοποιείται ο έλεγχος χ^2 όταν το μέγεθος του δείγματος είναι <50 για να δώσει καλύτερη προσέγγιση για το p-value. Σε έλεγχο ανεξαρτησίας για έναν πίνακα ενδεχομένων 2×2 μπορούμε να χρησιμοποιήσουμε είτε τον έλεγχο του Fisher είτε τον έλεγχο χ^2 (Chi-square).

Ο έλεγχος του Fisher είναι ιδανικός στις περιπτώσεις που στον πίνακα ενδεχομένων έχουμε πολύ μικρά μεγέθη (δηλαδή αριθμό μικρότερο του 5). Με τη χρήση του Fisher λαμβάνουμε πιο ακριβές p-value. Ο έλεγχος χ^2 έχει απλούστερους υπολογισμούς, αλλά καταλήγει σε ένα μόνο προσεγγιστικό p-value. Ο έλεγχος του Fisher δεν χρησιμοποιείται όταν το μέγεθος του N είναι μεγαλύτερο του 200. Όταν οι αριθμοί είναι μεγάλοι τα p-value που προκύπτουν τόσο από τον ακριβή έλεγχο του Fisher όσο από τον χ^2 έλεγχο είναι σχεδόν παρόμοια.

1.4. ΜΕΤΡΑ ΥΠΟΛΟΓΙΣΜΟΥ ΤΩΝ ΑΠΟΚΛΙΣΕΩΝ ΑΠΟ ΤΗΝ ΗWE

Εκτός από του ελέγχους σημαντικότητας των αποκλίσεων από την ισορροπία HW, υπάρχουν κάποια μέτρα που ποσοτικοποιούν τις αποκλίσεις (Shoemaker, Painter et al. 1998). Ένα από αυτά τα μέτρα είναι ο συντελεστής a , ο οποίος χρησιμοποιείται για να εκτιμήσουμε το μέγεθος των αποκλίσεων από τον νόμο H-W (C. Pereira and Rogatko 1984; Lindley 1988). Ο συντελεστής a ορίζεται ως εξής:

$$a_i = \frac{1}{2} \log(P_{AA} P_{BB} / P_{AB}^2), -\infty < a < \infty \quad (11)$$

$$a_i = \frac{1}{2} \log\left(\frac{4AA_{xi}BB_{xi}}{AB_{xi}^2}\right)$$

Όπου P_{AA} , P_{AB} και P_{BB} αποτελούν τις τρεις γονοτυπικές πιθανότητες. Αν και ο συντελεστής α_i μπορεί να ποικίλει θεωρητικά ανάμεσα στο θετικό και το αρνητικό άπειρο, οι απόλυτες τιμές πάνω από 1 είναι σπάνιες (Minelli, Thompson et al. 2008). Υπολογίζεται και η διακύμανση του α , στην οποία όταν ένας από τους παράγοντες της είναι μηδέν γίνεται πρόσθεση με $\frac{1}{2}$ σε κάθε αριθμό.

$$s_{\alpha_i}^2 = \frac{1}{4AA_{0i}} + \frac{1}{AB_{0i}} + \frac{1}{4BB_{0i}} \quad (12)$$

Σε αντίθεση με τον συντελεστή α , υπολογίζουμε και δυο άλλα μέτρα, που μετρούν την απόκλιση από την ισορροπία HW, τον συντελεστή ομομιξίας F (Weir 1970) και τον συντελεστή ανισορροπίας D (Hernandez and Weir 1989). Ο συντελεστής ομομιξίας ορίζεται ως εξής:

$$F = \frac{P_{BB}}{p_B} + \frac{P_{AA}}{p_A} - 1, \max[-p_B / p_A, -p_A / p_B] \leq f \leq 1 \quad (13)$$

$$F_{ji} = 1 - \frac{2AB_{ji}(AA_{ji} + AB_{ji} + BB_{ji})}{(2BB_{ji} + AB_{ji})(2AA_{ji} + AB_{ji})} = \frac{4BB_{ji}AA_{ji} - (AB_{ji})^2}{(2BB_{ji} + AB_{ji})(2AA_{ji} + AB_{ji})}$$

Όπου p_A και p_B είναι οι συχνότητες των αλληλομόρφων. Ο συντελεστής ανισορροπίας ορίζεται ως εξής:

$$D = P_{BB} - p_B^2, \max[-p_B^2, -p_A^2] \leq D \leq p_B, p_A \quad (14)$$

$$D_i = \frac{4AA_{0i}BB_{0i} - (AB_{0i})^2}{4(AA_{0i} + AB_{0i} + BB_{0i})^2}$$

1.5. Εγκυρότητα των μελετών γενετικής συσχέτισης

Η εγκυρότητα των μελετών γενετικής συσχέτισης εξαρτάται σημαντικά από την χρήση των κατάλληλων ελέγχων, τόσο για ομάδες ασθενών που δεν έχουν μολυνθεί από ομάδες άλλων πληθυσμών όσο και για ομάδες που έχουν μία συγκεκριμένη ασθένεια να ακολουθούν την ισορροπία HW. Το ίδιο ισχύει και για τις ομάδες ασθενών και μαρτύρων από μελέτες που όλα τα δείγματα που έχουν μια συγκεκριμένη ασθένεια, για παράδειγμα μελέτες που αξιολογούν διαφορετική θεραπεία ή αλλά αποτελέσματα κάθε φορά που ο κίνδυνος ασθένειας δεν επηρεάζεται από τον πολυμορφισμό που έχει αξιολογηθεί. Το HWE δεν είναι ένας θεωρητικός νόμος, οι αποκλίσεις μπορούν να δείχνουν σημαντικά προβλήματα, λάθη ή ανωμαλίες στα δείγματα δεδομένων που αναλύθηκαν (Trikalinos, Salanti et al. 2006).

Τα αποτελέσματα από την μελέτη γενετικής συσχέτισης μπορούν να γίνουν συμβατά αν η ισορροπία HW παραβιαστεί. Τα στοιχεία που έχουν συγκεντρωθεί δείχνουν ότι η αναφορά για HWE μπορεί να είναι άριστη σε επιστημονικά περιοδικά που δεν έχουν να κάνουν με την γενετική αλλά δεν υπάρχουν εμπειρικά δεδομένα σχετικά με την αναφορά για το HWE σε επιστημονικά περιοδικά γενετικής.

Επιπλέον θα ήταν ενδιαφέρον να δώσουμε μερικές αποδείξεις για την δύναμη μελετών πάνω στον γενετικό συσχέτισμό που γίνανε πρόσφατα για να εντοπίσουμε απόκλιση από το HWE. Ο στόχος αυτής της μελέτης είναι να εξετάσουμε το βαθμό στον οποίο το HWE εκτιμάται και με ποια μέσα είτε αναλύσεις σε σχέση με το HWE είναι ακριβής είτε οι μελέτες έχουν αρκετή δύναμη να εκτιμήσουν το HWE και πως οι αποκλίσεις αντιμετωπίζονται σε σύγχρονες μελέτες γενετικής συσχέτισης.

ΚΕΦΑΛΑΙΟ 2: ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ

2. ΜΕΤΑ-ΑΝΑΛΥΣΗ

2.1. Μέθοδος της μετά-ανάλυσης

Η μετά-ανάλυση είναι μια ποσοτική ανασκόπηση και σύνθεση των αποτελεσμάτων που προκύπτουν από τις αναμενόμενες και τις παρατηρούμενες μελέτες, ώστε να εξαχθεί ένα συνολικό αποτέλεσμα που αφορά την αποτελεσματικότητα μιας θεραπείας ή την επίδραση ενός πολυμορφισμού στην εμφάνιση μιας ασθένειας (Petiti 1994; Normand 1999).

Η μετά-ανάλυση συμβάλλει στην αποτελεσματική αντιμετώπιση των προβλημάτων που εμφανίζονται λόγω έλλειψης επαναληψιμότητας μεταξύ των ατομικών μελετών, καθώς αυξάνεται η στατιστική ισχύς σχετικά με τις μεμονωμένες μελέτες. Ακόμη, βελτιστοποιεί την ικανότητα ανίχνευσης της συσχέτισης ή μη μίας γενετικής παραλλαγής με μία ασθένεια καθώς μελετάται η συσχέτιση σε πολύ μεγαλύτερο αριθμό ατόμων από ότι οι μεμονωμένες μελέτες ξεχωριστά.

Η διαδικασία της μετά-ανάλυσης περιλαμβάνει τη συλλογή των δεδομένων από βιβλιογραφία (μελέτες γενετικής συσχέτισης, μελέτες ασθενών-μαρτύρων, επιλογή δεδομένων / μελετών που θα χρησιμοποιηθούν στη μετά-ανάλυση. Στη συνέχεια ακολουθείται η στατιστική ανάλυση των δεδομένων με τη χρήση υπολογιστικών πακέτων STATA, RevMan κ.α. για την εξαγωγή συμπερασμάτων.

2.2. Στατιστική ανάλυση

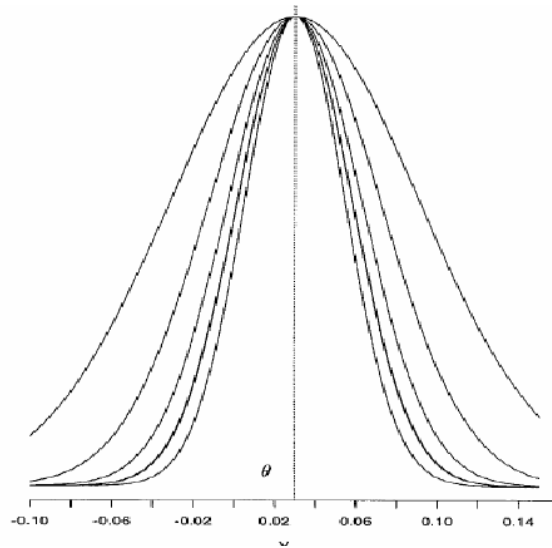
Όταν τα αποτελέσματα των επιμέρους μελετών διαφέρουν κατά πολύ μεταξύ τους, τότε ίσως δεν είναι αποτελεσματικό να συνδυαστούν για μετά-ανάλυση. Ένας τρόπος για να ελεγχθεί αυτό είναι να εξεταστεί στατιστικά ο βαθμός ομοιότητας των αποτελεσμάτων, δηλαδή να γίνει έλεγχος για την ετερογένεια μεταξύ των μελετών.

Αν ο έλεγχος δείξει ότι υπάρχει ομοιογένεια στα αποτελέσματα, τότε ισχύει η

παραδοχή ότι όλες οι μελέτες προέρχονται από έναν κοινό πληθυσμό με όμοια χαρακτηριστικά και αν οι μελέτες γίνονταν απεριόριστα μεγάλες, θα έδιναν όμοια αποτελέσματα. Συνεπώς, σε αυτή την περίπτωση, οι τυχόν διαφορές που παρατηρούνται οφείλονται στην τύχη και στη μετά-ανάλυση χρησιμοποιείται το μοντέλο σταθερών επιδράσεων (fixed – effect model). Αν όμως δειχθεί ότι υπάρχει ετερογένεια, τότε γίνεται η παραδοχή ότι οι μελέτες προέρχονται από διαφορετικούς πληθυσμούς, οι οποίοι εμφανίζουν διαφορετικές μέσες τιμές και διασπορές στην κατανομή των υπό μελέτη παραμέτρων. Σε αυτή την περίπτωση, υπάρχουν πραγματικές διαφορές ανάμεσα στα αποτελέσματα και χρησιμοποιείται το μοντέλο τυχαίων επιδράσεων (random – effect model). Σε όλες τις περιπτώσεις, η ετερογένεια δεν πρέπει να αγνοείται αντίθετα πρέπει να λαμβάνεται υπ' όψιν και να εξηγείται. Παρακάτω ακολουθεί μια σύντομη μαθηματική περιγραφή των δυο μοντέλων.

2.2.1 Μοντέλο σταθερών επιδράσεων (fixed effect model)

Στο μοντέλο σταθερών επιδράσεων, γίνεται η παραδοχή ότι όλες οι μελέτες προέρχονται από έναν κοινό πληθυσμό με όμοια χαρακτηριστικά (ομοιογένεια αποτελεσμάτων) και πως αν οι μελέτες γίνονταν απεριόριστα μεγάλες, θα έδιναν όμοια αποτελέσματα. Σε αυτή την περίπτωση, οι τυχόν διαφορές που παρατηρούνται οφείλονται στην τύχη. Έτσι, όλες οι ανεξάρτητες μελέτες ($Y_i, i=1,2,3\dots n$) έχουν ένα μέσο κοινό εκτιμητή του αποτελέσματος θ (π.χ. Odds Ratio) το οποίο είναι ανεξάρτητο του αριθμού των μελετών και της διασποράς αυτών ($s_i^2, i=1, 2, 3\dots n$).

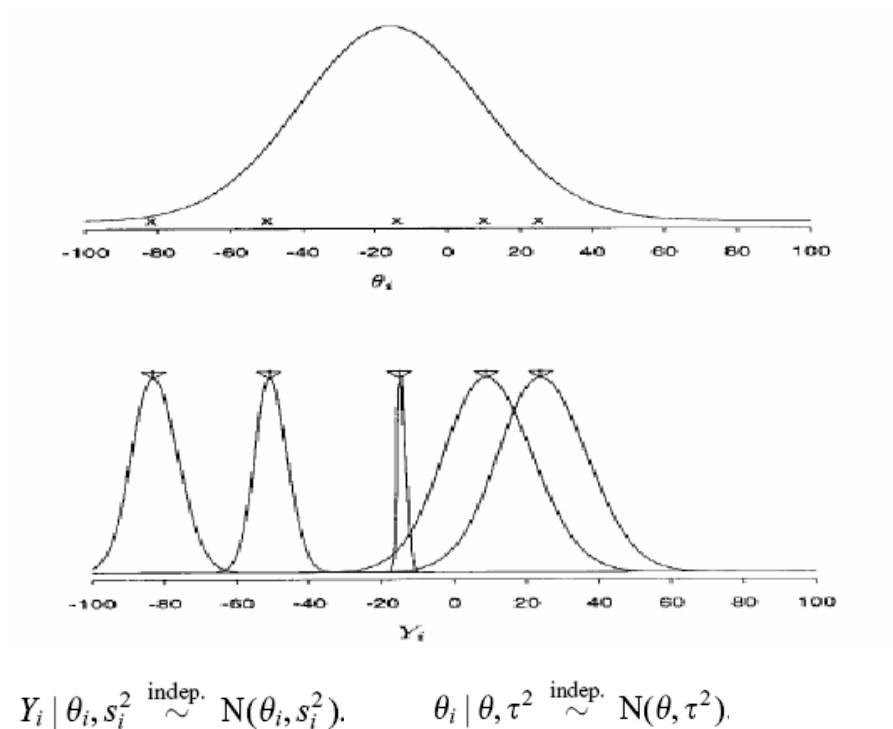


$$Y_i \stackrel{\text{indep.}}{\sim} N(\theta, s_i^2) \text{ for } i = 1, 2, \dots, k$$

Εικόνα 4: Το μοντέλο σταθερών επιδράσεων. Η κατανομή 5 υποθετικών επιδράσεων χρησιμοποιώντας το μοντέλο. Κάθε δείγμα Y_i έχει ένα κοινό μέγεθος επίδρασης θ . Η διαφορά ανάμεσα στις 5 μελέτες είναι η διαφορετική διακύμανση κάθε μελέτης s_i^2 , δηλαδή πόσο καλά υπολογίζει κάθε μελέτη το θ .

2.2.2. Μοντέλο τυχαίων επιδράσεων (random effect model)

Στο μοντέλο τυχαίων επιδράσεων, γίνεται η παραδοχή ότι οι μελέτες προέρχονται από διαφορετικούς πληθυσμούς, οι οποίοι εμφανίζουν διαφορετικές μέσες τιμές και διασπορές στην κατανομή των υπό μελέτη παραμέτρων (ετερογένεια αποτελεσμάτων). Σε αυτή την περίπτωση, υπάρχουν πραγματικές διαφορές ανάμεσα στα αποτελέσματα. Έτσι, η καθεμιά ανεξάρτητη μελέτη (Y_i , $i=1,2,3\dots n$) έχει ένα ξεχωριστό εκτιμητή του αποτελέσματος θ_i , $i=1,2,3\dots n$ (π.χ. Odds ratio) μια τιμή διασποράς (s_i^2 , $i=1, 2, 3\dots n$). Το κάθε θ_i θεωρείται ότι είναι δείγμα που προέρχεται από ένα σύνολο τιμών εκτιμητών με μέση τιμή θ και διασπορά τ^2 . Συνεπώς, σε αυτό το μοντέλο εισάγεται μια νέα μεταβλητή τ^2 που αντιπροσωπεύει την παρατηρούμενη ετερογένεια ανάμεσα στις ξεχωριστές μελέτες.



Εικόνα 5: Το μοντέλο τυχαίων επιδράσεων. Κατανομή 5 υποθετικών δειγμάτων κάνοντας χρήση μοντέλου τυχαίων επιδράσεων. Κάθε μέγεθος επίδρασης θ_i προέρχεται από τον υπερπληθυσμό με μέγεθος επίδρασης θ και διακύμανση τ^2 . Στο παράδειγμα κάθε ένα από τα μεγέθη επίδρασης δημιούργησαν τα 5 αποτελέσματα των μοντέλων.

2.2.3. Σχετικός λόγος πιθανοτήτων ODDS-RATIO (OR)

Τα τελευταία χρόνια το Odds Ratio χρησιμοποιούνται ευρέως στις ιατρικές μελέτες. Αυτό οφείλεται σε τρεις λόγους: i) παρέχουν μια τιμή μαζί με διάστημα εμπιστοσύνης για τη σχέση μεταξύ δύο δυαδικών (ΝΑΙ ή ΟΧΙ) μεταβλητών, ii) μας επιτρέπουν να εξετάσουμε τα αποτελέσματα απλών μεταβλητών σε αυτήν τη σχέση, χρησιμοποιώντας, λογιστική παλινδρόμηση (logistic regression), iii) ερμηνεύουν με ξεχωριστό τρόπο τις μελέτες ασθενών- μαρτύρων. Ο λόγος πιθανοτήτων (OR) είναι ένας τρόπος αναπαράστασης της πιθανότητας. Ο σχετικός λόγος πιθανοτήτων, Odds Ratio είναι η αναλογία της πιθανότητας να συμβεί ένα γεγονός που μας ενδιαφέρει,

προς την πιθανότητα να μη συμβεί και αποτελεί ένα χρήσιμο δείκτη της δύναμης της σχέσης (Bland and Altman 2000). Μπορούμε να υπολογίσουμε ένα τυπικό σφάλμα της απόδοσης αναλογίας λογαρίθμου και συνεπώς ένα διάστημα εμπιστοσύνης. Το τυπικό σφάλμα του λογαρίθμου του odds ratio $se(\log OR)$ υπολογίζεται απλά από την τετραγωνική ρίζα του αθροίσματος των αντιστρόφων των τεσσάρων συχνοτήτων. Το OR μας βοήθησε να υπολογίσουμε τη πιθανότητα των γονοτύπων και των αλληλόμορφων γονιδίων. Ο παρακάτω πίνακας (πίνακας 3) βοηθάει στην κατανόηση, του ποιες μεταβλητές συνδυάζει το OR και με ποιο τρόπο.

	Παράγοντας κινδύνου	
Ασθενείς (cases)	α	β
Μάρτυρες (controls)	γ	δ

Πίνακας 2: Αριθμός ασθενών και υγιών ατόμων που φέρουν ή όχι το αλληλόμορφο (παράγοντα κινδύνου) που προκαλεί την ασθένεια.

Ο υπολογισμός του OR δίνεται από τον τύπο και εκφράζει την συσχέτιση του αλληλόμορφου με την ασθένεια:

$$Odds - Ratio = \frac{\alpha \cdot \delta}{\beta \cdot \gamma} \quad (14)$$

$$se_{\log OR} = \sqrt{\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} + \frac{1}{\delta}} \quad (15)$$

Η ύπαρξη 2 αλληλομόρφων με 3 γονοτύπους οδηγεί στην ύπαρξη ενός πίνακα 3x2 (Πίνακας 4) (Sasieni 1997).

	γονότυποι		
	AA	AB	BB
Ασθενείς (cases)	α	β	ε
Μάρτυρες (controls)	γ	δ	ζ

Πίνακας 3: Αριθμός ετερόζυγων και ομόζυγων γονοτύπων σε ασθενείς και μάρτυρες.

Στη συγκεκριμένη περίπτωση είναι απαραίτητο να γίνει σύμπτυξη του πίνακα 3x2 σε έναν 2x2, για να μην αντιμετωπίσουμε δυσκολίες κατά τον υπολογισμό του Odds-Ratio. Η σύμπτυξη επιτυγχάνεται ενώνοντας τους γονοτύπους AA+AB ή AB+BB, χρησιμοποιώντας ένα μοντέλο υπολειπόμενης κληρονομικότητας. Συγκεκριμένα, βρήκαμε ποιο αλληλόμορφο ευθύνεται για την ασθένεια και έπειτα πραγματοποιήσαμε την σύμπτυξη των υπολοίπων (Bagos and Nikolopoulos 2007). Για παράδειγμα αν το A αλληλόμορφο ευθύνεται για την ασθένεια, ενώνουμε τους γονοτύπους AB και BB και τους συσχετίζουμε με τον AA (Πίνακας 5) (Thakkinstian, McElduff et al. 2005; Bagos 2008).

	γονότυποι	
	AA	AB+BB
Ασθενείς (cases)	α	β + ε
Μάρτυρες (controls)	γ	δ + ζ

Πίνακας 4: Χρησιμοποίηση του μοντέλου επικρατούς κληρονομικότητας.

	γονότυποι	
	BB	AB+AA
Ασθενείς (cases)	ϵ	$\beta + \alpha$
Μάρτυρες (controls)	ζ	$\delta + \gamma$

Πίνακας 5: Χρησιμοποίηση του μοντέλου υπολειπόμενης κληρονομικότητας.

	γονότυποι	
	2BB+AB	2AA+AB
Ασθενείς (cases)	$2\epsilon + \beta$	$2\alpha + \beta$
Μάρτυρες (controls)	$2\zeta + \delta$	$2\gamma + \delta$

Πίνακας 6: Χρησιμοποίηση του μοντέλου συνεπικρατούς κληρονομικότητας.

Ο λογάριθμος του OR (logOR) μπορεί να πάρει οποιαδήποτε τιμή και έχει σχεδόν κανονική κατανομή. Εκτός από τον υπολογισμό των λόγων των πιθανοτήτων, για να αξιολογήσουμε αν οι μελέτες με παραβίαση του HWE διαφέρουν από τις μελέτες που δεν παραβιάζονται η HWE υπολογίζεται η σταθερή τιμή του συντελεστής z , που ακολουθεί τυπική, κανονική κατανομή. Η τιμή του z είναι ο λόγος της διαφοράς των λογαρίθμων των Odds-Ratio στις μελέτες με παραβίαση και στις μελέτες που δεν παραβιάζεται η HWE προς την σταθερή απόκλιση αυτής της διαφοράς.

$$z = \frac{\log(\text{Odds-Ratio})}{Se_{\text{Odds-Ratio}}} \sim N(0,1) \quad (16)$$

2.4 Μετά-ανάλυση στις μελέτες γενετικής συσχέτισης

Στις μελέτες γενετικής συσχέτισης εξετάζεται μία περιοχή με δύο αλληλόμορφα (A, B) που ευθύνονται για μια ασθένεια. Επιλέγονται οι ασθενείς και οι μάρτυρες βασιζόμενοι στον πληθυσμό (cases-controls), και κατόπιν ο γονότυπος καθορίζεται και συγκρίνεται. Ο πίνακας 6 παρουσιάζει τα στοιχεία της μετά-ανάλυσης από K μελέτες, ερευνώντας την υποθετική συσχέτιση (Bagos 2008).

Study	Cases			Controls		
	<i>AA</i> _{1<i>i</i>}	<i>AB</i> _{1<i>i</i>}	<i>BB</i> _{1<i>i</i>}	<i>AA</i> _{0<i>i</i>}	<i>AB</i> _{0<i>i</i>}	<i>BB</i> _{0<i>i</i>}
1	<i>AA</i> ₁₁	<i>AB</i> ₁₁	<i>BB</i> ₁₁	<i>AA</i> ₀₁	<i>AB</i> ₀₁	<i>BB</i> ₀₁
2	<i>AA</i> ₁₂	<i>AB</i> ₁₂	<i>BB</i> ₁₂	<i>AA</i> ₀₂	<i>AB</i> ₀₂	<i>BB</i> ₀₂
...
<i>k</i>	<i>AA</i> _{1<i>k</i>}	<i>AB</i> _{1<i>k</i>}	<i>BB</i> _{1<i>k</i>}	<i>AA</i> _{0<i>k</i>}	<i>AB</i> _{0<i>k</i>}	<i>BB</i> _{0<i>k</i>}

Πίνακας 7: Ένα χαρακτηριστικό σχεδιάγραμμα των στοιχείων που χρησιμοποιούνται σε μια μετά-ανάλυση των ασθενών-μαρτύρων που περιλαμβάνουν έναν ενιαίο αλληλόμορφο γεωμετρικό τόπο. Η διανομή των διάφορων γονοτύπων (AA, AB και BB) παρατίθεται στις στήλες για τους ασθενείς-μάρτυρες, ενώ οι διαφορετικές μελέτες (*i*=1, 2... K) που περιλαμβάνονται στη μετά-ανάλυση παρατίθενται στις διαφορετικές γραμμές.

Στην πλειοψηφία τους οι μελέτες γενετικής συσχέτισης της μετά-ανάλυσης περιλαμβάνουν έναν τόπο με δύο αλληλόμορφα (A εναντίον B). Η προσέγγιση αποτελείται από τους γονοτύπους των δύο κατηγοριών που κάθε φορά υποθέτουν ένα γενετικό μοντέλο ενώ στη συνέχεια εκτελούνται συγκρίσεις με την βοήθεια του σχετικού λόγου των πιθανοτήτων Odds Ratio (OR) (Thakkinstian, McElduff et al. 2005). Για παράδειγμα αν εξετάσουμε την σύγκριση των AB+BB εναντίον των γονοτύπων AA (AB+BB vs AA), υποθέτουμε έτσι το επικρατές πρότυπο κληρονομικότητας (Sasieni 1997), που ορίζεται ως εξής:

$$\log OR_i^D = \log \left(\frac{(AB_{1i} + BB_{1i}) AA_{0i}}{AA_{1i} (AB_{0i} + BB_{0i})} \right) \quad (17)$$

Η διακύμανση του $\log OR$ για το κυρίαρχο πρότυπο κληρονομικότητας (s^2) είναι η ακόλουθη:

$$\text{var}(\log OR_i^D) = \frac{1}{AB_{1i} + BB_{1i}} + \frac{1}{AA_{0i}} + \frac{1}{AB_{0i} + BB_{0i}} + \frac{1}{AA_{1i}} \quad (18)$$

Όταν χρησιμοποιούμε την αντίθεση BB vs. $AB+AA$, αυτή αντιστοιχεί στο υπολειπόμενο πρότυπο κληρονομικότητας (Sasieni 1997) και θα έχει τη μορφή:

$$\log OR_i^R = \log \left(\frac{BB_{1i} (AA_{0i} + AB_{0i})}{(AA_{1i} + AB_{1i}) BB_{0i}} \right) \quad (19)$$

Και η αντίστοιχη διακύμανση του λόγου της πιθανότητας για το υπολειπόμενο πρότυπο κληρονομικότητας (s^2) ισούται με:

$$\text{var}(\log OR_i^R) = \frac{1}{BB_{1i}} + \frac{1}{AB_{0i} + AA_{0i}} + \frac{1}{BB_{0i}} + \frac{1}{AB_{1i} + AA_{1i}} \quad (20)$$

Μια άλλη προσέγγιση, που εκτελείται των συσχετίσεων, είναι να συγκριθούν οι συχνότητες των αλληλόμορφων γονιδίων μεταξύ των ασθενών-μαρτύρων (allele-based-contrast) (Sasieni 1997). Ο λόγος των πιθανοτήτων για το αλληλόμορφο πρότυπο κληρονομικότητας (AB vs $AA+AB$) είναι:

$$\log OR_i^{Allele} = \log \left(\frac{(2BB_{1i} + AB_{1i})(2AA_{0i} + AB_{0i})}{(2AA_{1i} + AB_{1i})(2BB_{0i} + AB_{0i})} \right) \quad (21)$$

με διακύμανση του λόγου των πιθανοτήτων για το αλληλόμορφο πρότυπο κληρονομικότητας (s^2) να ισούται με τη μαθηματική σχέση:

$$\text{var}(\log OR_i^{Allele}) = \frac{1}{2BB_{1i} + AB_{1i}} + \frac{1}{2AA_{1i} + AB_{1i}} + \frac{1}{2AA_{0i} + AB_{0i}} + \frac{1}{2BB_{0i} + AB_{0i}} \quad (22)$$

2.4.2. Τροποποίηση για την ενσωμάτωση των αποκλίσεων από την ισορροπία HW

Ο νόμος του Hardy-Weinberg επιτυγχάνεται όταν ο αριθμός των παρατηρούμενων γονοτύπων στην ομάδα των μαρτύρων της μελέτης i (AA_{0i} , AB_{0i} , BB_{0i}) καθορίζεται από τις συχνότητες των αλληλομόρφων γονιδίων (π_{0i} , $1-\pi_{0i}$). Αυτό συμβαίνει όταν είναι ίσοι με τους αναμενόμενους:

$$AA_{0i} = n_{0i} (\pi_{0i}^A)^2, AB_{0i} = 2n_{0i} \pi_{0i}^A \pi_{0i}^B \text{ και } BB_{0i} = n_{0i} (\pi_{0i}^B)^2, \quad (23)$$

Όπου n_{0i} είναι ο συνολικός αριθμός των ατόμων στην ομάδα των μαρτύρων ($AA_{0i} + AB_{0i} + BB_{0i}$). Οι συχνότητες των αλληλόμορφων συχνοτήτων τους μπορούν να υπολογιστούν από τους παρατηρούμενους γονοτύπους χρησιμοποιώντας την σχέση 8:

$$\pi_{0i}^A = \frac{2AA_{0i} + AB_{0i}}{2(AA_{0i} + AB_{0i} + BB_{0i})}, \text{ και } \pi_{0i}^B = 1 - \pi_{0i}^A = \frac{2BB_{0i} + AB_{0i}}{2(AA_{0i} + AB_{0i} + BB_{0i})} \quad (24)$$

Με βάση τον έλεγχο χ^2 του Pearson μπορούμε να διατυπωθούν εύκολα οι δοκιμές για την ανίχνευση των αποκλίσεων των παρατηρούμενων γονοτυπικών συχνοτήτων από τις αναμενόμενες υπό την ισορροπία Hardy-Weinberg. (Emigh 1980). Προτείνονται, διάφορα μέτρα για να μετρήσουν την ποσότητα της απόκλισης από την ισορροπία HWE (Shoemaker, Painter et al. 1998). Η πιο κοινή και συνηθέστερη μετρική μέθοδος είναι ο συντελεστής ομομιξίας F (inbreeding coefficient) (Weir 1970), ο συντελεστής ανισορροπίας D (Hernandez and BS. 1989) και ο συντελεστής α (Pereira and Rogatko 1984; Lindley 1988). Ο συντελεστής ομομιξίας F εκφράζεται από τις μετρήσεις των παρατηρούμενων γονοτύπων και δίνεται από την μαθηματική σχέση:

$$F_{ji} = 1 - \frac{2AB_{ji}(AA_{ji} + AB_{ji} + BB_{ji})}{(2BB_{ji} + AB_{ji})(2AA_{ji} + AB_{ji})} = \frac{4BB_{ji}AA_{ji} - (AB_{ji})^2}{(2BB_{ji} + AB_{ji})(2AA_{ji} + AB_{ji})} \quad (25)$$

Ο συντελεστής ομομιξίας F κυμαίνεται μεταξύ -1 και 1 και εξαρτάται από τις συχνότητες των αλληλόμορφων γονιδίων (Brown 1970; Mendell and Simon 1984). Ο συντελεστής ανισορροπίας D της ανισορροπίας είναι ανεξάρτητος από τις συχνότητες των αλληλομόρφων και εκφράζεται από τις μετρήσεις των παρατηρούμενων γονοτύπων και δίνεται από την σχέση 10:

$$D_i = \frac{4AA_{0i}BB_{0i} - (AB_{0i})^2}{4(AA_{0i} + AB_{0i} + BB_{0i})^2} \quad (26)$$

Ο συντελεστής α είναι ο στατιστικός λόγος και υπολογίζεται από την σχέση:

$$\alpha_i = \frac{1}{2} \log \left(\frac{4AA_{0i}BB_{0i}}{AB_{0i}^2} \right) \quad (27)$$

Η διακύμανση του συντελεστή α (s^2) δίνεται από την σχέση:

$$s_{\alpha_i}^2 = \frac{1}{4AA_{0i}} + \frac{1}{AB_{0i}} + \frac{1}{4BB_{0i}} \quad (28)$$

Στην περίπτωση αποκλίσεων από τη ισορροπία HWE, τα προαναφερθέντα μέτρα της συσχέτισης χρειάζονται τροποποίηση (Lathrop 1983; Shaid and Jacobsen 1999; Trikalinos, Salanti et al. 2006). Στην περίπτωση των γονοτυπικών συγκρίσεων, προτείνεται η λύση της αντικατάστασης των παρατηρούμενων γονοτύπων στους ασθενείς από τις αναμενόμενες μετρήσεις (Lathrop 1983; Trikalinos, Salanti et al. 2006). Κατά συνέπεια για το επικρατές μοντέλο κληρονομικότητας θα έχουμε τον τύπο για τα $\log OR_{i_adjusted}$ που δίνεται στην σχέση 14:

$$\log OR_{i_adjusted} = \log \left(\frac{(AB_{1i} + BB_{1i}) AA_{0i}^{Expected}}{AA_{1i} (AB_{0i}^{Expected} + BB_{0i}^{Expected})} \right) \quad (29)$$

Η διακύμανση του τροποποιημένου λόγου των πιθανοτήτων (s^2) δίνεται από την σχέση:

$$\text{var}(\log OR_i^{D-adjusted}) = \frac{1}{AB_{1i} + BB_{1i}} + \frac{1}{AA_{1i}} + \frac{4}{(1 - \pi_i^B)(1 + \pi_i^B)^2 N_i^B} \quad (30)$$

όπου π_i^B μπορεί να υπολογιστεί από την σχέση (8) και N_i^B είναι ο συνολικός αριθμός των αλληλόμορφων B που είναι ίσος με $2BB_{0i} + AB_{0i}$. Οι αναμενόμενες μετρήσεις γονοτύπων υπό την ισορροπία HW μπορούν να υπολογιστούν όπως εκτιμήσαμε πριν από τους τύπους:

$$AA_{0i}^{Expected} = n_{0i} (\pi_{0i}^A)^2, \quad AB_{0i}^{Expected} = 2n_{0i} \pi_{0i}^A \pi_{0i}^B \quad \text{και} \quad BB_{0i}^{Expected} = n_{0i} (\pi_{0i}^B)^2 \quad (31)$$

Όμοια στην περίπτωση για το υπολειπόμενο πρότυπο κληρονομικότητας υπολογίζουμε τον τροποποιημένο λόγο των πιθανοτήτων που δίνεται από την σχέση:

$$\log OR_i^{R-adjusted} = \log \left(\frac{BB_{1i} (AA_{0i}^{Expected} + AB_{0i}^{Expected})}{(AA_{1i} + AB_{1i}) BB_{0i}^{Expected}} \right) \quad (32)$$

Η διακύμανση για το υπολειπόμενο πρότυπο με τροποποίηση (s^2) ισούται με:

$$\text{var}(\log OR_i^{R-adjusted}) = \frac{1}{BB_{1i}} + \frac{1}{AB_{1i} + AA_{1i}} + \frac{4}{(1 - \pi_i^A)(1 + \pi_i^A)^2 N_i^A} \quad (33)$$

Όμοια, το π_i^A υπολογίζεται από την σχέση (8) και το N_i^A είναι ο συνολικός αριθμός του αλληλόμορφου B το οποίο είναι ίσο με: $2AA_{0i} + AB_{0i}$. Επίσης χρησιμοποιείται ο λόγος των πιθανοτήτων της σύγκρισης για το συνεπικρατές πρότυπο κληρονομικότητας χωρίς τροποποίηση και έτσι η σχέση (5) παραμένει αμετάβλητη.

Η διακύμανση του όμως τροποποιείται παρόλα αυτά με δύο τρόπους και δίνεται από την σχέση 34 από την μέθοδο του Sato (Sato, Suganami et al. 2006):

$$\begin{aligned} \text{var}(\log OR_i^{Allele_adjusted}) = & \left(\frac{1}{2BB_{1i} + AB_{1i}} + \frac{1}{2AA_{1i} + AB_{1i}} \right) (1 + F_{1i}) \\ & + \left(\frac{1}{2AA_{0i} + AB_{0i}} + \frac{1}{2BB_{0i} + AB_{0i}} \right) (1 + F_{0i}) \end{aligned} \quad (34)$$

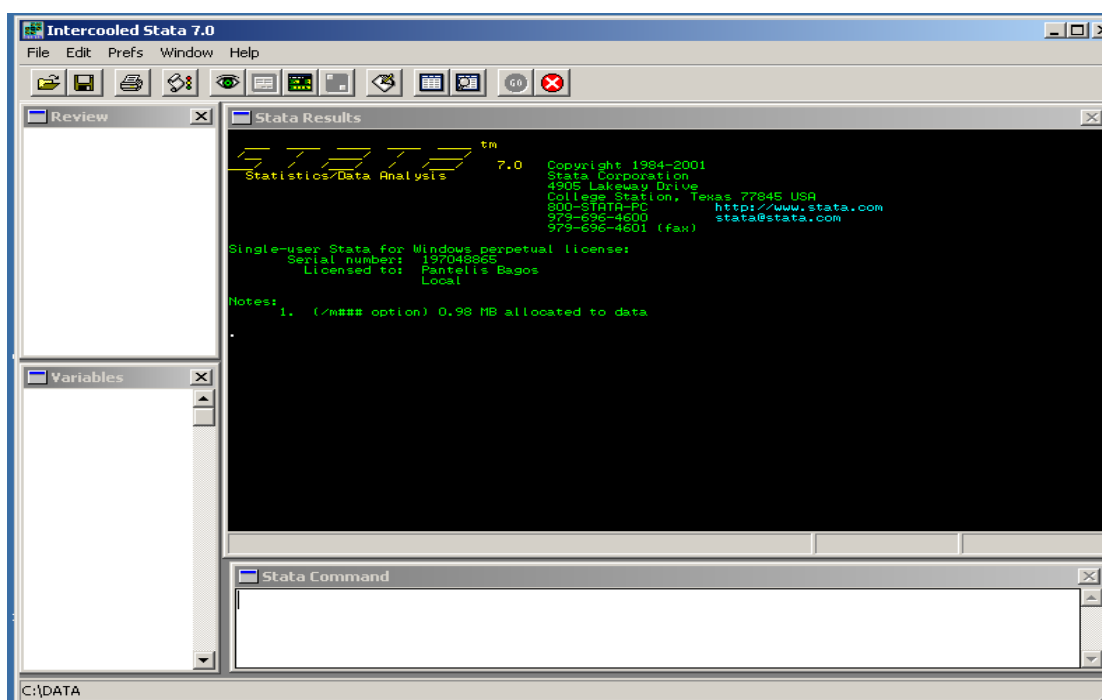
Και την διακύμανση με τροποποίηση της σχέσης 35, από τη μέθοδο του Ιωαννίδη:

$$\text{var}(\log OR_i^{Allele_adjusted}) = \left(\frac{1}{2BB_{1i} + AB_{1i}} + \frac{1}{2AA_{1i} + AB_{1i}} + \frac{1}{2AA_{0i} + AB_{0i}} + \frac{1}{2BB_{0i} + AB_{0i}} \right) (1 + F) \quad (35)$$

Στην παραπάνω σχέση τα F_{1i} και F_{0i} είναι οι συντελεστές ομομιξίας για τους ασθενείς και μάρτυρες αντίστοιχα. Έχουν προταθεί δύο σχέσεις για την διακύμανση της σχέσης 16. Στην πρώτη περίπτωση η τροποποίηση εκφράζεται με τον συντελεστή ανισορροπίας D αντί για τον συντελεστή ομομιξίας F (Zintzaras 2008). Στην δεύτερη περίπτωση οι τροποποιήσεις περιλαμβάνουν τον πολλαπλασιασμό της διακύμανσης της σχέσης 16 με τον παράγοντα, F_{i+1} όπου F_i είναι ο συντελεστής ομομιξίας του δείγματος ασθενών-μαρτύρων της i μελέτης (Trikalinos, Salanti et al. 2006). Οι προαναφερθείσες τροποποιήσεις είναι απλές και εφαρμόζονται στις μεθόδους της μετά-ανάλυσης. Υπάρχουν δύο ζητήματα εντούτοις που πρέπει να αντιμετωπιστούν. Αρχικά, οι αποκλίσεις από την ισορροπία HWE δεν μπορούν να προσδιοριστούν εύκολα ακόμα και όταν εφαρμόζονται οι τροποποιήσεις και αφετέρου, το γενετικό μοντέλο της κληρονομικότητας πρέπει να διευκρινιστεί εκ των προτέρων.

2.5. Στατιστικό Πακέτο STATA (version 10) Software

Το στατικό πακέτο STATA είναι ένα πανίσχυρο πρόγραμμα ικανό για επεξεργασία, παρουσίαση και ανάλυση αριθμητικών δεδομένων. Διαθέτει πανίσχυρες ιδιότητες για στατιστική επεξεργασία δεδομένων, γραφικές παραστάσεις. Το βασικό χαρακτηριστικό του είναι ότι διαθέτει μια δική του γλώσσα προγραμματισμού ούτως ώστε κάθε χρήστης να μπορεί να πραγματοποιήσει πολύπλοκες αναλύσεις οι οποίες ακόμα δεν παρέχονται από την επίσημη έκδοση. Το σπουδαιότερο από όλα είναι το γεγονός ότι οι συνεισφορές ανεξάρτητων χρηστών μαζεύονται σε ειδικούς ιστότοπους από όπου μπορεί να γίνει αναζήτηση μέσω του ίδιου του προγράμματος. Με αυτό τον τρόπο μπορούμε πολλές φορές να βρούμε χρήσιμες ρουτίνες που να καλύπτουν και τις πιο απαιτητικές ανάγκες.



Εικόνα 6: Το interface του προγράμματος φαίνεται στην παραπάνω εικόνα, με το κυρίως παράθυρο, κάτω την γραμμή εντολών και αριστερά τις μεταβλητές που υπάρχουν στην μνήμη.

2.5.1. Γενική περιγραφή του STATA

Στο παράθυρο **Stata Commands** πληκτρολογούνται οι εντολές. Το **Review** δείχνει τις προηγούμενες εντολές που εκτελέσαμε, το **Variables** δείχνει τις μεταβλητές που υπάρχουν στην μνήμη και το **Stata Results** είναι το βασικό παράθυρο των αποτελεσμάτων. Από τα εικονίδια της γραμμής εργαλείων, πιο σημαντικά είναι τα **Data Editor** (Επεξεργασία δεδομένων), **Data Browser** (φυλλομέτρηση δεδομένων) και **Do-file Editor** (Επεξεργασία εντολών του προγράμματος- do files ή ado files).

Το πρόγραμμα διαθέτει λειτουργία βοήθειας, η οποία καλείται είτε από τη γραμμή εργαλείων (Help-> Search ...) είτε από τη γραμμή εντολών ως εξής : help [subject] ή search [subject]

Τα παρακάτω κουμπιά ενεργοποιούν το data editor το οποίο είναι ένα «spreadsheet» παράθυρο που μπορεί να χρησιμοποιηθεί για την εισαγωγή δεδομένων στο Stata ή την αναζήτηση και την πρόσθεση είδη υπάρχοντων δεδομένων του Stata. Με το πάτημα του κουμπιού καθορίζουμε την λειτουργία. Το Stata είναι ένα στατιστικό λογισμικό πακέτο εντολών. Μερικές εντολές μπορούν να εκτελεστούν είτε από το Menu είτε από την γραμμή εργαλείων. Όλες οι εντολές μπορούν να πληκτρολογηθούν στο παράθυρο γραμμής εντολών χρησιμοποιώντας την συντακτική γλώσσα του Stata.



Εικόνα 7: Η γραμμή εργαλείων για άνοιγμα αρχείου, για αποθήκευση, για εισαγωγή δεδομένων στον data editor, για άνοιγμα του do-file.

Η βασική σύνταξη της γλώσσας του Stata είναι η ακόλουθη:

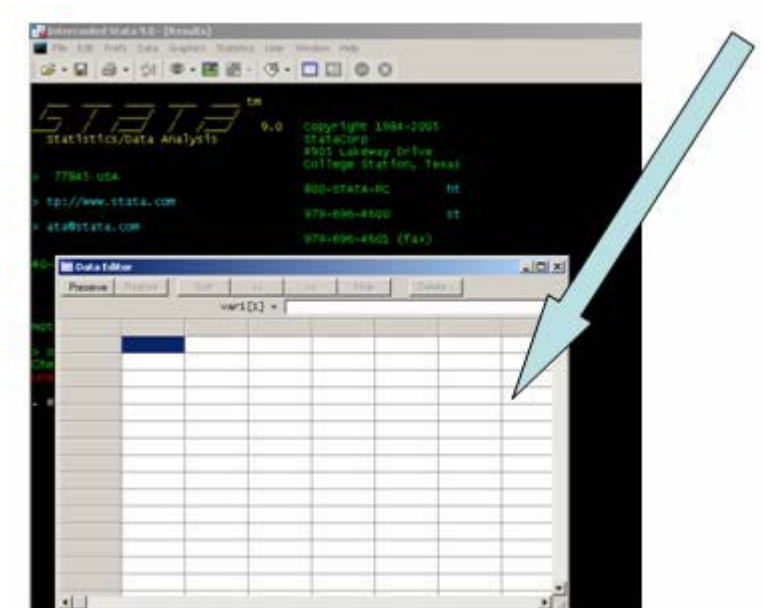
```
[by varlist:] command [varlist] [=exp] [if exp] [in range]
[weight] [,options]
```

Όπου οι αγκύλες δείχνει τους χαρακτήρες. Οι αγκύλες δεν μπαίνουν στην είσοδο της εντολής. Όλες οι εντολές του Stata πληκτρολογούνται σε πεζά-μικρά γράμματα. Το [varlist] δείχνει μία λίστα ονομάτων, το command μία εντολή του Stata, το [=exp] υποδηλώνει μία αλγεβρική έκφραση, το [in range] μία σειρά παρατηρήσεων, το [weight] μία έκφραση στάθμισης και το [,options] μία λίστα από επιλογές. Δεδομένου ότι όλα εκτός από την εντολές που εσωκλείονται στις αγκύλες, η απλούστερη μορφή της γλώσσας θα ήταν να εκδοθεί ακριβώς μια εντολή. Παραδείγματος χάριν, εάν πληκτρολογήσουμε: summarize αυτό θα μας έδινε τη μέση, σταθερή απόκλιση και τη σειρά όλων των μεταβλητών στο φύλλο εργασίας του Stata.

