



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ
ΙΑΤΡΙΚΗ ΣΧΟΛΗ**

**ΤΟΜΕΑΣ ΚΟΙΝΩΝΙΚΗΣ ΙΑΤΡΙΚΗΣ ΚΑΙ ΨΥΧΙΚΗΣ ΥΓΕΙΑΣ
ΕΡΓΑΣΤΗΡΙΟ ΥΓΙΕΙΝΗΣ ΚΑΙ ΕΠΙΔΗΜΙΟΛΟΓΙΑΣ**

**Μελέτες πολλαπλών γενετικών παραγόντων
κινδύνου σε πολυσύνθετα νοσήματα**

Νικόλαος Α. Πατσόπουλος
Ιατρός

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ



ΙΩΑΝΝΙΝΑ 2008



Η έγκριση της διδακτορικής διατριβής από την Ιατρική Σχολή του Πανεπιστημίου Ιωαννίνων δεν υποδηλώνει αποδοχή των γνώμων του συγγραφέα (Ν. 5343/32, άρθρο 202, παράγραφος 2).



Ημερομηνία αίτησης του κ. Πατσόπουλου Νικολάου: 14-9-2005

Ημερομηνία ορισμού Τριμελούς Συμβουλευτικής Επιτροπής: 563^α/4-10-2005

Μέλη Τριμελούς Συμβουλευτικής Επιτροπής:

Επβλέπων

Ιωαννίδης Ιωάννης, Καθηγητής Υγιεινής

Μέλη

Τσατσούλης Αγαθοκλής Καθηγητής Παθολογίας-Ενδοκρινολογίας
Αλαμάνος Ιωάννης, Επίκουρος Καθηγητής Υγιεινής

Ημερομηνία ορισμού θέματος: 17-10-2005

«Μελέτες πολλαπλών γενετικών παραγόντων κινδύνου σε πολυσύνθετα νοσήματα »

ΔΙΟΡΙΣΜΟΣ ΕΠΤΑΜΕΛΟΥΣ ΕΞΕΤΑΣΤΙΚΗΣ ΕΠΙΤΡΟΠΗΣ : 648^ο/18-11-2008

1. Ιωαννίδης Ιωάννης Καθηγητής Υγιεινής Ιατρικής Σχολής Πανεπιστημίου Ιωαννίνων
2. Κυρίτσης Αθανάσιος Καθηγητής Νευρολογίας Ιατρικής Σχολής Πανεπιστημίου Ιωαννίνων
3. Παυλίδης Νικόλαος Καθηγητής Παθολογίας-Ογκολογίας Ιατρικής Σχολής Πανεπιστημίου Ιωαννίνων
4. Τσατσούλης Αγαθοκλής Καθηγητής Παθολογίας- Ενδοκρινολογίας Ιατρικής Σχολής Πανεπιστημίου Ιωαννίνων
5. Αλαμάνος Ιωάννης Αναπληρωτής Καθηγητής Υγιεινής, Ιατρικής Σχολής Παν/μίου Πατρών
6. Δημολιάτης Ιωάννης Επίκουρος Καθηγητής Υγιεινής Ιατρικής Σχολής Πανεπιστημίου Ιωαννίνων
7. Σκαπινάκης Πέτρος Επίκουρος Καθηγητής Ψυχιατρικής Ιατρικής Σχολής Πανεπιστημίου Ιωαννίνων

Έγκριση Διδακτορικής Διατριβής με βαθμό «ΑΡΙΣΤΑ» στις 4-12-2008



ΠΡΟΕΔΡΟΣ ΤΗΣ ΙΑΤΡΙΚΗΣ ΣΧΟΛΗΣ

Γουδέβενος Ιωάννης

Καθηγητής Παθολογίας-Καρδιολογίας

Η Γραμματεας της Σχολής

ΕΥΑΓΓΕΛΙΑ ΤΣΑΤΣΙΔΑ



ΕΥΧΑΡΙΣΤΙΕΣ

Ευχαριστώ όλους όσους βοήθησαν στην παρούσα διδακτορική διατριβή.