2.5.2 Εισαγωγή δεδομένων

Η εισαγωγή δεδομένων μπορεί να γίνει απευθείας στον editor (δακτυλογράφηση ή αντιγραφή-επικόλληση) είτε από υπάρχον αρχείο.

meta	pmid	study	AA1	AB1	BB1	AA0	AB0	BB0	GENE	Disease	Author	year
1	16229232	1	204	177	49	188	177	49	VDR Tag I	pulmonary tuberculosis	Lewis	2005
1	16229232	2	79	90	31	51	42	15	VDR Tag I	pulmonary tuberculosis	Lewis	2005
1	16229232	3	21	18	2	35	68	13	VDR Tag I	pulmonary tuberculosis	Lewis	2005
1	16229232	4	325	30	3	96	10	0	VDR Tag I	pulmonary tuberculosis	Lewis	2005
1	16229232	5	174	132	37	331	253	50	VDR Tag I	pulmonary tuberculosis	Lewis	2005
1	16229232	6	105	12	3	203	32	5	VDR Tag I	pulmonary tuberculosis	Lewis	2005
1	16229232	7	90	10	0	169	31	1	VDR Tag I	pulmonary tuberculosis	Lewis	2005
1	16229232	8	13	23	10	27	27	10	VDR Tag I	pulmonary tuberculosis	Lewis	2005
2	16229232	1	24	14	1	74	39	3	VDR Fek I	pulmonary tuberculosis	Lewis	2005
2	16229232	2	78	36	6	43	29	8	VDR Fek I	pulmonary tuberculosis	Lewis	2005
2	16229232	3	258	138	20	444	242	32	VDR Fek I	pulmonary tuberculosis	Lewis	2005
2	16229232	4	29	63	28	85	120	35	VDR Fek I	pulmonary tuberculosis	Lewis	2005
2	16229232	5	9	32	69	14	78	109	VDR Fek I	pulmonary tuberculosis	Lewis	2005
2	16229232	6	28	15	3	38	23	3	VDR Fek I	pulmonary tuberculosis	Lewis	2005
3	16172806	1	66	117	48	76	130	30	BDNF G196A	Parkinson	Zintzaras	2005
3	16172806	2	171	79	7	209	85	12	BDNF G196A	Parkinson	Zintzaras	2005
3	16172806	3	32	49	26	24	65	24	BDNF G196A	Parkinson	Zintzaras	2005
3	16172806	4	112	128	51	86	141	64	BDNF G196A	Parkinson	Zintzaras	2005
3	16172806	5	12	96	98	4	103	88	BDNF G196A	Parkinson	Zintzaras	2005
3	16172806	6	92	171	64	88	140	47	BDNF G196A	Parkinson	Zintzaras	2005
4	15887731	1	4	5	3	9	11	19	B2ADR Arg/Gly16	Asthma	Thakkinstian	2005
4	15887731	2	56	163	84	101	318	185	B2ADR Arg/Gly16	Asthma	Thakkinstian	2005
4	15887731	3	38	97	38	51	62	24	B2ADR Arg/Gly16	Asthma	Thakkinstian	2005
4	15887731	4	52	54	22	38	64	34	B2ADR Arg/Gly16	Asthma	Thakkinstian	2005
4	15887731	5	29	47	78	17	39	35	B2ADR Arg/Gly16	Asthma	Thakkinstian	2005
4	15887731	6	14	50	53	74	263	180	B2ADR Arg/Gly16	Asthma	Thakkinstian	2005
4	15887731	7	5	19	27	7	16	33	B2ADR Arg/Gly16	Asthma	Thakkinstian	2005
4	15887731	8	45	151	127	21	85	75	B2ADR Arg/Gly16	Asthma	Thakkinstian	2005
4	15887731	9	13	40	19	25	51	35	B2ADR Arg/Gly16	Asthma	Thakkinstian	2005
4	15887731	10	5	18	15	35	108	88	B2ADR Arg/Gly16	Asthma	Thakkinstian	2005
4	15887731	11	25	38	13	22	37	11	B2ADR Arg/Gly16	Asthma	Thakkinstian	2005
4	15887731	12	7	24	7	34	67	54	B2ADR Arg/Gly16	Asthma	Thakkinstian	2005
4	15887731	13	34	35	11	27	25	17	B2ADR Arg/Gly16	Asthma	Thakkinstian	2005
4	15887731	14	11	54	37	28	147	142	B2ADR Arg/Gly16	Asthma	Thakkinstian	2005



Εικόνα 8: Εισαγωγή δεδομένων στον Data Editor του STATA 10

Το άνοιγμα αρχείων και η αποθήκευση γίνεται με τις εντολές:

- `help use, save`

Η εντολή `clear` σβήνει τα δεδομένα (μεταβλητές) από τη μνήμη.

Οι μεταβλητές του Stata περιλαμβάνουν εντολές, συναρτήσεις και τελεστές :

Εντολές:

- `generate, egen` (δημιουργία νέας μεταβλητής),
- `replace` (αλλαγή στο περιεχόμενο μιας ήδη υπάρχουσας),
- `set` (καθορισμός διαφόρων παραμέτρων)
- `set obs` (number of observations)
- `set type` (variable type – float, double etc)

Τύποι μεταβλητών:

`help datatypes`

- αριθμητικές (numeric): `byte, int, long, float, double`
- συμβολοσειρές (strings)
- `encode, decode, recode` (ειδικές εντολές για αλλαγές στις μεταβλητές)

Συναρτήσεις:

-Mathematical-Statistical-String-Random numbers-Date-Special-Time series-Matrix

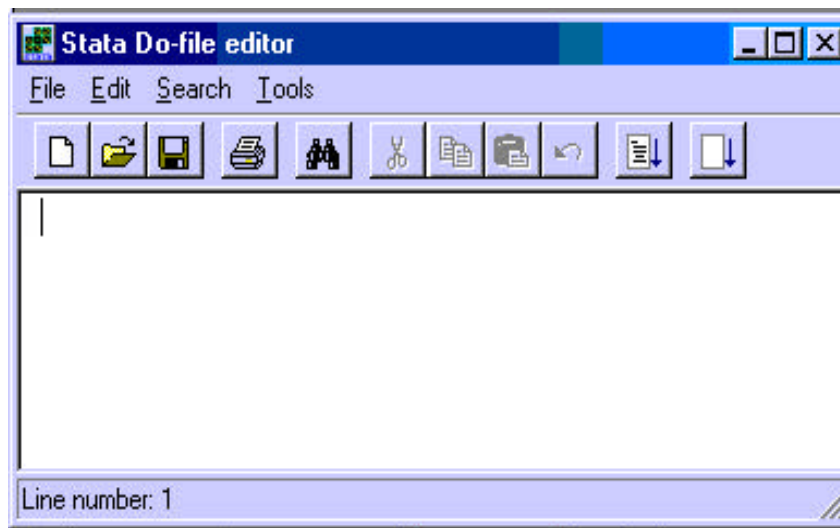
Τελεστές-Λογικές εκφράσεις:

- Αριθμητικοί:
+, -, /, *, ^,
- Λογικοί:
~ (not), | (or), & (and)

για έλεγχο χρησιμοποιούμε το **διπλό ίσον** (==)

Χρήση του do-file editor:

Ο do-file editor επιτρέπει την υποβολή διάφορων εντολών σε Stata με τη δακτυλογράφηση στον editor και έπειτα το πάτημα των κουμπιών (εικόνα 10). Μπορούμε να σώσουμε τις εντολές για να επιτρέπεται η αναπαραγωγή της διεργασίας χωρίς να πρέπει να δακτυλογραφηθεί πάλι κάθε μια από τις εντολές. Για να ανοίξει το αρχείο do- file editor γράφουμε τον τύπο στο παράθυρο την εντολή: doedit.



Εικόνα 9 : Παράθυρο του stata do-file editor για την πληκτρολόγηση των εντολών. Υπάρχει η επιλογή button για run των εντολών και do current-file.

ΚΕΦΑΛΑΙΟ 3: ΑΠΟΤΕΛΕΣΜΑΤΑ

3. ΑΠΟΤΕΛΕΣΜΑΤΑ

Κύριος σκοπός της έρευνας μας, είναι να πάρουμε ένα μεγάλο δείγμα μελετών, να μελετήσουμε όλους τους τρόπους διόρθωσης και να εξάγουμε συμπεράσματα . Για την έρευνα μας χρησιμοποιήσαμε το στατιστικό πακέτο Stata, το πρόγραμμα με το οποίο μας βοηθάει να κάνουμε έλεγχο της ισορροπίας HW για 71 μετά-αναλύσεις. Οι μετά-αναλύσεις αυτές είναι του έτους 2005-2006 και περιέχουν αναλυτικά όλους τους γονοτύπους. Όλες οι μετά-αναλύσεις πληρούν τους κανονισμούς που πρέπει να προϋποθέτει ένας πληθυσμός τον οποίο μελετάμε.

Αναγνωρίσαμε τις μελέτες γενετικής συσχέτισης μέσω λεπτομερούς ερευνάς όλων των θεμάτων που δημοσιεύονται σε διάφορες εφημερίδες γενετικής. Τα άτομα αξιολογήθηκαν με βάση τον συσχετισμό ανάμεσα σε ένα συγκεκριμένο πολυμορφισμό. Οι αναλύσεις συγκεντρώθηκαν σε μελέτες και συσχετισμούς όπου η διασπορά του γονότυπου δινόταν ώστε να μπορούμε να ελέγξουμε για ισορροπία HW. Κάθε μία από τις 71 μετά-αναλύσεις περιλαμβάνουν όλες τις απαραίτητες πληροφορίες ενός πληθυσμού. Υπάρχει ο αριθμός της μετά-ανάλυσης, meta, το πλήθος study της κάθε μετά-ανάλυσης που στην περίπτωση μας είναι από 1 έως 14 μελέτες, ένας κωδικός της Pub MED, το PMID, οι γονότυποι aa₀,ab₀,bb₀ για τους μάρτυρες (controls) και ab₁,bb₁,aa₁ για τους ασθενείς (cases) με όλες τις συχνότητες των γονοτύπων, το συγκεκριμένο γονίδιο, gene, που ευθύνεται για την συγκεκριμένη ασθένεια, το όνομα της ασθένειας (disease), το έτος (year) και τον συγγραφέα (author) που δημοσίευσε το συγκεκριμένο γονίδιο (Πίνακας 4).

Έχοντας συλλέξει όλες αυτές τις πληροφορίες δημιουργήσαμε ένα πρόγραμμα σε γλώσσα Stata για την εξαγωγή αποτελεσμάτων για το πώς η απόκλιση στην ισορροπία HW έχει επιπτώσεις στις μελέτες γενετικής συσχέτισης ενός πληθυσμού.

3.1. ΠΙΝΑΚΑΣ ΜΕ ΤΙΣ 71 ΜΕΤΑ-ΑΝΑΛΥΣΕΙΣ

meta	pmid	st	gene	disease	author	year	cases	controls	aa0t	aa1t	ab0t	ab1t	bb0t	bb1t
1	16229232	6	VDR Taq I	pulmonary tuberculosis	Lewis	2005	1636	1873	1100	1011	630	490	143	135
2	16229232	5	VDR Fok I	pulmonary tuberculosis	Lewis	2005	841	1419	698	426	531	298	190	117
3	16172806	3	BDNF G196A	Parkinson	Zintzaras	2005	1419	1406	487	485	654	640	265	294
4	15987731	9	B2ADR Arg/Gly16	Asthma	Thakkestian	2005	1665	2714	489	336	1293	795	932	534
5	15987731	12	B2ADR Gln/Glu27	Asthma	Thakkestian	2005	1495	2587	1190	772	1054	548	343	175
6	15961581	8	CTLA-4 +49A/G	type I diabetes	Kavvoura	2005	4745	5829	1997	1350	2641	2134	1191	1261
7	15947278	3	MTHFR C677T	Ischemic Stroke	Cronin	2005	4933	7082	3146	2060	3090	2185	846	688
8	15902512	3	MTHFR C677T	diabetic retinopathy	Zintzaras	2005	435	617	249	144	285	207	83	84
9	15861039	5	CYP2D6(LF/Wt)	tardive dyskinesia(schizophrenia)	Patsopoulos	2005	220	339	15	16	113	84	211	120
10	15861039	3	CYP2D6(10/Wt)	tardive dyskinesia(schizophrenia)	Patsopoulos	2005	168	288	74	56	139	87	75	25
11	15857569	6	HTR2A T102C	tardive dyskinesia	Lerer	2005	266	379	82	37	203	130	94	99
12	15849806	9	CYP1A1 Ile-Val	esophageal cancer	Yang	2005	754	1563	984	367	499	293	80	94
13	15849806	2	CYP1A1 Mspl	esophageal cancer	Yang	2005	336	851	461	148	302	152	88	36
14	15849806	3	CYP2E1	esophageal cancer	Yang	2005	420	1156	704	281	405	110	47	29
15	15849806	6	GSTP1 Ile105Val	esophageal cancer	Yang	2005	399	680	422	241	226	138	32	20
16	15830182	20	ACE I/D	diabetic nephropathy	D. P. K. Ng	2005	8663	6064	1608	2643	2774	4052	1682	1968
17	15802297	9	STK15 T91A	Cancer (breast)	Ewart-Toland	2005	6695	5743	2528	3006	2165	2478	1050	1211
18	15729744	4	MTHFR C677T	Schizophrenia	Lewis	2005	1119	1308	520	402	623	521	165	196
19	15682272	5	GLUT1 XbaI(0.9/1.1)	diabetic nephropathy	Zintzaras	2005	710	750	325	268	345	344	80	98
20	15592940	6	NEUROD1 Ala45Thr	type I diabetes	Kavvoura	2005	1213	1422	922	653	411	453	89	107
21	15592940	7	NEUROD1 Ala45Thr	type II diabetes	Kavvoura	2005	1844	1720	1217	1188	409	520	94	136
22	16133179	4	CTLA-4 +49A/G	rheumatoid arthritis	Han	2005	2315	2537	720	530	1178	1075	639	710
23	17194198	21	ACE I/D	Preeclampsia	Serrano	2006	2568	3828	952	592	1902	1131	974	845
24	17105984	8	MTFHR C677T	Breast Cancer	Lewis	2006	6373	8434	3757	2656	3633	2846	1044	871
25	17074966	2	MTFHR C677T	Depression	Gilbody	2006	1280	11051	5277	557	4732	562	1042	161

26	17074966	8	MTFHR C677T	Schizophrenia	Gilbody	2006	2762	3363	1574	1197	1444	1181	345	384
27	17074966	4	MTFHR C677T	Bipolar Disorder	Gilbody	2006	550	1098	539	263	456	227	103	60
28	16905645	8	VDR Fok I	Type I Diabetes	Guo	2006	1424	3301	1240	536	1645	690	416	198
29	16905645	13	VDR Bsm I	Type I Diabetes	Guo	2006	1571	4207	521	218	1702	650	1984	703
30	16905645	9	VDR Apa I	Type I Diabetes	Guo	2006	1101	2815	794	313	1380	522	641	266
31	16905558	3	CFH Y402H	age-related macular degeneration	Thakkinstian	2006	4863	3326	1319	892	1574	2308	433	1663
32	16897583	9	MTFHR C677T	acute lymphoblastic leukemia	Zintzaras	2006	1566	1958	862	717	838	671	258	178
33	16897583	3	MTFHR A1289C	acute lymphoblastic leukemia	Zintzaras	2006	1377	1682	782	676	721	603	179	98
34	16773678	3	MDR1 C3435T	Cronn's Disease	Annese	2006	1310	1535	399	362	770	664	366	284
35	16773678	2	MDR1 C3435T	Ulcerative Colitis	Annese	2006	1386	1558	403	317	788	691	367	378
36	16758123	5	MTFHR C677T	Gastric Cancer	Zintzaras	2006	1584	2785	864	464	1342	747	579	373
37	16758123	4	MTFHR A1289C	Gastric Cancer	Zintzaras	2006	760	1624	1156	536	445	212	23	12
38	16740590	5	Factor XIII Val34Leu	Venous Thromboembolism	Wells	2006	3065	4909	2864	1834	1743	1089	302	142
39	17171228	10	MTFHR C677T	Ischemic Stroke	Banerjee	2006	3074	3315	1362	1004	1409	1452	544	618
40	17171228	5	ACE I/D	Ischemic Stroke	Banerjee	2006	988	1441	551	371	681	448	209	169
41	17112651	2	eNOS/G894T (Glu298Asp)	Preeclampsia	Medica	2006	1225	2215	1655	835	509	328	51	62
42	17112651	6	AGT/T704C (Met235Thr)	Preeclampsia	Medica	2006	1122	2252	439	168	844	335	969	619
43	17063279	11	XRCC3 Thr241Met	Breast Cancer	Sang-Ah	2006	12159	11552	4835	4470	5236	5459	1481	2230
44	16820192	4	ITGA2 C807T	Ischemic Stroke	Nikolopoulos	2006	546	905	369	223	416	246	120	77
45	16513329	10	DRD3 Ser9Gly	tardive dyskinesia	Bakker	2006	695	915	465	318	369	305	81	72
46	16517578	8	TNF-alpha -308 G/A	Asthma	Gao	2006	2063	2650	1914	1447	659	537	77	79
47	16797381	4	Uteroglobin (G38A)	IgA nephropathy	Yong	2006	930	768	281	336	353	430	134	164
48	16940230	15	eNOS/G894T (Glu298Asp)	hypertension	Zintzaras	2006	6683	8486	5548	4076	2523	2183	415	424
49	16940230	9	eNoS (4 b/a)	hypertension	Zintzaras	2006	1487	1611	1217	1062	353	394	41	31
50	16940230	4	eNOS (T786C)	hypertension	Zintzaras	2006	2089	3456	2383	1399	909	576	164	114
51	17020975	1	RNASEL (Arg462Gln)	prostate cancer	Huihua Li	2006	3237	1713	639	1276	824	1514	250	447
52	17020975	3	RNASEL (Asp541Glu)	prostate cancer	Huihua Li	2006	3139	1958	441	636	925	1551	592	952
53	17264402	14	VDR Bsm I	osteoporosis	Zintzaras	2006	898	1594	294	207	795	403	505	288
54	17264402	6	VDR Taq I	osteoporosis	Zintzaras	2006	500	776	276	143	390	267	110	90
55	17264402	1	VDR Apa I	osteoporosis	Zintzaras	2006	540	1188	336	169	638	292	214	79

56	16691128	4	MTFHR C677T	Schizophrenia	Zintzaras	2006	2380	2852	1264	1002	1265	1041	323	337
57	16691128	3	MTFHR C677T	Bipolar Disorder	Zintzaras	2006	442	973	404	203	456	188	113	51
58	16691128	1	MTFHR C677T	Non Bipolar Depression	Zintzaras	2006	542	1062	444	233	493	244	125	65
59	16691128	4	MTFHR A1298C	Schizophrenia	Zintzaras	2006	1111	1454	729	513	605	484	120	114
60	16630166	1	MTFHR C677T	Breast Cancer	Zintzaras	2006	5467	7336	3325	2323	3138	2410	873	734
61	16630166	3	MTFHR A1298C	Breast Cancer	Zintzaras	2006	3768	5276	3100	2238	1839	1274	337	256
62	16449362	3	PADI4_94 & PADI4_89	Rheumatoid Arthritis	Iwamoto	2006	5493	4128	1518	1752	1927	2692	683	1049
63	16435105	3	PPARGC1A Gly482Ser	Type II Diabetes	Barroso	2006	3718	4816	2105	1636	2110	1615	601	467
64	16571112	1	NAT2 c.590G>A	Parkinson Disease	Borlak	2006	473	620	326	258	238	173	56	42
65	16600714	8	VDR Taq I	Prostate Cancer	Berndt	2006	2727	3686	1792	1321	1460	1110	434	296
66	16600714	3	VDR Apa I	Prostate Cancer	Berndt	2006	973	1139	384	301	487	415	268	257
67	16600714	3	Poly(A)	Prostate Cancer	Berndt	2006	1435	1683	720	600	715	630	248	205
68	16600714	7	VDR Bsm I	Prostate Cancer	Berndt	2006	2616	3153	1509	1258	1115	903	529	455
69	16600714	7	VDR Fok I	Prostate Cancer	Berndt	2006	1678	1599	627	661	723	758	249	259
70	16509765	2	GSTP1 Ile105Val	Lung Cancer	Ye	2006	6219	7607	3746	3060	3134	2517	727	642
71	16509765	1	GSTM3 3bp deletion in intron6	Lung Cancer	Ye	2006	1240	1179	874	886	270	316	35	38

Πίνακας 4: Παρουσιάζονται οι 71 μετά-αναλύσεις κρατώντας μόνο την τελευταία μελέτη από την κάθε μία. Η μετά-ανάλυση περιλαμβάνει ένα κωδικό PMID, το σύνολο των γονοτύπων με τις συχνότητες τους, την ασθένεια του πληθυσμού και το συγκεκριμένο γονίδιο που ευθύνεται για αυτήν. Επίσης το σύνολο των ασθενών και των μαρτύρων.

3.2. ΜΕΘΟΔΟΛΟΓΙΑ ΠΡΟΓΡΑΜΜΑΤΟΣ

Σε αυτή την εργασία χρησιμοποιείται το πρόγραμμα Stata version 10, ένα software για την εύκολη και γρήγορη διεκπεραίωση και εξαγωγή των στατιστικών συμπερασμάτων. Κύριος σκοπός είναι να γίνει η μετά-ανάλυση και στις 71 μελέτες με τρεις επιλογές. Η πρώτη επιλογή είναι η μετά-ανάλυση χωρίς τροποποίηση, δηλαδή εδώ υπολογίζονται οι λόγοι των πιθανοτήτων (log odds-ratio) αναθέτοντας σε μεταβλητές b_x και των διακυμάνσεων τους (variances) σε V_x για όλες τις συγκρίσεις γονοτύπων για το επικρατές, το υπολειπόμενο και το συνεπικρατές πρότυπο κληρονομικότητας. Σε αυτή την περίπτωση οι μεταβλητές που ορίζονται στην μετά-ανάλυση χωρίς τροποποίηση είναι: $b_{1,2,3}$ με διακυμάνσεις $V_{1,2,3}$. Οι τιμές των συντελεστών και των σταθερών αποκλίσεων για κάθε μετά-ανάλυση αποθηκεύονται στις μεταβλητές που έχουν οριστεί ως $b_{1overall}$, $V_{1overall}$.

Η δεύτερη επιλογή είναι η μετά-ανάλυση με τροποποίηση. Εδώ οι λόγοι των πιθανοτήτων και των διακυμάνσεων είναι τροποποιημένοι με τις δύο μεθόδους για τροποποίηση, του Sato και του Ιωαννίδη (odds ratio adjusted, variances adjusted) για όλες τις συγκρίσεις γονοτύπων. Γι' αυτό ορίζονται οι μεταβλητές $b_{4,5,6,7}$ με διακυμάνσεις $V_{4,5,6,7}$. Η τρίτη επιλογή είναι η μετά-ανάλυση χωρίς τροποποιήσεις αλλά με το να αφαιρούνται από το σύνολο των μελετών εκείνες που δεν είναι σε ισορροπία HW. Χρησιμοποιούνται οι μεταβλητές $b_{8,9,10}$ και οι διακυμάνσεις $V_{8,9,10}$. Αφού διεκπεραιωθεί η μετά-ανάλυση και με τους τρεις τρόπους, υπολογίζονται οι διαφορές των λόγων πιθανοτήτων, των διακυμάνσεων, του συντελεστή z , των τροποποιημένων από των μη τροποποιημένων και οι διαφορές των τροποποιημένων από τις μελέτες που είναι σε ισορροπία HW, με σκοπό την εξαγωγή γραφημάτων. Τα γραφήματα που εξάγονται, είναι σε συνάρτηση της διαφοράς των συντελεστών με το συνολικό ποσοστό των μελετών που βρίσκονται σε ισορροπία HW και με συνάρτηση του μέτρου ποσοτικοποίησης των αποκλίσεων α . Στη συνέχεια με την εντολή regression γίνεται παλινδρόμηση για όλες τις διαφορές. Στον πίνακα της παλινδρόμησης, μας ενδιαφέρει ο συντελεστής παλινδρόμησης β_1 , από την εξίσωση $y = \beta_0 + \beta_1 x$. Στα γραφήματα βγάζουμε συμπεράσματα για το ποιες μελέτες αποκλίνουν και με ποιο τρόπο επηρεάζουν τα αποτελέσματα της μετά-ανάλυσης.

3.3. ΠΡΟΓΡΑΜΜΑ STATA

Όλα τα δεδομένα του Πίνακα 4 για τις 71 μετά-αναλύσεις τα εισάγουμε στο data editor του Stata. Αφού κάνουμε αποθήκευση των δεδομένων ανοίγουμε το do-file editor για την καταγραφή των εντολών σε γλώσσα Stata. Αρχικά δημιουργούμε μεταβλητές που περιέχουν όλα τα αθροίσματα των ασθενών και μαρτύρων για όλες τις μελέτες. Επίσης υπολογίζεται το συνολικό άθροισμα όλων αλληλομόρφων των aa_0 , aa_1 , ab_0 , ab_1 , bb_0 , bb_1 για τους ασθενείς και του μάρτυρες και το πλήθος των μελετών κάθε μετά-ανάλυσης (εντολή 1).

Στη συνέχεια υπολογίζουμε τους σχετικούς λόγους των πιθανοτήτων (Odds Ratio). Έχουμε 3 είδη συγκρίσεων ανάλογα με το μοντέλο κληρονομικότητας, για τα αλληλόμορφα (A, B): εξετάζοντας τα AB+BB έναντι των γονοτύπων AA, για το επικρατές πρότυπο κληρονομικότητας, τα AA+AB έναντι των BB, για το υπολειπόμενο μοντέλο κληρονομικότητας, και τα AB (2BB+AB) έναντι (2AA+AB), για το συνεπικρατές πρότυπο κληρονομικότητας, μεταξύ των ασθενών-μαρτύρων (εντολή 2).

Κατόπιν, υπολογίζονται οι διακυμάνσεις για όλα τα πρότυπα κληρονομικότητας. Ακολουθεί ο έλεγχος σε περίπτωση μηδενισμού του κλάσματος στους λόγους των πιθανοτήτων και στις διακυμάνσεις και γίνεται αντικατάσταση (εντολή 3). Στη συνέχεια υπολογίζεται ο συντελεστής α (σχέση 27) και η διακύμανση του (σχέση 28). Αφού υπολογιστούν γίνεται διόρθωση και αντικατάσταση σε περίπτωση μηδενισμού του παρονομαστή (εντολή 5). Μετέπειτα, υπολογίζονται τα τυπικά σφάλματα όλων των διακυμάνσεων. Χρησιμοποιώντας τον νόμο Hardy-Weinberg υπολογίζουμε τον συντελεστή ομομιξίας F για τους ασθενείς-μάρτυρες αφού υπολογίσουμε τις πιθανότητες των αλληλομόρφων, υπολογίζουμε το F_0 , τον συντελεστή για τα AA_0 , AB_0 , BB_0 για του ασθενείς, το F_1 για τα AA_1 , AB_1 , BB_1 στους μάρτυρες και μετά το συνολικό F (σχέση 25).

Στη συνέχεια γίνεται η τροποποίηση για την ενσωμάτωση των αποκλίσεων από την ισορροπία HW με δύο τρόπους, με την μέθοδο του Ιωαννίδη και την μέθοδο του Sato (σχέση 34,35 εντολή 10). Αρχικά βρίσκουμε τον αριθμό των ατόμων της ομάδας των μαρτύρων. Εδώ για να επιτευχθεί η ισορροπία HW πρέπει το πλήθος των

γονοτύπων στους μάρτυρες σε κάθε μελέτη να καθορίζεται από τις συχνότητες των αλληλομόρφων (σχέση 23,24) όταν είναι ίσοι με τις αναμενόμενες τιμές (εντολή 9).

Οι μέθοδοι για τις τροποποιήσεις των αποκλίσεων από το HWE, βοηθούν για τον υπολογισμό των τροποποιημένων σχετικών λόγων των πιθανοτήτων (σχέση 29, 32) και των τροποποιημένων διακυμάνσεων (σχέση 30, 33, 34, 35). Οι διακυμάνσεις για το συνεπικρατές πρότυπο κληρονομικότητας, υπολογίζονται με δύο μεθόδους, του Sato και του Ιωαννίδης (εντολή 10,13). Αναθέτουμε όλους τους λόγους των πιθανοτήτων και των διακυμάνσεων με και χωρίς τροποποιήσεις σε μεταβλητές b_1 έως b_7 και αντίστοιχα V_1 έως V_7 για να πραγματοποιηθεί η μέθοδος της μετά-ανάλυσης. Τα V_6 και V_7 αντιστοιχούν στα τροποποιημένα των δύο μεθόδων για τις διακυμάνσεις του συνεπικρατούς μοντέλου κληρονομικότητας (εντολή 14).

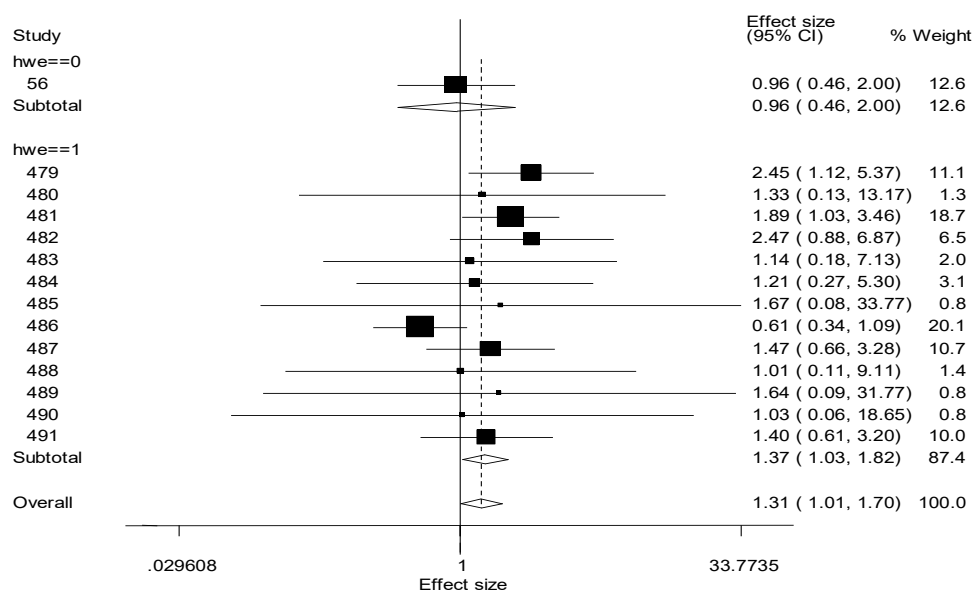
Δημιουργείται ένας πίνακας όπου διαβάζει όλες τις 75 μετά-αναλύσεις από τον αριθμό των μελετών και υπολογίζει αν οι μελέτες βρίσκονται σε ισορροπία με τον έλεγχο του exact test του Fisher (σχέση 10). Σε αυτές που έχουν HWE μικρότερο από 0.05 p-value τις θέτουμε ίσο με μονάδα και σε αυτές που δεν είναι σε ισορροπία να ισούνται με μηδέν (εντολή 15). Αρχικοποιούμε τις μεταβλητές $b_{1\text{overall}}$ - $b_{10\text{overall}}$ και V_1 - $V_{10\text{overall}}$ με κενές τιμές (εντολή 16), καθώς και την μεταβλητή a_{overall} και e_{overall} . Πραγματοποιείται η μέθοδος της μετά-ανάλυσης με την εντολή *mvmeta* και τα αποτελέσματα που θέλουμε για κάθε επανάληψη τα αναθέτουμε σε αυτές τις νέες τιμές μεταβλητών με κατάληξη $_{\text{overall}}$ όταν υπολογίζονται. Σε αυτή την εργασία χρησιμοποιείται το πρόγραμμα Stata για την εύκολη και γρήγορη διεκπεραίωση για την εξαγωγή των στατιστικών συμπερασμάτων.

Κύριος σκοπός είναι να γίνει η μετά-ανάλυση και στις 75 μελέτες με τρεις επιλογές. Η πρώτη επιλογή είναι η μετά-ανάλυση χωρίς τροποποίηση, δηλαδή εδώ παίρνουμε τους λόγους των πιθανοτήτων αναθέτοντας σε μεταβλητές b_x και των διακυμάνσεων τους variances σε μεταβλητές V_x . Σε αυτή την περίπτωση οι μεταβλητές που ορίζονται στην μετά-ανάλυση είναι: b_1, b_2, b_3 με διακυμάνσεις V_1, V_2, V_3 . Η δεύτερη επιλογή είναι η μετά-ανάλυση με τροποποίηση. Εδώ οι λόγοι των πιθανοτήτων και των διακυμάνσεων είναι τροποποιημένοι με τις δύο μεθόδους για τροποποίηση, του Sato και του Ιωαννίδη για όλα τα πρότυπα κληρονομικότητας (σχέση 34, 35). Για αυτό το λόγο ορίζονται οι μεταβλητές b_4, b_5, b_6, b_7 με διακυμάνσεις V_4, V_5, V_6, V_7 . Η τρίτη επιλογή είναι η μετά-ανάλυση χωρίς τροποποιήσεις αλλά να

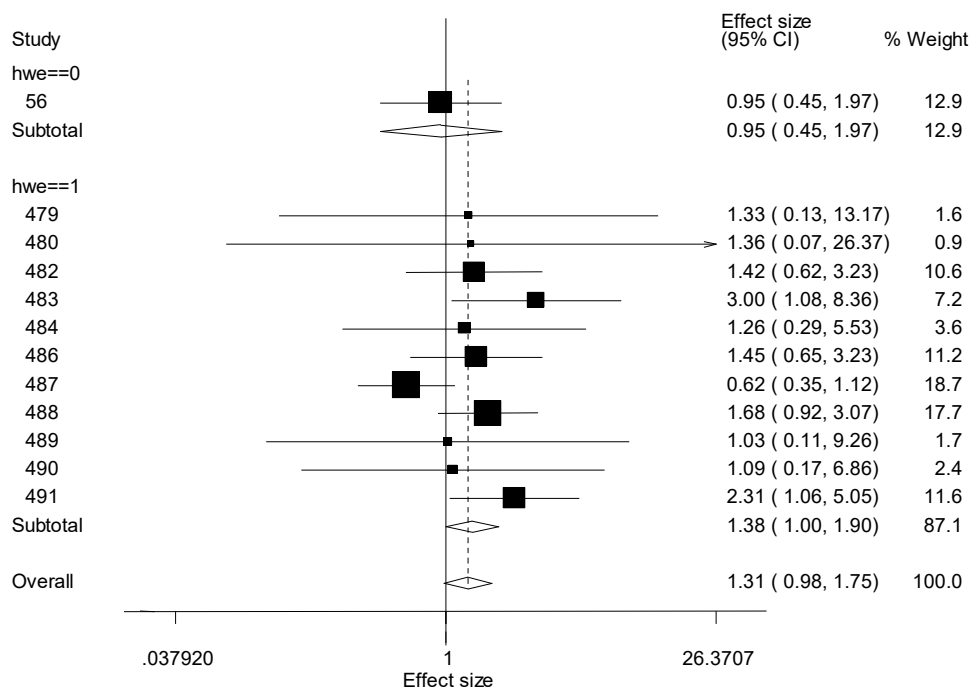
αφαιρούνται οι μελέτες αυτές που δεν είναι σε ισορροπία (HWD).. Χρησιμοποιούνται οι μεταβλητές b_8, b_9, b_{10} και οι διακυμάνσεις V_8, V_9, V_{10} . Τέλος υπάρχει και μία τέταρτη επιλογή μετά-ανάλυσης, η οποία γίνεται στο συντελεστή απόκλισης α , για το λόγο αυτό ορίζεται και η μεταβλητή $\alpha_{overall}$.

Αφού διεκπεραιωθεί η μετά-ανάλυση και με τους τρεις τρόπους, γίνονται συγκρίσεις μεταξύ των τροποποιημένων και των μη τροποποιημένων, των τροποποιημένων και αυτών των που είναι σε ισορροπία HW, για του σχετικούς λόγους πιθανοτήτων, των διακυμάνσεων και των συντελεστών z . Δεύτερη σύγκριση είναι, η διαφορές των σχετικών λόγων πιθανοτήτων, των διακυμάνσεων και των z για όλα τα πρότυπα κληρονομικότητας, σε συνάρτηση με το συνολικό ποσοστό των μελετών που είναι σε ισορροπία HW και σε συνάρτηση του συντελεστή απόκλισης α (σχέση 27). Φτιάχνουμε μία μεταβλητή $last$ η οποία είναι ίση με 1 στην τελευταία μελέτη κάθε μετά-ανάλυσης. Μας ενδιαφέρει να κρατήσουμε την τελευταία γραμμή για κάθε μία από τις 75 μετά-αναλύσεις για να έχουμε 75 παρατηρήσεις (εντολή 18).

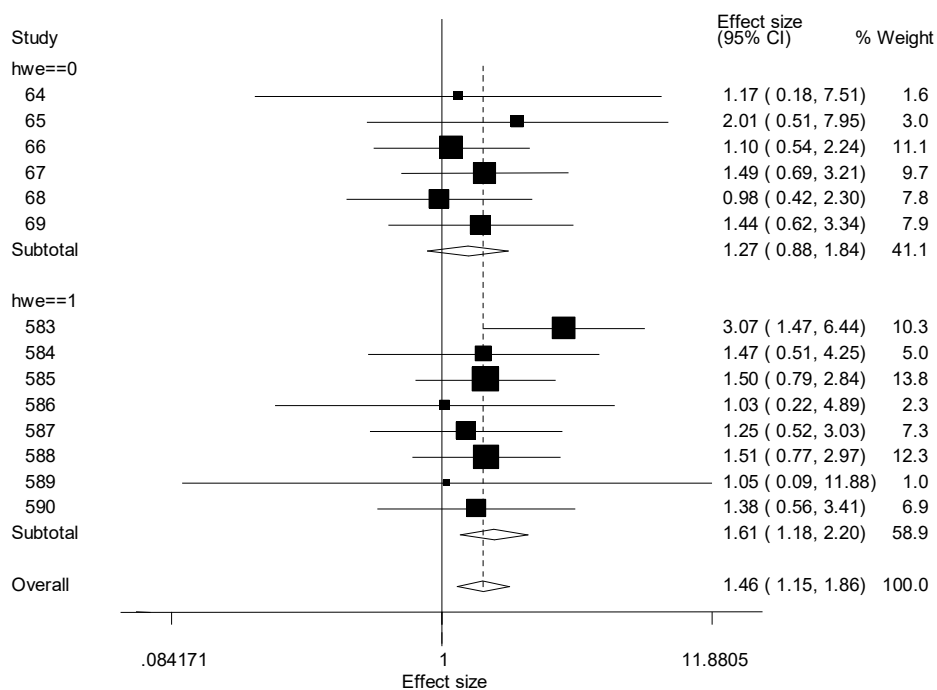
Παρακάτω είναι τα διαγράμματα forest-plot έπειτα από την εκτέλεση κάποια συγκεκριμένης μετά-ανάλυσης. Στα αριστερά παρουσιάζονται οι μελέτες της μετά-ανάλυσης διαχωρίζοντας αυτές που είναι σε ισορροπία HW και αυτές που δεν είναι. Στα δεξιά του κάθε forest plot παρουσιάζονται τα OR και τα διαστήματα εμπιστοσύνης της κάθε μελέτης.



Εικόνα 10: μετά-ανάλυση χωρίς τροποποίηση, της μελέτης με γονίδιο AGT/T704C (Met235Thr), υπεύθυνο για την ασθένεια Preeclampsia.



Εικόνα 11: μετά-ανάλυση με τροποποίηση, της μελέτης για το γονίδιο AGT/T704C (Met235Thr), σε συσχέτιση με την ασθένεια Preeclampsia.



Εικόνα 12: μετά-ανάλυση χωρίς τροποποίηση, της μελέτης για το γονίδιο VDR Bsm I σε συσχέτιση με την ασθένεια της οστεοπόρωσης.

Study	ES	[95% Conf. Interval]		% Weight	
-----+-----					
hwe==0					
446	-0.043	-0.779	0.693	12.62	
Sub-total					
D+L pooled ES	-0.043	-0.779	0.693	12.62	
-----+-----					
hwe==1					
440	0.898	0.115	1.680	11.15	
441	0.284	-2.010	2.578	1.30	
442	0.638	0.034	1.242	18.71	
443	0.902	-0.122	1.927	6.51	
444	0.127	-1.710	1.964	2.03	
445	0.188	-1.292	1.667	3.12	
447	0.513	-2.493	3.520	0.76	
448	-0.495	-1.078	0.088	20.09	
449	0.386	-0.414	1.187	10.66	
450	0.014	-2.183	2.210	1.42	
451	0.497	-2.465	3.459	0.78	
452	0.033	-2.859	2.926	0.82	
453	0.337	-0.488	1.162	10.04	
Sub-total					
D+L pooled ES	0.316	0.033	0.599	87.38	
-----+-----					
Overall					
D+L pooled ES	0.269	0.007	0.530	100.00	
-----+-----					
Test(s) of heterogeneity:					
	Heterogeneity statistic	degrees of freedom	P	I-squared**	Tau-squared
hwe==0	0.00	0	.	.%	0.0000
hwe==1	12.14	12	0.434	1.2%	0.0035
Overall	12.93	13	0.453	0.0%	0.0000
Overall Test for heterogeneity between sub-groups :					
	0.79	1	0.375		
** I-squared: the variation in ES attributable to heterogeneity)					
Significance test(s) of ES=0					
hwe==0	z= 0.11	p = 0.910			
hwe==1	z= 2.19	p = 0.029			
Overall	z= 2.01	p = 0.044			
-----+-----					

Εικόνα 13: Στο παραπάνω πίνακα φαίνονται τα αποτελέσματα των εκτιμητών χρησιμοποιώντας την εντολή metan. Αριστερά εμφανίζονται οι μελέτες ανάλογα με το αν είναι σε ισορροπία ή όχι, στο κέντρο φαίνονται οι εκτιμητές log(OR) (ES) και το διάστημα εμπιστοσύνης (CI) και δεξιά εμφανίζεται το ποσοστό συμβολής κάθε μελέτης. Από κάτω παρατηρούμε τα αποτελέσματα των μεταβλητών P (Q-Cochran), I^2 , τ^2 .

Study	ES	[95% Conf. Interval]		% Weight	
-----+-----					
hwe==0					
56		0.945	0.453	1.973	12.87
Sub-total					
D+L pooled ES		0.945	0.453	1.973	12.87
-----+-----					
hwe==1					
479		2.310	1.056	5.055	11.59
480		1.329	0.134	13.170	1.56
481		1.675	0.916	3.066	17.67
482		3.002	1.077	8.363	7.22
483		1.093	0.174	6.862	2.40
484		1.261	0.287	5.534	3.64
486		0.623	0.348	1.116	18.66
487		1.453	0.652	3.235	11.16
488		1.030	0.115	9.258	1.69
489		1.364	0.071	26.371	0.94
491		1.416	0.621	3.231	10.60
Sub-total					
D+L pooled ES		1.380	1.003	1.897	87.13
-----+-----					
Overall					
D+L pooled ES		1.309	0.980	1.747	100.00
-----+-----					
Test(s) of heterogeneity:					
Heterogeneity statistic	degrees of freedom	P	I-squared**	Tau-squared	
hwe==0	0.00		0	.	0.0000
hwe==1	11.56		10	0.316	13.5%
Overall	12.34		11	0.339	10.8%
Overall Test for heterogeneity between sub-groups :					
0.78	1	0.376			
** I-squared: the variation in ES attributable to heterogeneity)					
Significance test(s) of ES=1					
hwe==0	z=	0.15	p =	0.880	
hwe==1	z=	1.98	p =	0.048	
Overall	z=	1.82	p =	0.068	
-----+-----					

Εικόνα 14: Στο παραπάνω πίνακα φαίνονται τα αποτελέσματα των εκτιμητών χρησιμοποιώντας την εντολή metan για μετά-ανάλυση με την επιλογή της τροποποίησης (adjusted).

3.5. ΕΝΤΟΛΕΣ ΣΤΟ STATA

Οι εντολές που χρησιμοποιούμε στο πρόγραμμα STATA για την εξαγωγή των αποτελεσμάτων είναι οι ακόλουθες:

```
set more off      #για εκτέλεση όλων των δεδομένων χωρίς να πατάμε κάθε
#φορά space
qui sum meta

local max=r(max)

qui gen case=aa1+ab1+bb1

qui gen control=aa0+ab0+bb0

qui bysort meta: egen cases=sum( case)  #υπολογισμός αθροισμάτων

qui bysort meta: egen controls=sum( control)

qui bysort meta: egen aa0total=sum( aa0)

qui bysort meta: egen aa1total=sum( aa1)

qui bysort meta: egen ab0total=sum( ab0)

qui bysort meta: egen ab1total=sum( ab1)

qui bysort meta: egen bb0total=sum( bb0)

qui bysort meta: egen bb1total=sum( bb1)

qui bysort meta: egen studies=max(study) #πλήθος κάθε μελέτης      (1)

gen logor_dom=log((ab1+bb1)/aa1)/((ab0+bb0)/aa0)

gen logor_res=log(((bb1)/(ab1+aa1))/((bb0)/(ab0+aa0)))

gen logor_allele=log(((2*bb1+ab1)/(2*aa1+ab1))/
((2*bb0+ab0)/(2*aa0+ab0)))      #υπολογισμός odds-ratio      (2)

gen var_dom=(1/aa0+1/(ab0+bb0)+1/aa1+1/(ab1+bb1))

gen var_res=(1/bb0 +1/(ab0+aa0) +1/bb1 +1/(ab1+aa1))

gen var_allele=(1/(2*bb1+ab1)+1/(2*aa1+ab1)+      #υπολογισμός variances
+1/(2*bb0+ab)+1/(2*aa0+ab0))

replace logor_dom=log(((ab1+bb1+0.5)/(aa1+0.5))/#έλεγχος και διόρθωση
```

```

((ab0+bb0+0.5)/(aa0+0.5))
if aa0==0 |bb0+ab0==0|aa1==0 |bb1+ab1==0

replace logor_res=log(((bb1+0.5)/(aa1+ab1+0.5))/
((bb0+0.5)/(aa0+ab1+0.5)))
if ab0+aa0==0 |bb0==0|ab1+aa1==0 |bb1==0

replace logor_allele=log(((2*bb1+ab1+0.5)/(ab1+2*aa1+0.5))/
(2*bb0+ab0+0.5)/(ab0+2*aa0+0.5))) if ab0+2*aa0==0
|2*bb0+ab0==0|ab1+2*aa1==0 |2*bb1+ab1==0

replace var_dom= 1/(aa0+0.5) +1/(ab0+bb0+0.5) +1/(aa1+0.5)
+1/(ab1+bb1+0.5) if aa0==0 |ab0+bb0==0|aa1==0 |ab1+bb1==0

replace var_res= 1/(ab0+aa0+0.5) +1/(bb0+0.5) +1/(ab1+aa1+0.5)
+1/(bb1+0.5) if ab0+aa0==0 |bb0==0|ab1+aa1==0 |bb1==0

replace var_allele=(1/(2*bb1+ab1+0.5) +1/(ab1+2*aa1+0.5)
+1/(2*bb0+ab0+0.5) +1/(ab0+2*aa0+0.5)) if ab0+2*aa0==0
|2*bb0+ab0==0|ab1+2*aa1==0 |2*bb1+ab1==0

gen a1=0.5*log(((4*aa0*bb0)/(ab0*ab0)) #υπολογισμός συντελεστή α

gen e11=(0.25)*(1/aa0+1/bb0+4/ab0) #διακύμανση του α

replace a1=0.5*log(((4*(aa0+0.5)*(bb0+0.5))/((ab0+0.5)*(ab0+0.5)))) if
aa0==0 |bb0==0|ab0==0

replace e11=(0.25)*(1/(aa0+0.5)+1/(bb0+0.5)+4/(ab0+0.5)) if aa0==0
|bb0==0 |ab0==0 (5)

gen se_dom=sqrt(var_dom) #υπολογισμός stadar-error
gen se_res=sqrt(var_res)
gen se_allele=sqrt(var_allele) (6)

gen pA1=(2*aa1+ab1)/(2*aa1+2*ab1+2*bb1)
gen pB1=(2*bb1+ab1)/(2*aa1+2*ab1+2*bb1)

gen pAA1=aa1/(aa1+ab1+bb1)
gen pBB1=bb1/(aa1+ab1+bb1)

gen F1=pAA1/pA1+pBB1/pB1-1

gen pA0=(2*aa0+ab0)/(2*aa0+2*ab0+2*bb0)
gen pB0=(2*bb0+ab0)/(2*aa0+2*ab0+2*bb0)

gen pAA0=aa0/(aa0+ab0+bb0)
gen pBB0=bb0/(aa0+ab0+bb0)

gen F0=pAA0/pA0+pBB0/pB0-1

gen pA=(2*(aa1+aa0)+(ab1+ab0))/

```

```

(2*(aa1+aa0)+2*(ab1+ab0)+2*(bb1+bb0))

gen pB=(2*(bb1+bb0)+(ab1+ab0))/
(2*(aa1+aa0)+2*(ab1+ab0)+2*(bb1+bb0))

gen pAA=(aa1+aa0)/((aa1+ab1+bb1)+(aa0+ab0+bb0))

gen pBB=(bb1+bb0)/((aa1+ab1+bb1)+(aa0+ab0+bb0))

gen F=pAA/pA+pBB/pB-1 #συντελεστής ομομοιξίας (7)

gen N0=aa0+ab0+bb0 #αριθμός της ομάδας μαρτύρων

gen pa0=(2*aa0+ab0)/(2*(aa0+ab0+bb0)) #συχνότητες αλληλομόρφων
από τους γονότυπους των
παρατηρούμενων

gen pb0=(2*bb0+ab0)/(2*(aa0+ab0+bb0)) # 1-pa0 (8)

gen f0i=(4*bb0*aa0 - ab0*ab0)/((2*bb0 + ab0)*(2*aa0+ab0))

gen f1i=(4*bb1*aa1 - ab1*ab1)/((2*bb1 + ab1)*(2*aa1+ab1))

gen falli=(4*(bb1+bb0)*(aa1+aa0)-(ab0+ab1)*(ab1+ab0))/((2*(bb1+bb0) +
(ab1+ab0))*(2*(aa1+aa0)+(ab1+ab0))) #μέθοδος Ιωαννίδη

gen aa0_adj=N0*pa0^2 #adjustments για τους γονότυπους

gen ab0_adj=N0*2*pa0*pb0 (9)

gen bb0_adj=N0*pb0^2

gen f1s=1-(2*(aa1+ab1+bb1)*ab1)/((2*aa1+ab1)*(2*bb1+ab1)) #μέθοδος
Sato

gen f0s=1-(2*(aa0+ab0+bb0)*ab0)/((2*aa0+ab0)*(2*bb0+ab0))

gen var_allele_adj=((1/(2*aa1+ab1))+1/(2*bb1+ab1))*
(1+f1s))+((1/(2*aa0+ab0))+1/(2*bb0+ab0))*(1+f0s) (10)

replace
var_allele_adj=((1/(2*aa1+ab1+0.5))+1/(2*bb1+ab1+0.5)))+((1/(2*aa
0+ab0+0.5))+1/(2*bb0+ab0+0.5)) if f1s==. |f0s==.

gen logor_dom_adj=log((ab1+bb1)/aa1)/ #odds dominant_adjustment
((ab0_adj+bb0_adj)/aa0_adj)

gen logor_res_adj=log((bb1)/(ab1+aa1))/( #ressesive
(bb0_adj)/(ab0_adj+aa0_adj))

replace logor_dom_adj=log((ab1+bb1+0.5)/aa1+0.5)/ #αντικατάσταση
((ab0_adj+bb0_adj+0.5)/aa0_adj+0.5) if aa0_adj==0
|bb0_adj+ab0_adj==0|aa1==0 |bb1+ab1==0

replace logor_res_adj=log((bb1+0.5)/(ab1+aa1+0.5))/(
(bb0_adj+0.5)/(ab0_adj+aa0_adj+0.5)) if ab0_adj+aa0_adj==0
|bb0_adj==0|ab1+aa1==0 |bb1==0

```

```

gen var_dom_adj=1/aa1 +1/(ab1+bb1) + #variances adjustments
4/((1pb0)*(1+pb0)*(1+pb0)*(2*bb0+ab0))

gen var_res_adj=1/bb1 +1/(ab1+aa1) + 4/((1-
pa0)*(1+pa0)*(1+pa0)*(2*aa0+ab0))

gen var_allele_adj=(1/(2*bb1+ab1)+1/(2*aa1+ab1)+ #allele_adj από τη
1/(2*bb0+ab0)+1/(2*aa0+ab0))*(1+falli) μέθοδο του Ιωαννίδη

replace var_dom_adj=1/(aa1+0.5) +1/(ab1+bb1+0.5) + 4/((1-pb0-
0.000001)*(1+pb0+0.000001)*(1+pb0+0.000001)*(2*bb0+ab0+0.5)) if
aa1==0 |ab1+bb1==0 |pb0==0 |pb0==0 |2*bb0+ab0==0 (11)

replace var_res_adj=1/(bb1+0.5) +1/(ab1+aa1+0.5) + 4/((1-pa0-
0.000001)*(1+pa0+0.000001)*(1+pa0+0.000001)*(2*aa0+ab0+0.5)) if
bb1==0 |ab1+aa1==0 |pa0==1 |pa0==0 |2*aa0+ab0==0 (12)

replace var_allele_adj=(1/(2*bb1+ab1+0.5)+1/(2*aa1+ab1+0.5)+
1/(2*bb0+ab0+0.5)+1/(2*aa0+ab0+0.5)) if falli ==. |ab0+2*aa0==0
|2*bb0+ab0==0|ab1+2*aa1==0 |2*bb1+ab1==0 (13)

gen b1=logor_dom #ορίζουμε μεταβλητές για όλα τα odds ratio
gen b2=logor_res και τα variances
gen b3=logor_allele
gen V11=var_dom
gen V22=var_res
gen V33=var_allele
gen b4=logor_dom_adj
gen b5=logor_res_adj
gen b6=logor_allele
gen b7=b6
gen V44=var_dom_adj
gen V55=var_res_adj
gen V66=var_allele_adj
gen V77=var_allele_adj (14)

gen st=_n
gen dis0=.
gen p_hwe0=.
sum st
local n=r(N)
sort st

forvalues i=1(1)`n'{ #δημιουργία πίνακα
    sort st
    local a= aa0[`i']
    local b= ab0[`i']
    local c= bb0[`i']
    qui genhwi `a' `b' `c'
    qui replace dis0=r(D) if st == `i'
    qui replace p_hwe0 = r(p_exact) if st== `i'
}

gen hwe=1 if p_hwe0>0.05 (15)
replace hwe=0 if hwe==.

gen b1overall=. #αρχικοποίηση κενών μεταβλητών
gen b2overall=.

```

```

gen b3overall=.
gen b4overall=.
gen b5overall=.
gen b6overall=.
gen b7overall=.
gen V1overall=.
gen V2overall=.
gen V3overall=.
gen V4overall=.
gen V5overall=.
gen V6overall=.
gen V7overall=.
gen a_overall=.
gen e_overall=.
gen b8overall=.
gen b9overall=.
gen b10overall=.
gen V8overall=.
gen V9overall=.
gen V10overall=.

```

(16)

```

local N 75
forvalues i=1(1)`N' {
    #μετά-ανάλυση

capture noisily mvmeta b V if meta== `i',vars(b1) mm (17)
    matrix c=e(b)
    matrix q=e(V)
    replace b1overall=c[1,1] if meta==`i' #χωρίς τροποποίηση
    replace V1overall=q[1,1] if meta==`i'

capture noisily mvmeta b V if meta== `i' & hwe==1,vars(b1) mm
    matrix c=e(b)
    matrix q=e(V)
    replace b8overall=c[1,1] if meta==`i' #αφαίρεση των μελετών
    replace V8overall=q[1,1] if meta==`i' που δεν είναι σε HWE

capture noisily mvmeta b V if meta== `i',vars(b2) mm
    matrix c=e(b)
    matrix q=e(V)
    replace b2overall=c[1,1] if meta==`i'
    replace V2overall=q[1,1] if meta==`i'

capture noisily mvmeta b V if meta== `i' & hwe==1,vars(b2) mm
    matrix c=e(b)
    matrix q=e(V)
    replace b9overall=c[1,1] if meta==`i'
    replace V9overall=q[1,1] if meta==`i'

capture noisily mvmeta b V if meta == `i',vars(b3) mm
    matrix c=e(b)
    matrix q=e(V)
    replace b3overall=c[1,1] if meta==`i'
    replace V3overall=q[1,1] if meta==`i'

capture noisily mvmeta b V if meta == `i' & hwe==1,vars(b3) mm
    matrix c=e(b)
    matrix q=e(V)
    replace b10overall=c[1,1] if meta==`i'

```



```

    replace V10overall=q[1,1] if meta=='i'

capture noisily mvmeta b V if meta=='i',vars(b4) mm
    matrix c=e(b) #με τροποποίηση
    matrix q=e(V)
    replace b4overall=c[1,1] if meta=='i'
    replace V4overall=q[1,1] if meta=='i'

capture noisily mvmeta b V if meta == `i',vars(b5) mm
    matrix c=e(b)
    matrix q=e(V)
    replace b5overall=c[1,1] if meta=='i'
    replace V5overall=q[1,1] if meta=='i'

capture noisily mvmeta b V if meta=='i',vars(b6) mm
    matrix c=e(b)
    matrix q=e(V)
    replace b6overall=c[1,1] if meta=='i'
    replace V6overall=q[1,1] if meta=='i'

capture noisily mvmeta b V if meta == `i',vars(b7) mm
    matrix c=e(b)
    matrix q=e(V)
    replace b7overall=c[1,1] if meta=='i'
    replace V7overall=q[1,1] if meta=='i'

capture noisily mvmeta a e if meta=='i', vars(a1) mm
    matrix c=e(b)
    matrix q=e(V)
    replace a_overall=c[1,1] if meta=='i'
    replace e_overall=q[1,1] if meta=='i'
}

qui bysort meta: gen last=_n==_N

qui bysort meta: egen total_hwe=mean(hwe)

keep if last==1 (18)

sum b1overall b4overall #αθροισμα
sum b1overall b8overall
sum b2overall b5overall
sum b2overall b9overall
sum b3overall b6overall
sum b3overall b7overall
sum b3overall b10overall

gen diffb1b4=b1overall-b4overall
gen diffb1b8=b1overall-b8overall
gen diffb2b5=b2overall-b5overall
gen diffb2b9=b2overall-b9overall
gen diffb3b6=b3overall-b6overall
gen diffb3b7=b3overall-b7overall (19)
gen diffb3b10=b3overall-b10overall
gen diffV1V4=V1overall-V4overall
gen diffV1V8=V1overall-V8overall
gen diffV2V5=V2overall-V5overall

```

```

gen diffV2V9=V2overall-V9overall
gen diffV3V6=V3overall-V6overall
gen diffV3V7=V3overall-V7overall
gen diffV3V10=V3overall-V10overall          (20)
gen z1=b1overall/sqrt(V1overall)
gen z2=b2overall/sqrt(V2overall)
gen z3=b3overall/sqrt(V3overall)
gen z4=b4overall/sqrt(V4overall)
gen z5=b5overall/sqrt(V5overall)
gen z6=b6overall/sqrt(V6overall)
gen z7=b7overall/sqrt(V7overall)
gen z8=b8overall/sqrt(V8overall)
gen z9=b9overall/sqrt(V9overall)          (21)
gen z10=b10overall/sqrt(V10overall)
gen diffz1z4=z1-z4
gen diffz1z8=z1-z8
gen diffz2z5=z2-z5
gen diffz2z9=z2-z9
gen diffz3z6=z3-z6
gen diffz3z7=z3-z7
gen diffz3z10=z3-z10          (22)

```

```

gr7 diffb1b4 total_hwe, xlab ylab          #γραφήματα
gr7 diffb1b8 total_hwe, xlab ylab
gr7 diffb2b5 total_hwe, xlab ylab
gr7 diffb2b9 total_hwe, xlab ylab
gr7 diffb3b6 total_hwe, xlab ylab
gr7 diffb3b7 total_hwe, xlab ylab          (23)
gr7 diffb3b10 total_hwe, xlab ylab

```

```

gr7 b1overall b4overall if b1overall> -50,xlab ylab
gr7 b1overall b8overall if b1overall> -50,xlab ylab
gr7 b2overall b5overall if b2overall> -50,xlab ylab
gr7 b2overall b9overall if b2overall> -50,xlab ylab
gr7 b3overall b6overall if b3overall> -50,xlab ylab
gr7 b3overall b7overall if b3overall> -50,xlab ylab
gr7 b3overall b10overall if b3overall> -50,xlab ylab          (24)

```

```

gr7 diffb1b4 a_overall, xlab ylab
gr7 diffb1b8 a_overall, xlab ylab
gr7 diffb2b5 a_overall, xlab ylab
gr7 diffb2b9 a_overall, xlab ylab
gr7 diffb3b6 a_overall, xlab ylab
gr7 diffb3b7 a_overall, xlab ylab
gr7 diffb3b10 a_overall, xlab ylab          (25)

```

```

gr7 diffV1V4 total_hwe, xlab ylab
gr7 diffV1V8 total_hwe, xlab ylab
gr7 diffV2V5 total_hwe, xlab ylab
gr7 diffV2V9 total_hwe, xlab ylab
gr7 diffV3V6 total_hwe, xlab ylab
gr7 diffV3V7 total_hwe, xlab ylab
gr7 diffV3V10 total_hwe, xlab ylab          (26)

```

```

gr7 V1overall V4overall if V1overall> -50,xlab ylab
gr7 V1overall V8overall if V1overall> -50,xlab ylab
gr7 V2overall V5overall if V2overall> -50,xlab ylab
gr7 V2overall V9overall if V2overall> -50,xlab ylab
gr7 V3overall V6overall if V3overall> -50,xlab ylab
gr7 V3overall V7overall if V3overall> -50,xlab ylab
gr7 V3overall V10overall if b3overall> -50,xlab ylab          (27)
gr7 diffV1V4 a_overall, xlab ylab

```

```

gr7 diffV1V8 a_overall, xlab ylab
gr7 diffV2V5 a_overall, xlab ylab
gr7 diffV2V9 a_overall, xlab ylab
gr7 diffV3V6 a_overall, xlab ylab
gr7 diffV3V7 a_overall, xlab ylab
gr7 diffV3V10 a_overall, xlab ylab (28)
gr7 diffz1z4 total_hwe, xlab ylab
gr7 diffz1z8 total_hwe, xlab ylab
gr7 diffz2z5 total_hwe, xlab ylab
gr7 diffz2z9 total_hwe, xlab ylab
gr7 diffz3z6 total_hwe, xlab ylab
gr7 diffz3z7 total_hwe, xlab ylab
gr7 diffz3z10 total_hwe, xlab ylab (29)
gr7 z1 z4, xlab ylab xline(-1.96,1.96) yline(-1.96,1.96)
gr7 z1 z8, xlab ylab xline(-1.96,1.96) yline(-1.96,1.96)
gr7 z2 z5, xlab ylab xline(-1.96,1.96) yline(-1.96,1.96)
gr7 z2 z9, xlab ylab xline(-1.96,1.96) yline(-1.96,1.96)
gr7 z3 z6, xlab ylab xline(-1.96,1.96) yline(-1.96,1.96)
gr7 z3 z7, xlab ylab xline(-1.96,1.96) yline(-1.96,1.96)
gr7 z3 z10, xlab ylab xline(-1.96,1.96) yline(-1.96,1.96) (30)
reg b1overall b4overall
reg b1overall b8overall
reg b2overall b5overall
reg b2overall b9overall
reg b3overall b6overall
reg b3overall b7overall
reg b3overall b10overall (31)
reg V1overall V4overall
reg V1overall V8overall
reg V2overall V5overall
reg V2overall V9overall
reg V3overall V6overall
reg V3overall V7overall
reg V3overall V10overall (32)
reg diffb1b4 total_hwe #παλινδρόμηση διαφορών odds-ratio
reg diffb1b8 total_hwe με total_hwe
reg diffb2b5 total_hwe
reg diffb2b9 total_hwe
reg diffb3b6 total_hwe
reg diffb3b7 total_hwe
reg diffb3b10 total_hwe (33)
reg diffV1V4 total_hwe # παλινδρόμηση διακυμάνσεων με x
reg diffV1V8 total_hwe με total_hwe
reg diffV2V5 total_hwe
reg diffV2V9 total_hwe
reg diffV3V6 total_hwe
reg diffV3V7 total_hwe
reg diffV3V10 total_hwe (34)
reg diffb1b4 a_overall #σε συνάρτηση του α
reg diffb1b8 a_overall
reg diffb2b5 a_overall
reg diffb2b9 a_overall
reg diffb3b6 a_overall
reg diffb3b7 a_overall
reg diffb3b10 a_overall (35)
reg diffV1V4 a_overall
reg diffV1V8 a_overall
reg diffV2V5 a_overall
reg diffV2V9 a_overall

```

```

reg diffV3V6 a_overall
reg diffV3V7 a_overall
reg diffV3V10 a_overall
reg diffz1z4 total_hwe
reg diffz1z8 total_hwe
reg diffz2z5 total_hwe
reg diffz2z9 total_hwe
reg diffz3z6 total_hwe
reg diffz3z7 total_hwe
reg diffz3z10 total_hwe
reg z1 z4
reg z1 z8
reg z2 z5
reg z2 z9
reg z3 z6
reg z3 z7
reg z3 z10

```

(36)

(37)

#γραφήματα συγκρίσεων μεταξύ των τροποποιημένων και αυτών που αφαιρούνται όσες μελέτες δεν είναι σε ισορροπία

```

gr7 b4overall b8overall if b4overall >-50, xlab ylab
gr7 b5overall b9overall if b5overall >-50, xlab ylab
gr7 b6overall b10overall if b6overall >-50, xlab ylab
gr7 b7overall b10overall if b7overall >-50, xlab ylab
reg b4overall b8overall
reg b5overall b9overall
reg b6overall b10overall
reg b7overall b10overall
gen diffb4b8=b4overall-b8overall
gen diffb5b9=b5overall-b9overall
gen diffb6b10=b6overall-b10overall
gen diffb7b10=b7overall-b10overall
gr7 diffb4b8 total_hwe, xlab ylab
gr7 diffb5b9 total_hwe, xlab ylab
gr7 diffb6b10 total_hwe, xlab ylab
gr7 diffb7b10 total_hwe, xlab ylab
reg diffb4b8 total_hwe
reg diffb5b9 total_hwe
reg diffb6b10 total_hwe
reg diffb7b10 total_hwe

gr7 diffb4b8 a_overall, xlab ylab
gr7 diffb5b9 a_overall, xlab ylab
gr7 diffb6b10 a_overall, xlab ylab
gr7 diffb7b10 a_overall, xlab ylab
reg diffb4b8 a_overall
reg diffb5b9 a_overall
reg diffb6b10 a_overall
reg diffb7b10 a_overall

gr7 V4overall V8overall if V4overall> -50,xlab ylab
gr7 V5overall V9overall if V5overall> -50,xlab ylab
gr7 V6overall V10overall if V6overall> -50,xlab ylab
gr7 V7overall V10overall if V7overall> -50,xlab ylab
reg V4overall V8overall
reg V5overall V9overall
reg V6overall V10overall
reg V7overall V10overall

```

```

gen diffV4V8=V4overall-V8overall
gen diffV5V9=V5overall-V9overall
gen diffV6V10=V6overall-V10overall
gen diffV7V10=V7overall-V10overall
gr7 diffV4V8 total_hwe, xlab ylab
gr7 diffV5V9 total_hwe, xlab ylab
gr7 diffV6V10 total_hwe, xlab ylab
gr7 diffV7V10 total_hwe, xlab ylab
reg diffV4V8 total_hwe
reg diffV5V9 total_hwe
reg diffV6V10 total_hwe
reg diffV7V10 total_hwe

gr7 diffV4V8 a_overall, xlab ylab
gr7 diffV5V9 a_overall, xlab ylab
gr7 diffV6V10 a_overall, xlab ylab
gr7 diffV7V10 a_overall, xlab ylab
reg diffV4V8 a_overall
reg diffV5V9 a_overall
reg diffV6V10 a_overall
reg diffV7V10 a_overall

gr7 z4 z8, xlab ylab xline(-1.96,1.96) yline(-1.96,1.96)
gr7 z5 z9, xlab ylab xline(-1.96,1.96) yline(-1.96,1.96)
gr7 z6 z10, xlab ylab xline(-1.96,1.96) yline(-1.96,1.96)
gr7 z7 z10, xlab ylab xline(-1.96,1.96) yline(-1.96,1.96)

reg z4 z8
reg z5 z9
reg z6 z10
reg z7 z10

gen diffz4z8=z4-z8
gen diffz5z9=z5-z9
gen diffz6z10=z6-z10
gen diffz7z10=z7-z10

gr7 diffz4z8 total_hwe, xlab ylab
gr7 diffz5z9 total_hwe, xlab ylab
gr7 diffz6z10 total_hwe, xlab ylab
gr7 diffz7z10 total_hwe, xlab ylab

reg diffz4z8 total_hwe
reg diffz5z9 total_hwe
reg diffz6z10 total_hwe
reg diffz7z10 total_hwe

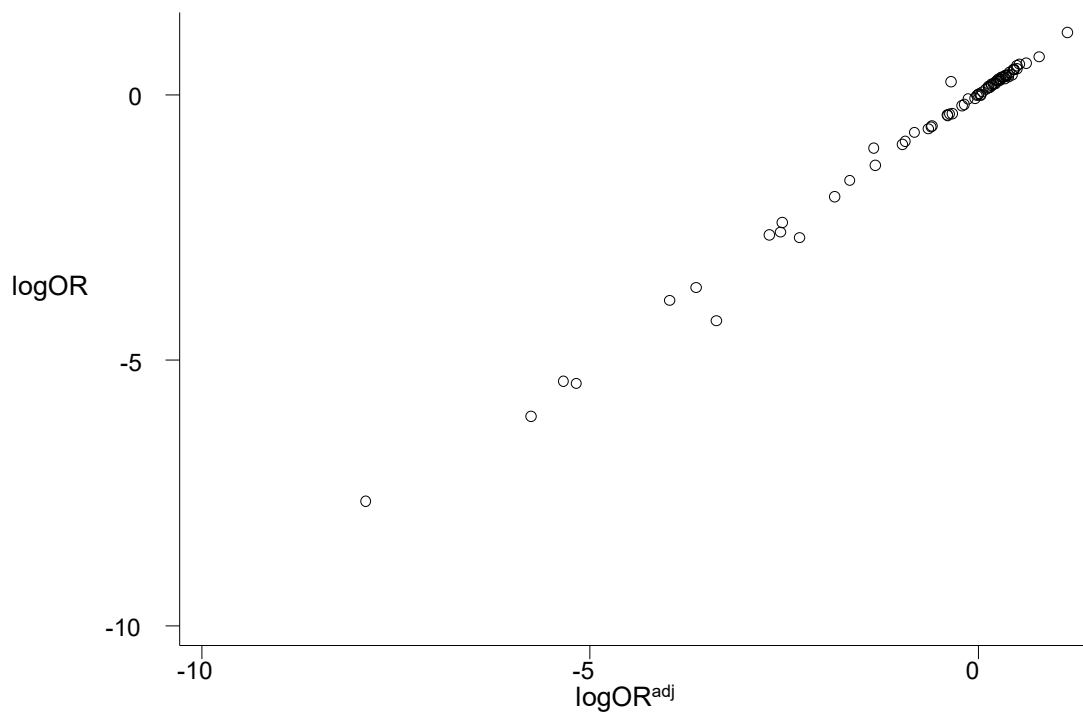
# εύρεση μελετών που αποκλίνουν
gen z1cat=1 if z1>1.96
replace z1cat=1 if z1<-1.96
gen z4cat=1 if z4<-1.96
replace z4cat=1 if z4>1.96
replace z1cat=0 if z1cat==.
replace z4cat=0 if z4cat==.
list meta if z1cat==0 & z4cat==1

```

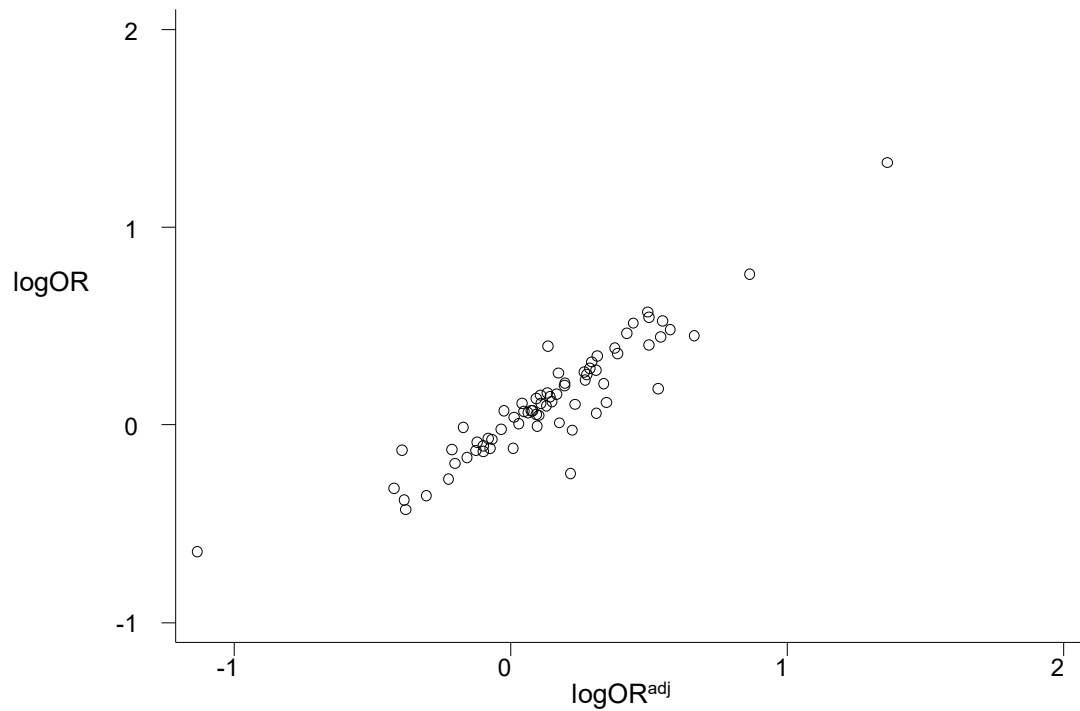
(38)

3.6. ΣΥΝΟΠΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

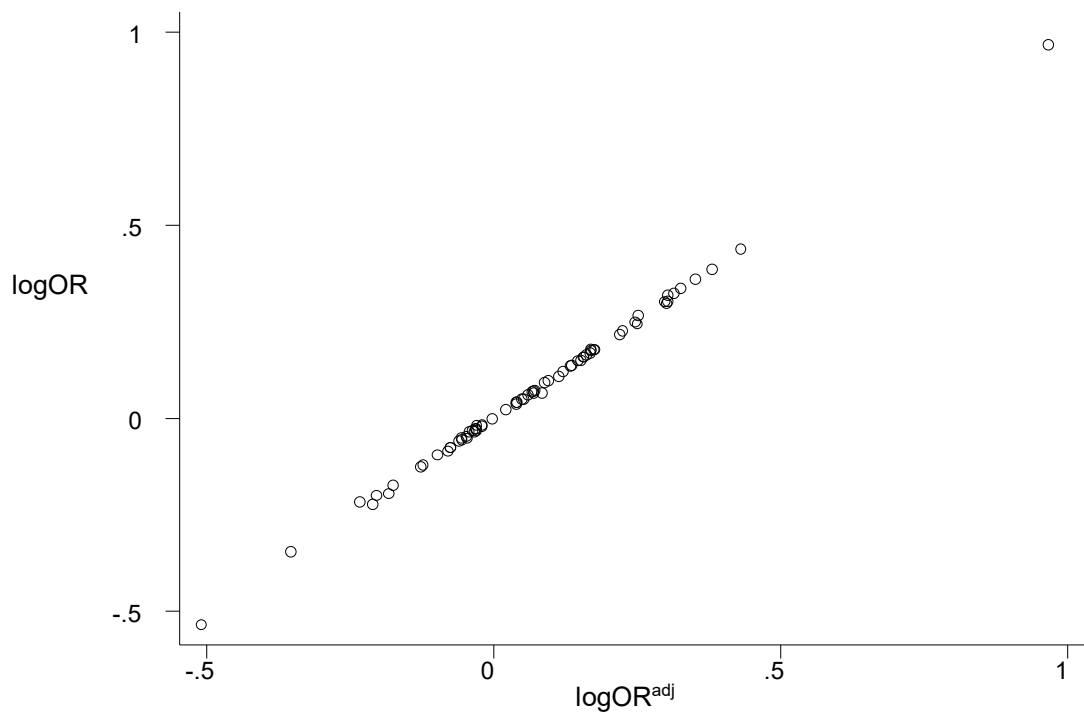
Αποτελέσματα της σύγκρισης μεταξύ των log odds ratio (σχέση 2), μεταξύ των variances (σχέση 3) και μεταξύ των συντελεστών z (σχέση 1) για τις μελέτες όπου έχουν υποστεί τροποποίηση, χωρίς τροποποίηση και στις μελέτες που έχουμε αφαιρέσει αυτές που δεν είναι σε ισορροπία HWE. Επίσης εξάγουμε αποτελέσματα καθώς, συγκρίνουμε την διαφορά των log odds ratio, την διαφορά των διακυμάνσεων (variances), την διαφορά των σταθερών συντελεστών z , σε συνάρτηση με το ποσοστό των μελετών που είναι σε ισορροπία (proportion of studies in HWE) και σε συνάρτηση του συντελεστή α . Μας ενδιαφέρει ο συντελεστής παλινδρόμησης β_1 , από την εξίσωση $y=\beta_0+\beta_1x$.



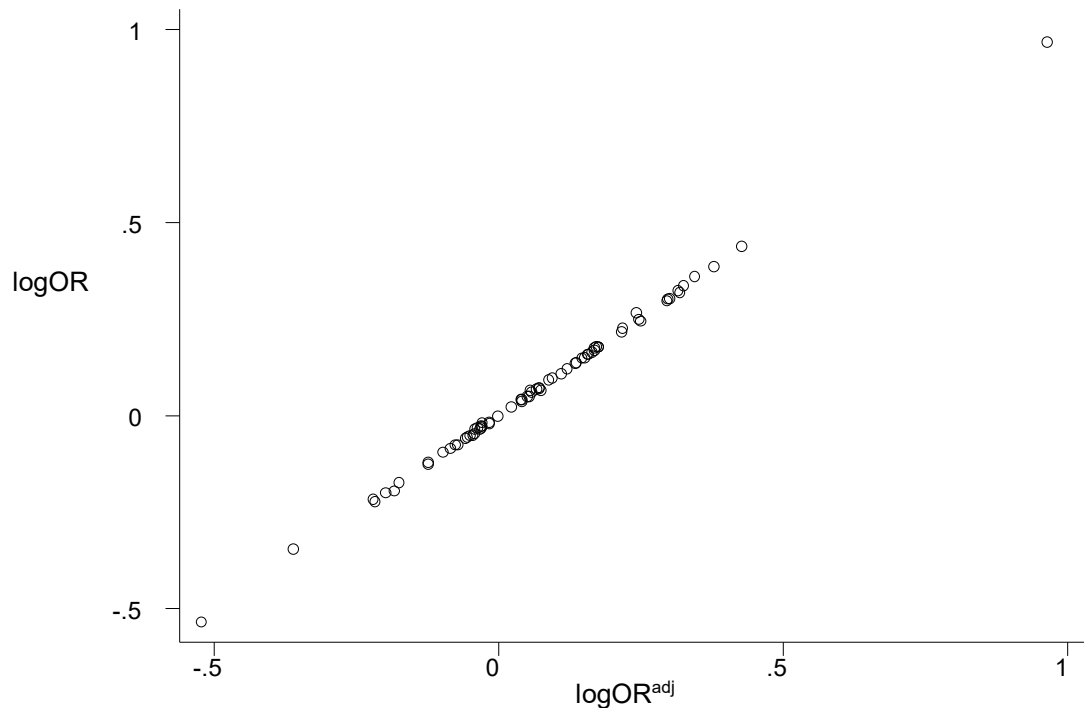
Εικόνα 15: Στον οριζόντιο άξονα είναι ο λόγος των πιθανοτήτων με τροποποίηση (σχέση 29) και στον κάθετο άξονα ο λόγος των πιθανοτήτων χωρίς τροποποίηση (σχέση 17) για το επικρατές μοντέλο κληρονομικότητας. Εδώ ο συντελεστής, β_1 της παλινδρόμησης ισούται με 1,025 και $p\text{-value}<10^{-4}$.



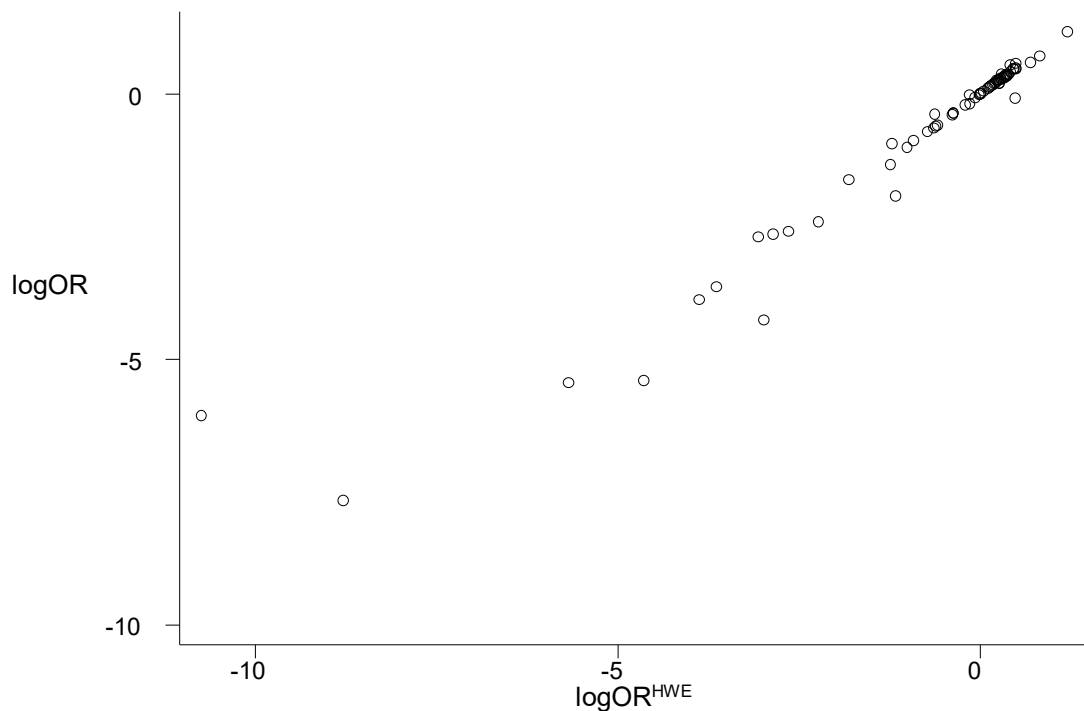
Εικόνα 16: Στον οριζόντιο άξονα είναι ο λόγος των πιθανοτήτων με τροποποίηση (σχέση 32) και στον κάθετο άξονα ο λόγος των πιθανοτήτων χωρίς τροποποίηση (σχέση 19), για το υπολειπόμενο μοντέλο κληρονομικότητας. Εδώ ο συντελεστής, β_1 της παλινδρόμησης ισούται με 0,73 και $p\text{-value} < 10^{-4}$



Εικόνα 17: Στον οριζόντιο άξονα είναι ο λόγος των πιθανοτήτων με τροποποίηση και στον κάθετο άξονα ο λόγος των πιθανοτήτων χωρίς τροποποίηση (σχέση 21), για το συνεπικρατές μοντέλο κληρονομικότητας, με τη μέθοδο του Sato.

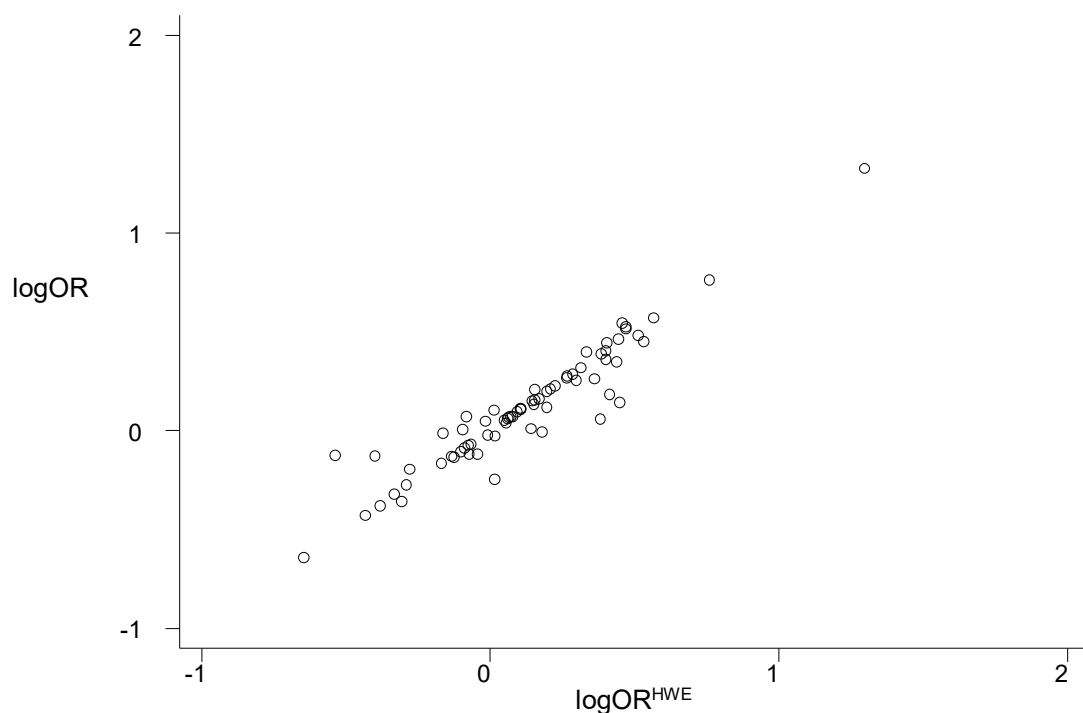


Εικόνα 18: Στον οριζόντιο άξονα είναι ο λόγος των πιθανοτήτων με τροποποίηση και στον κάθετο άξονα ο λόγος των πιθανοτήτων χωρίς τροποποίηση (σχέση 21), για το συνεπικρατές μοντέλο κληρονομικότητας, με τη μέθοδο του Ιωαννίδη. Εδώ ο συντελεστής, βI της παλινδρόμησης ισούται με 1.01.

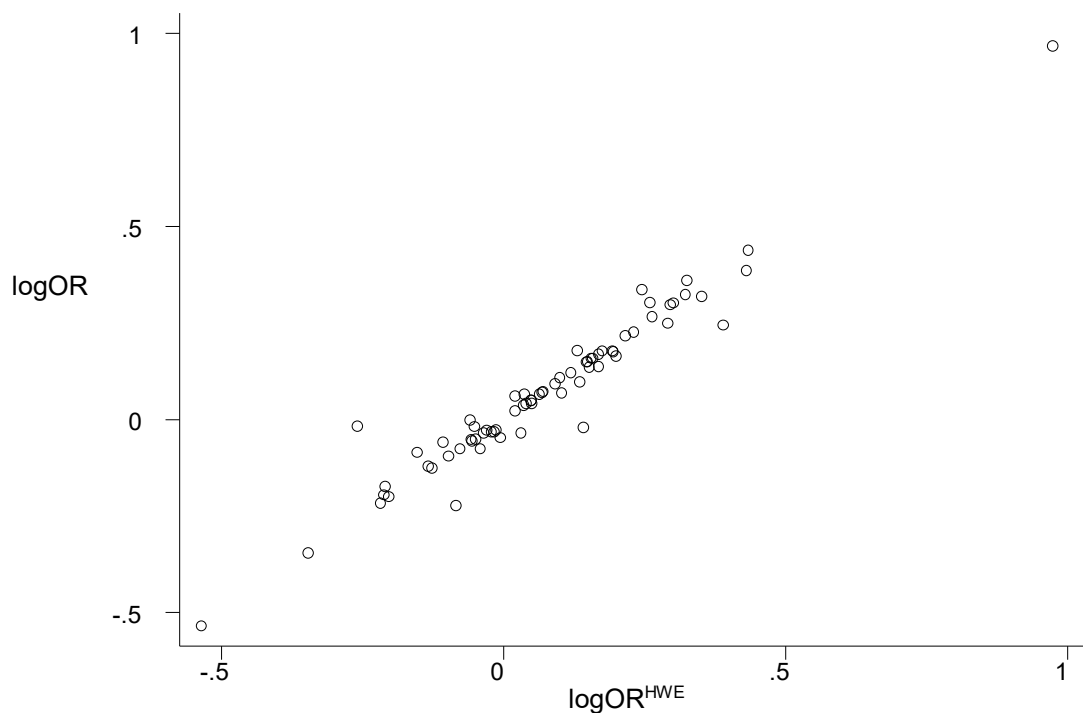


Εικόνα 19: Στον οριζόντιο άξονα είναι τα log odds ratio των μελετών των οποίων έχουν αφαιρεθεί αυτές που δεν είναι σε ισορροπία HWE και στον κάθετο άξονα τα log odds ratio χωρίς τροποποίηση

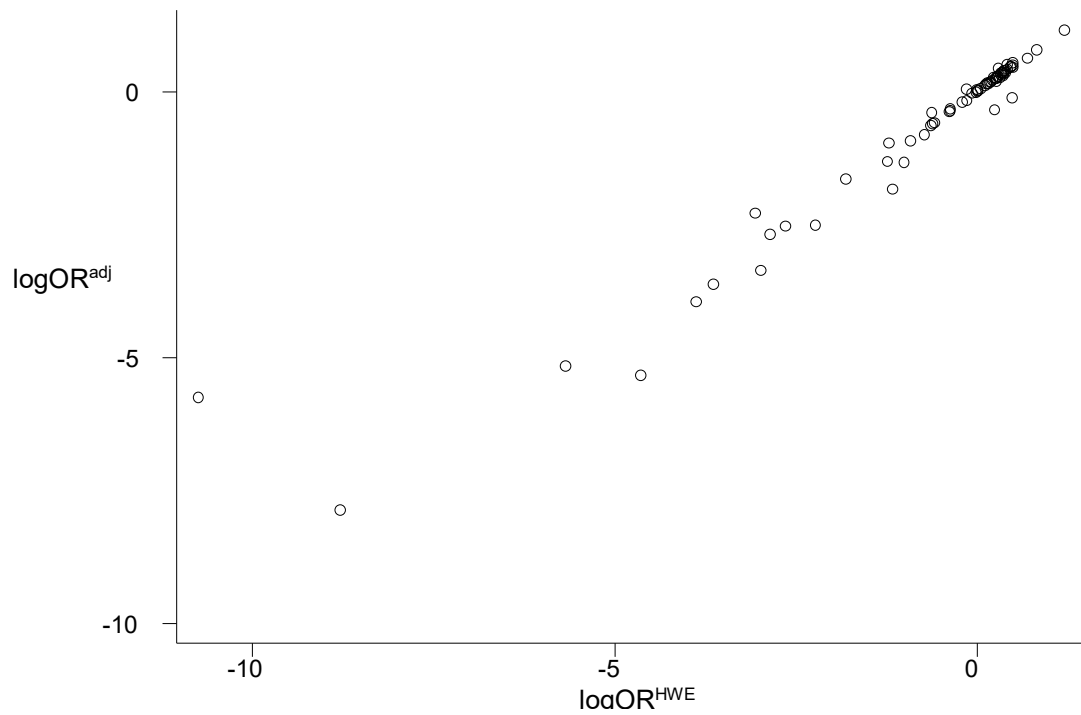
για το επικρατές πρότυπο κληρονομικότητας (σχέση 17). Εδώ ο συντελεστής, β_1 της παλινδρόμησης 0,99 και το $p\text{-value} < 10^{-4}$



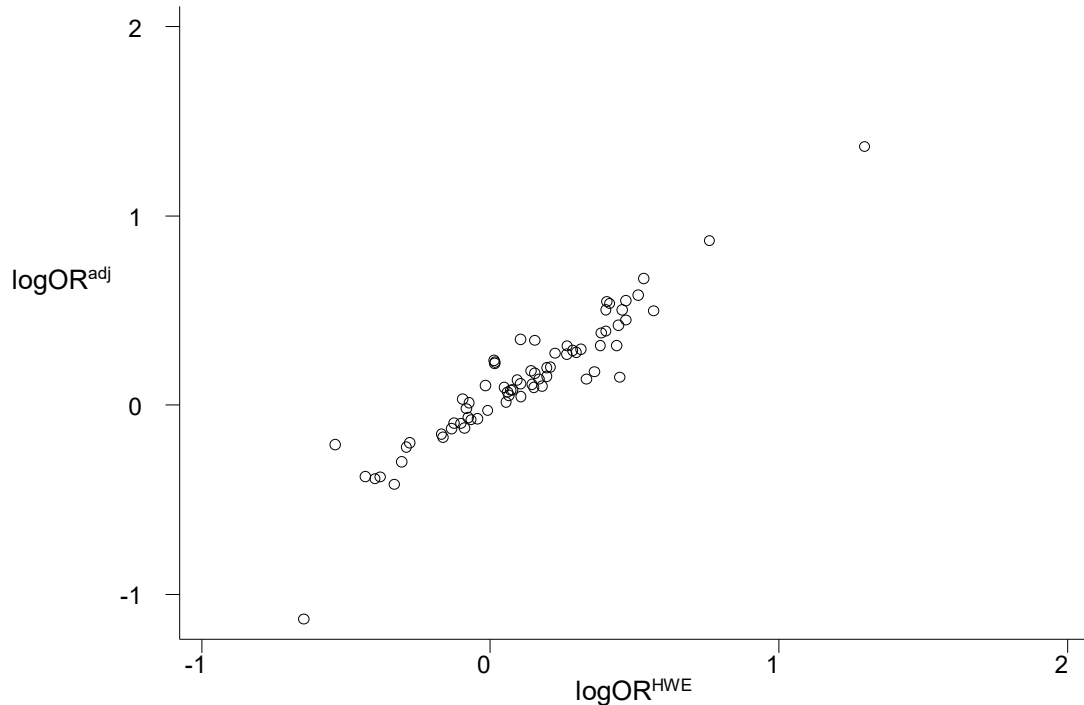
Εικόνα 20: Στον οριζόντιο άξονα είναι τα log odds ratio των μελετών των οποίων έχουν αφαιρεθεί αυτές που δεν είναι σε ισορροπία HW και στον κάθετο άξονα τα log odds ratio χωρίς τροποποίηση, για το υπολειπόμενο πρότυπο κληρονομικότητας (σχέση 19). Εδώ ο συντελεστής, β_1 της παλινδρόμησης είναι 0,99.



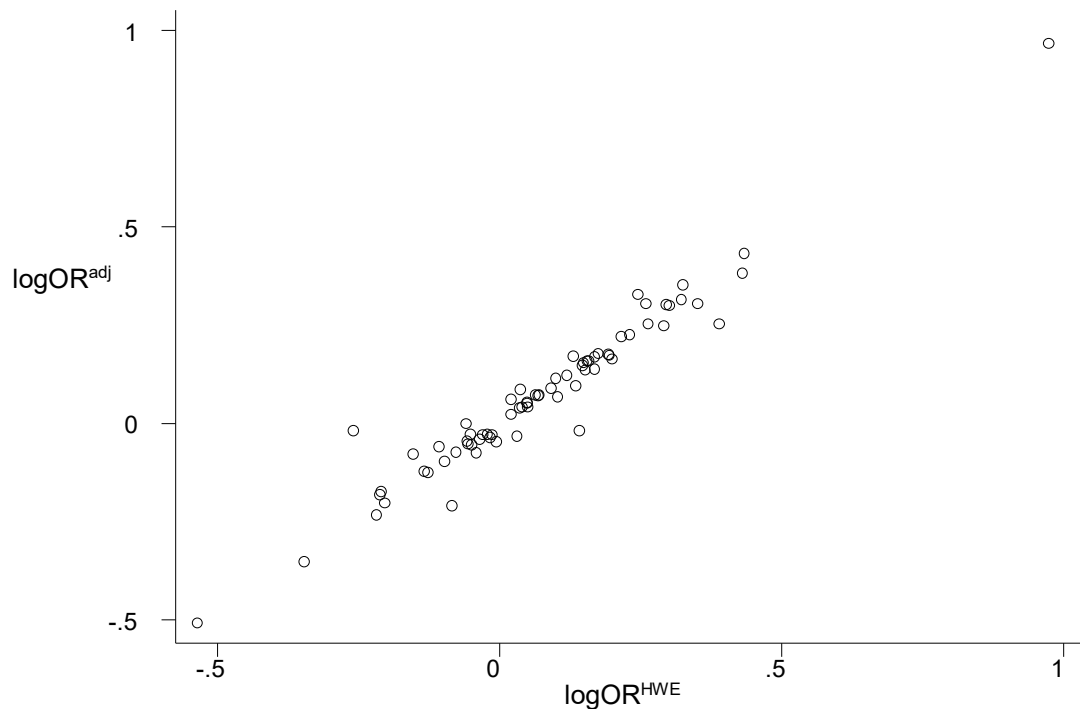
Εικόνα 21: Στον οριζόντιο άξονα είναι τα log odds ratio των μελετών των οποίων έχουν αφαιρεθεί αυτές που δεν είναι σε ισορροπία HW και στον κάθετο άξονα τα log odds ratio χωρίς τροποποίηση, για το συνεπικρατές πρότυπο κληρονομικότητας (σχέση 21). Εδώ ο συντελεστής, β_1 ισούται με 0.954.



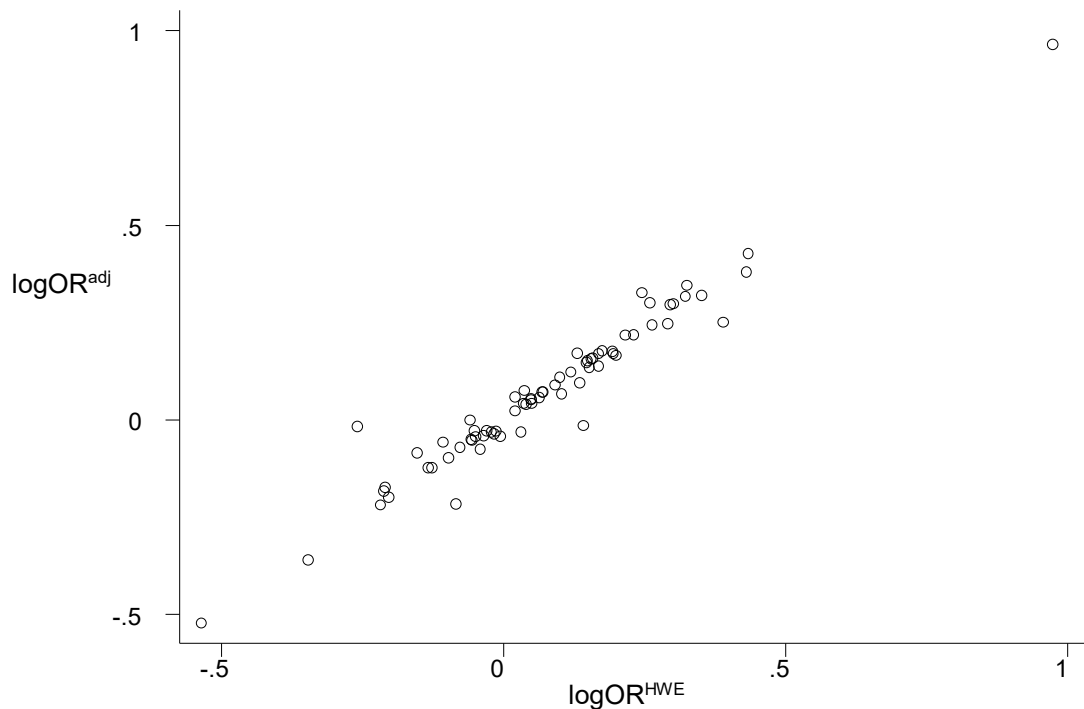
Εικόνα 22: Στον οριζόντιο άξονα είναι τα log odds ratio των μελετών των οποίων έχουν αφαιρεθεί αυτές που δεν είναι σε ισορροπία HW και στον κάθετο άξονα τα log odds ratio με τροποποίηση, για το επικρατές πρότυπο κληρονομικότητας (σχέση 29). Εδώ ο συντελεστής, $\beta 1$ ισούται με 0.9746



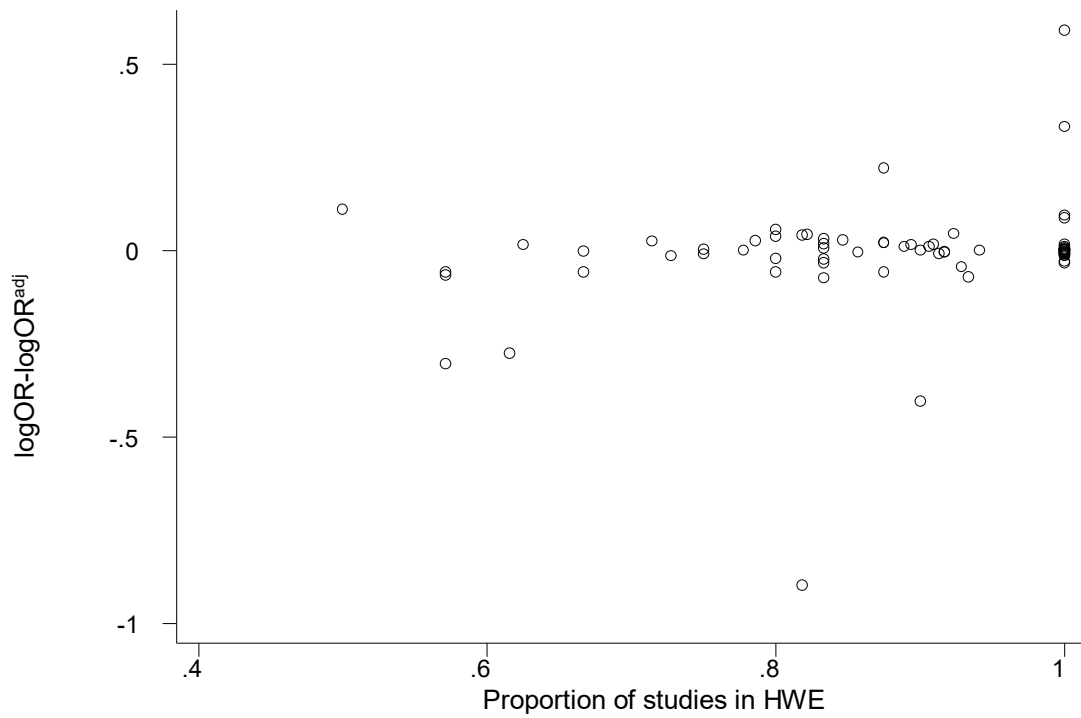
Εικόνα 23: Στον οριζόντιο άξονα είναι τα log odds ratio των μελετών των οποίων έχουν αφαιρεθεί αυτές που δεν είναι σε ισορροπία HW και στον κάθετο άξονα τα log odds ratio με τροποποίηση, για το υπολειπόμενο πρότυπο κληρονομικότητας (σχέση 32). Εδώ ο συντελεστής, $\beta 1$ ισούται με 0.9962



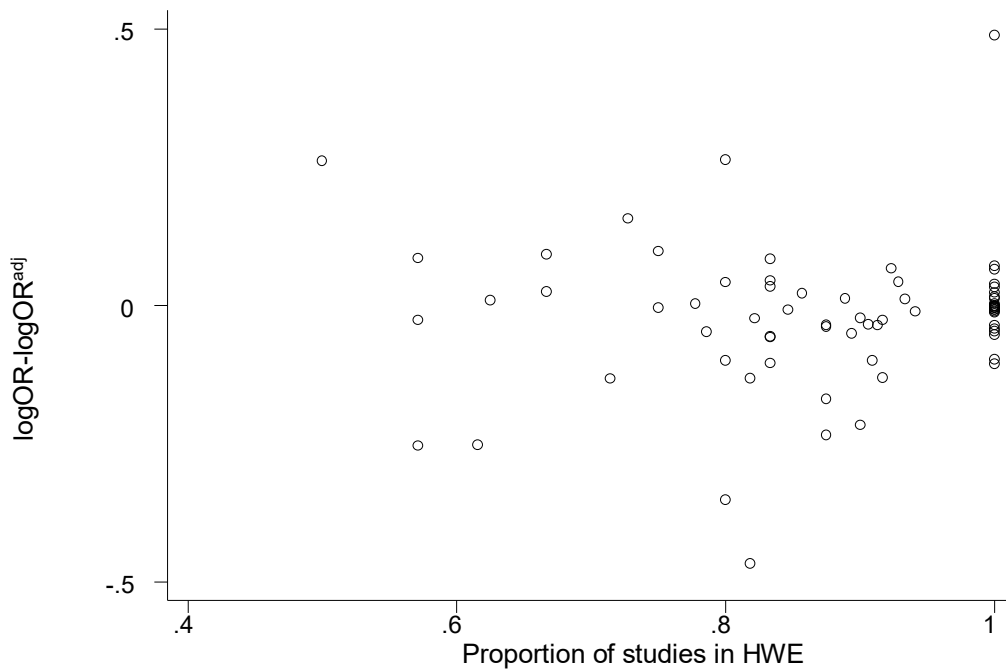
Εικόνα 24: Στον οριζόντιο άξονα είναι τα log odds ratio των μελετών των οποίων έχουν αφαιρεθεί αυτές που δεν είναι σε ισορροπία HW και στον κάθετο άξονα τα log odds ratio με τροποποίηση (σχέση 21) , για το συνεπικρατές πρότυπο κληρονομικότητας με τη μέθοδο του Sato.



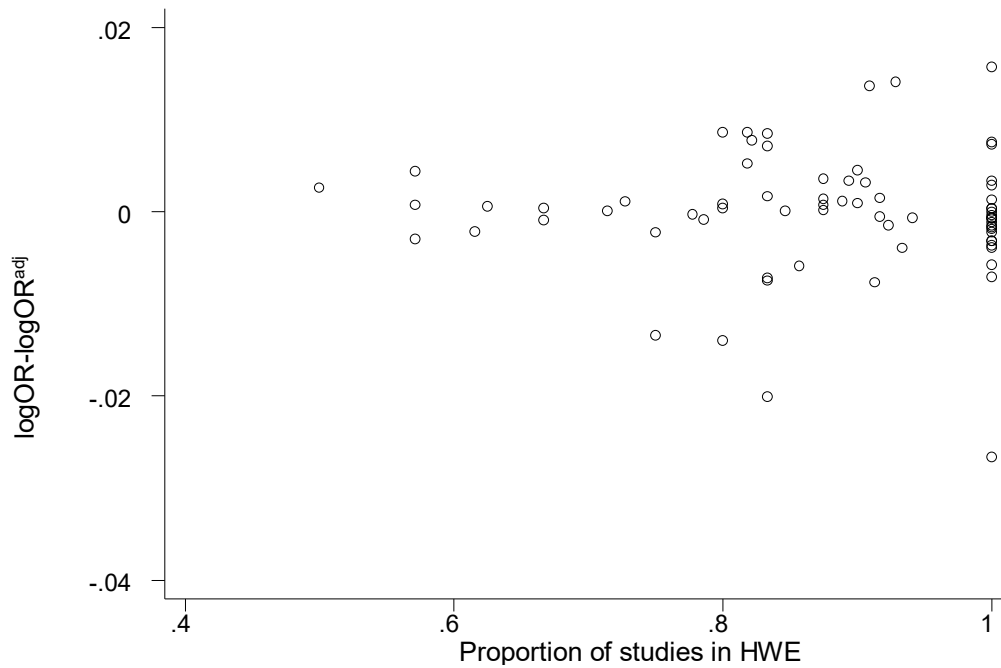
Εικόνα 25: Στον οριζόντιο άξονα είναι τα log odds ratio των μελετών των οποίων έχουν αφαιρεθεί αυτές που δεν είναι σε ισορροπία HW και στον κάθετο άξονα τα log odds ratio με τροποποίηση (σχέση 21), για το συνεπικρατές πρότυπο κληρονομικότητας, με τη μέθοδο του Ιωαννίδη. Εδώ ο συντελεστής, β_1 ισούται με 0.9403



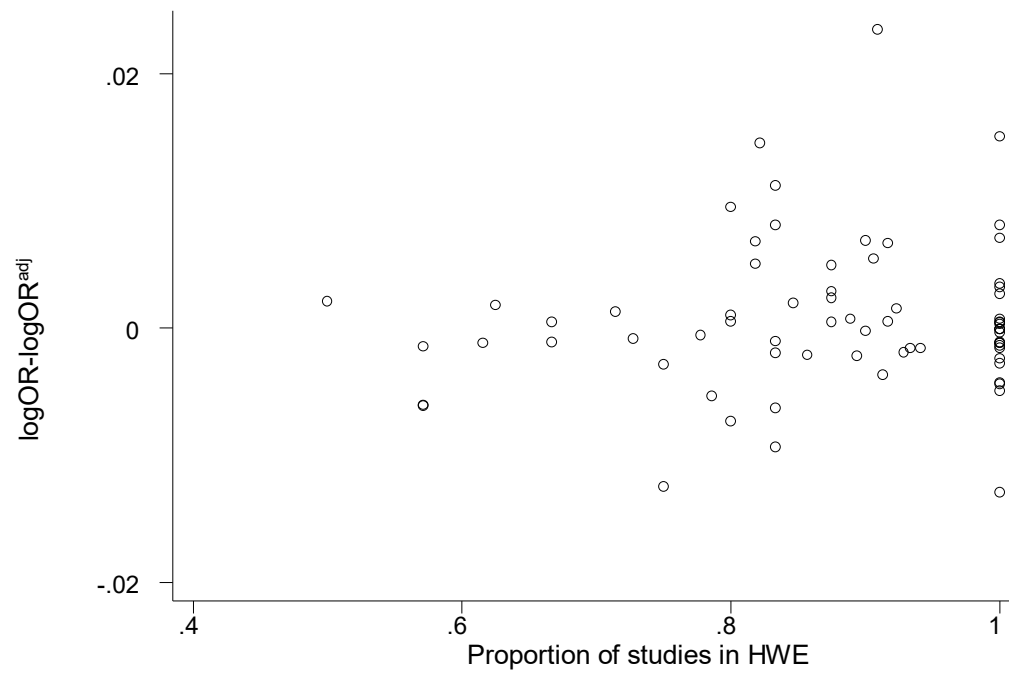
Εικόνα 26: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HWE και στον κάθετο άξονα η διαφορά των log odds ratio αυτών χωρίς τροποποίηση με αυτών που έχουν τροποποίηση (adjusted), για το επικρατές μοντέλο κληρονομικότητας (σχέση 17, 29).



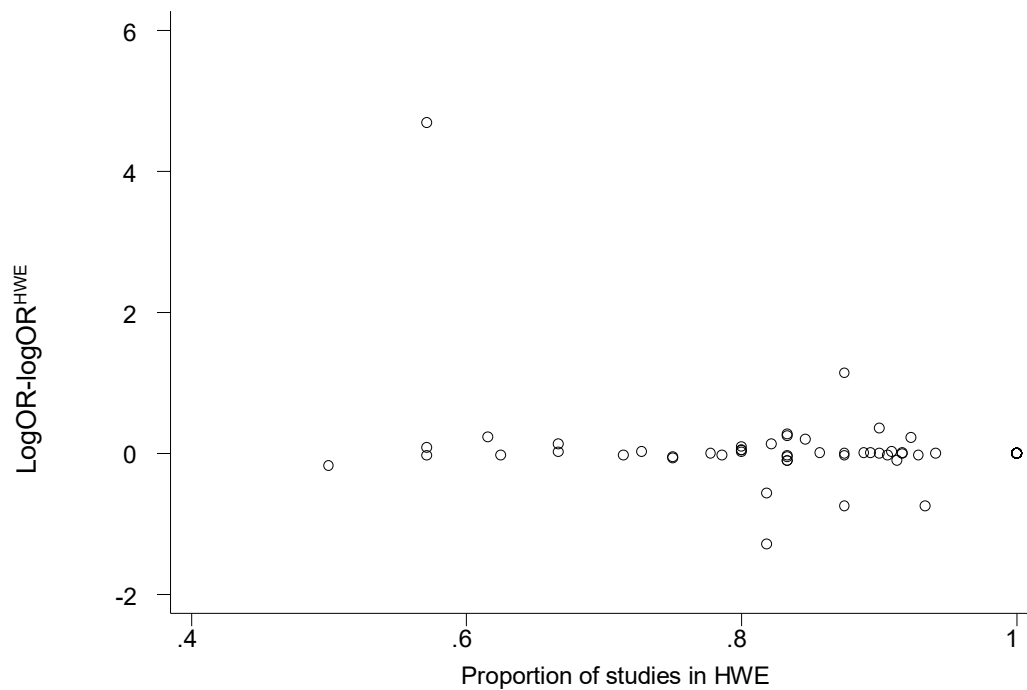
Εικόνα 27: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HW και στον κάθετο άξονα η διαφορά των log odds ratio αυτών χωρίς τροποποίηση με αυτών που έχουν τροποποίηση (adjusted), για το υπολειπόμενο μοντέλο κληρονομικότητας (σχέση 19, 32).



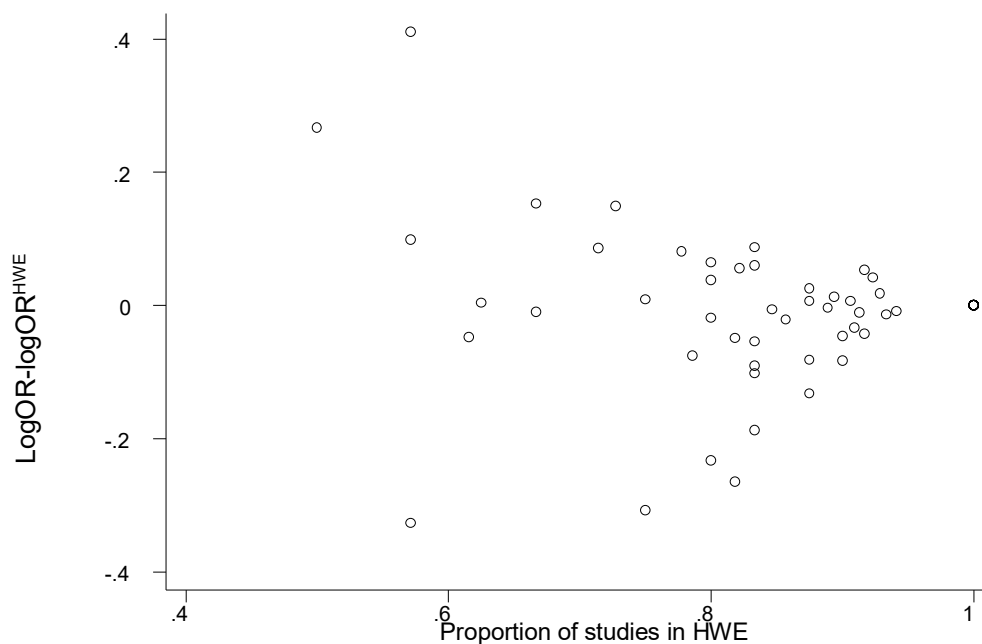
Εικόνα 28: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HW και στον κάθετο άξονα η διαφορά των log odds ratio αυτών χωρίς τροποποίηση με αυτών που έχουν τροποποίηση (adjusted), για το συνεπικρατές μοντέλο κληρονομικότητας με τη μέθοδο του Sato (σχέση 21, 34).



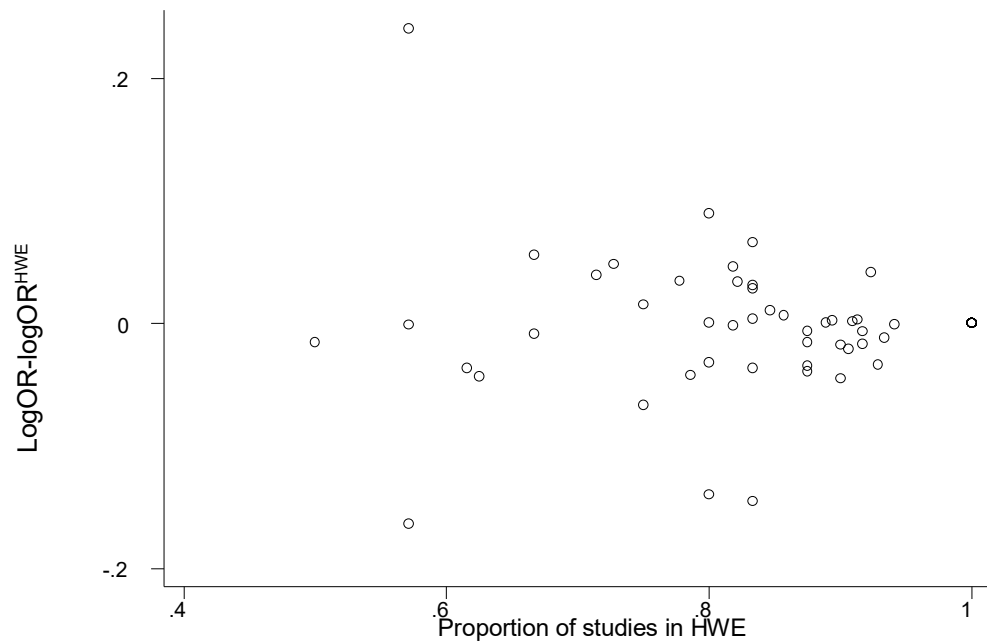
Εικόνα 29: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HW και στον κάθετο άξονα η διαφορά των log odds ratio αυτών χωρίς τροποποίηση με αυτών που έχουν τροποποίηση (adjusted), για το συνεπικρατές μοντέλο κληρονομικότητας με τη μέθοδο του Ιωαννίδη (σχέση 21,35). Εδώ ο συντελεστής παλινδρόμησης β_1 ισούται 0.0797.



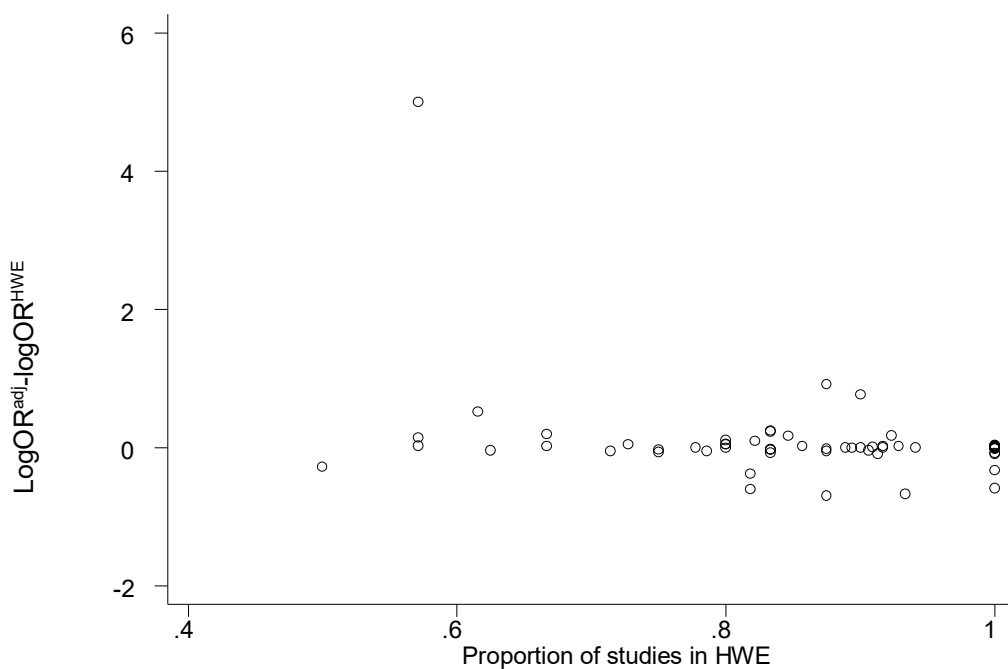
Εικόνα 30: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HW και στον κάθετο άξονα η διαφορά των log odds ratio, αυτών χωρίς τροποποίηση από τις μελέτες στις οποίες έχουν αφαιρεθεί αυτές που δεν είναι σε ισορροπία HW για το επικρατές μοντέλο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης είναι β_1 1.046338 και $p\text{-value} < 10^{-4}$.



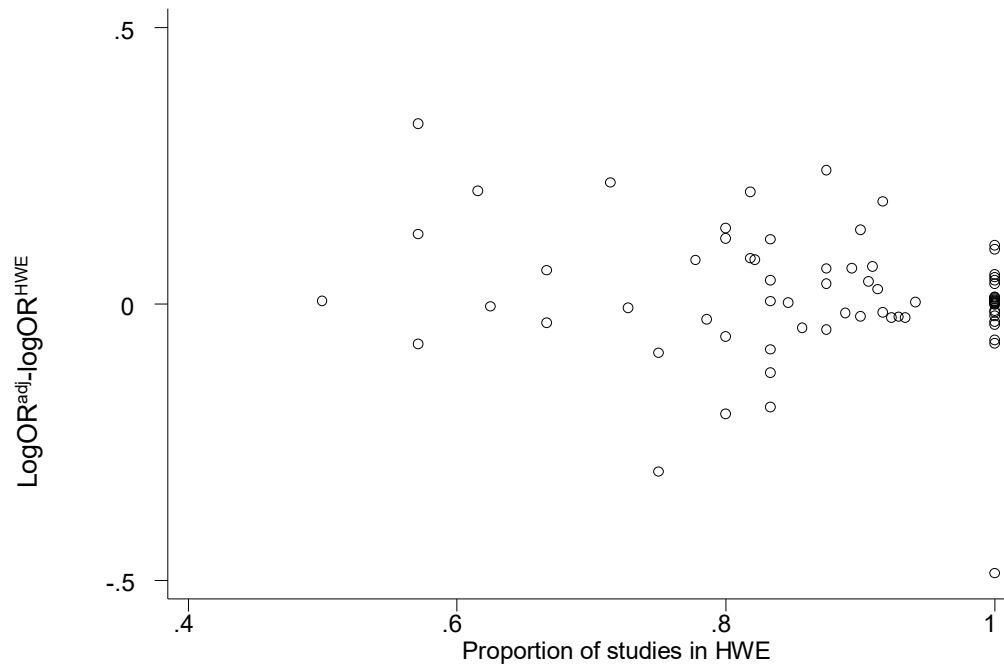
Εικόνα 31: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HW και στον κάθετο άξονα η διαφορά των log odds ratio, αυτών χωρίς τροποποίηση από τις μελέτες στις οποίες έχουν αφαιρεθεί αυτές που δεν είναι σε ισορροπία HW για το υπολειπόμενο μοντέλο κληρονομικότητας.



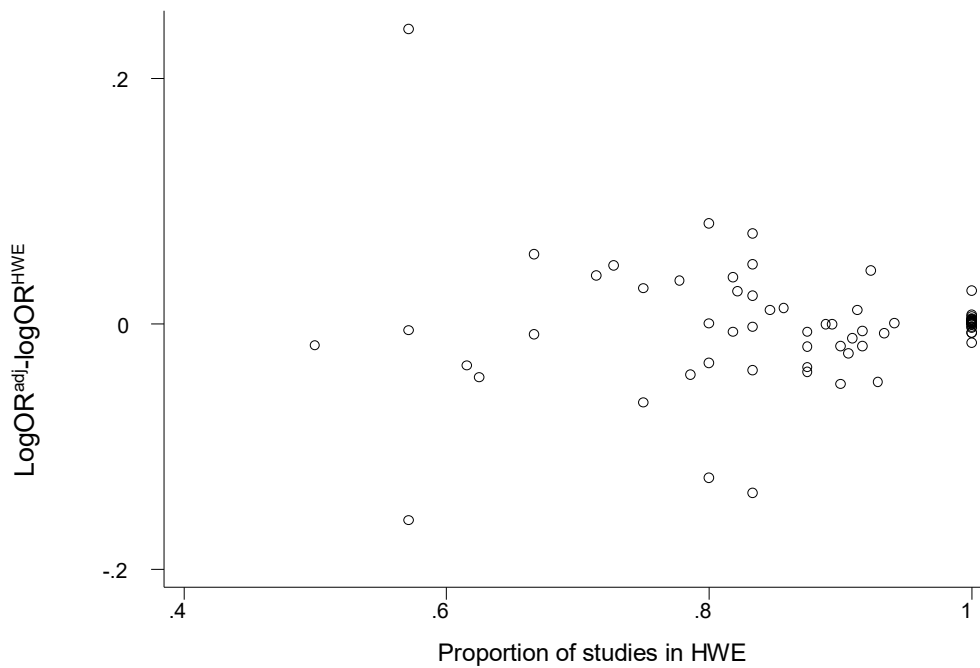
Εικόνα 32: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HW και στον κάθετο άξονα η διαφορά των log odds ratio, αυτών χωρίς τροποποίηση από τις μελέτες στις οποίες έχουν αφαιρεθεί αυτές που δεν είναι σε ισορροπία HW, για το συνεπικρατές μοντέλο κληρονομικότητας.



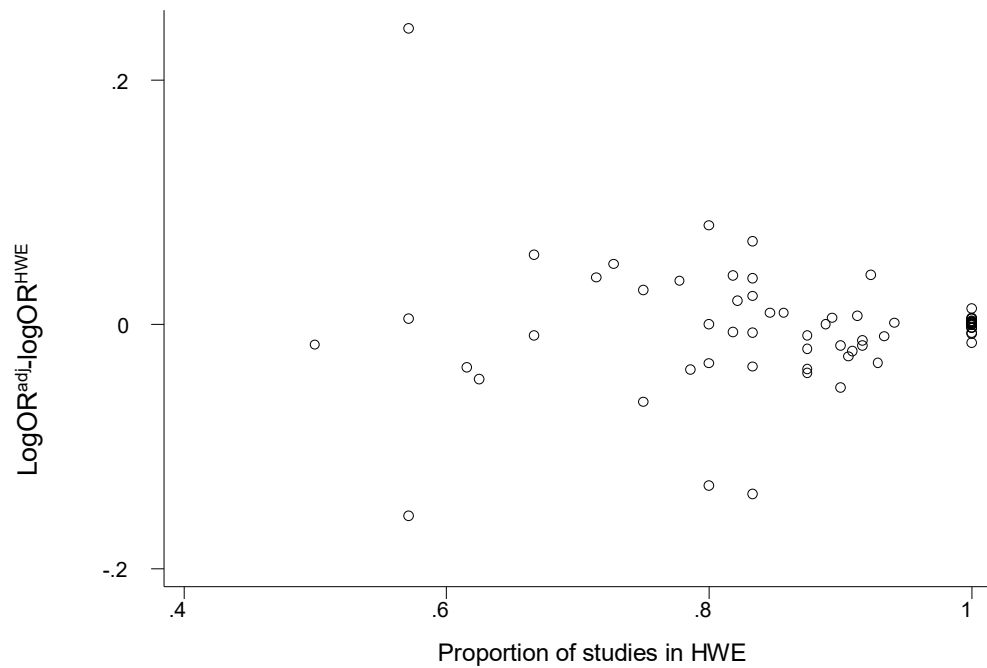
Εικόνα 33: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HW και στον κάθετο άξονα η διαφορά των log odds ratio, αυτών με τροποποίηση από τις μελέτες στις οποίες έχουν αφαιρεθεί όσες δεν είναι σε ισορροπία HW, για το επικρατές μοντέλο κληρονομικότητας.



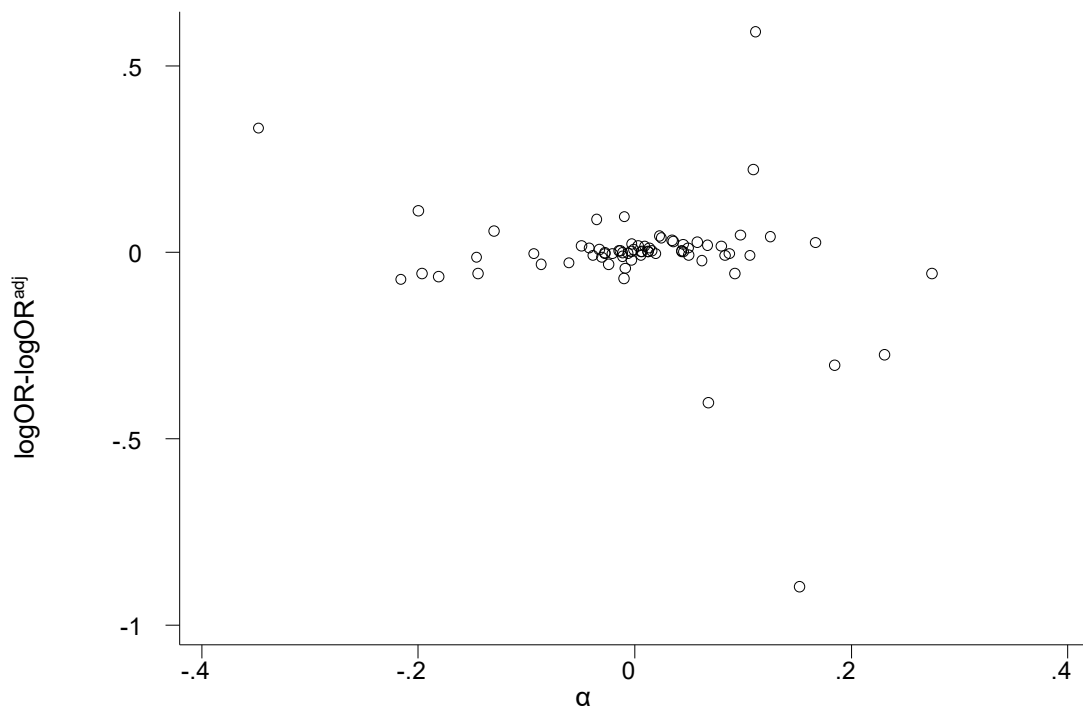
Εικόνα 34: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HW και στον κάθετο άξονα η διαφορά των log odds ratio, αυτών με τροποποίηση από τις μελέτες στις οποίες έχουν αφαιρεθεί όσες δεν είναι σε ισορροπία HW, για το υπολειπόμενο μοντέλο κληρονομικότητας.



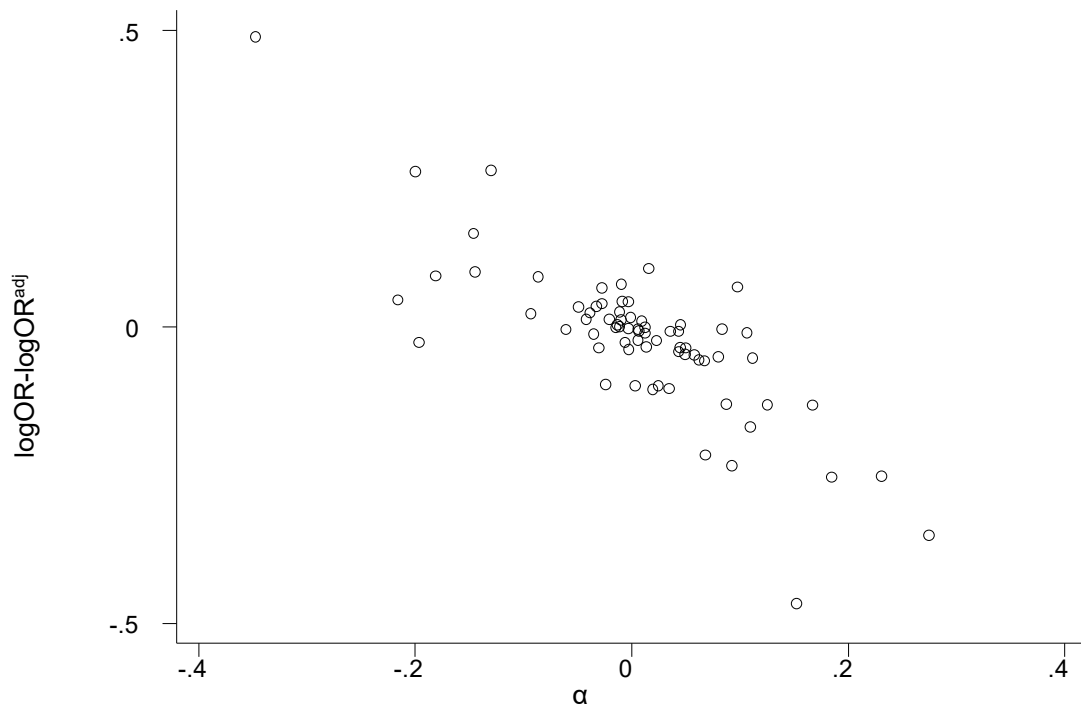
Εικόνα 35: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HW και στον κάθετο άξονα η διαφορά των log odds ratio, αυτών με τροποποίηση από τις μελέτες στις οποίες έχουν αφαιρεθεί όσες δεν είναι σε ισορροπία HW, για το συνεπικρατές μοντέλο κληρονομικότητας, με τη μέθοδο του Sato. Εδώ ο συντελεστής παλινδρόμησης β_1 είναι -0.0154



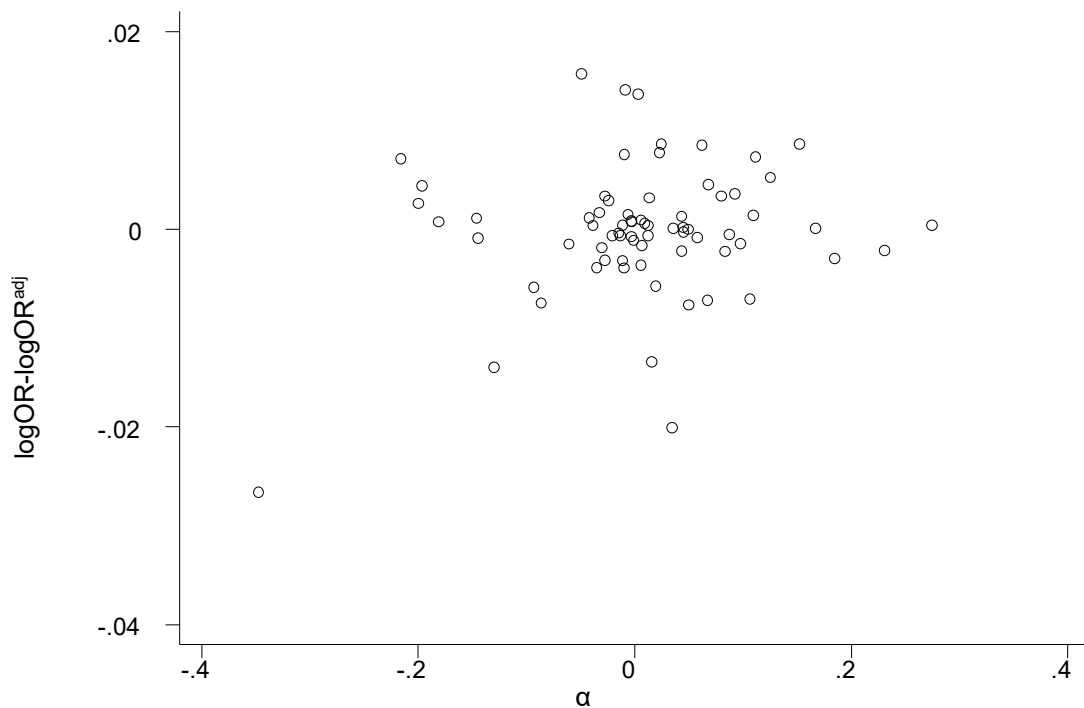
Εικόνα 36: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HW και στον κάθετο άξονα η διαφορά των log odds ratio, αυτών με τροποποίηση από τις μελέτες στις οποίες έχουν αφαιρεθεί όσες δεν είναι σε ισορροπία HW, για το συνεπικρατές μοντέλο κληρονομικότητας, με τη μέθοδο του Ιωαννίδη. Εδώ ο συντελεστής παλινδρόμησης β_1 είναι -0.02339



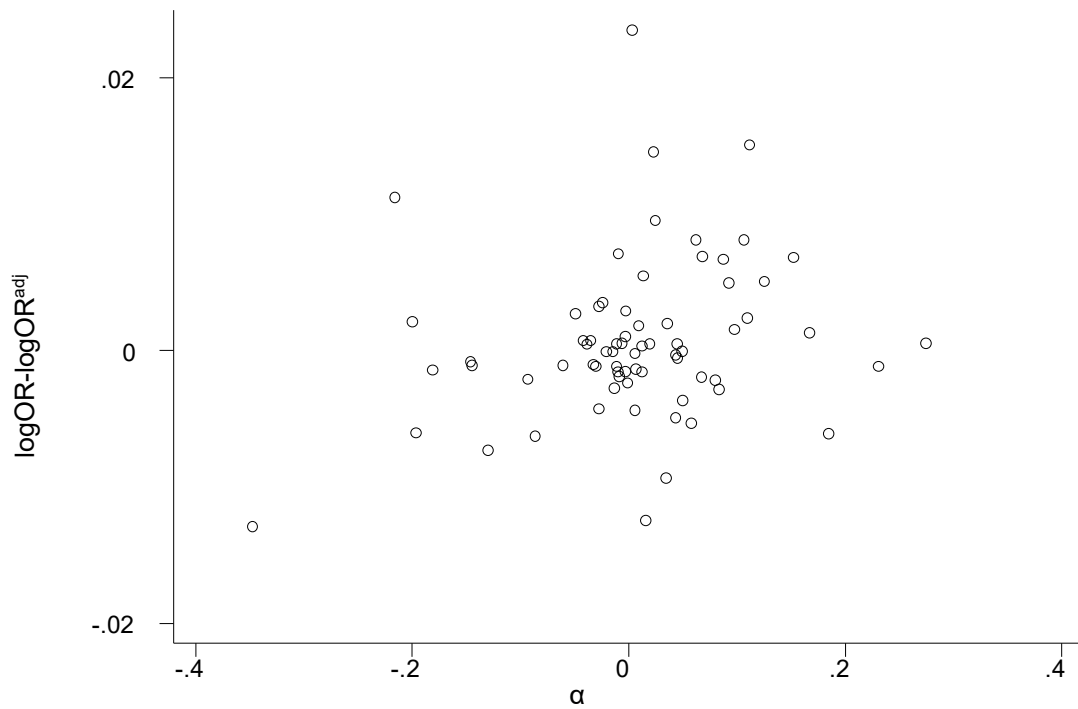
Εικόνα 37: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α (σχέση 12) και στον κάθετο άξονα η διαφορά των log odds ratio, χωρίς τροποποίηση από αυτές που έχουν τροποποιήσεις για το επικρατές πρότυπο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης β_1 είναι -1.085794 και $p\text{-value} < 10^{-4}$



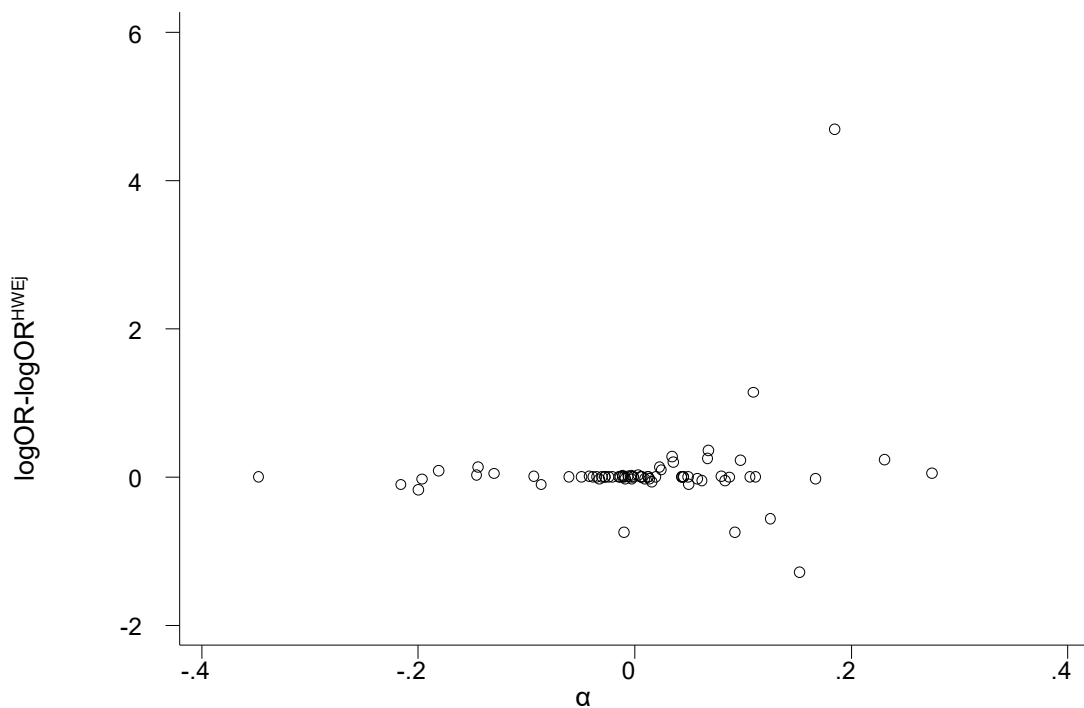
Εικόνα 38: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α (σχέση 27) και στον κάθετο άξονα η διαφορά των log odds ratio , χωρίς τροποποίηση από αυτές που έχουν τροποποιήσει, για το υπολειπόμενο πρότυπο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης είναι -0.77879 .



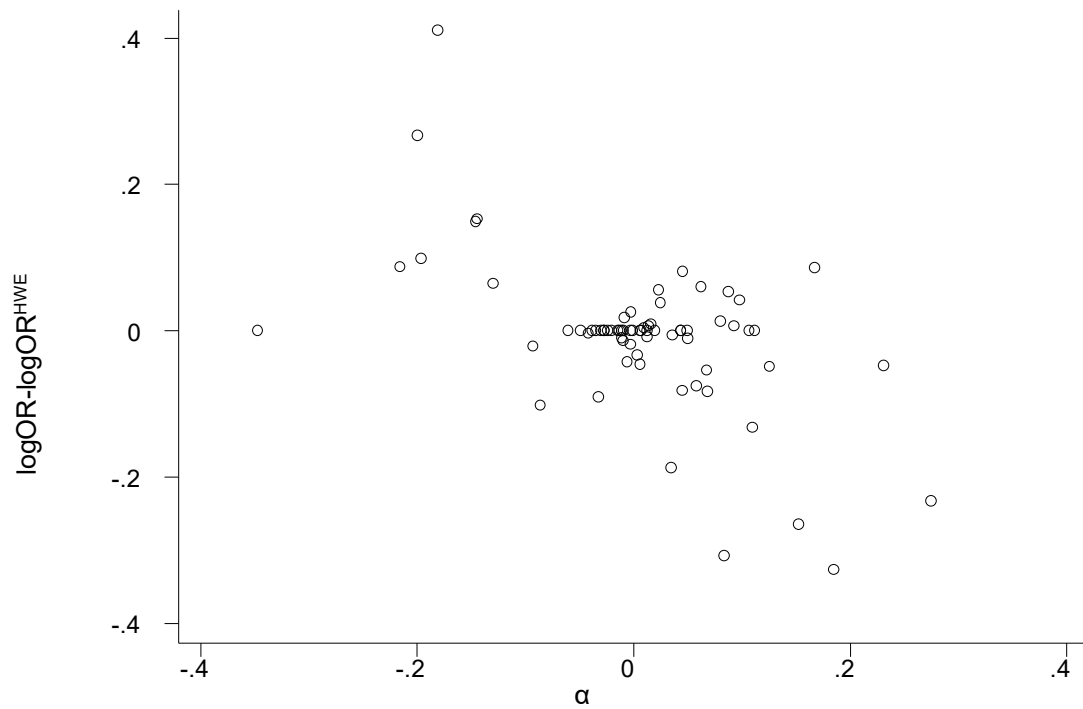
Εικόνα 39: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α (σχέση 27) και στον κάθετο άξονα η διαφορά των log odds ratio , χωρίς τροποποίηση από αυτές που έχουν τροποποιήσει, για το συνεπικρατές πρότυπο κληρονομικότητας, με τη μέθοδο του Sato. Εδώ ο συντελεστής παλινδρόμησης είναι 0.00476



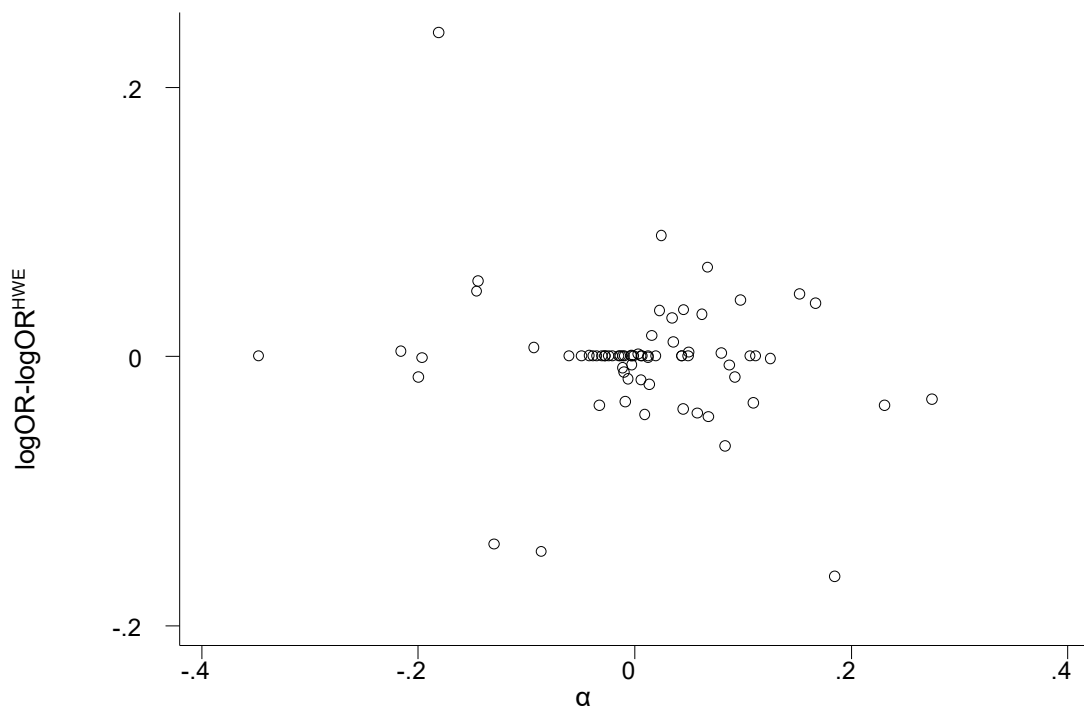
Εικόνα 40: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α (σχέση 27) και στον κάθετο άξονα η διαφορά των log odds ratio , χωρίς τροποποίηση από αυτές που έχουν τροποποιήσει, για το συνεπικρατές πρότυπο κληρονομικότητας, με τη μέθοδο του Ιωαννίδη. Εδώ ο συντελεστής παλινδρόμησης είναι 0.0093538.



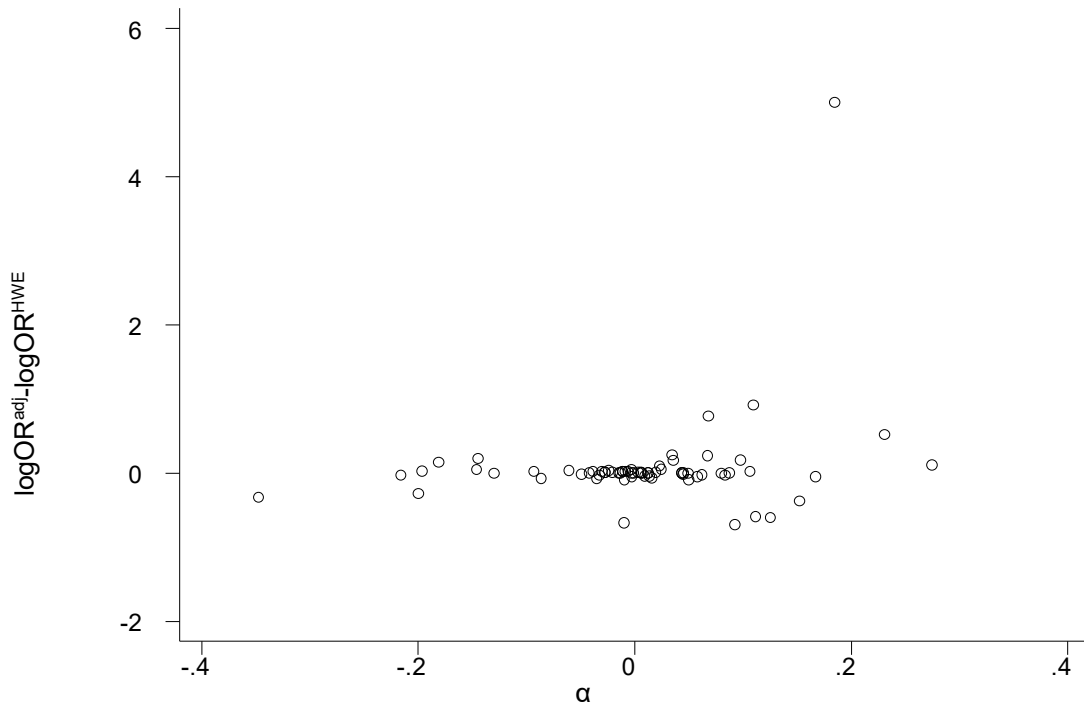
Εικόνα 41: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α (σχέση 27) και στον κάθετο άξονα η διαφορά των log odds ratio , χωρίς τροποποίηση από αυτές που έχουν αφαιρεθεί όσες δεν είναι σε HWE για το επικρατές πρότυπο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης είναι β_1 είναι 0.137916.



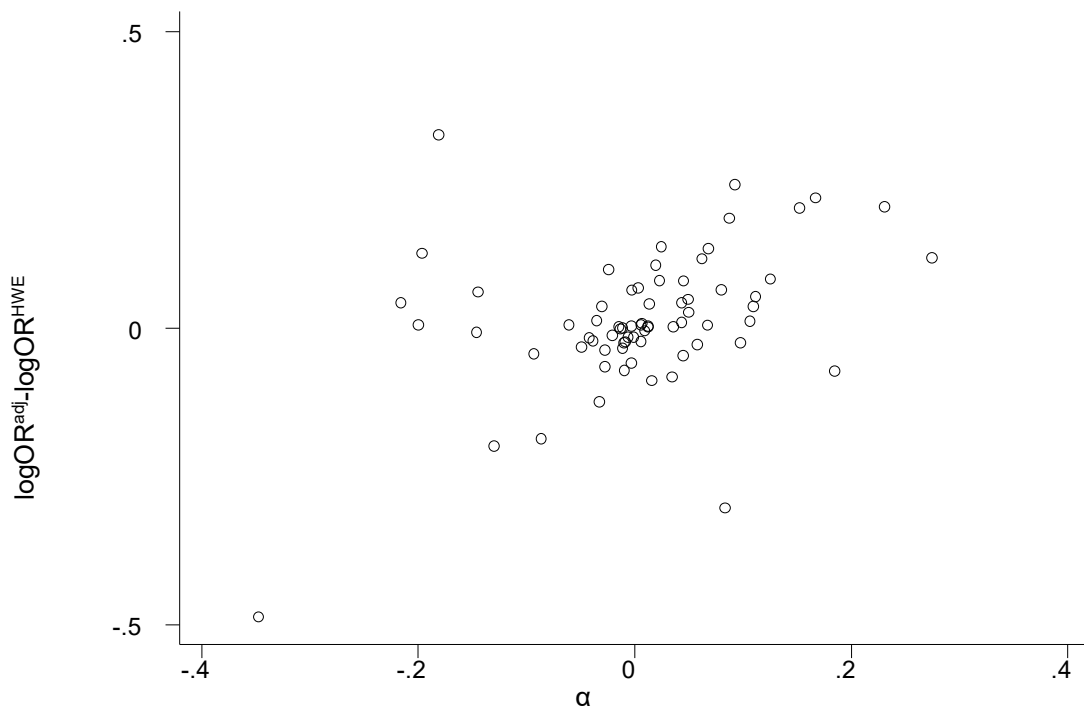
Εικόνα 42: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α (σχέση 27) και στον κάθετο άξονα η διαφορά των log odds ratio , χωρίς τροποποίηση από αυτές που έχουν αφαιρεθεί όσες δεν είναι σε HWE, για το υπολειπόμενο πρότυπο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης β_1 είναι - 0.111761



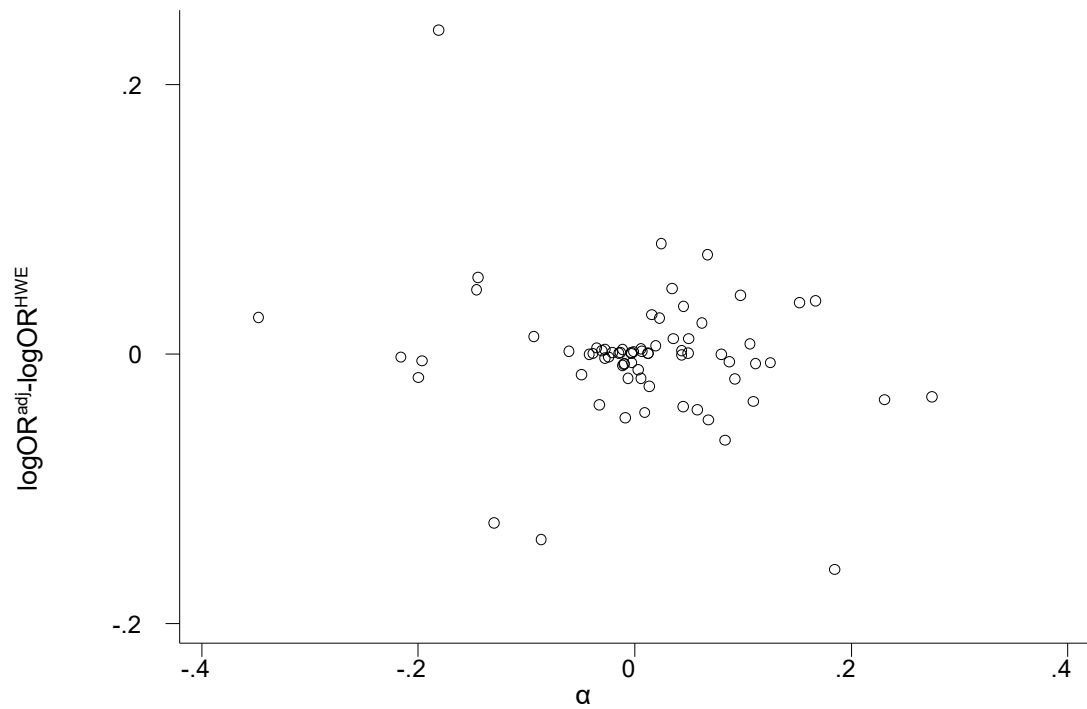
Εικόνα 43: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α (σχέση 27) και στον κάθετο άξονα η διαφορά των log odds ratio , χωρίς τροποποίηση από αυτές που έχουν αφαιρεθεί όσες δεν είναι σε HWE, για το συνεπικρατές πρότυπο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης β_1 είναι - 0.0166189 και p-value=0.23.



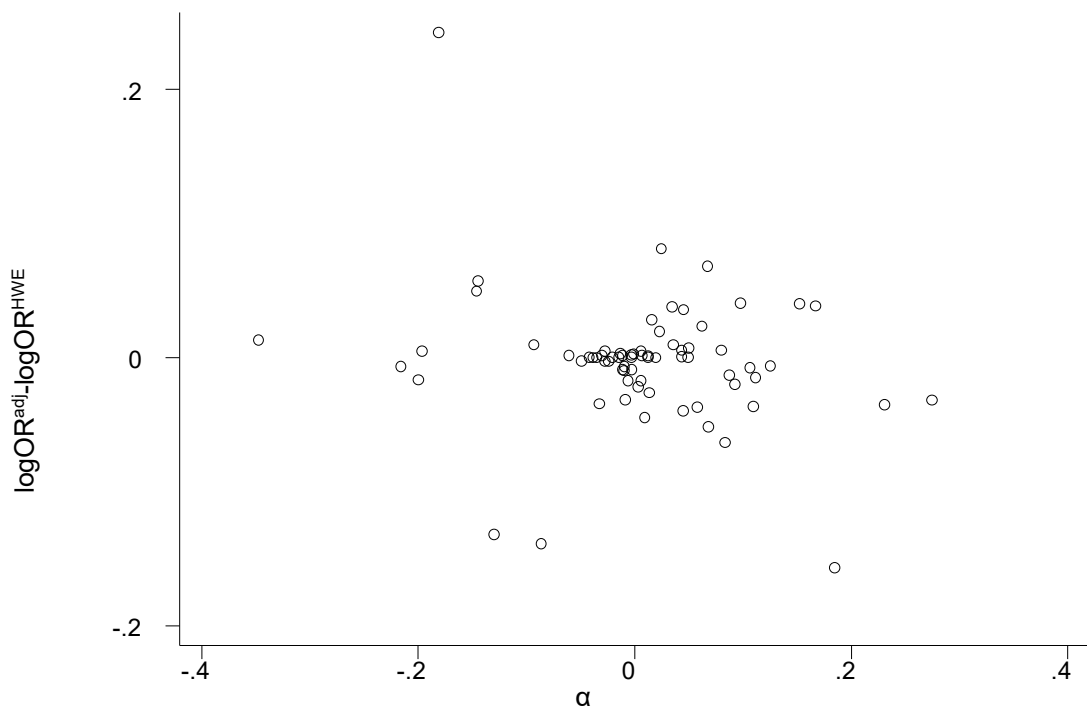
Εικόνα 44: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α (σχέση 27) και στον κάθετο άξονα η διαφορά των log odds ratio , με τροποποίηση από αυτές που έχουν αφαιρεθεί όσες δεν είναι σε HWE, για το επικρατές πρότυπο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης β_1 είναι 1.9958



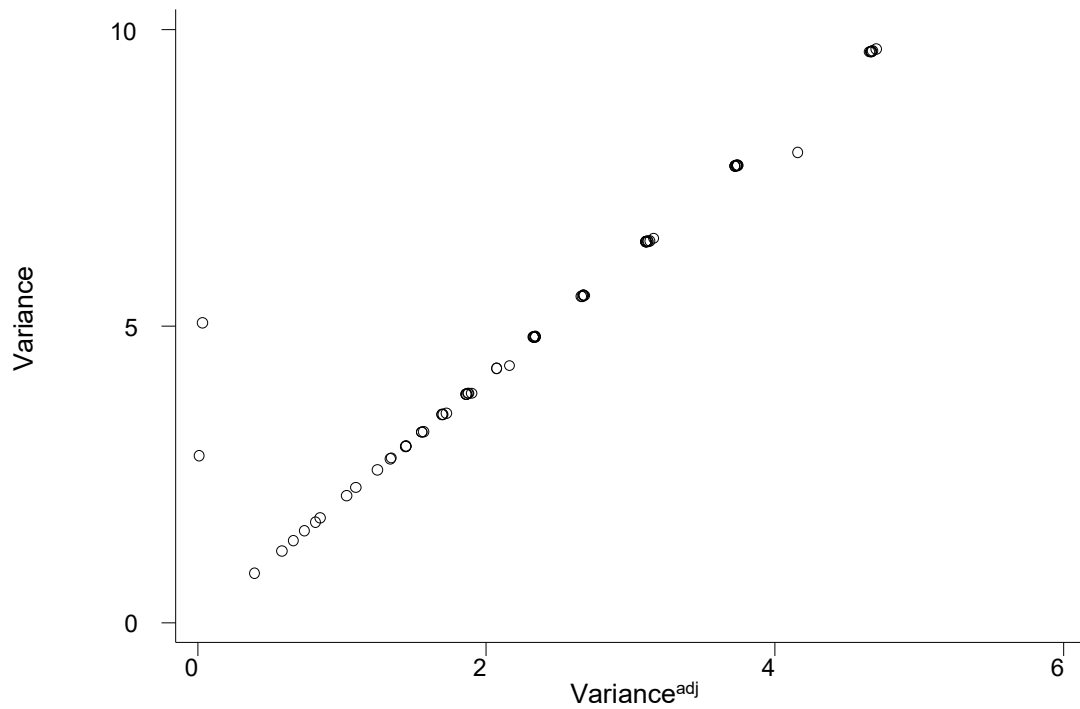
Εικόνα 45: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α (σχέση 27) και στον κάθετο άξονα η διαφορά των log odds ratio , με τροποποίηση από αυτές που έχουν αφαιρεθεί όσες δεν είναι σε HWE, για το υπολειπόμενο πρότυπο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης β_1 είναι 0.6670224



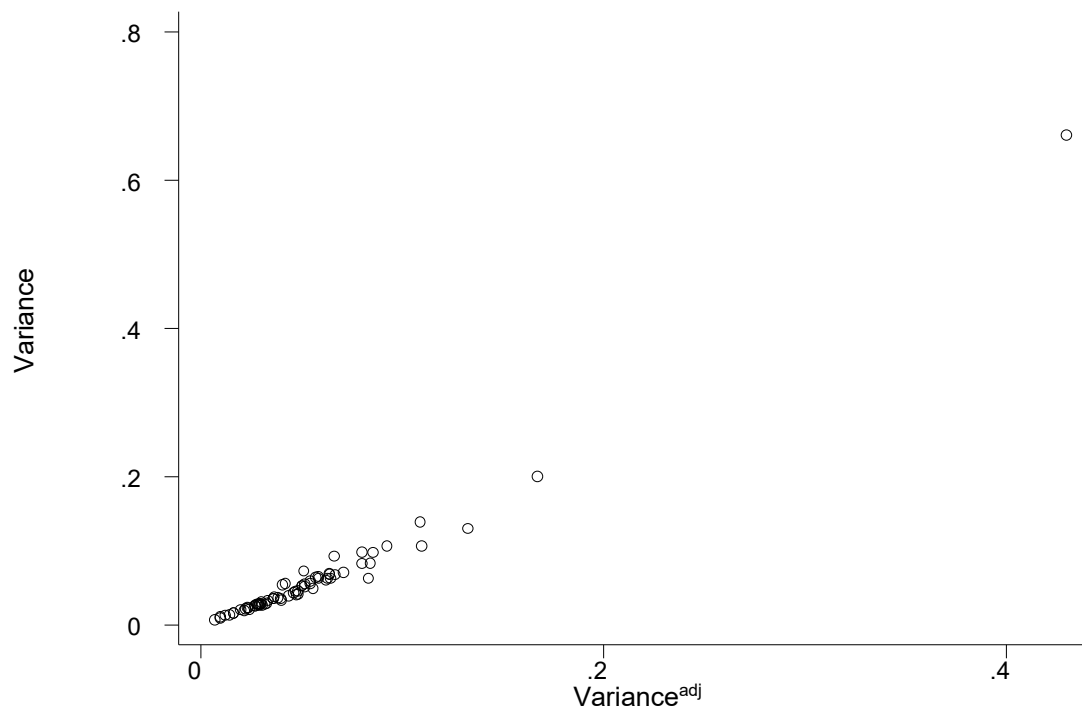
Εικόνα 46: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α (σχέση 27) και στον κάθετο άξονα η διαφορά των log odds ratio , με τροποποίηση από αυτές που έχουν αφαιρεθεί όσες δεν είναι σε HWE, για το συνεπικρατές πρότυπο κληρονομικότητας, με τη μέθοδο του Sato . Εδώ ο συντελεστής παλινδρόμησης β_1 είναι -0.0213793



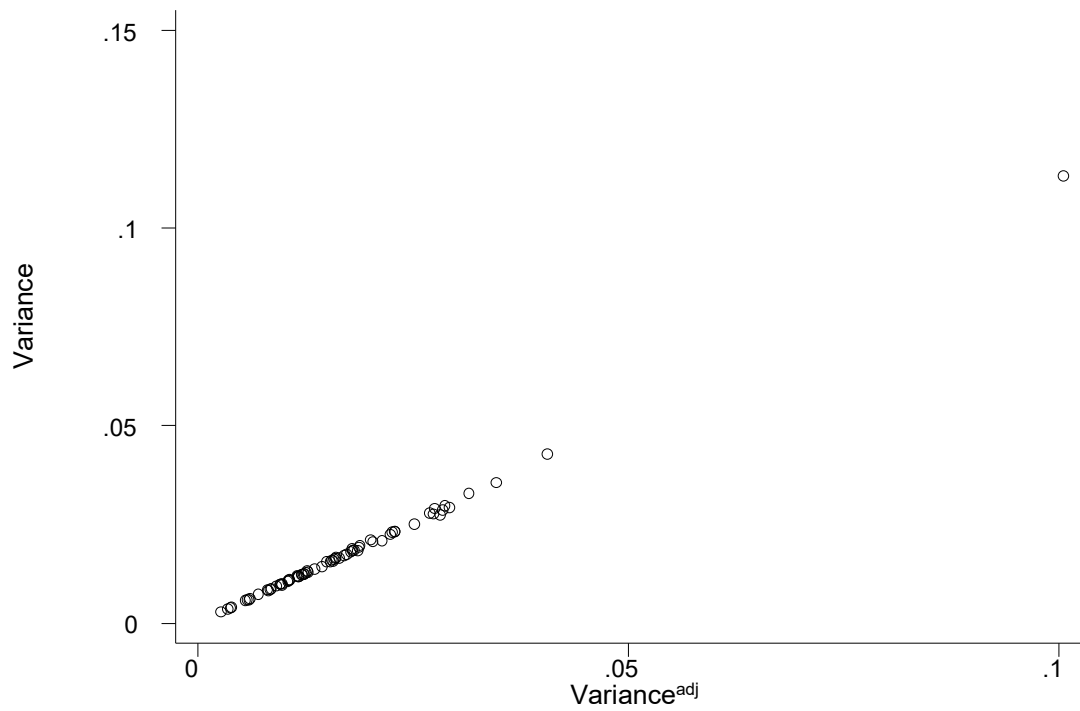
Εικόνα 47: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α (σχέση 27) και στον κάθετο άξονα η διαφορά των log odds ratio , με τροποποίηση από αυτές που έχουν αφαιρεθεί όσες δεν είναι σε HWE, για το συνεπικρατές πρότυπο κληρονομικότητας, με τη μέθοδο του Ιωαννίδη. Εδώ ο συντελεστής παλινδρόμησης β_1 είναι -0.0259727



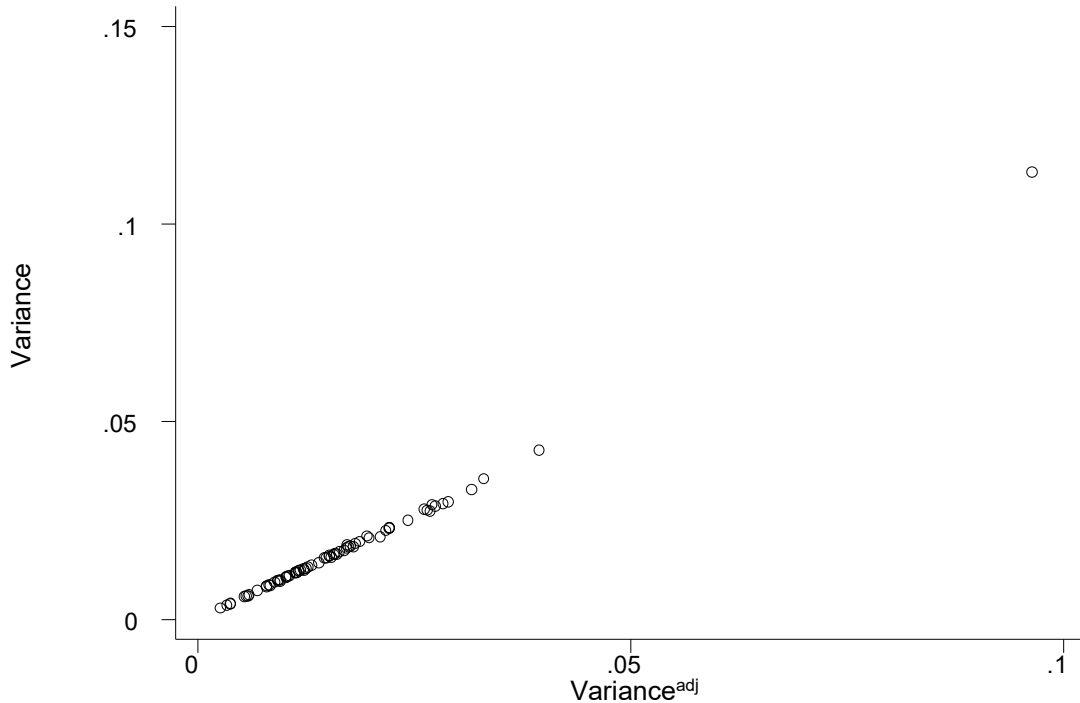
Εικόνα 48: Στον οριζόντιο άξονα είναι η διακύμανση με τροποποίηση (σχέση 30) και στον κάθετο άξονα η διακύμανση χωρίς τροποποίηση (σχέση 3) για το επικρατές πρότυπο κληρονομικότητας.



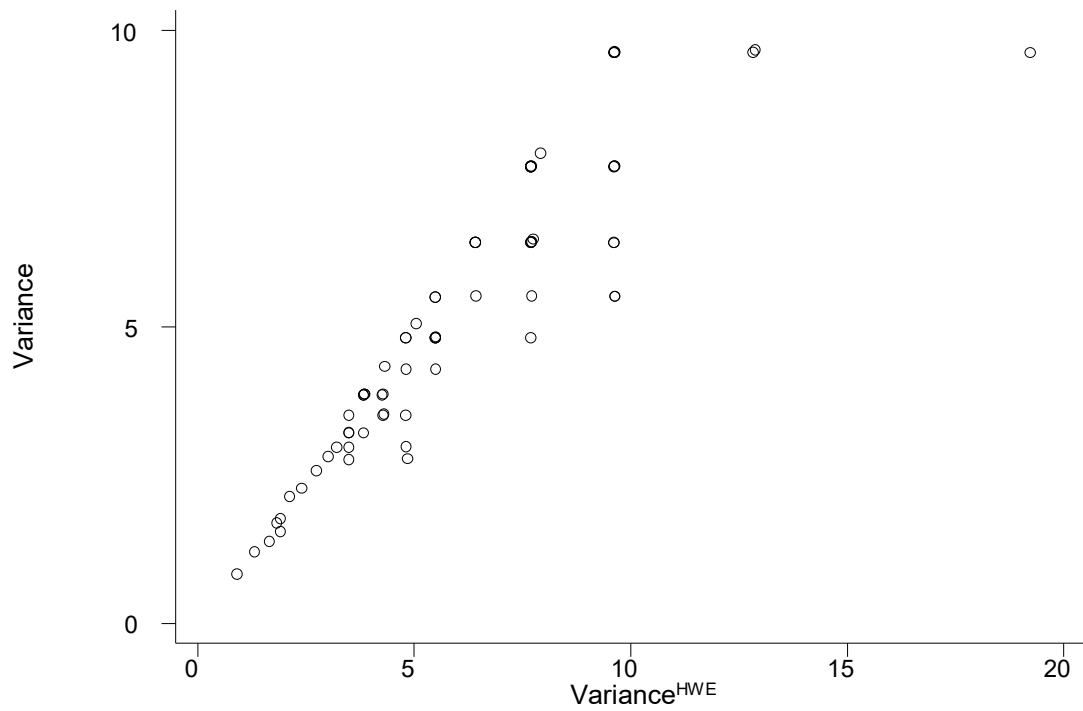
Εικόνα 49: Στον οριζόντιο άξονα είναι η διακύμανση με τροποποίηση και στον κάθετο άξονα η διακύμανση χωρίς τροποποίηση, για το υπολειπόμενο πρότυπο κληρονομικότητας (σχέση 20, 33).



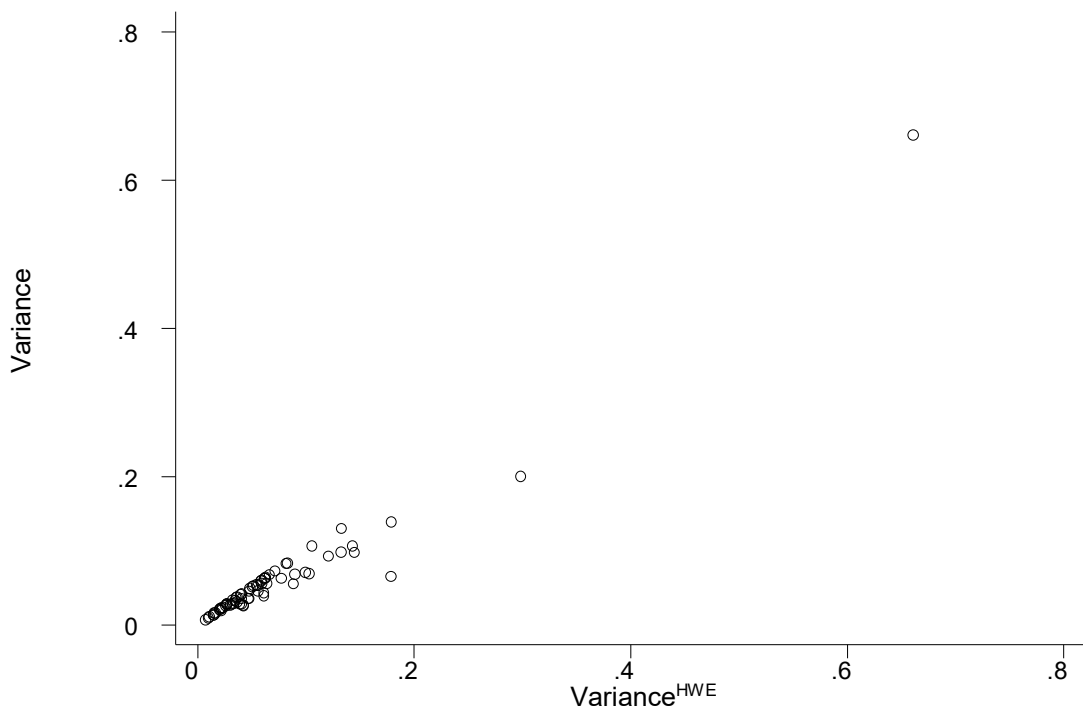
Εικόνα 50: Στον οριζόντιο άξονα είναι η διακύμανση με τροποποίηση και στον κάθετο άξονα η διακύμανση χωρίς τροποποίηση, για το συνεπικρατές πρότυπο κληρονομικότητας, με τη μέθοδο του Sato (σχέση 22, 34). Εδώ ο συντελεστής, $\beta 1$ της παλινδρόμησης ισούται με 0.2507945 και το p-value=0.



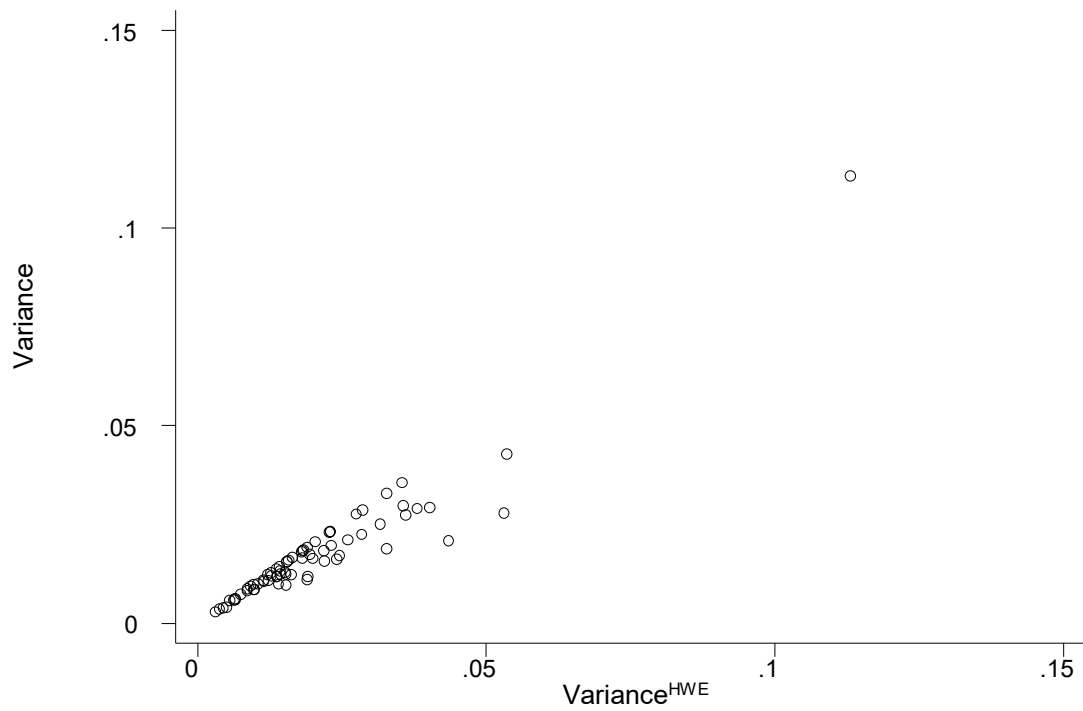
Εικόνα 51: Στον οριζόντιο άξονα είναι η διακύμανση με τροποποίηση και στον κάθετο άξονα η διακύμανση χωρίς τροποποίηση, για το συνεπικρατές πρότυπο κληρονομικότητας, με τη μέθοδο του Ιωαννίδη (σχέση 35). Εδώ ο συντελεστής, $\beta 1$ της παλινδρόμησης ισούται με 1.115144 και το p-value=0.



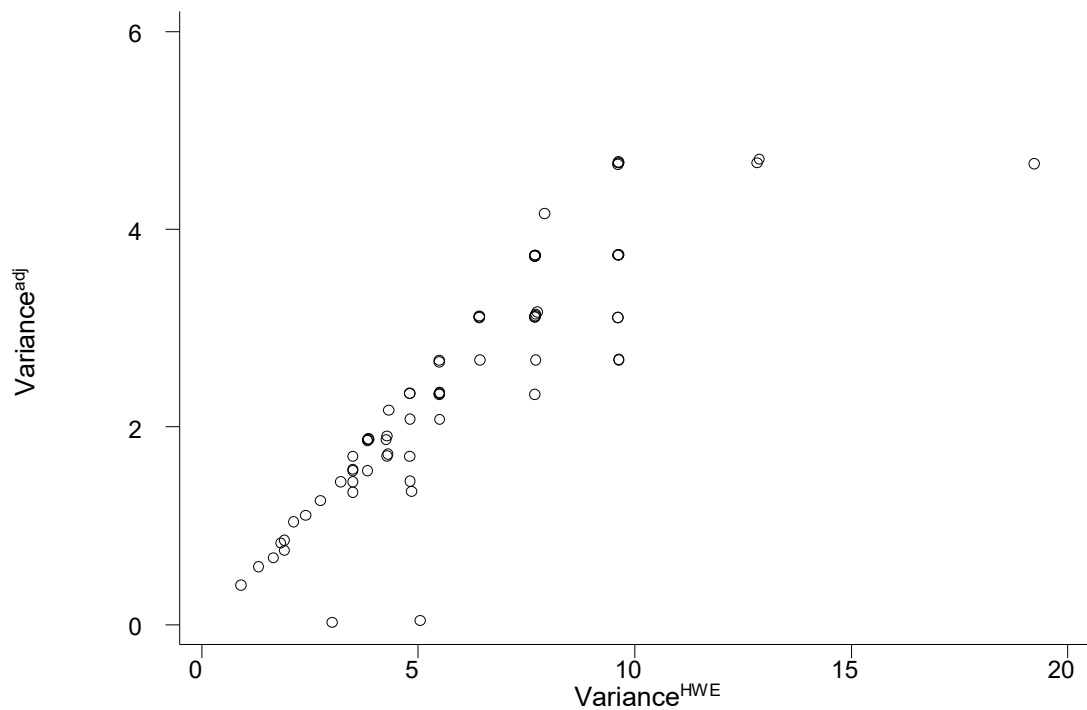
Εικόνα 52: Στον κάθετο άξονα η διακύμανση χωρίς τροποποίηση (σχέση 18) και στον οριζόντιο άξονα οι διακύμανση των μελετών που έχουμε αφαιρέσει αυτές που δεν είναι σε HWE για το επικρατές πρότυπο κληρονομικότητας. Εδώ ο συντελεστής, $\beta 1$ της παλινδρόμησης είναι 0,7290.



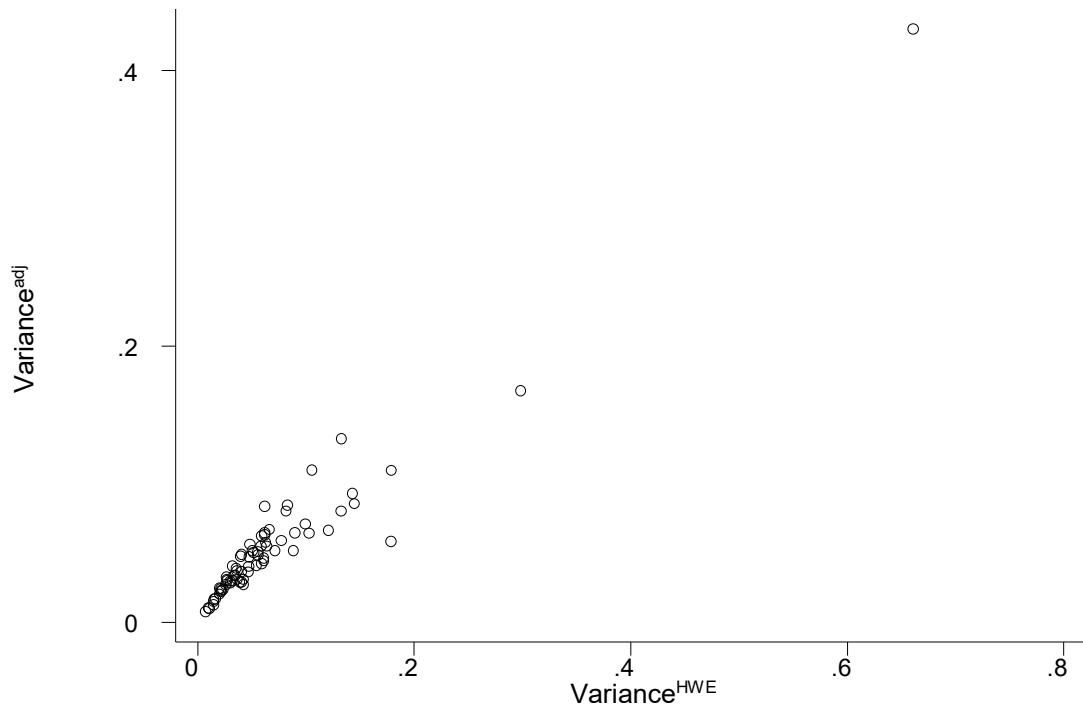
Εικόνα 53: Στον κάθετο άξονα η διακύμανση χωρίς τροποποίηση (σχέση 20) και στον οριζόντιο άξονα η διακύμανση των μελετών, που έχουμε αφαιρέσει αυτές που δεν είναι σε HWE για το υπολειπόμενο πρότυπο κληρονομικότητας. Εδώ ο συντελεστής, $\beta 1$ της παλινδρόμησης είναι 0,7290.



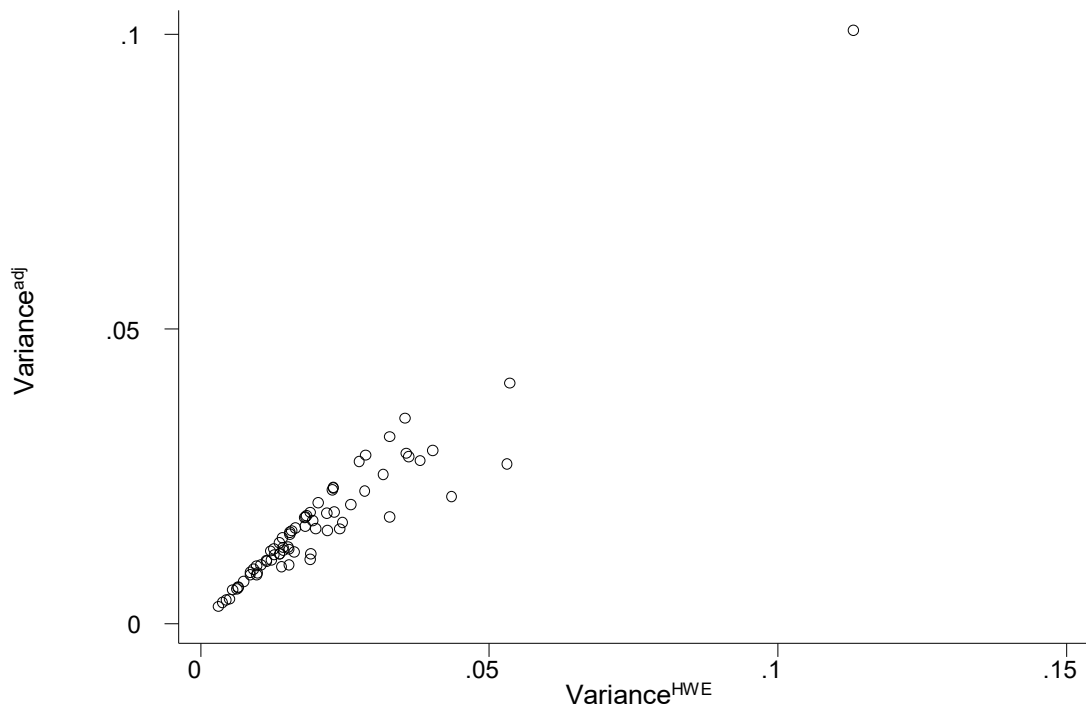
Εικόνα 54: Στον κάθετο άξονα η διακύμανση χωρίς τροποποίηση (σχέση 22) και στον οριζόντιο άξονα η διακύμανση των μελετών, που έχουμε αφαιρέσει αυτές που δεν είναι σε HWE, για το συνεπικρατές πρότυπο κληρονομικότητας. Εδώ ο συντελεστής, βI της παλινδρόμησης είναι 0.8416411.



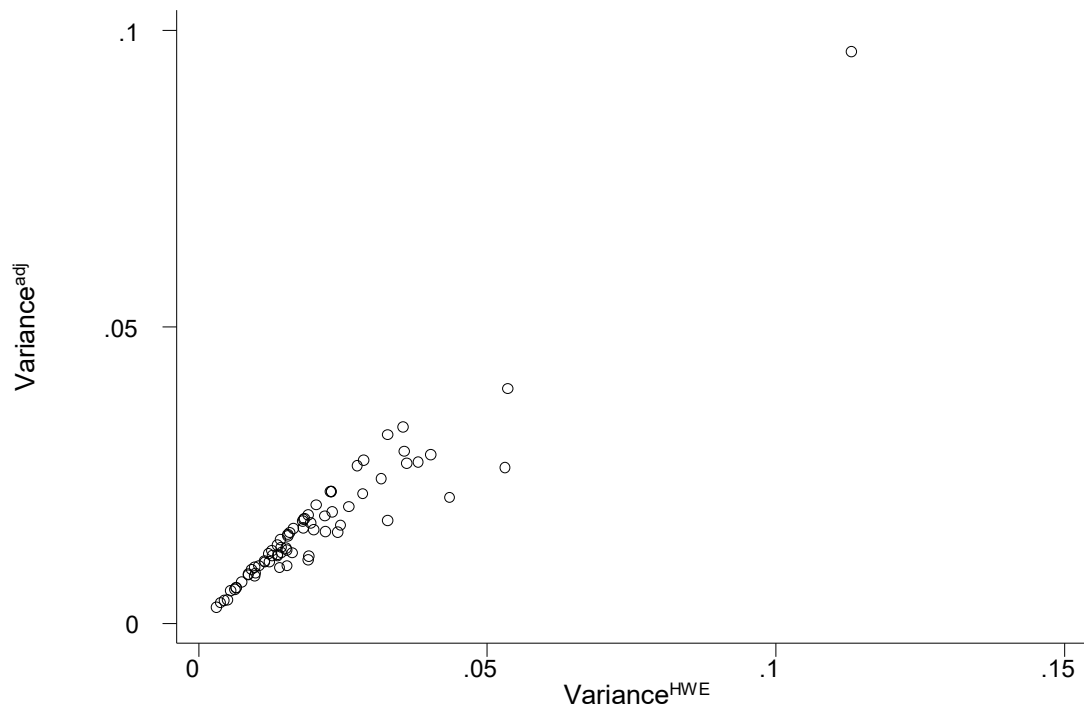
Εικόνα 55: Στον κάθετο άξονα η διακύμανση χωρίς τροποποίηση και στον οριζόντιο άξονα η διακύμανση των μελετών, που έχουν αφαιρεθεί όσες δεν είναι σε HWE, για το επικρατές πρότυπο κληρονομικότητας (σχέση 30). Εδώ ο συντελεστής, βI της παλινδρόμησης είναι 0.3562217



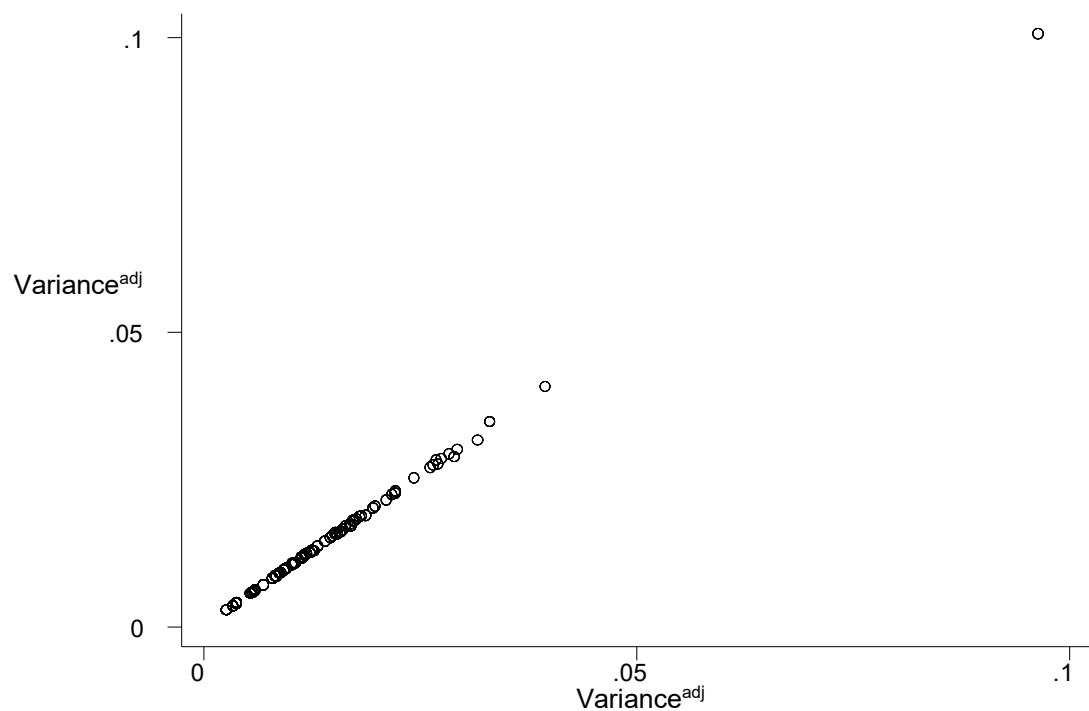
Εικόνα 56: Στον κάθετο άξονα η διακύμανση χωρίς τροποποίηση και στον οριζόντιο άξονα η διακύμανση των μελετών, που έχουν αφαιρεθεί όσες δεν είναι σε HWE, για το υπολειπόμενο πρότυπο κληρονομικότητας (σχέση 33). Εδώ ο συντελεστής, βI της παλινδρόμησης είναι 0.4207785



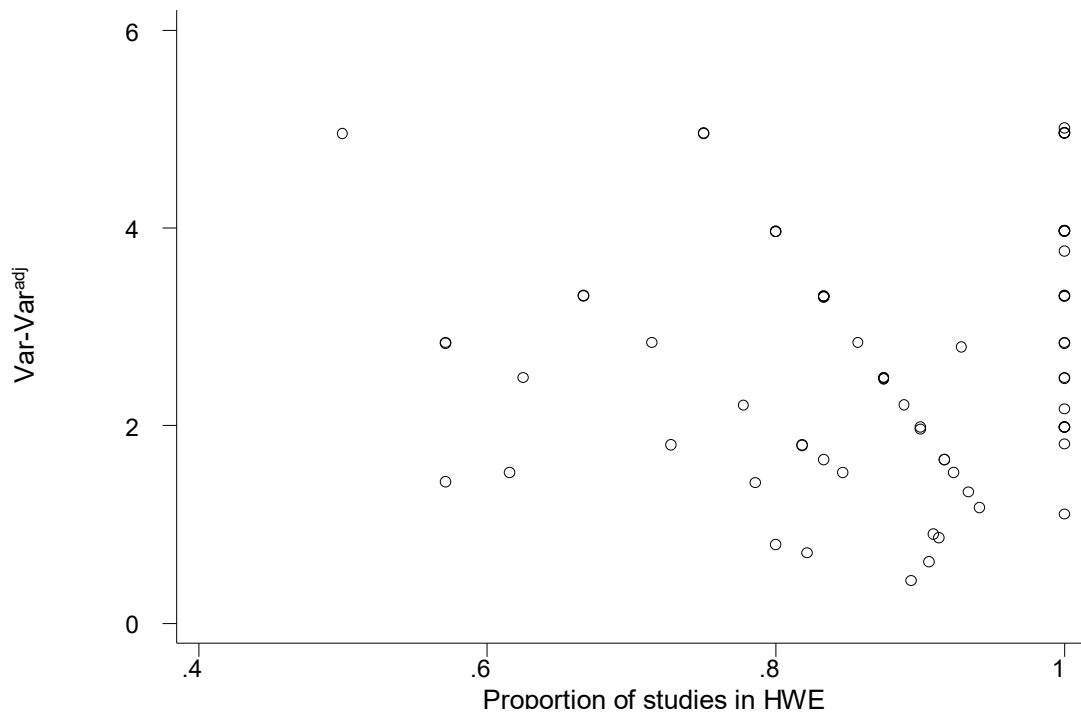
Εικόνα 57: Στον κάθετο άξονα η διακύμανση χωρίς τροποποίηση και στον οριζόντιο άξονα η διακύμανση των μελετών, που έχουν αφαιρεθεί όσες δεν είναι σε HWE, για το συνεπικρατές πρότυπο κληρονομικότητας (σχέση 34). Εδώ ο συντελεστής, βI της παλινδρόμησης είναι 0.78391



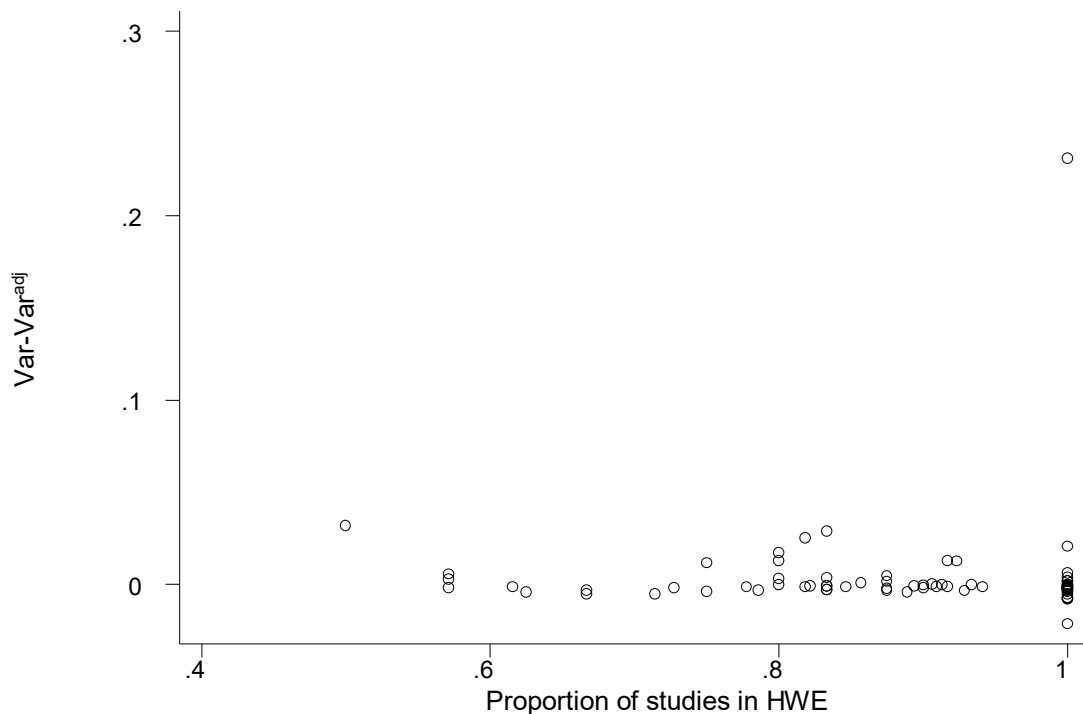
Εικόνα 58: Στον κάθετο άξονα η διακύμανση χωρίς τροποποίηση και στον οριζόντιο άξονα η διακύμανση των μελετών, που έχουν αφαιρεθεί όσες δεν είναι σε HWE, για το συνεπικρατές πρότυπο κληρονομικότητας (σχέση 35). Εδώ ο συντελεστής, $\beta 1$ της παλινδρόμησης είναι 0.756788



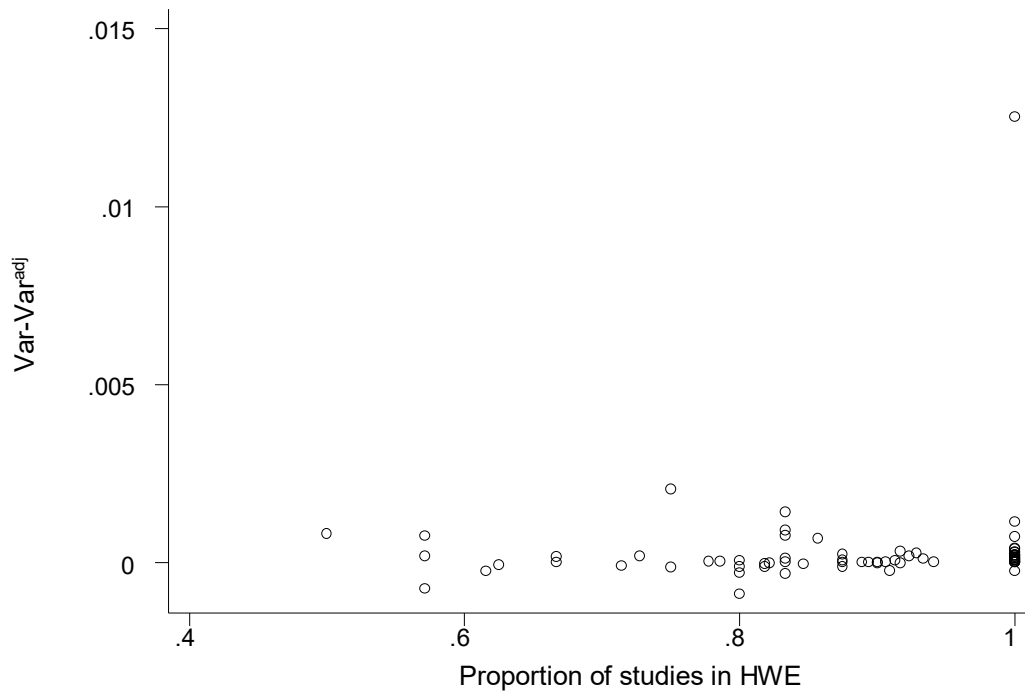
Εικόνα 59: Στον κάθετο άξονα η διακύμανση με τροποποίηση με τη μέθοδο του Sato (σχέση 34) και στον οριζόντιο άξονα η διακύμανση των μελετών με τροποποίηση με τη μέθοδο του Ιωαννίδη (σχέση 35), για το συνεπικρατές πρότυπο κληρονομικότητας .



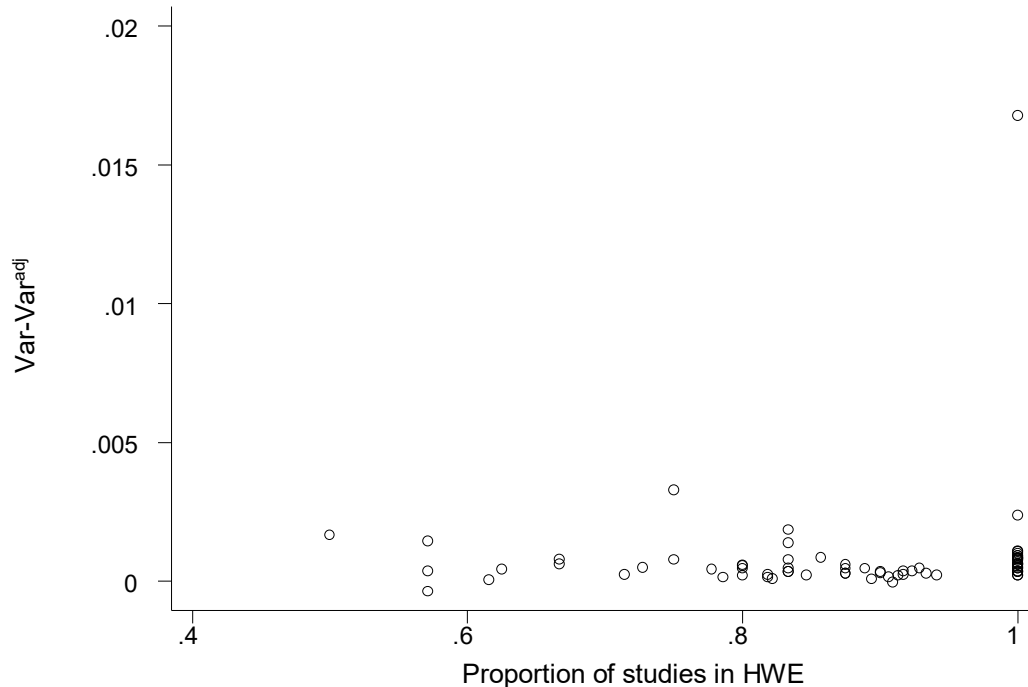
Εικόνα 60: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HWE και στον κάθετο άξονα η διαφορά των διακυμάνσεων, χωρίς τροποποίηση από αυτές που έχουν τροποποιήσεις, για το επικρατές πρότυπο κληρονομικότητας (σχέση 18,30). Εδώ ο συντελεστής παλινδρόμησης β_1 είναι -0.2981 και $p\text{-value} < 10^{-4}$



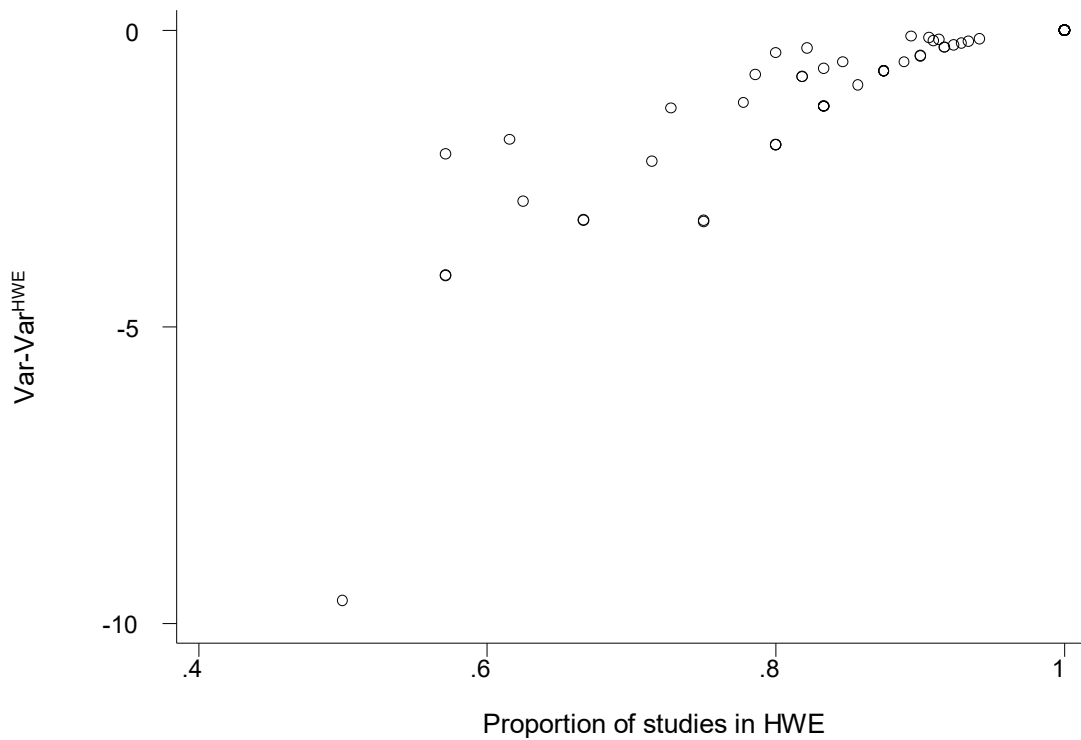
Εικόνα 61: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HWE και στον κάθετο άξονα η διαφορά των διακυμάνσεων, χωρίς τροποποίηση από αυτές που έχουν τροποποιήσεις, για το υπολειπόμενο πρότυπο κληρονομικότητας (σχέση 20,33). Εδώ ο συντελεστής β_1 είναι 0.01733 , $p\text{-value} = 0.026$.



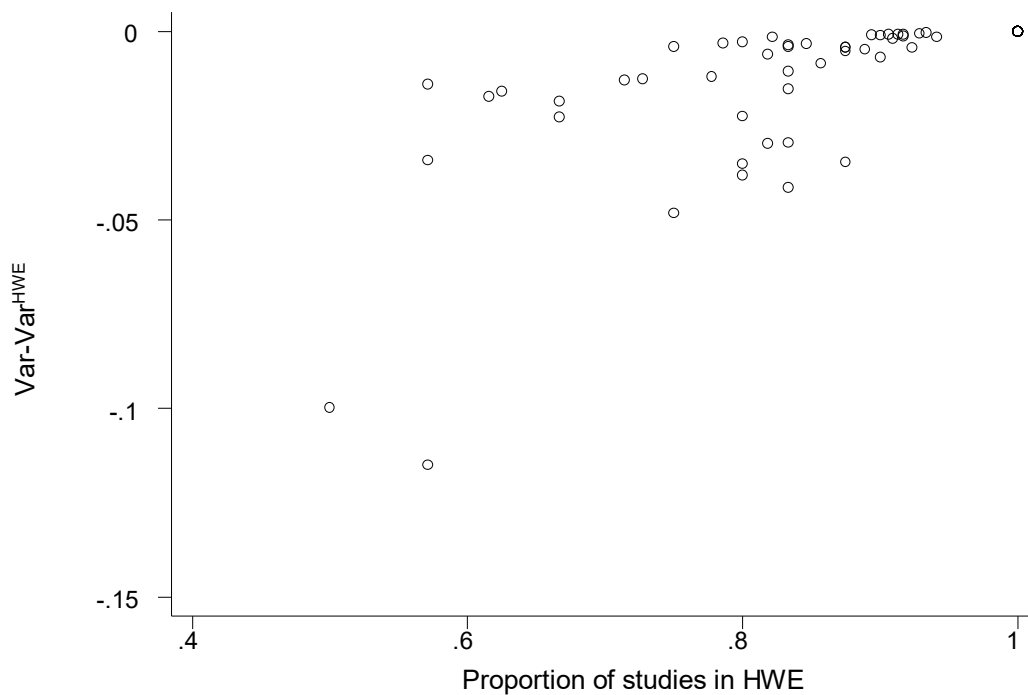
Εικόνα 62: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HWE και στον κάθετο άξονα η διαφορά των διακυμάνσεων, χωρίς τροποποίηση από αυτές που έχουν τροποποιήσεις, για το συνεπικρατές πρότυπο κληρονομικότητας, με τη μέθοδο του Sato (σχέση 22,34). Εδώ ο συντελεστής παλινδρόμησης είναι 0,0008719 και το p-value=0.008.



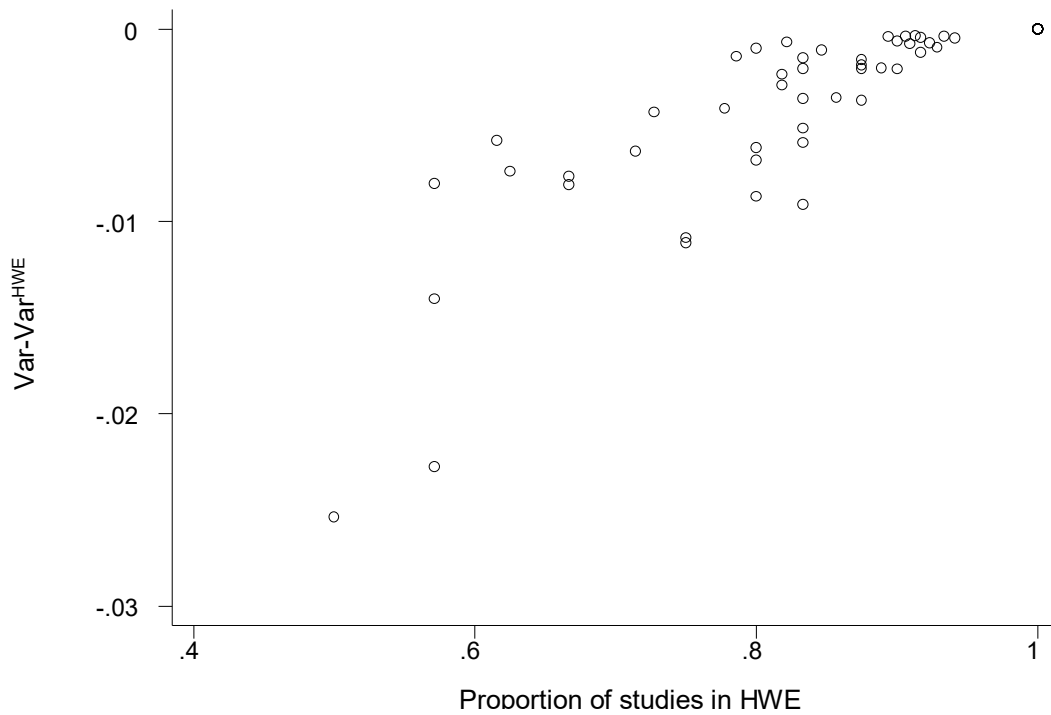
Εικόνα 63: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HWE και στον κάθετο άξονα η διαφορά των διακυμάνσεων, χωρίς τροποποίηση από αυτές που έχουν τροποποιήσεις, για το συνεπικρατές πρότυπο κληρονομικότητας, με τη μέθοδο του Ιωαννίδη (σχέση 35). Εδώ ο συντελεστής παλινδρόμησης είναι 0.009795 και το p-value=0.026.



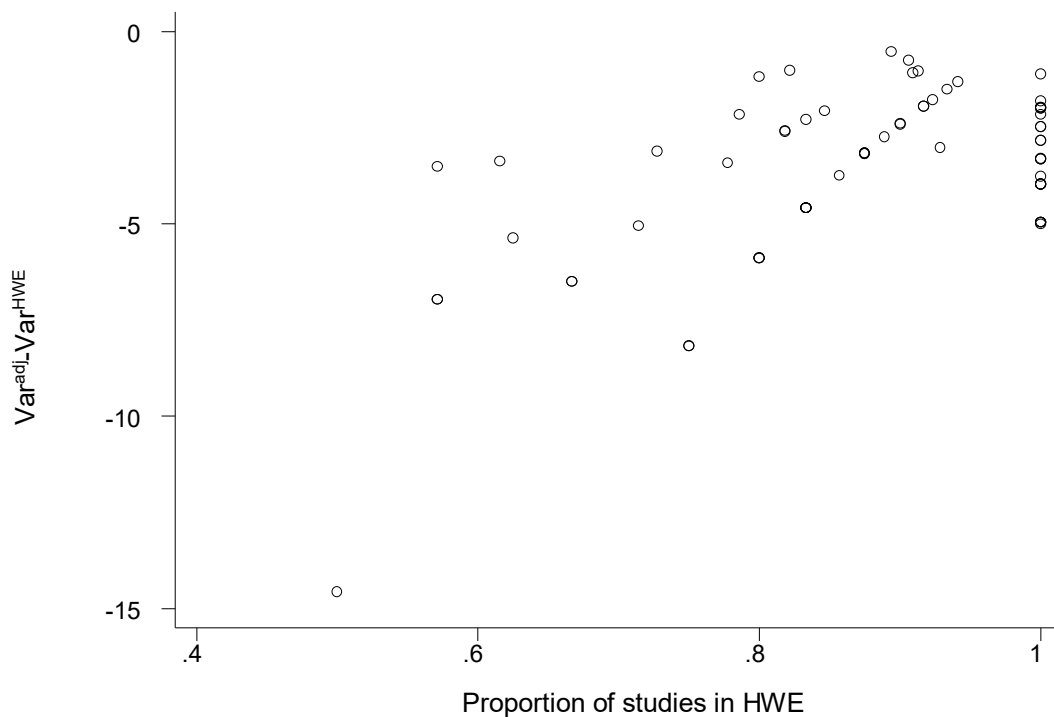
Εικόνα 64 : Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HWE και στον κάθετο άξονα η διαφορά των διακυμάνσεων, χωρίς τροποποίηση από αυτές που είναι σε HWE, για το επικρατές πρότυπο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης β_1 είναι 7.93



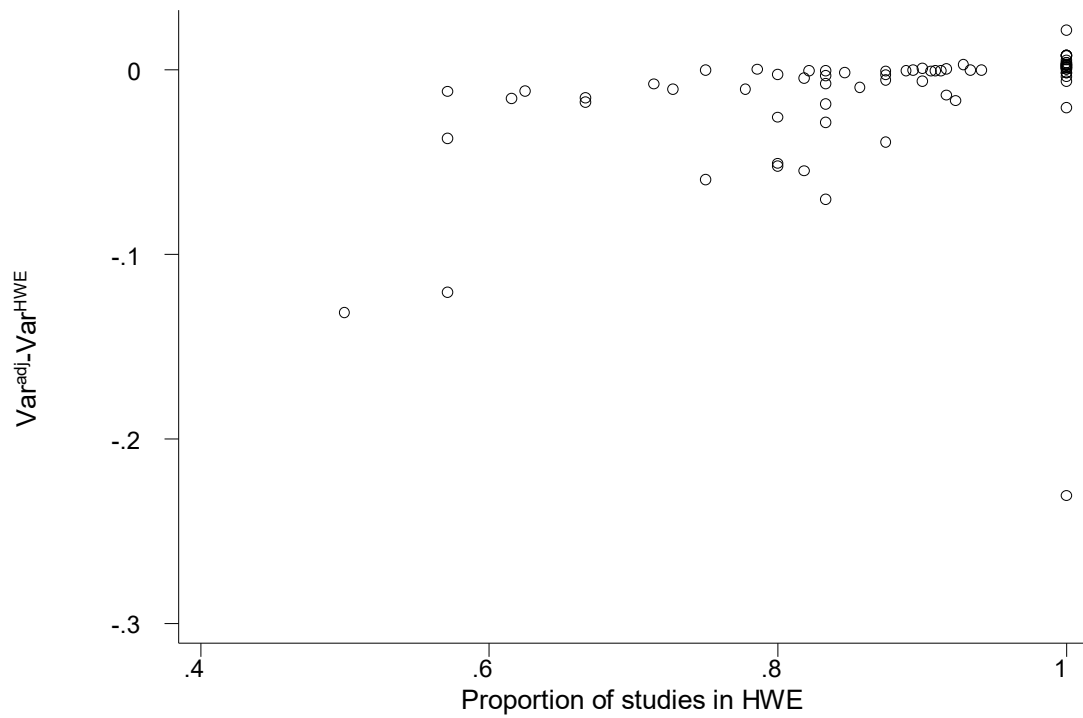
Εικόνα 65: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HW και στον κάθετο άξονα η διαφορά των διακυμάνσεων, χωρίς τροποποίηση από αυτές που είναι σε HWE, για το υπολειπόμενο πρότυπο κληρονομικότητας. Εδώ ο συντελεστής β_1 είναι 0.11064.



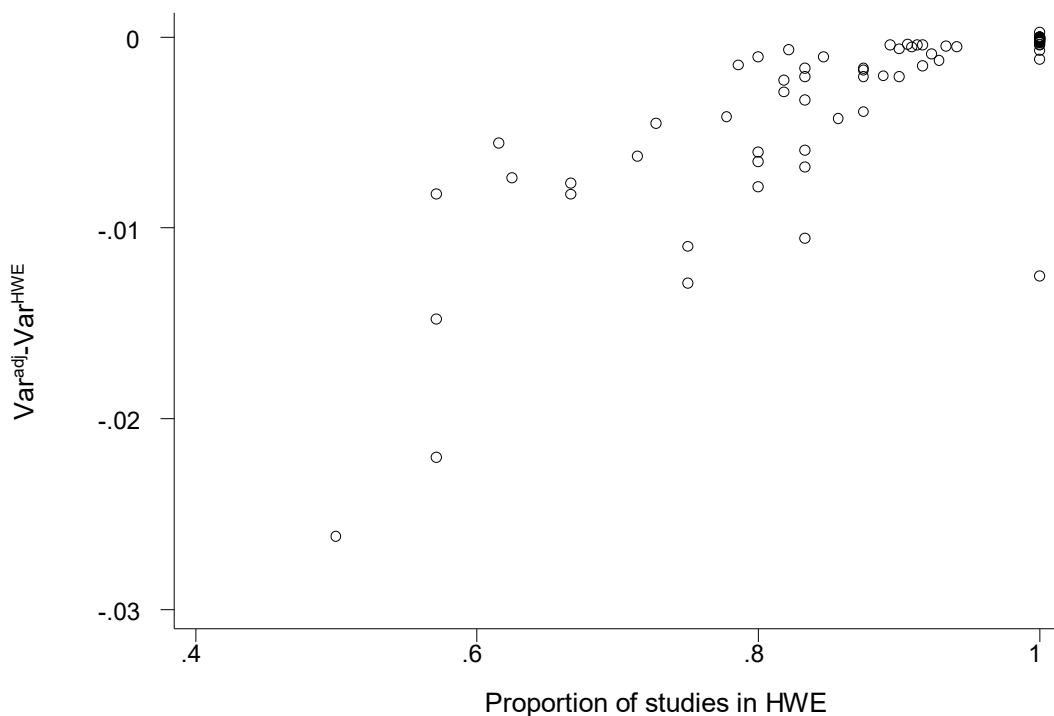
Εικόνα 66: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HW και στον κάθετο άξονα η διαφορά των διακυμάνσεων, χωρίς τροποποίηση από αυτές που είναι σε HWE, για το συνεπικρατές πρότυπο κληρονομικότητας. Εδώ ο συντελεστής βI είναι 0.00264416.



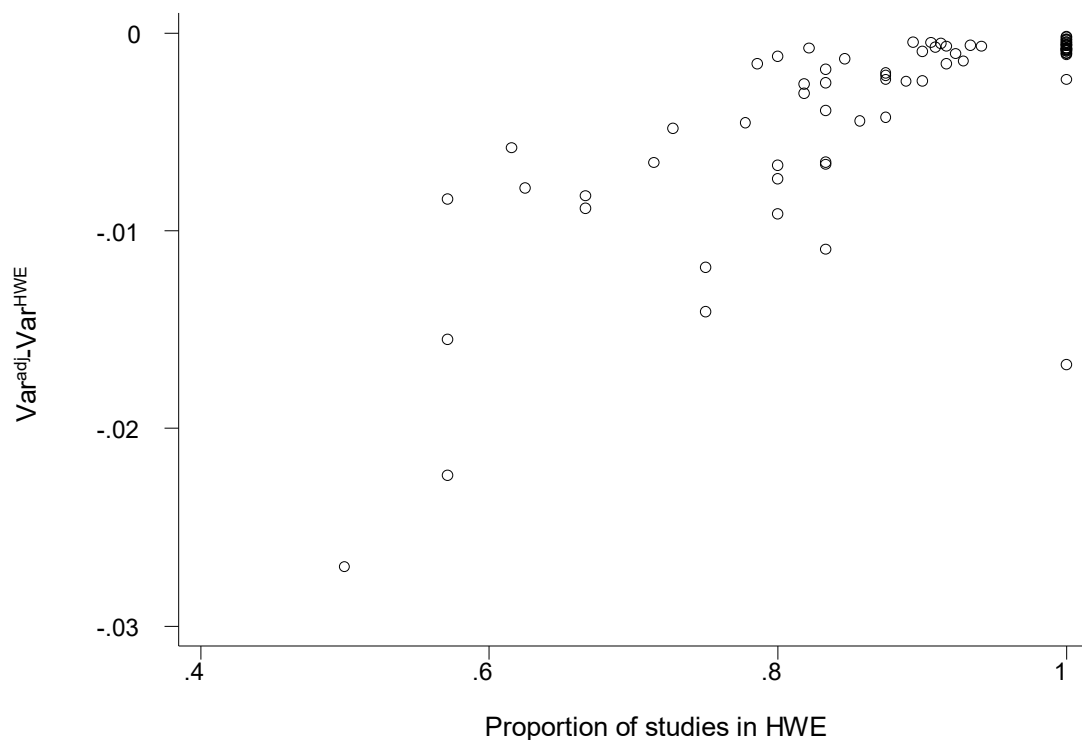
Εικόνα 67: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HW και στον κάθετο άξονα η διαφορά των διακυμάνσεων, με τροποποίηση από αυτές που είναι σε HWE, για το επικρατές πρότυπο κληρονομικότητας. Εδώ ο συντελεστής βI είναι 8.234



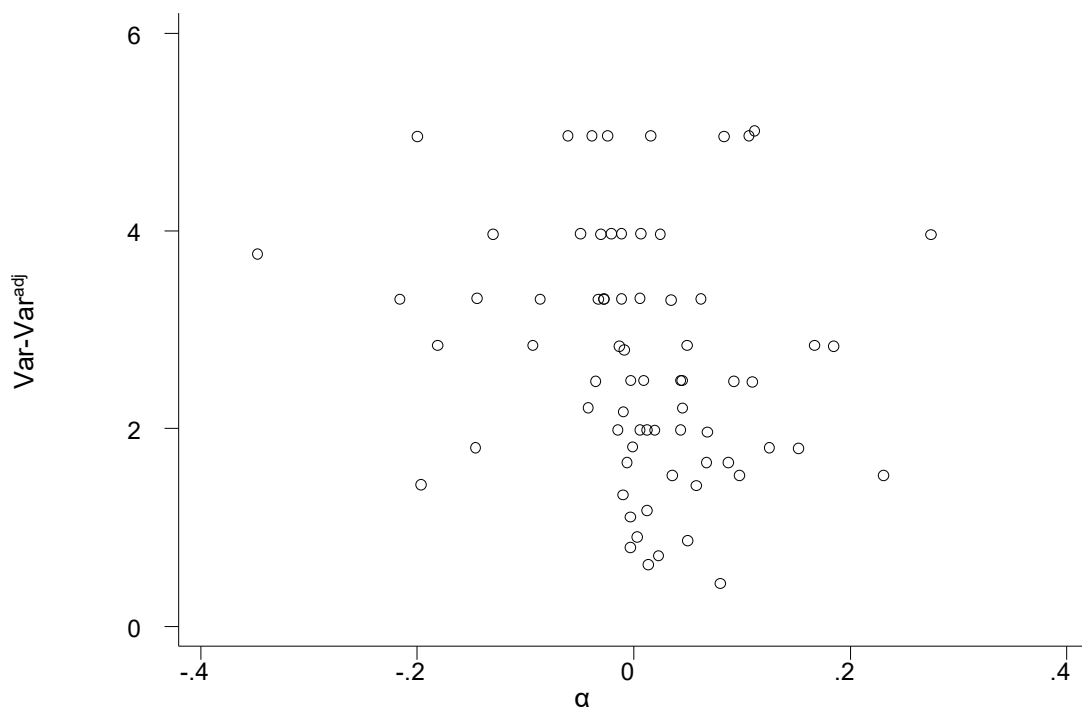
Εικόνα 68: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HW και στον κάθετο άξονα η διαφορά των διακυμάνσεων, με τροποποίηση από αυτές που είναι σε HWE, για το υπολειπόμενο πρότυπο κληρονομικότητας. Εδώ ο συντελεστής βI είναι 0.0933124



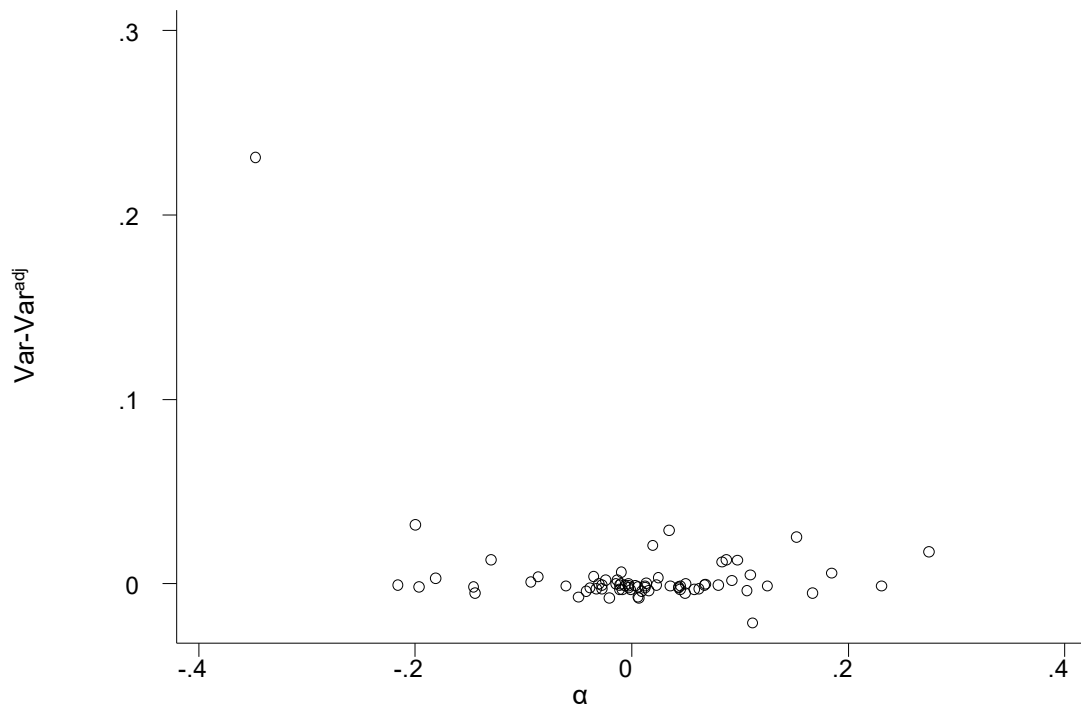
Εικόνα 69: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HW και στον κάθετο άξονα η διαφορά των διακυμάνσεων, με τροποποίηση από αυτές που είναι σε HWE, για το συνεπικρατές πρότυπο κληρονομικότητας. Εδώ ο συντελεστής βI είναι 0.0255697



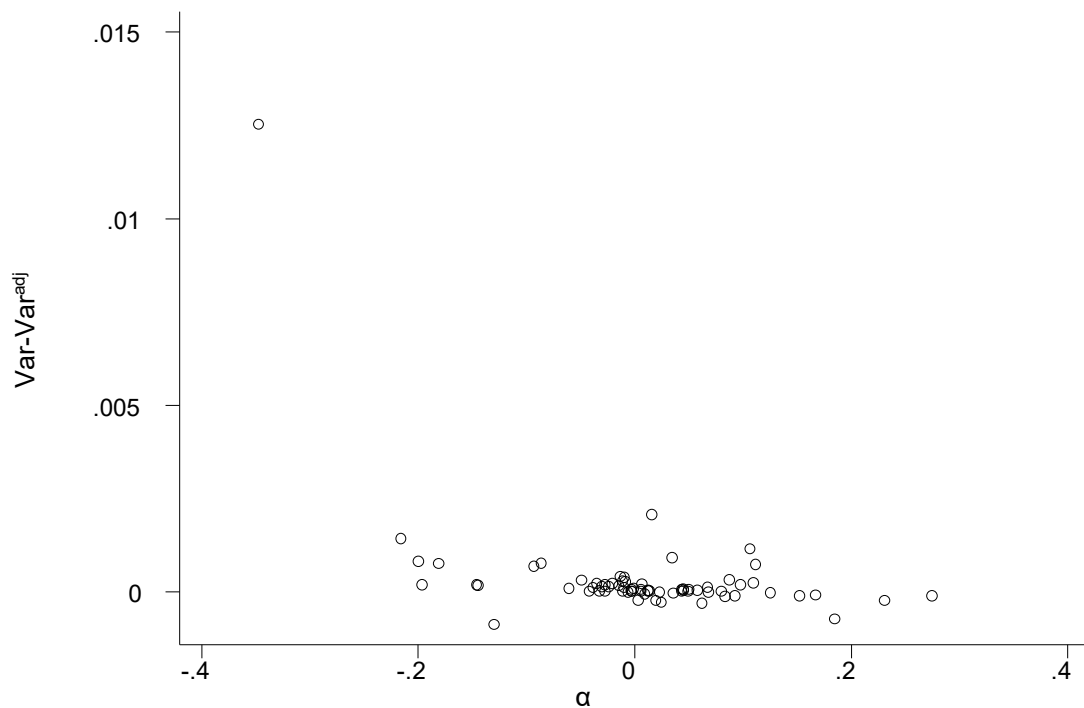
Εικόνα 70: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία HW και στον κάθετο άξονα η διαφορά των διακυμάνσεων, με τροποποίηση από αυτές που είναι σε HWE, για το συνεπικρατές πρότυπο κληρονομικότητας. Εδώ ο συντελεστής β_1 είναι 0.0254621



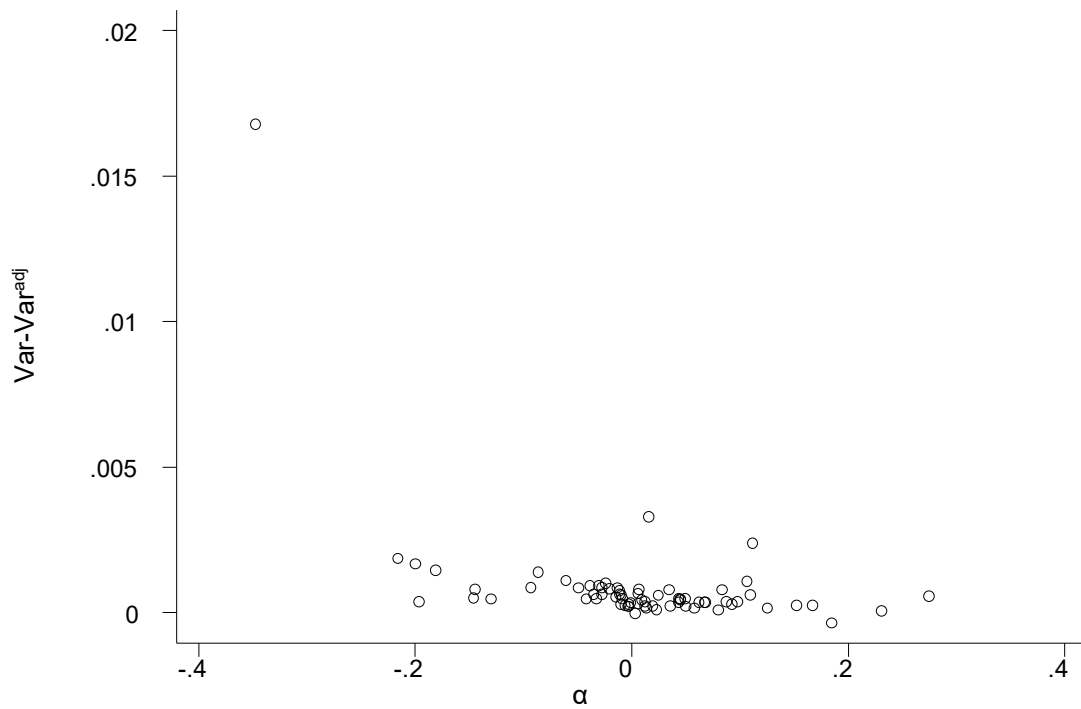
Εικόνα 71: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α και στον κάθετο άξονα η διαφορά των διακυμάνσεων χωρίς τροποποίηση από αυτές που έχουν υποστεί τροποποίηση για το επικρατές πρότυπο κληρονομικότητας (σχέση 18,27,30). Εδώ ο συντελεστής παλινδρόμησης β_1 είναι -1.35942 και $p\text{-value} < 10^{-4}$



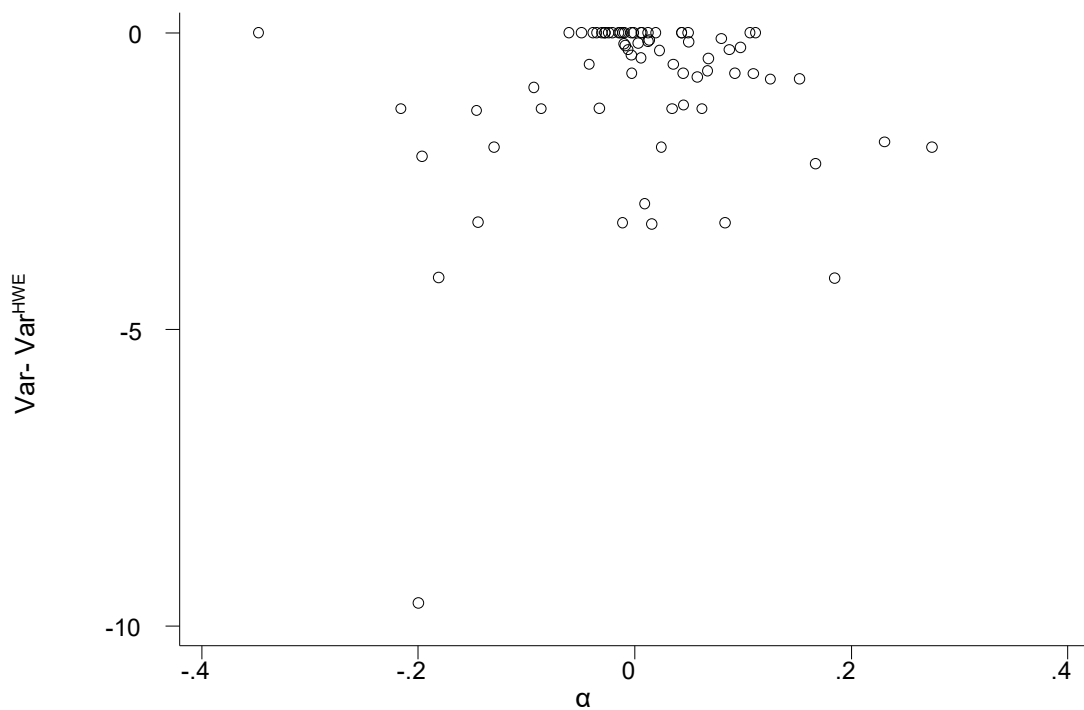
Εικόνα 72: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α και στον κάθετο άξονα η διαφορά των διακυμάνσεων χωρίς τροποποίηση από αυτές που έχουν υποστεί τροποποίηση, για το υπολειπόμενο πρότυπο κληρονομικότητας (σχέση 20,27,33). Εδώ ο συντελεστής παλινδρόμησης β_1 είναι 0.0399423.



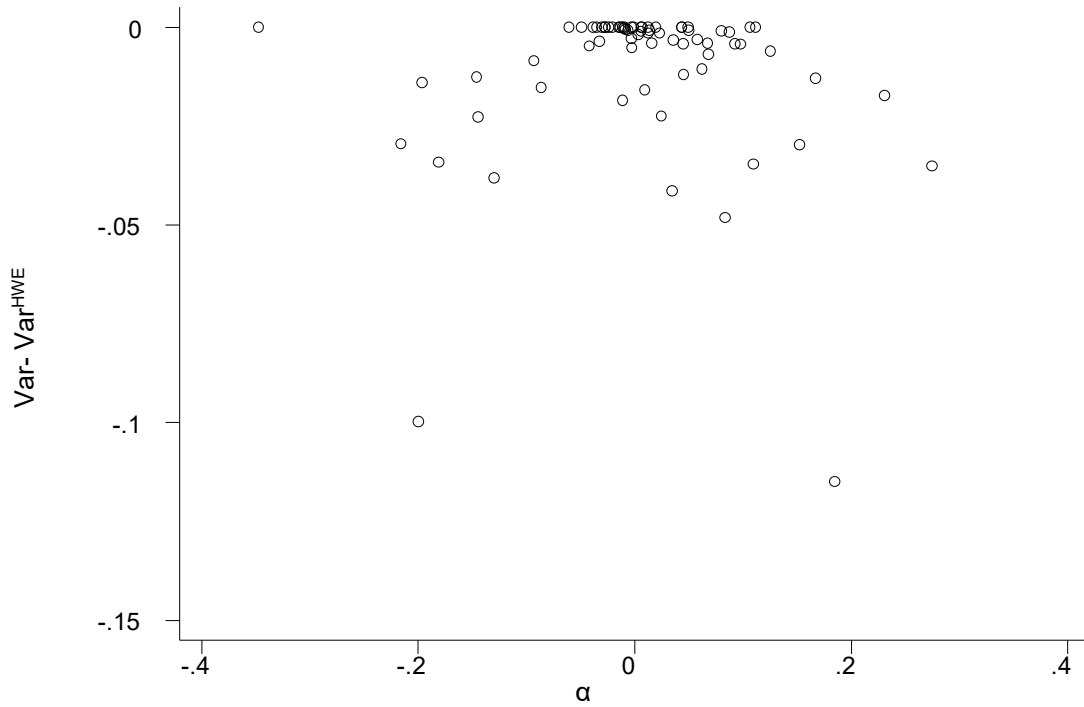
Εικόνα 73: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α και στον κάθετο άξονα η διαφορά των διακυμάνσεων χωρίς τροποποίηση από αυτές που έχουν υποστεί τροποποίηση, για το συνεπικρατές πρότυπο κληρονομικότητας, με τη μέθοδο του Sato (σχέση 22,27,34). Εδώ ο συντελεστής παλινδρόμησης β_1 είναι -0.0010963.



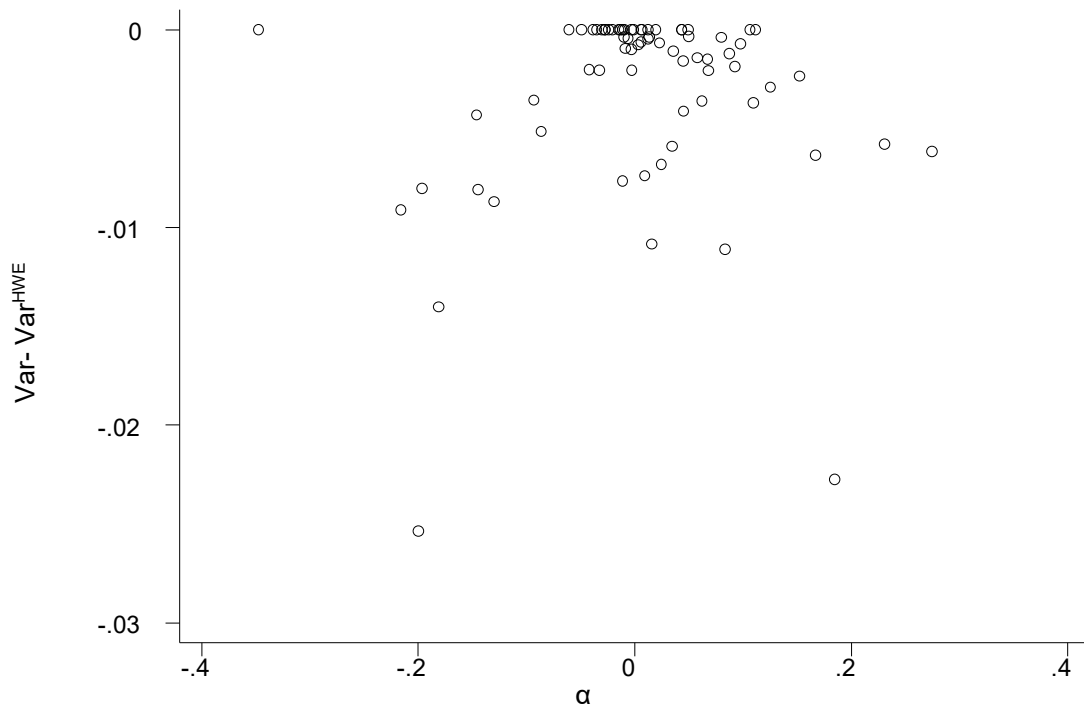
Εικόνα 74: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α και στον κάθετο άξονα η διαφορά των διακυμάνσεων χωρίς τροποποίηση από αυτές που έχουν υποστεί τροποποίηση, για το συνεπικρατές πρότυπο κληρονομικότητας, με τη μέθοδο του Ιωαννίδη (σχέση 22,27,35). Εδώ ο συντελεστής παλινδρόμησης β_1 είναι -0.016442.



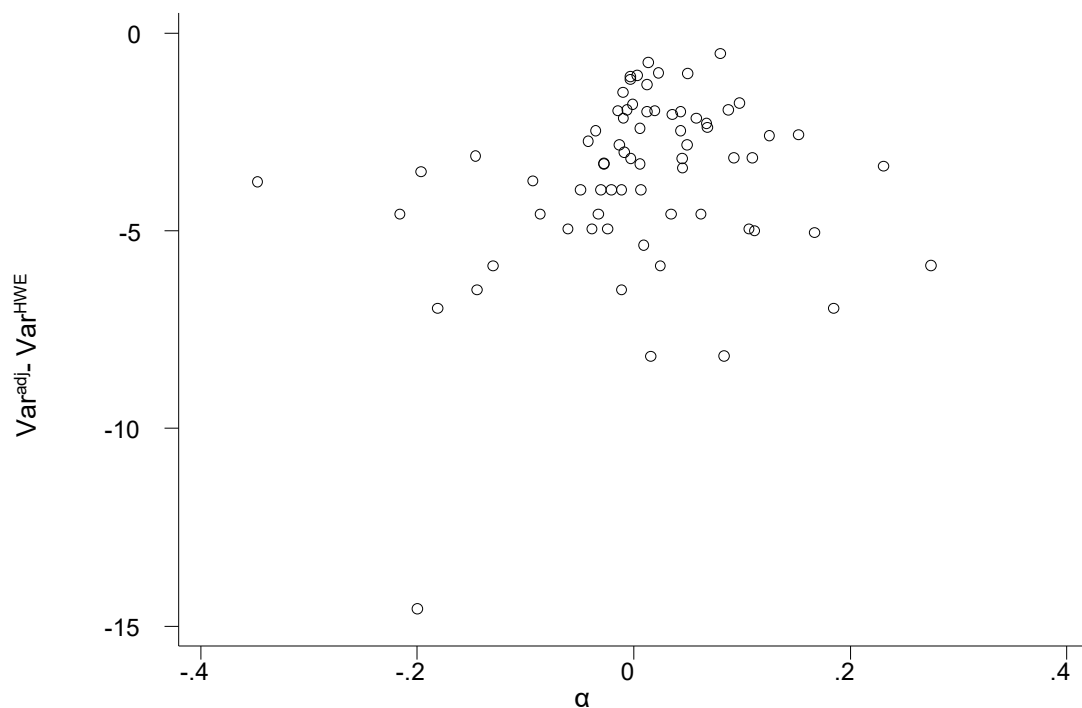
Εικόνα 75: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α και στον κάθετο άξονα η διαφορά των διακυμάνσεων, χωρίς τροποποίηση από τις διακυμάνσεις των μελετών που έχουν αφαιρεθεί αυτές που δεν είναι σε HWE, για το επικρατές πρότυπο κληρονομικότητας (σχέση 18,27). Εδώ ο συντελεστής παλινδρόμησης είναι -0.777



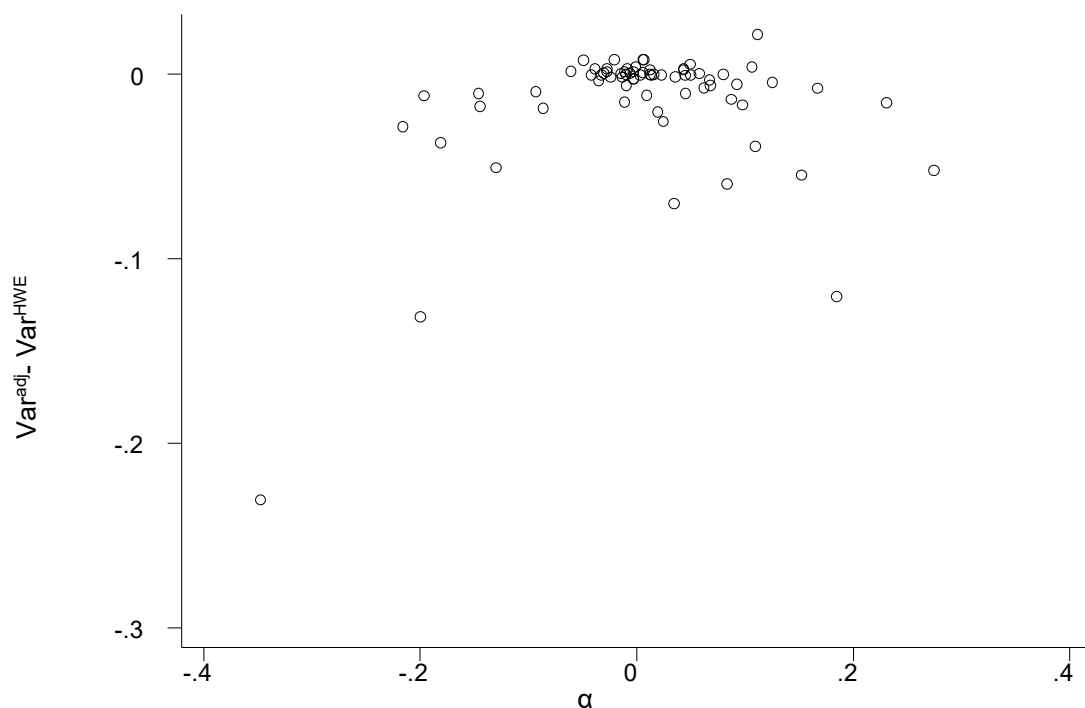
Εικόνα 76: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α και στον κάθετο άξονα η διαφορά των διακυμάνσεων, χωρίς τροποποίηση από τις διακυμάνσεις των μελετών που έχουν αφαιρεθεί αυτές που δεν είναι σε HWE, για το υπολειπόμενο πρότυπο κληρονομικότητας (σχέση 20,27). Εδώ ο συντελεστής παλινδρόμησης β_1 είναι -0.0167395 και $p\text{-value}=0.015$.



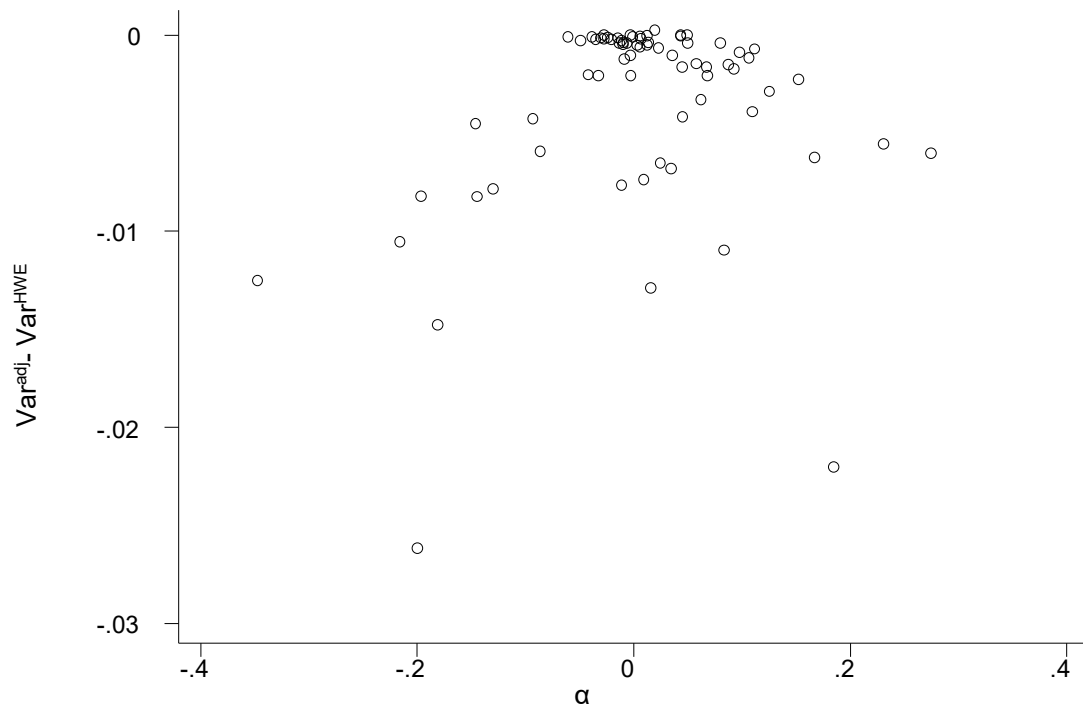
Εικόνα 77: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α και στον κάθετο άξονα η διαφορά των διακυμάνσεων, χωρίς τροποποίηση από τις διακυμάνσεις των μελετών που έχουν αφαιρεθεί αυτές που δεν είναι σε HWE, για το συνεπικρατές πρότυπο κληρονομικότητας (σχέση 22,27). Εδώ ο συντελεστής παλινδρόμησης β_1 είναι -0.0024417 .



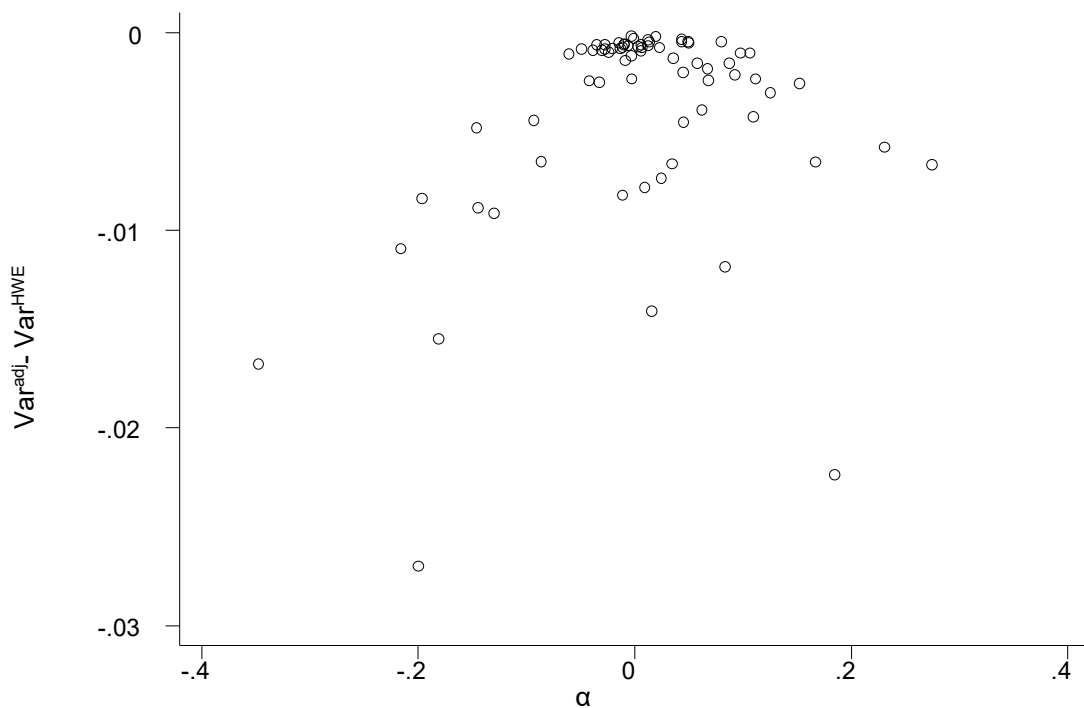
Εικόνα 78: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α και στον κάθετο άξονα η διαφορά των διακυμάνσεων, με τροποποίηση από τις διακυμάνσεις των μελετών που έχουν αφαιρεθεί αυτές που δεν είναι σε HWE, για το επικρατές πρότυπο κληρονομικότητας (σχέση 22,30). Εδώ ο συντελεστής παλινδρόμησης β_1 είναι 2.1365



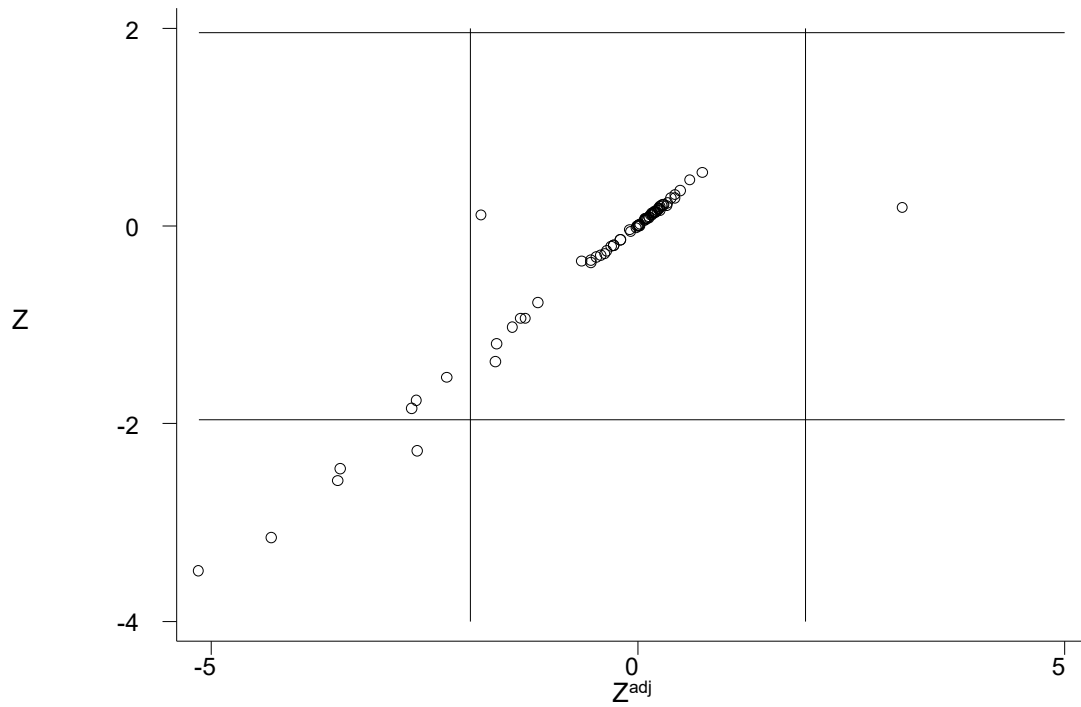
Εικόνα 79: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α και στον κάθετο άξονα η διαφορά των διακυμάνσεων, με τροποποίηση από τις διακυμάνσεις των μελετών που έχουν αφαιρεθεί αυτές που δεν είναι σε HWE, για το υπολειπόμενο πρότυπο κληρονομικότητας (σχέση 27,33). Εδώ ο συντελεστής παλινδρόμησης β_1 είναι -0.056681



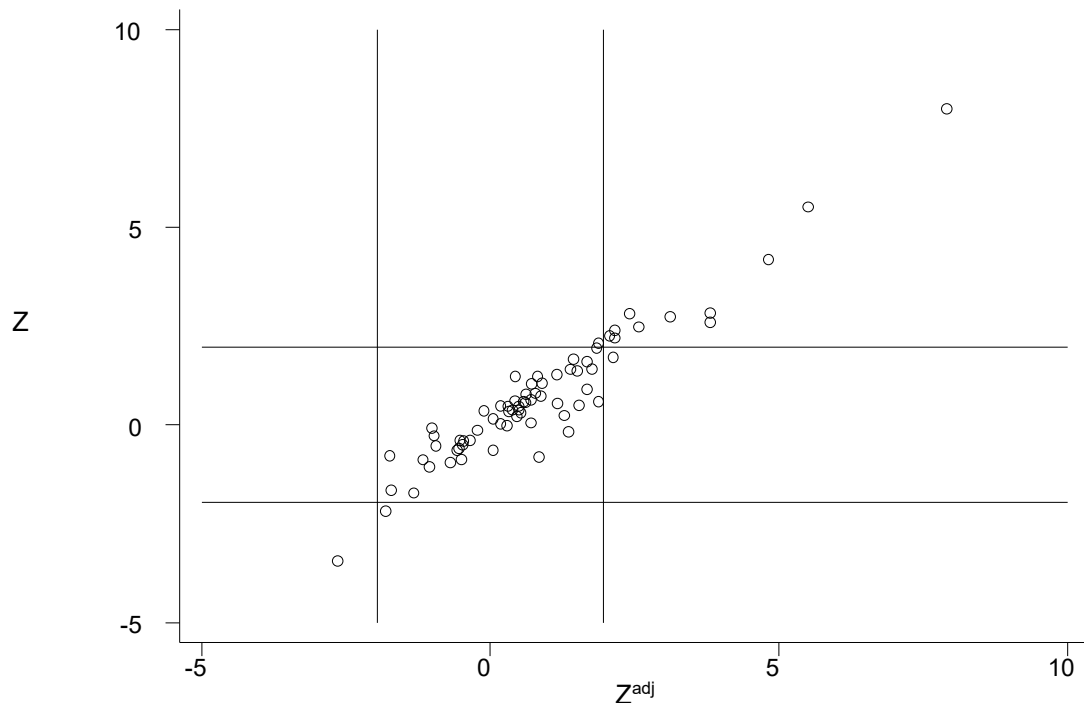
Εικόνα 80: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α και στον κάθετο άξονα η διαφορά των διακυμάνσεων, με τροποποίηση από τις διακυμάνσεις των μελετών που έχουν αφαιρεθεί αυτές που δεν είναι σε HWE, για το συνεπικρατές πρότυπο κληρονομικότητας (σχέση 27,34). Εδώ ο συντελεστής παλινδρόμησης β_1 είναι 0.003538



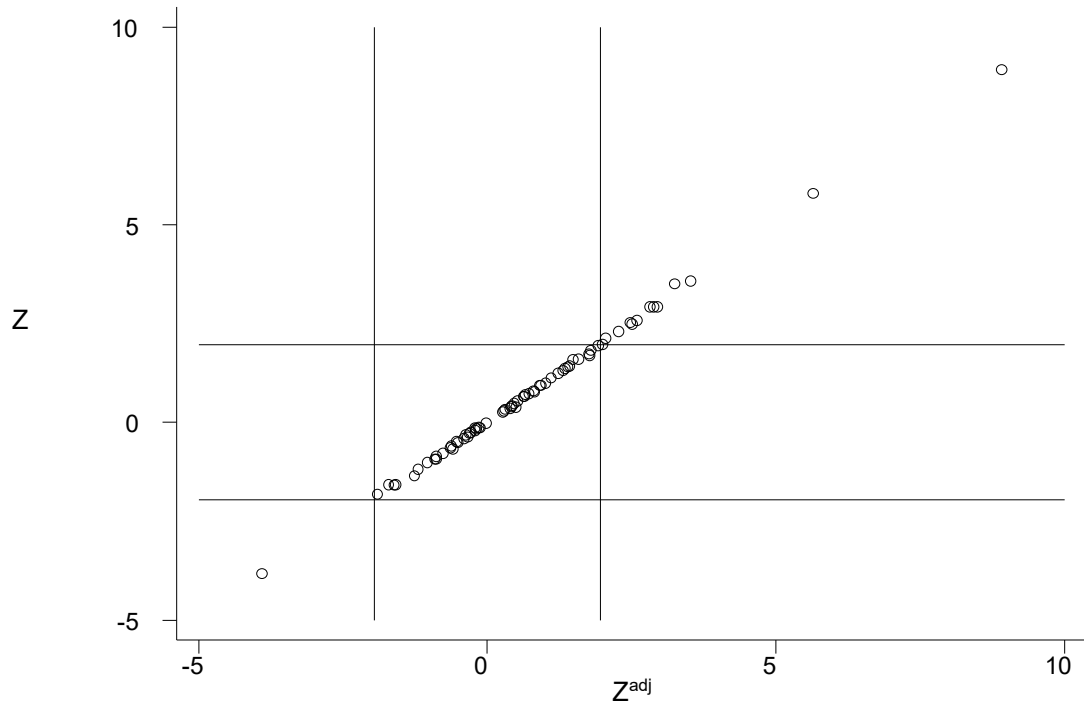
Εικόνα 81: Στον οριζόντιο άξονα είναι ο συντελεστής απόκλισης α και στον κάθετο άξονα η διαφορά των διακυμάνσεων, με τροποποίηση από τις διακυμάνσεις των μελετών που έχουν αφαιρεθεί αυτές που δεν είναι σε HWE, για το συνεπικρατές πρότυπο κληρονομικότητας (σχέση 35). Εδώ ο συντελεστής παλινδρόμησης β_1 είναι 0.0040859



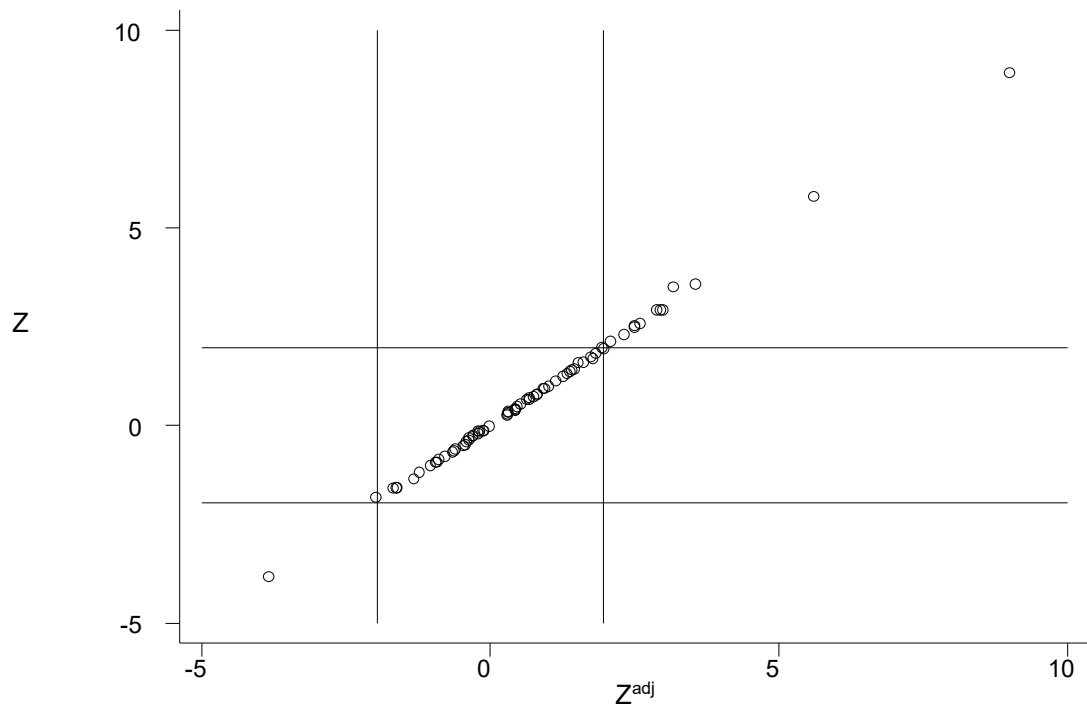
Εικόνα 82: Στον οριζόντιο άξονα είναι ο συντελεστής z , ο λόγος του log odds ratio προς την διακύμανση του log odds ratio με τροποποίηση και στον κάθετο άξονα ο συντελεστής z χωρίς τροποποίηση για το επικρατές μοντέλο κληρονομικότητας (σχέση 16). Εδώ ο συντελεστής παλινδρόμησης β_1 είναι 0.7334025, $p\text{-value} < 10^{-4}$



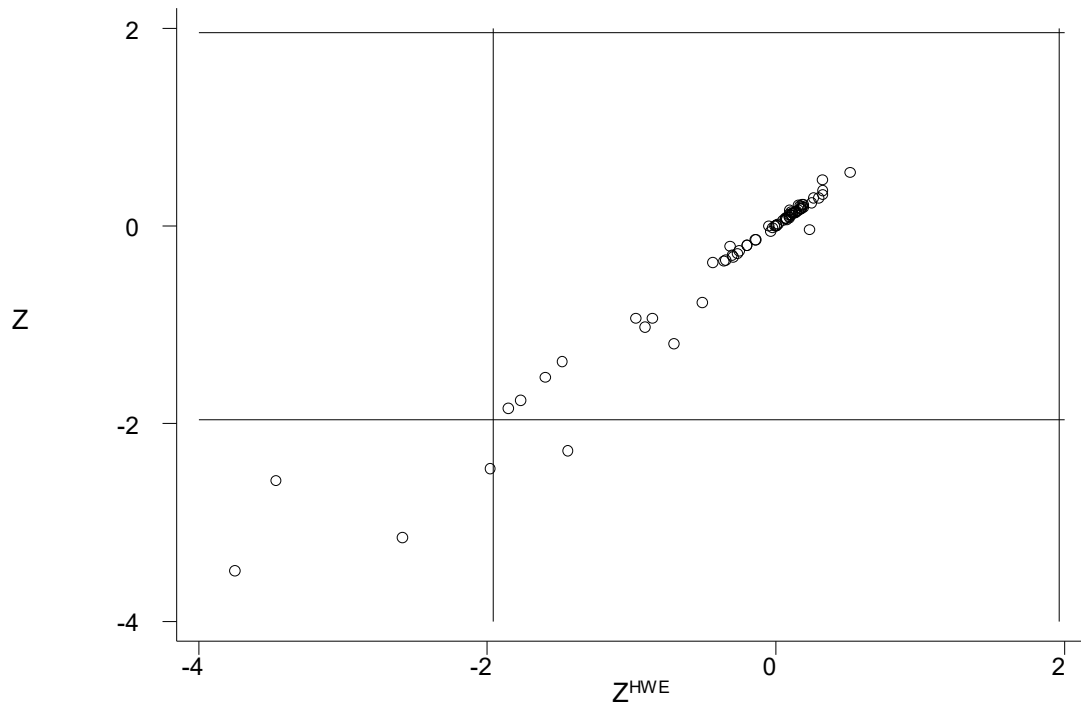
Εικόνα 83: Στον οριζόντιο άξονα είναι ο συντελεστής z , ο λόγος του log odds ratio προς την διακύμανση του log odds ratio με τροποποίηση και στον κάθετο άξονα ο συντελεστής z χωρίς τροποποίηση, για το υπολειπόμενο πρότυπο κληρονομικότητας (σχέση 16). Εδώ ο συντελεστής παλινδρόμησης είναι 0.9469288.



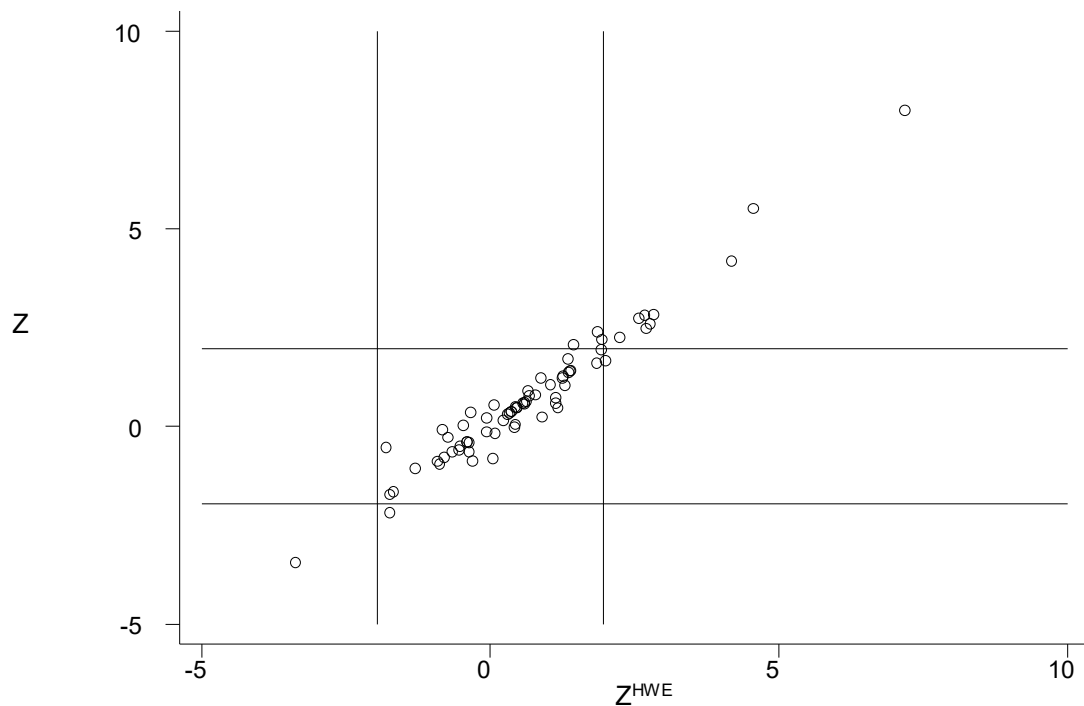
Εικόνα 84: Στον οριζόντιο άξονα είναι ο συντελεστής z , ο λόγος του log odds ratio προς την διακύμανση του log odds ratio με τροποποίηση και στον κάθετο άξονα ο συντελεστής z χωρίς τροποποίηση για το συνεπικρατές μοντέλο κληρονομικότητας, με τη μέθοδο του Sato (σχέση 16). Εδώ ο συντελεστής παλινδρόμησης είναι 1.007728, p -value=0.



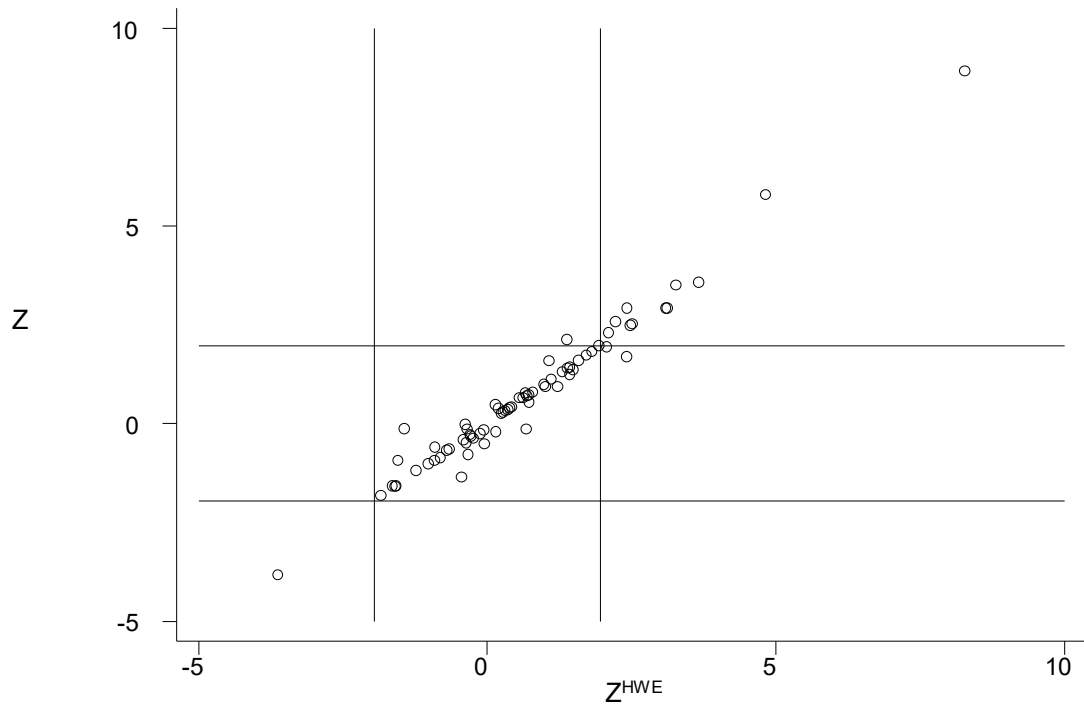
Εικόνα 85: Στον οριζόντιο άξονα είναι ο συντελεστής z , ο λόγος του log odds ratio προς την διακύμανση του log odds ratio με τροποποίηση και στον κάθετο άξονα ο συντελεστής z χωρίς τροποποίηση για το συνεπικρατές μοντέλο κληρονομικότητας, με τη μέθοδο του Ιωαννίδη (σχέση 16). Εδώ ο συντελεστής παλινδρόμησης είναι 0.0038191.



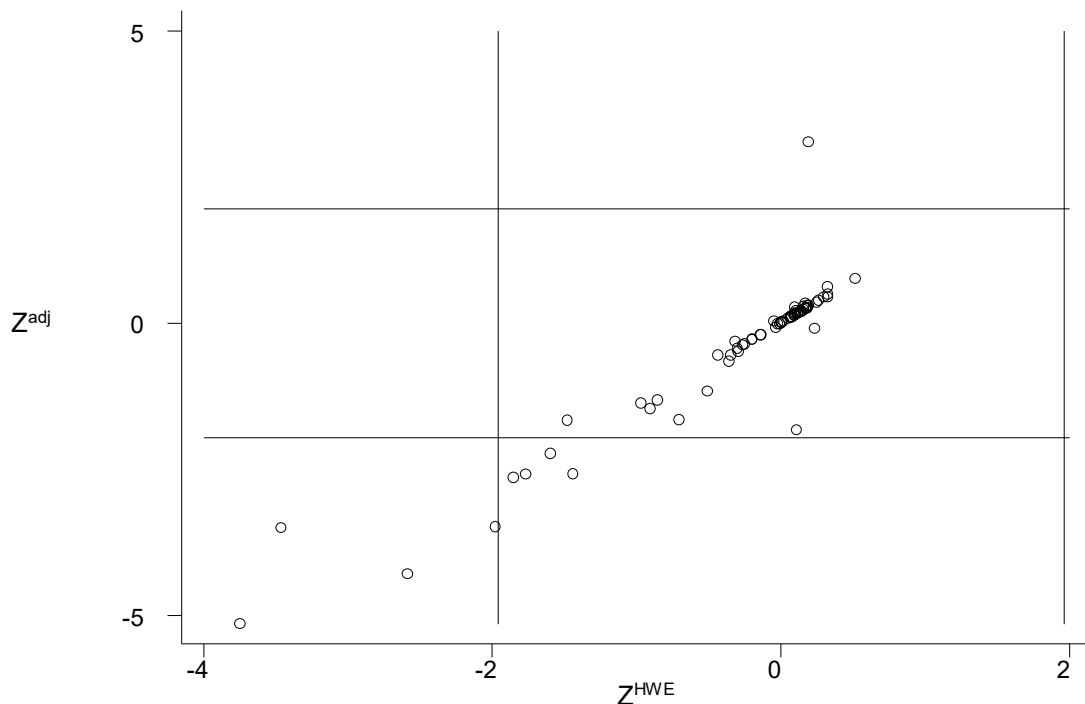
Εικόνα 86: Στον οριζόντιο άξονα είναι ο συντελεστής z , ο λόγος του log odds ratio προς την διακύμανση του log odds ratio που είναι σε ισορροπία HWE και στον κάθετο άξονα ο συντελεστής z χωρίς τροποποίηση, για το επικρατές μοντέλο κληρονομικότητας (σχέση 16) . Εδώ ο συντελεστής παλινδρόμησης β_1 είναι 0.9998176.



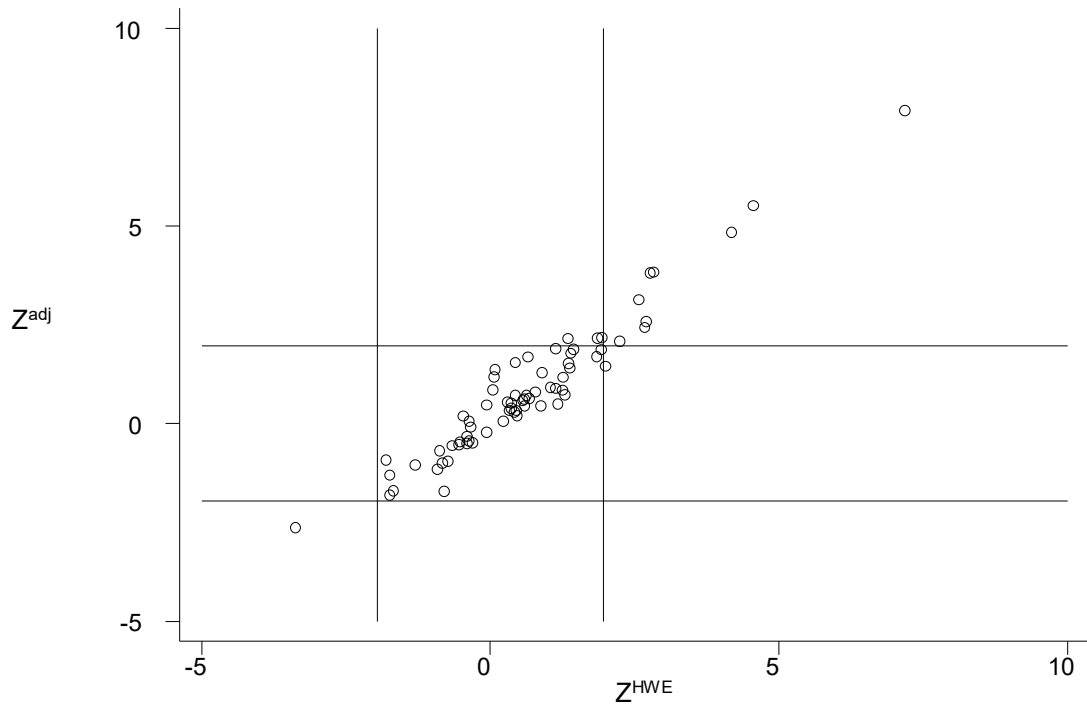
Εικόνα 87: Στον οριζόντιο άξονα είναι ο συντελεστής z , ο λόγος του log odds ratio προς την διακύμανση του log odds ratio που είναι σε ισορροπία HW και στον κάθετο άξονα ο συντελεστής z χωρίς τροποποίηση, για το υπολειπόμενο μοντέλο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης β_1 είναι 1.007728.



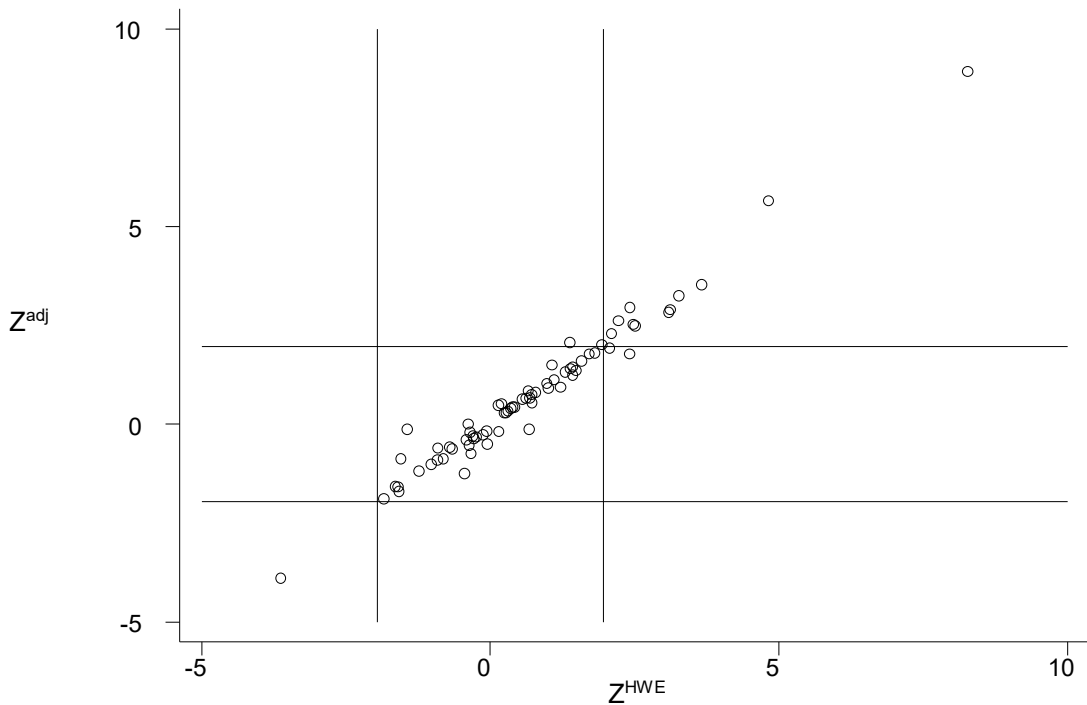
Εικόνα 88: Στον οριζόντιο άξονα είναι ο συντελεστής z , ο λόγος του log odds ratio προς την διακύμανση του log odds ratio που είναι σε ισορροπία HW και στον κάθετο άξονα ο συντελεστής z χωρίς τροποποίηση, για το συνεπικρατές μοντέλο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης β είναι 1.047699.



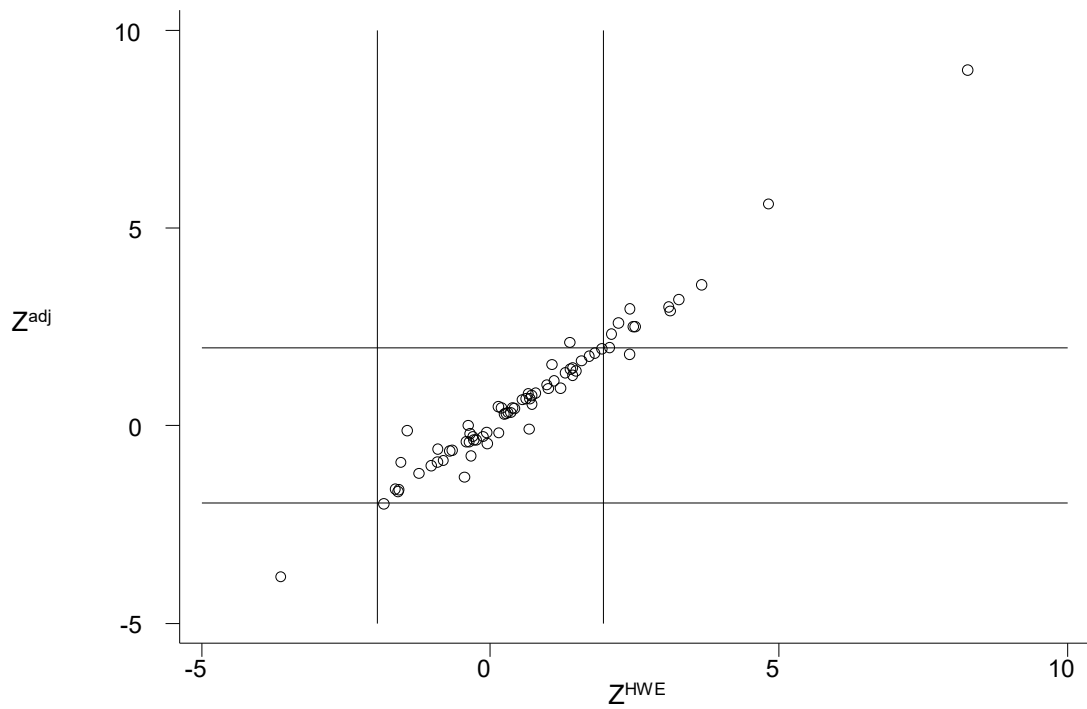
Εικόνα 89: Στον οριζόντιο άξονα είναι ο συντελεστής z , ο λόγος του log odds ratio προς την διακύμανση του log odds ratio των μελετών που έχουν αφαιρεθεί από αυτές όσες δεν είναι σε HWE. Στον κάθετο άξονα ο συντελεστής z με τροποποίηση, για το επικρατές μοντέλο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης β είναι 1.362513



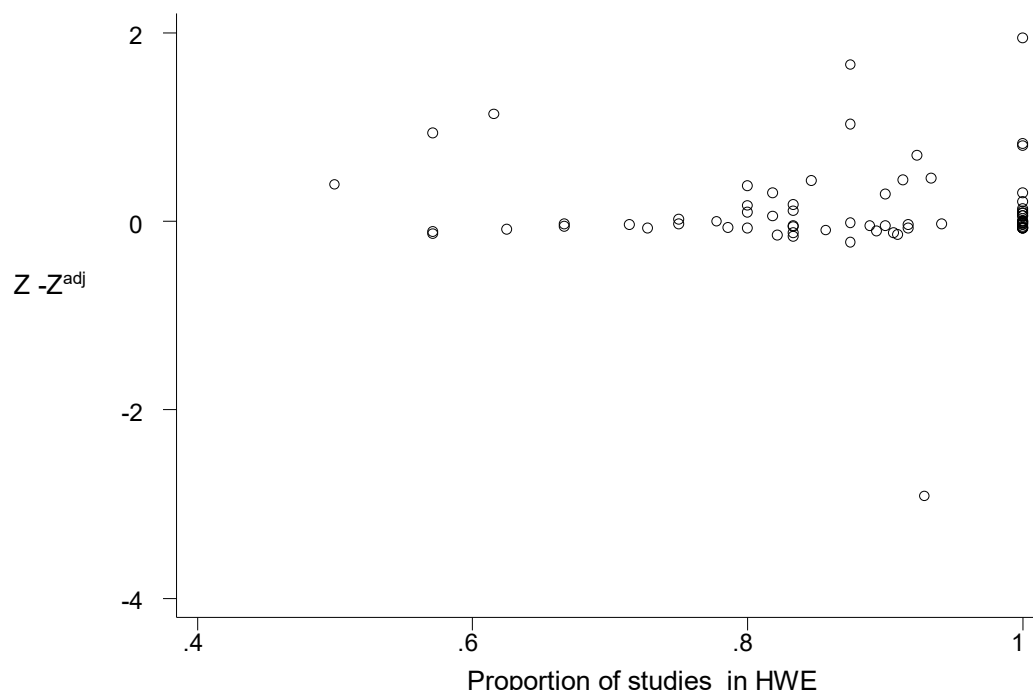
Εικόνα 90: Στον οριζόντιο άξονα είναι ο συντελεστής z , ο λόγος του log odds ratio προς την διακύμανση του log odds ratio των μελετών που έχουν αφαιρεθεί από αυτές όσες δεν είναι σε HWE. Στον κάθετο άξονα ο συντελεστής z με τροποποίηση, για το υπολειπόμενο μοντέλο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης β είναι 1.007933



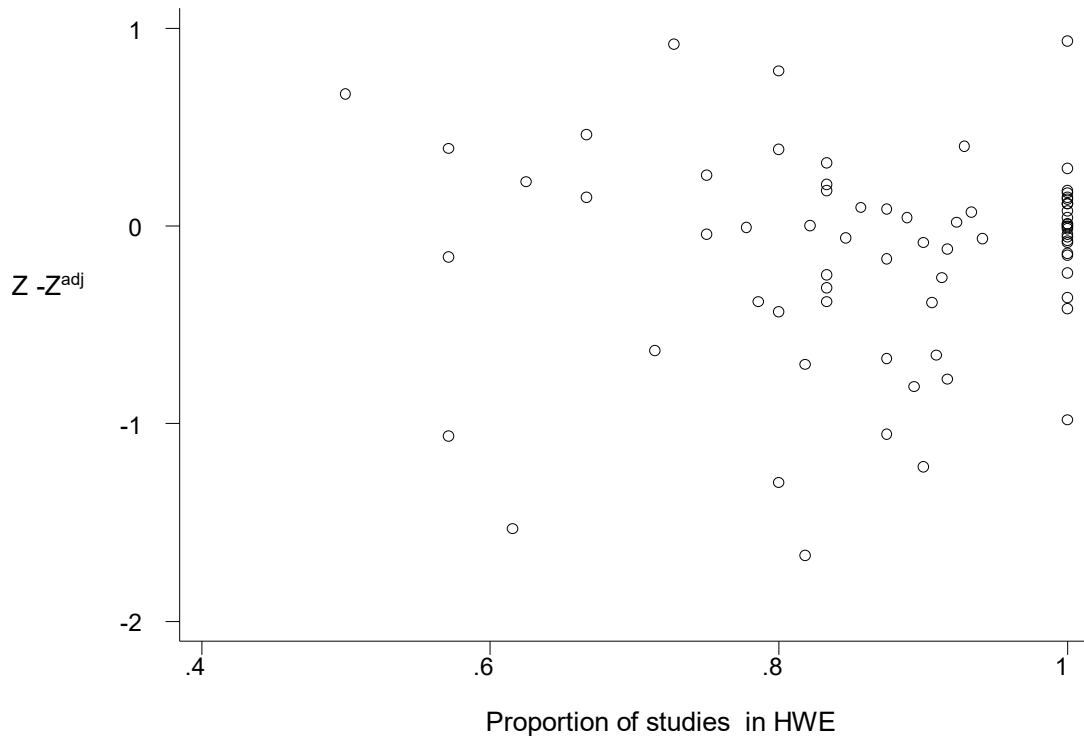
Εικόνα 91: Στον οριζόντιο άξονα είναι ο συντελεστής z , ο λόγος του log odds ratio προς την διακύμανση του log odds ratio των μελετών που έχουν αφαιρεθεί από αυτές όσες δεν είναι σε HWE. Στον κάθετο άξονα ο συντελεστής z με τροποποίηση, για το συνεπικρατές μοντέλο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης β είναι 1.03958



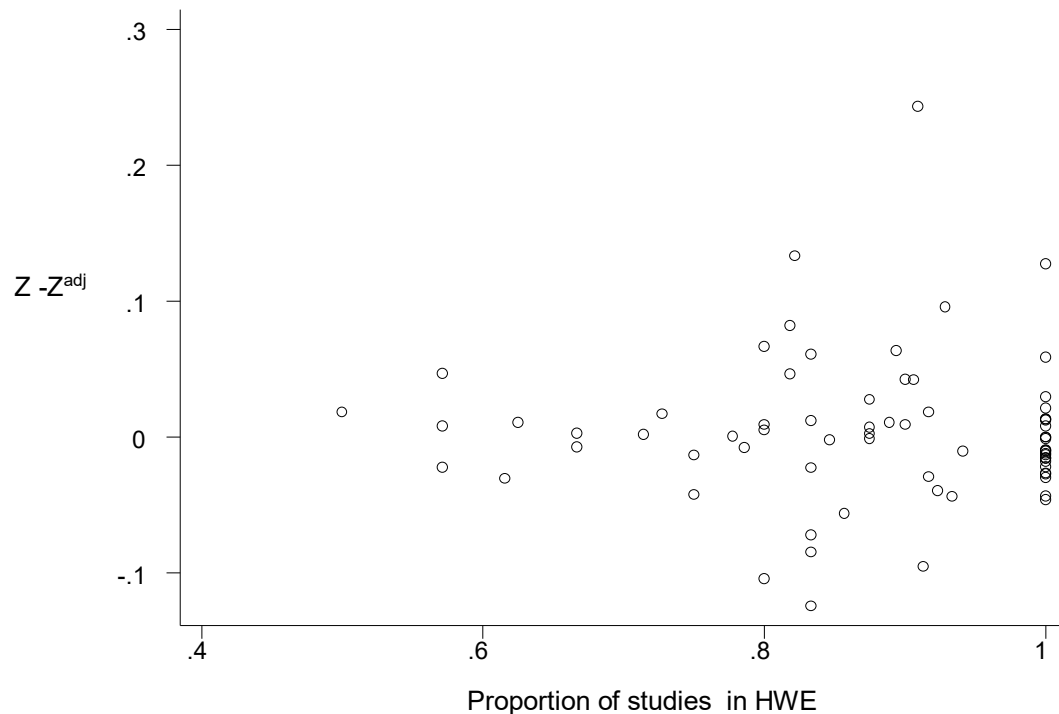
Εικόνα 92: Στον οριζόντιο άξονα είναι ο συντελεστής z, ο λόγος του log odds ratio προς την διακύμανση του log odds ratio των μελετών που έχουν αφαιρεθεί από αυτές όσες δεν είναι σε HWE. Στον κάθετο άξονα ο συντελεστής z με τροποποίηση, για το συνεπικρατές μοντέλο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης β_1 είναι 1.038555



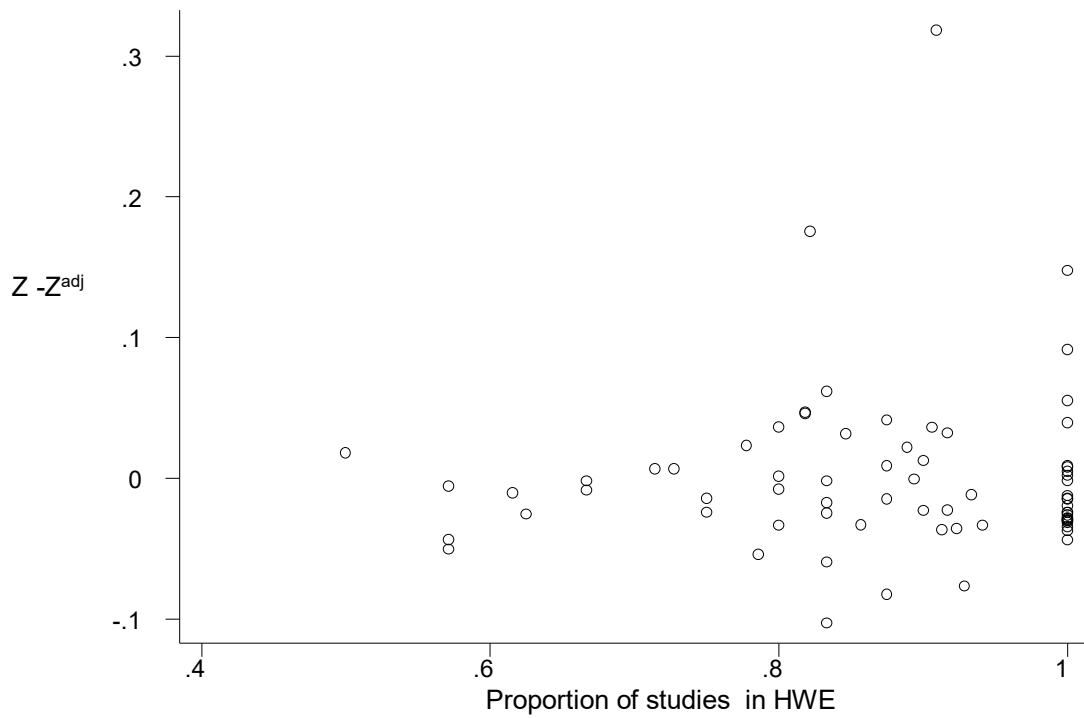
Εικόνα 93: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία και στον κάθετο άξονα η διαφορά των συντελεστών z, δηλαδή η διαφορά των λόγων των log OR προς την διακύμανση, αυτών χωρίς τροποποίηση από αυτές που έχουν υποστεί τροποποίηση, για το επικρατές πρότυπο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης, β_1 είναι 8.44004, $p\text{-value} < 10^{-4}$



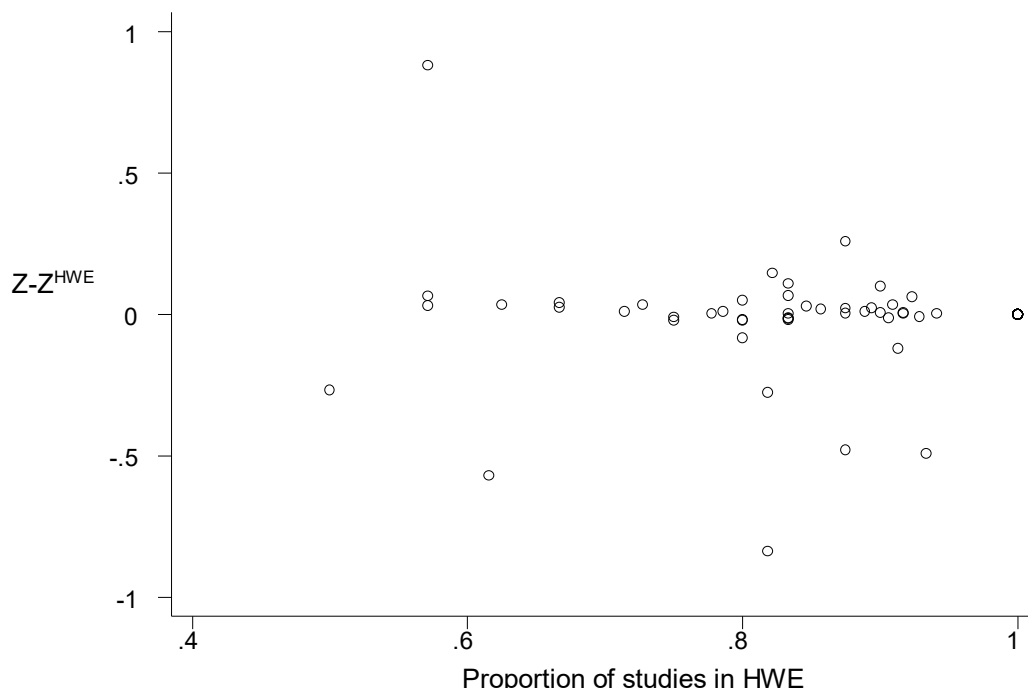
Εικόνα 94: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία και στον κάθετο άξονα η διαφορά των συντελεστών z , δηλαδή η διαφορά των λόγων των $\log OR$ προς την διακύμανση, αυτών χωρίς τροποποίηση από αυτές που έχουν υποστεί τροποποίηση, για το υπολειπόμενο πρότυπο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης, β_1 είναι -4.420961 , $p\text{-value}=0.040$.



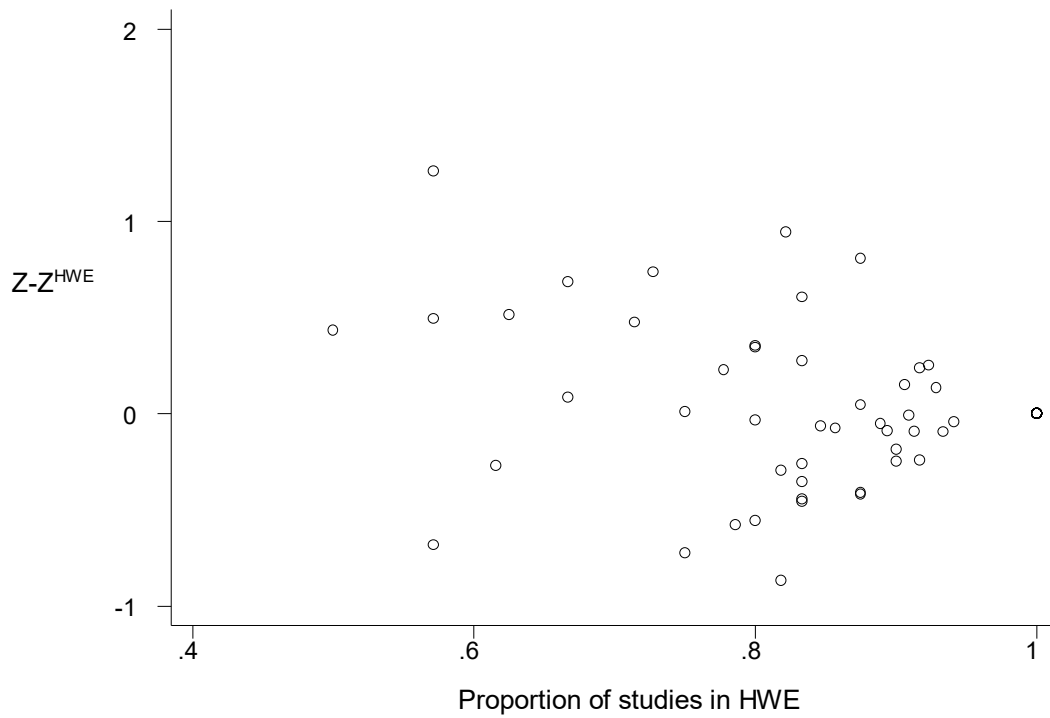
Εικόνα 95: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία και στον κάθετο άξονα η διαφορά των συντελεστών z , δηλαδή η διαφορά των λόγων των $\log OR$ προς την διακύμανση, αυτών χωρίς τροποποίηση από αυτές που έχουν υποστεί τροποποίηση, για το συνεπικρατές πρότυπο κληρονομικότητας, με τη μέθοδο του Sato. Εδώ ο συντελεστής παλινδρόμησης, β_1 είναι -0.0033467 , $p\text{-value}=0.866$.



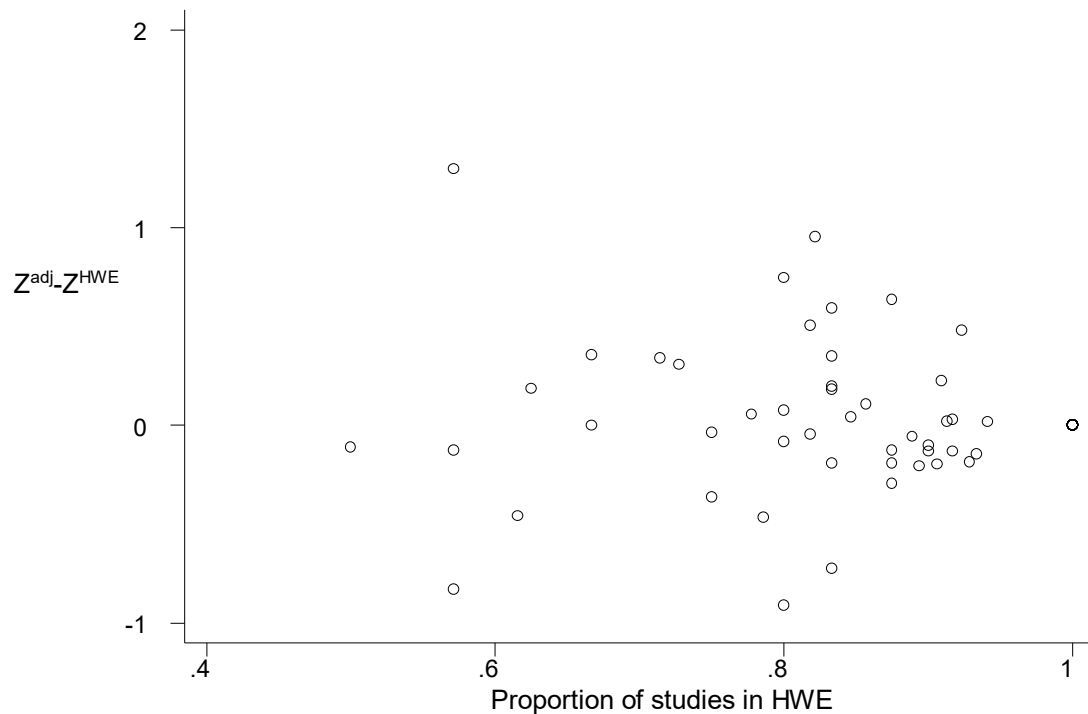
Εικόνα 96: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία και στον κάθετο άξονα η διαφορά των συντελεστών z, δηλαδή η διαφορά των λόγων των log OR προς την διακύμανση, αυτών χωρίς τροποποίηση από αυτές που έχουν υποστεί τροποποίηση, για το συνεπικρατές πρότυπο κληρονομικότητας, με τη μέθοδο του Ιωαννίδη. Εδώ ο συντελεστής παλινδρόμησης, β_1 είναι 0.036069, p-value=0.121



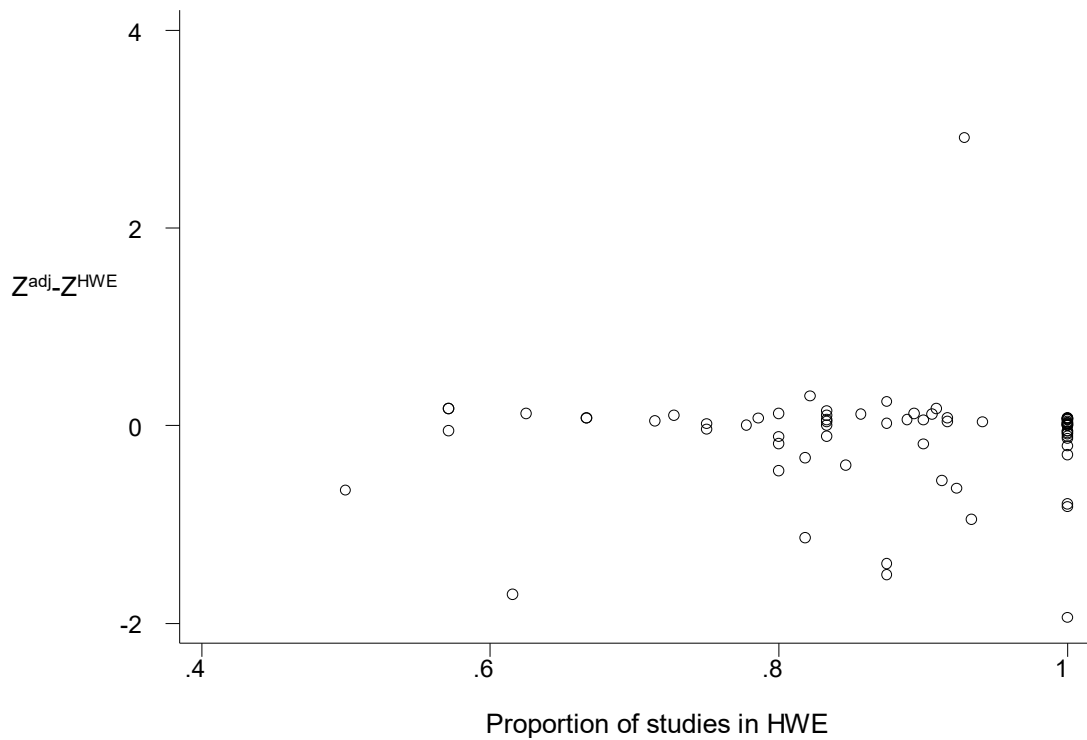
Εικόνα 97: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία και στον κάθετο άξονα η διαφορά των συντελεστών z, δηλαδή η διαφορά των λόγων των log OR προς την διακύμανση, αυτών χωρίς τροποποίηση από αυτές που έχουν αφαιρεθεί όταν δεν είναι σε ισορροπία HW, για το επικρατές πρότυπο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης β_1 είναι 0.0264668, p-value=0.65.



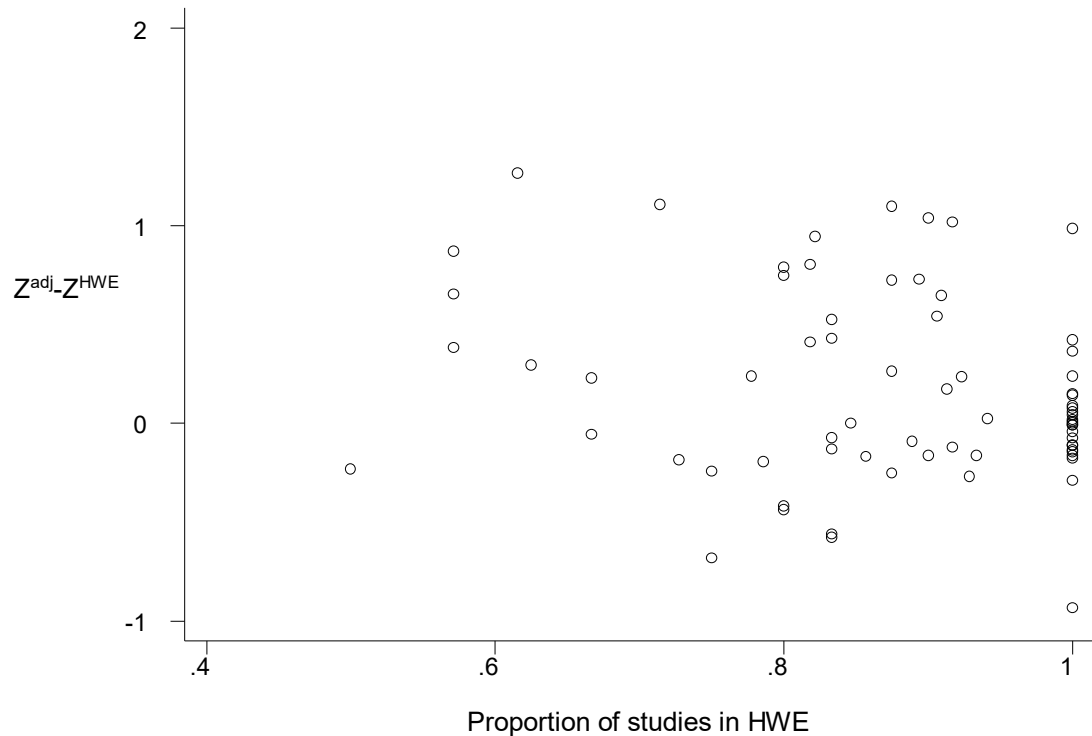
Εικόνα 98: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία και στον κάθετο άξονα η διαφορά των συντελεστών z , δηλαδή η διαφορά των λόγων των $\log OR$ προς την διακύμανση, αυτών χωρίς τροποποίηση από αυτές που έχουν αφαιρεθεί όταν δεν είναι σε ισορροπία HW, για το υπολειπόμενο πρότυπο κληρονομικότητας. Ο συντελεστής παλινδρόμησης β_1 είναι -0.6300376 .



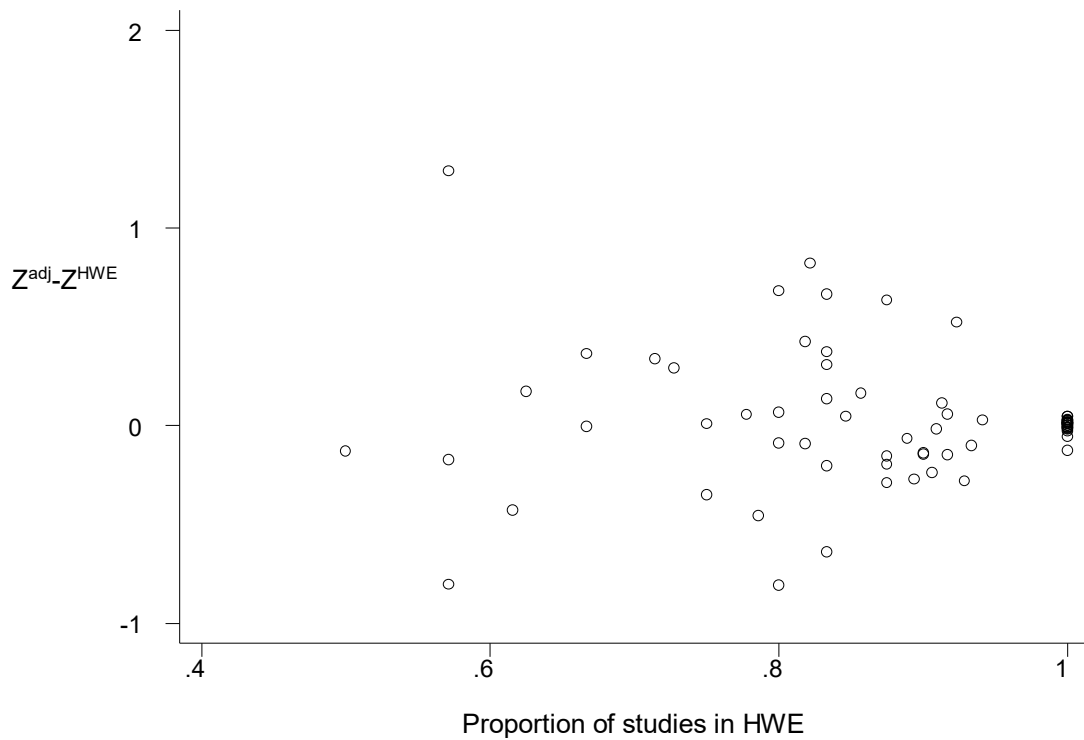
Εικόνα 99: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία και στον κάθετο άξονα η διαφορά των συντελεστών z , δηλαδή η διαφορά των λόγων των $\log OR$ προς την διακύμανση, αυτών χωρίς τροποποίηση από αυτές που έχουν αφαιρεθεί όταν δεν είναι σε ισορροπία HW, για το συνεπικρατές πρότυπο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης β_1 είναι -0.2101224 , $p\text{-value}=0.045$.



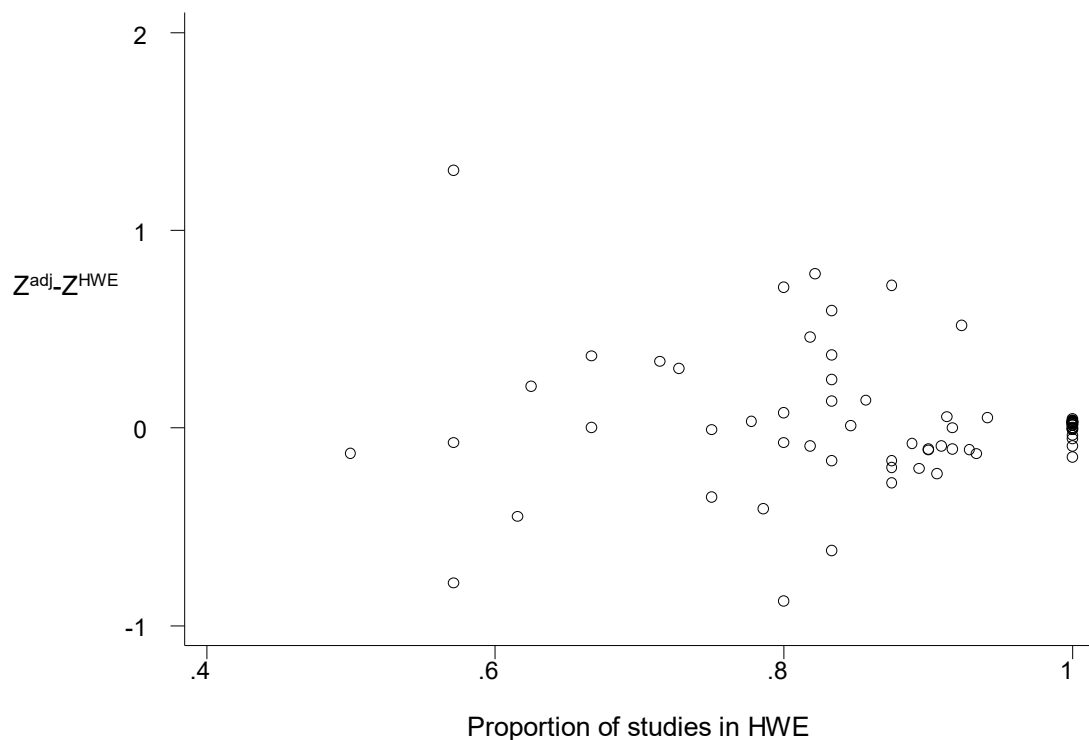
Εικόνα 100: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία και στον κάθετο άξονα η διαφορά των συντελεστών z , δηλαδή η διαφορά των λόγων των $\log OR$ προς την διακύμανση, αυτών με τροποποίηση από αυτές τις μελέτες που έχουν αφαιρεθεί όσες δεν είναι σε ισορροπία HW, για το επικρατές πρότυπο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης $\beta 1$ είναι -8.413574



Εικόνα 101: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία και στον κάθετο άξονα η διαφορά των συντελεστών z , δηλαδή η διαφορά των λόγων των $\log OR$ προς την διακύμανση, αυτών με τροποποίηση από αυτές τις μελέτες που έχουν αφαιρεθεί όσες δεν είναι σε ισορροπία HW, για το υπολειπόμενο πρότυπο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης $\beta 1$ είναι -0.1879416



Εικόνα 102: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία και στον κάθετο άξονα η διαφορά των συντελεστών z , δηλαδή η διαφορά των λόγων των $\log OR$ προς την διακύμανση, αυτών με τροποποίηση από αυτές τις μελέτες που έχουν αφαιρεθεί όσες δεν είναι σε ισορροπία HW, για το συνεπικρατές πρότυπο κληρονομικότητας. Εδώ ο συντελεστής παλινδρόμησης β_1 είναι -0.2067757



Εικόνα 103: Στον οριζόντιο άξονα είναι το ποσοστό των μελετών που είναι σε ισορροπία και στον κάθετο άξονα η διαφορά των συντελεστών z , δηλαδή η διαφορά των λόγων των $\log OR$ προς την διακύμανση, αυτών με τροποποίηση από αυτές τις μελέτες που έχουν αφαιρεθεί όσες δεν είναι σε ισορροπία HW, για το συνεπικρατές πρότυπο κληρονομικότητας (μέθοδος Ιωαννίδη). Εδώ ο συντελεστής παλινδρόμησης β_1 είναι -0.2461914

ΚΕΦΑΛΑΙΟ 4: ΣΥΜΠΕΡΑΣΜΑΤΑ

4.1. ΣΥΜΠΕΡΑΣΜΑΤΑ

Ο κλάδος της Γενετικής Επιδημιολογίας αναπτύσσεται ταχύτατα τα τελευταία χρόνια, καθώς ολοένα και περισσότερες ασθένειες εντοπίζονται και αναλύονται σε επίπεδο γονιδίων και μεντελικής κληρονομικότητας. Τα γονίδια, αλλά και οι περιβαλλοντικοί παράγοντες διαδραματίζουν πρωταρχικό ρόλο στην εκδήλωση νοσημάτων και η ανάγκη για θεραπεία και πρόληψη καθιστά επιτακτική την ανάγκη για έρευνα. Υπό αυτό το πρίσμα, πραγματοποιήθηκε μια συστηματική ανασκόπηση για την ύπαρξη ισορροπίας Hardy-Weinberg στις μετά-αναλύσεις. Σημαντικός φόβος είναι ότι οι συσχετίσεις ενδέχεται να είναι εσφαλμένες και η κατανομή των γονοτύπων στις ομάδες μαρτύρων να αποκλίνει από την ισορροπία. Η HWE εξαρτάται από κάποιες προϋποθέσεις σχετικά με τον έλεγχο του πληθυσμού οι οποίες συμβάλλουν στην ύπαρξη αποκλίσεων από την ισορροπία HW. Για τον λόγο αυτό διάφορες προσεγγίσεις και στατιστικές μέθοδοι έχουν προταθεί προκειμένου να εξεταστούν τέτοιες αποκλίσεις σε μία μετά-ανάλυση (Lathrop 1983 Salanti, 2007 Sato, Shaid and Jacobsen 1999 Tricalinos et al 2000; Minelli, Thompson et al. 2008).

Στην παρούσα εργασία, έγινε λεπτομερής η υλοποίηση των μεθόδων για την ανίχνευση αποκλίσεων από την ισορροπία HW. Τα αποτελέσματα της μετά-ανάλυσης γίνονται με τρεις τρόπους επιλογής: η πρώτη, κάνοντας απλή μετά-ανάλυση χωρίς τροποποιήσεις, η δεύτερη επιλογή, με τροποποιήσεις για τις αποκλίσεις, και η τρίτη η μέθοδος της αφαίρεσης μελετών, όσων δεν ήταν σε ισορροπία HW ώστε να πραγματοποιηθεί μετά-ανάλυση μόνο σε αυτές που είναι σε HWE. Επιλέγονται διάφοροι τρόποι σύγκρισης, μία είναι σε συνάρτηση του ποσοστού όλων των μελετών που είναι σε ισορροπία και μία δεύτερη προσέγγιση σε συνάρτηση του συντελεστή απόκλισης a (σχέση 27). Συγκρίνεται η διαφορά μεταξύ των λόγων πιθανοτήτων, των διακυμάνσεων και του συντελεστή z (σχέση 16).

Με τις υπάρχουσες μεθοδολογίες οι διορθώσεις των εκτιμητών αποφέρουν αλλαγές στα τελικά αποτελέσματα της μετά-ανάλυσης. Στις Εικόνες 29, 30, 31, 32, στο σημείο με συντεταγμένες (1,0) όπου οι διαφορές των λόγων των πιθανοτήτων με τροποποίηση και χωρίς τροποποίηση είναι ίσος με 0, δηλαδή $\log OR - \log OR^{adj} = 0$ και το συνολικό ποσοστό ισορροπίας (proportion of studies in HWE) ισούται με 1, μας επιβεβαιώνει ότι εκείνες οι μελέτες που βρίσκονται στο σημείο είναι σε ισορροπία.

Αυτές που απομακρύνονται από το σημείο αυτό είναι λογικό να αποκλίνουν. Το ίδιο συμπέρασμα παρατηρούμε και στις εικόνες 40, 41, 42, 43 όπου οι διαφορές του λόγου των πιθανοτήτων σε συνάρτηση με τον συντελεστή α δίνει το ίδιο συμπέρασμα για αυτή στις 3 μελέτες. Στο σημείο με συντεταγμένες (0,0) οι συγκεντρωμένες μελέτες αποκλίνουν από την ισορροπία HWE για $\alpha=0$. Σε αυτό το σημείο επιβεβαιώνουμε ότι η μελέτες όντως αποκλίνουν.

Η επιλογή της σύγκρισης των συντελεστών z (σχέση 16) με τους τρεις τρόπους μετά-ανάλυσης για όλα τα πρότυπα κληρονομικότητας, έχει ως σημαντικό πλεονέκτημα την γρήγορη και αποτελεσματική διαπίστωση των μελετών εκείνων που αποκλίνουν από την ισορροπία HW. Στις εικόνες 84 έως 94, οι μελέτες που βρίσκονται εντός των ορίων (-1.96, 1.96) στον άξονα x και y δεν έχουν στατιστική σημαντικότητα. Στην εικόνα 85 οι μελέτες με $z^{adj} < -1.96$ και $z < -1.96$ έχουν στατιστική σημαντικότητα, ενώ αυτές με $z^{adj} > 1.96$ και $z > 1.96$ δεν ήταν σημαντικές και έγιναν. Με την εντολή 38 βρίσκουμε όσες μελέτες αποκλίνουν, όσες δηλαδή βρίσκονται εκτός ορίων.

Για παράδειγμα οι μελέτες οι οποίες αλλάζουν τα αποτελέσματα τους όταν αλλάζει ο τρόπος ανάλυσης, για το υπολειπόμενο πρότυπο κληρονομικότητας είναι δύο, η 18^η μετά-ανάλυση, για το γονίδιο CYPZE1 σε συσχέτιση με την ασθένεια esophageal cancer (εικόνα 89,92) και η 8^η μετά-ανάλυση για το γονίδιο VDRBsmI, σε συσχέτιση με την ασθένεια της οστεοπόρωσης (εικόνα 85). Για το επικρατές πρότυπο κληρονομικότητας έχουμε τέσσερις μελέτες που αποκλίνουν με αριθμό meta 20,42,43 και 48, με γονίδια αντίστοιχα για την 20^η το STK15 T91A, σε συσχέτιση με την ασθένεια cancer (colorectal), για την 42^η το γονίδιο eNOS/G894T (Glu298Asp), σε συσχέτιση με την ασθένεια Preeclampsia, για την 43^η με γονίδιο AGT/T704C (Met235Thr) σε συσχέτιση με την ασθένεια Preeclampsia και για την 48^η για το γονίδιο Uteroglobin (G38A) σε συσχέτιση με την ασθένεια Igt nephropathy (εικόνα 84). Τέλος, για το συνεπικρατές πρότυπο κληρονομικότητας 3 μελέτες αποκλίνουν με αριθμό meta 9,19,26. Η 19^η μετά-ανάλυση για το γονίδιο GSTD1 Ile105Val, σε συσχέτιση με την ασθένεια esophageal, η 9^η για το γονίδιο VDRTaql σε συσχέτιση με την ασθένεια pulmonary tuberculosis (εικόνα 87) και η 26^η για το MTFHR C677T γονίδιο, σε συσχέτιση με τον καρκίνο του στήθους (εικόνα 90, 93, 94).

Τα αποτελέσματα της στατιστικής σημαντικότητας (z) αλλάζουν σε ποσοστό 5,63% για το επικρατές μοντέλο, 2,8% για το υπολειπόμενο μοντέλο και 4,22% για το συνεπικρατές μοντέλο κληρονομικότητας. Συνολικά 9 από τις 71 μελέτες (12,7%) η στατιστική σημαντικότητα αλλάζει.

Τα συμπεράσματα από την παρούσα έρευνα μας είναι ότι οι αποκλίσεις από την ισορροπία HWE πρέπει να λαμβάνονται υπόψη. Αξιοσημείωτο ενδιαφέρον παρουσιάζει η σύγκριση των διαφορετικών μεθόδων που γίνονται για τις διακυμάνσεις. Τα τρία είδη διακυμάνσεων είναι οι τροποποιημένες, οι μη τροποποιημένες και αυτές που έχουν αφαιρεθεί όσες δεν είναι σε ισορροπία HW. Έχοντας υπολογίσει τα 3 είδη μπορούμε να τις συγκρίνουμε μεταξύ τους και να δούμε ποια από αυτές μικραίνει ή όχι. Αν η διαφορά μεταξύ τροποποιημένων και μη-τροποποιημένων διακυμάνσεων $V-V^{adj}$ είναι μεγαλύτερη του μηδενός (εικόνα 62,63,64,65) τότε η τροποποιημένη διακύμανση είναι μικρότερη από αυτές χωρίς τροποποίηση.

Καταλήγουμε στο συμπέρασμα ότι οι τροποποιημένες δίνουν στατιστικά σημαντικότερα αποτελέσματα. Το ίδιο διαπιστώνουμε για τις τροποποιημένες διακυμάνσεις σε σύγκριση με εκείνες που είχαν αφαιρεθεί όσες δεν ήταν σε ισορροπία HW, $Var^{adj} < Var^{HWE}$, (εικόνα 69,70,71,72). Αντιθέτως αν παρατηρήσουμε την διαφορά των μη τροποποιημένων και αυτών που αρχικά είχαν αφαιρεθεί όσες δεν είναι σε ισορροπία HW συμπεραίνουμε ότι $Var < Var^{HWE}$ (εικόνα 66,67,68), δηλαδή οι μη τροποποιημένες δίνουν στατιστικά σημαντικά αποτελέσματα. Το τελικό συμπέρασμα σχετικά με τις μεθόδους ρύθμισης των διακυμάνσεων είναι ότι η επιλογή να αφαιρούνται οι μελέτες που δεν είναι σε ισορροπία οδηγεί σε απώλεια στατιστικής ισχύος.

Αξιοσημείωτο ενδιαφέρον παρουσιάζουν οι δύο μέθοδοι ρύθμισης του Sato και του Ιωαννίδη (σχέση 34,35). Οι δύο τροποποιημένες διακυμάνσεις δεν διαφέρουν μεταξύ τους διότι όλα τα σημεία βρίσκονται στην ευθεία $x=y$ που περνάει από την αρχή των αξόνων. Επίσης κάνοντας συγκρίσεις του συνολικού ποσοστού που βρίσκονται σε ισορροπία HW (εικόνα 71,72) ή σε σύγκριση με τον συντελεστή απόκλισης α (εικόνα 75,76) μεταξύ της διαφοράς των τροποποιημένων διακυμάνσεων, είτε με τη μέθοδο του Sato είτε με τη μέθοδο του Ιωαννίδη, εξάγονται τα ίδια αποτελέσματα (εικόνα 64,65). Εξετάζοντας και τις τρεις επιλογές, συμπεραίνουμε ότι η μέθοδος της τροποποίησης είναι περισσότερο αξιόπιστη σε σύγκριση με τις υπόλοιπες.

Τέλος, στην παρούσα εργασία μας υπολογίζαμε ξεχωριστά τους εκτιμητές (odds ratio, variance, z) και είχαμε εξάγει αποτελέσματα για κάθε πρότυπο κληρονομικότητας (επικρατές, υπολειπόμενο, συνεπικρατές). Γι' αυτόν το λόγο θα ήταν χρήσιμη η ύπαρξη μιας μαθηματικής εξίσωσης που θα υπολογίζει τους διάφορους εκτιμητές που υπολογίζονται στην μετά-ανάλυση ανεξάρτητα από το μοντέλο κληρονομικότητας.

Σκοπός της πτυχιακής εργασίας, είναι να συμπεριλάβουμε όλες τις δυνατές επιλογές για μετά-ανάλυση, ώστε να τις συγκρίνουμε και να ερευνήσουμε και να παρουσιάσουμε τις δυνατότητες που προσφέρει η κάθε μία, καθώς τα αποτελέσματά αλλάζουν στατιστικώς σημαντικά. Στα πρακτικά, οι ερευνητές θα πρέπει να εφαρμόζουν και να δοκιμάζουν όλες τις επιλογές για την έρευνα τους όσο αφορά τις αποκλίσεις από την ισορροπία HW.

ΚΕΦΑΛΑΙΟ 5: ΒΙΒΛΙΟΓΡΑΦΙΑ

BIBΛΙΟΓΡΑΦΙΑ

- Bagos, P. G. (2008). "A unification of multivariate methods for meta-analysis of genetic association studies." Stat Appl Genet Mol Biol **7**(1): Article31.
- Bagos, P. G. and G. K. Nikolopoulos (2007). "A method for meta-analysis of case-control genetic association studies using logistic regression." Stat Appl Genet Mol Biol **6**: Article17.
- Bland, J. M. and D. G. Altman (2000). "Statistics notes. The odds ratio." BMJ **320**(7247): 1468.
- Brown, A. H. (1970). "The estimation of Wright's fixation index from genotypic frequencies." Genetica **41**(3): 399-406.
- Pereira, C. and A. Rogatko (1984). "The Hardy-Weinberg equilibrium under a Bayesian perspective." Rev Bras Genet **4**: 489-707.
- Emigh (1980). "comparison of tests for Hardy-Weinberg equilibrium. ." Biometrics **36**(627-42).
- Emigh, T. (1980). "A comparison of tests for Hardy-Weinberg equilibrium." Biometrics **36**: 627-642.
- Hardy, G. H. (1908). "Mendelian Proportions in a Mixed Population." Science **28**(706): 49-50.
- Hernandez, J. L. and W. BS. (1989). "A disequilibrium coefficient approach to Hardy-Weinberg testing." Biometrics **45**(1): 53-70.
- Hernandez, J. L. and B. S. Weir (1989). "A disequilibrium coefficient approach to Hardy-Weinberg testing." Biometrics **45**(1): 53-70.
- Khoury, M. J. (1997). "Genetic epidemiology and the future of disease prevention and public health." Epidemiol Rev **19**(1): 175-180.
- Khoury, M. J. and J. Little (2000). "Human genome epidemiologic reviews: the beginning of something HuGE." Am J Epidemiol **151**(1): 2-3.
- Lathrop, G. M. (1983). "Estimating genotype relative risks." Tissue Antigens **22**(2): 160-166.
- Lindley (1988). "Statistical inference concerning Hardy-Weinberg equilibrium." Bayesian Stat **3**: 307-326.
- Lindley, D. (1988). "Statistical inference concerning Hardy-Weinberg equilibrium." Bayesian Stat **3**: 307-326.

- Thomson, M.W. and R.RMcInnes Ιατρική Γενετική. Thomson & Thomson: 162-188.
- Mendell, N. R. and G. A. Simon (1984). "A general expression for the variance-covariance matrix of estimates of gene frequency: the effects of departures from Hardy-Weinberg equilibrium." Ann Hum Genet **48**(Pt 3): 283-286.
- Minelli, C., J. R. Thompson, et al. (2008). "How should we use information about HWE in the meta-analyses of genetic association studies?" Int J Epidemiol **37**(1): 136-146.
- Normand, S. L. (1999). "Meta-analysis: formulating, evaluating, combining, and reporting." Stat Med **18**(3): 321-359.
- Pereira and Rogatko (1984). "The Hardy-Weinberg equilibrium under a Bayesian perspective." Rev Bras Genet **4**: 689-707.
- Petiti, D. (1994). "Meta-analysis Decision Analysis and Cost-Effectiveness Analysis." Oxford University Press.
- Sasieni, P. D. (1997). "From genotypes to genes: doubling the sample size." Biometrics **53**(4): 1253-1261.
- Sato, Y., H. Suganami, et al. (2006). "The confidence interval of allelic odds ratios under the Hardy-Weinberg disequilibrium." J Hum Genet **51**(9): 772-780.
- Shaid and S. J. Jacobsen (1999). "Biased tests of association: comparisons of allele frequencies when departing from Hardy-Weinberg proportions." Am J Epidemiol **149**(8): 706-711.
- Shoemaker, J., I. Painter, et al. (1998). "A Bayesian characterization of Hardy-Weinberg disequilibrium." Genetics **149**(4): 2079-2088.
- Thakkinstian, A., P. McElduff, et al. (2005). "A method for meta-analysis of molecular association studies." Stat Med **24**(9): 1291-1306.
- Trikalinos, T. A., G. Salanti, et al. (2006). "Impact of violations and deviations in Hardy-Weinberg equilibrium on postulated gene-disease associations." Am J Epidemiol **163**(4): 300-309.
- Weir, B. S. (1970). "Equilibria under Inbreeding and Selection." Genetics **65**(2): 371-378.
- Weir, B. S. (1970). "Equilibria under Inbreeding and Selection." Genetics **65**.
- Weiss, S. T., E. K. Silverman, et al. (2001). "Case-control association studies in pharmacogenetics." Pharmacogenomics J **1**(3): 157-158.
- Wigginton, J. E., D. J. Cutler, et al. (2005). "A note on exact tests of Hardy-Weinberg equilibrium." Am J Hum Genet **76**(5): 887-893.

- Xu, J., A. Turner, et al. (2002). "Positive results in association studies are associated with departure from Hardy-Weinberg equilibrium: hint for genotyping error?" Hum Genet **111**(6): 573-574.
- Zhou, J. J., K. Lange, et al. (2009). "A heterozygote-homozygote test of Hardy-Weinberg equilibrium." Eur J Hum Genet **17**(11): 1495-1500.
- Zintzaras, E. (2008). "Variance estimation of allele-based odds ratio in the absence of Hardy-Weinberg equilibrium." Eur J Epidemiol **23**(5): 323-326.

6. ΠΑΡΑΡΤΗΜΑ

meta	pmid	study	aa1	ab1	bb1	aa0	ab0	bb0	gene	disease	author	year
1	16229232	6	105	12	3	203	32	5	VDR Taq I	pulmonary tuberculosis	Lewis	2005
2	16229232	5	9	32	59	14	78	109	VDR Fok I	pulmonary tuberculosis	Lewis	2005
3	16172806	3	32	49	26	24	55	24	BDNF G196A	Parkinson	Zintzaras	2005
4	15987731	9	13	40	19	25	51	35	B2ADR Arg/Gly16	Asthma	Thakkinstian	2005
5	15987731	12	23	12	2	107	36	12	B2ADR Gln/Glu27	Asthma	Thakkinstian	2005
6	15961581	8	21	42	54	22	47	72	CTLA-4 +49A/G	type I diabetes	Kavvoura	2005
7	15947278	3	16	27	13	18	14	4	MTHFR C677T	Ischemic Stroke	Cronin	2005
8	15902512	3	33	50	15	100	132	36	MTHFR C677T	diabetic retinopathy	Zintzaras	2005
9	15861039	5	0	12	3	0	14	8	CYP2D6(LF/Wt)	tardive dyskinesia(schizophrenia)	Patsopoulos	2005
10	15861039	3	22	46	11	20	36	13	CYP2D6(10/Wt)	tardive dyskinesia(schizophrenia)	Patsopoulos	2005
11	15857569	6	5	13	3	14	18	10	HTR2A T102C	tardive dyskinesia	Lerer	2005
12	15849806	9	21	56	50	31	48	22	CYP1A1 Ile-Val	esophageal cancer	Yang	2005
13	15849806	2	22	12	0	207	37	3	CYP1A1 MspI	esophageal cancer	Yang	2005
14	15849806	3	36	6	3	20	22	3	CYP2E1	esophageal cancer	Yang	2005
15	15849806	6	19	22	4	26	12	7	GSTP1 Ile105Val	esophageal cancer	Yang	2005
16	15830182	20	35	68	31	57	130	39	ACE I/D	diabetic nephropathy	D. P. K. Ng	2005
17	15802297	9	121	491	490	149	503	534	STK15 T91A	Cancer (breast)	Ewart-Toland	2005
18	15729744	4	81	98	31	79	106	33	MTHFR C677T	Schizophrenia	Lewis	2005
19	15682272	5	62	58	12	53	92	17	GLUT1 XbaI(0.9/1.1)	diabetic nephropathy	Zintzaras	2005
20	15592940	6	63	17	0	109	12	0	NEUROD1 Ala45Thr	type I diabetes	Kavvoura	2005
21	15592940	7	85	107	31	78	68	25	NEUROD1 Ala45Thr	type II diabetes	Kavvoura	2005
22	16133179	4	10	27	23	20	62	68	CTLA-4 +49A/G	rheumatoid arthritis	Han	2005
23	17194198	21	35	48	13	46	40	10	ACE I/D	Preeclampsia	Serrano	2006
24	17105984	8	58	96	32	50	80	17	MTFHR C677T	Breast Cancer	Lewis	2006
25	17074966	2	22	15	2	11	9	0	MTFHR C677T	Depression	Gilbody	2006
26	17074966	8	199	186	41	306	260	62	MTFHR C677T	Schizophrenia	Gilbody	2006
27	17074966	4	108	73	19	210	79	11	MTFHR C677T	Bipolar Disorder	Gilbody	2006
28	16905645	8	69	68	18	105	142	28	VDR Fok I	Type I Diabetes	Guo	2006

29	16905645	13	15	40	16	17	44	27	VDR Bsm I	Type I Diabetes	Guo	2006
30	16905645	9	15	37	19	28	43	17	VDR Apa I	Type I Diabetes	Guo	2006
31	16905558	3	192	454	306	181	169	53	CFH Y402H	age-related macular degeneration	Thakkinstian	2006
32	16897583	9	195	201	47	184	152	43	MTFHR C677T	acute lymphoblastic leukemia	Zintzaras	2006
33	16897583	3	122	108	17	279	249	69	MTFHR A1289C	acute lymphoblastic leukemia	Zintzaras	2006
34	16773678	3	56	140	72	82	190	98	MDR1 C3435T	Cronn's Disease	Annese	2006
35	16773678	2	26	77	46	39	76	34	MDR1 C3435T	Ulcerative Colitis	Annese	2006
36	16758123	5	105	171	44	113	172	28	MTFHR C677T	Gastric Cancer	Zintzaras	2006
37	16758123	4	219	93	8	230	73	10	MTFHR A1289C	Gastric Cancer	Zintzaras	2006
38	16740590	5	144	111	18	158	111	19	Factor XIII Val34Leu	Venous Thromboembolism	Wells	2006
39	17171228	10	30	49	21	32	44	24	MTFHR C677T	Ischemic Stroke	Banerjee	2006
40	17171228	5	128	166	71	116	154	49	ACE I/D	Ischemic Stroke	Banerjee	2006
41	17112651	2	86	26	0	294	41	0	eNOS/G894T (Glu298Asp)	Preeclampsia	Medica	2006
42	17112651	6	12	21	10	22	43	19	AGT/T704C (Met235Thr)	Preeclampsia	Medica	2006
43	17063279	11	253	906	815	980	1040	265	XRCC3 Thr241Met	Breast Cancer	Sang-Ah	2006
44	16820192	4	79	77	23	59	79	34	ITGA2 C807T	Ischemic Stroke	Nikolopoulos	2006
45	16513329	10	60	46	11	89	88	23	DRD3 Ser9Gly	tardive dyskinesia	Bakker	2006
46	16517578	8	47	52	26	44	41	11	TNF-alpha -308 G/A	Asthma	Gao	2006
47	16797381	4	87	109	43	69	95	32	Uteroglobin (G38A)	IgA nephropathy	Yong	2006
48	16940230	15	57	31	3	45	36	6	eNOS/G894T (Glu298Asp)	hypertension	Zintzaras	2006
49	16940230	9	158	37	0	151	27	4	eNoS (4 b/a)	hypertension	Zintzaras	2006
50	16940230	4	26	50	29	27	46	21	eNOS (T786C)	hypertension	Zintzaras	2006
51	17020975	1	88	106	39	69	84	23	RNASEL (Arg462Gln)	prostate cancer	Huihua Li	2006
52	17020975	3	233	668	462	162	372	257	RNASEL (Asp541Glu)	prostate cancer	Huihua Li	2006
53	17264402	14	18	54	3	17	42	7	VDR Bsm I	osteoporosis	Zintzaras	2006
54	17264402	6	19	13	3	8	27	9	VDR Taq I	osteoporosis	Zintzaras	2006
55	17264402	1	12	19	9	38	59	31	VDR Apa I	osteoporosis	Zintzaras	2006
56	16691128	4	15	16	3	45	35	17	MTFHR C677T	Schizophrenia	Zintzaras	2006
57	16691128	3	99	60	8	80	33	7	MTFHR C677T	Bipolar Disorder	Zintzaras	2006
58	16691128	1	15	20	5	154	214	51	MTFHR C677T	Non Bipolar Depression	Zintzaras	2006
59	16691128	4	177	209	40	292	272	64	MTFHR A1298C	Schizophrenia	Zintzaras	2006
60	16630166	1	30	19	5	25	21	11	MTFHR C677T	Breast Cancer	Zintzaras	2006
61	16630166	3	94	102	27	133	127	38	MTFHR A1298C	Breast Cancer	Zintzaras	2006
62	16449362	3	243	341	145	164	208	72	PADI4_94 & PADI4_89	Rheumatoid Arthritis	Iwamoto	2006
63	16435105	3	119	207	91	178	249	110	PPARGC1A Gly482Ser	Type II Diabetes	Barroso	2006

64	16571112	1	63	48	10	64	53	4	NAT2 c.590G>A	Parkinson Disease	Borlak	2006
65	16600714	8	71	85	34	81	87	22	VDR Taq I	Prostate Cancer	Berndt	2006
66	16600714	3	21	69	40	16	53	31	VDR Apa I	Prostate Cancer	Berndt	2006
67	16600714	3	1	1	1	4	1	1	Poly(A)	Prostate Cancer	Berndt	2006
68	16600714	7	135	185	52	201	300	90	VDR Bsm I	Prostate Cancer	Berndt	2006
69	16600714	7	51	95	41	87	153	62	VDR Fok I	Prostate Cancer	Berndt	2006
70	16509765	2	196	187	62	299	255	68	GSTP1 Ile105Val	Lung Cancer	Ye	2006
71	16509765	1	284	93	12	256	84	11	GSTM3 3bp deletion in intron 6	Lung Cancer	Ye	2006

Πίνακας 5: Οι 71 μετά-αναλύσεις, του έτους 2005-2006, με αριθμό μετά-ανάλυσης, τον κωδικό PMID, οι γονότυποι σθενών μαρτύρων, η ασθένεια και το γονίδιο που ευθύνεται για αυτήν.