ΠΕΡΙΕΧΟΜΕΝΑ

ΕΥΧΑΡΙΣΤΙΕΣ.....	ix
ΠΕΡΙΕΧΟΜΕΝΑ.....	xi
ΠΙΝΑΚΕΣ.....	xiii
ΕΙΚΟΝΕΣ.....	xv
ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ.....	xvii
ΚΕΦΑΛΑΙΟ 1: ΓΕΝΙΚΟ ΜΕΡΟΣ.....	1
1.1 Εισαγωγή στη Γενετική Επιδημιολογία.....	1
1.2 Το ανθρώπινο γονιδίωμα και η ποικιλότητά του.....	2
1.3 Η περίπτωση των πολυσύνθετων νοσημάτων.....	3
1.4 Μελέτες γενετικής συσχέτισης.....	4
1.5 Προβλήματα στη γενετική επιδημιολογία.....	6
1.6 Αλληλεπιδράσεις γονιδίων και αναλύσεις υποομάδων σε σύνθετα νοσήματα: το παράδειγμα της σύγκρισης φύλων.....	8
1.7 Σκοπός.....	9
ΚΕΦΑΛΑΙΟ 2: ΕΙΔΙΚΟ ΜΕΡΟΣ.....	11
2.1 Εμπειρική αποτίμηση της περιγραφής των μελετών γενετικής επιδημιολογίας.....	11
2.1.1 Σκοπός.....	11
2.1.2 Μέθοδοι.....	11
2.1.3 Αποτελέσματα.....	13
2.1.3.1 Άρθρα ανθρώπινης γενετικής επιδημιολογίας δημοσιευμένα από το 2001 έως το 2003.....	13
2.1.3.2 Άρθρα με πολλαπλούς γενετικούς παράγοντες.....	19
2.1.3.3 Σύγκριση των άρθρων του 2001-2003 με εκείνα του 2006.....	20
2.1.4 Συζήτηση.....	22
2.2 Αναλύσεις υποομάδων σε πολυπαραγοντικά νοσήματα: το παράδειγμα της σύγκρισης φύλων.....	25
2.2.1 Σκοπός.....	25
2.2.2 Μέθοδοι.....	25
2.2.2.1 Επιλογή των μελετών.....	25
2.2.2.2 Εξαγωγή των δεδομένων.....	26
2.2.2.3 Επανεξέταση των ισχυρισμών.....	27
2.2.2.4 Επιβεβαίωση των στατιστικά σημαντικών αλληλεπιδράσεων φύλου-γονιδίων από άλλες μελέτες.....	28
2.2.3 Αποτελέσματα.....	29
2.2.3.1 Επιλέξιμο δείγμα μελετών και ισχυρισμοί με βάση το φύλο.....	29
2.2.3.2 Αναφερόμενη εκ των προτέρων αξιολόγηση των ισχυρισμών.....	37
2.2.3.3 Συγκρίσεις μεταξύ γενετικών παραλλαγών και φαινοτύπων.....	37
2.2.3.4 Πληθυσμοί υπό ανάλυση στους ισχυρισμούς.....	38
2.2.3.5 Υποστήριξη των ισχυρισμών.....	38
2.2.3.6 Επανεξέταση των ισχυρισμών.....	40
2.2.3.7 Επιβεβαίωση των ισχυρισμών με την καλύτερη εσωτερική εγκυρότητα.....	44
2.2.4 Συζήτηση.....	44
ΣΥΜΠΕΡΑΣΜΑΤΑ.....	49
ΠΕΡΙΛΗΨΗ.....	51
ABSTRACT.....	53
ΑΝΑΦΟΡΕΣ.....	55



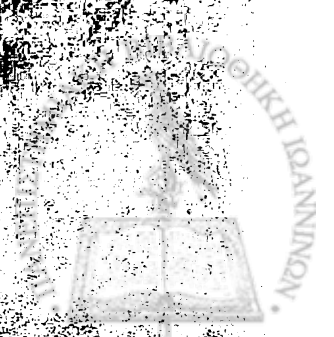
ΠΙΝΑΚΕΣ

Πίνακας 2.1.1: Περιγραφόμενα χαρακτηριστικά των σχεδιασμών μελέτης για 315 τυχαία επιλεγμένα άρθρα γενετικής επιδημιολογίας (2001-2003).....	14
Πίνακας 2.1.2: Περιγραφόμενα χαρακτηριστικά των μεθόδων γονοτύπησης για 315 τυχαία επιλεγμένα άρθρα γενετικής επιδημιολογίας (2001-2003).....	15
Πίνακας 2.1.3: Περιγραφόμενα χαρακτηριστικά περί πληθυσμιακής διαστρωμάτωσης για 315 τυχαία επιλεγμένα άρθρα γενετικής επιδημιολογίας (2001-2003).....	16
Πίνακας 2.1.4: Περιγραφόμενα χαρακτηριστικά των αναλυτικών μεθόδων και των αποτελεσμάτων για 315 τυχαία επιλεγμένα άρθρα γενετικής επιδημιολογίας (2001-2003).....	17
Πίνακας 2.1.5: Περιγραφόμενα χαρακτηριστικά της ανάλυσης πολλαπλών γενετικών παραγόντων για 315 τυχαία επιλεγμένα άρθρα γενετικής επιδημιολογίας (2001-2003).....	19
Πίνακας 2.1.6: Σύγκριση των περιγραφόμενων χαρακτηριστικών άρθρων γενετικής επιδημιολογίας από δυο διαφορετικές χρονικές περιόδους (2001-2003 vs. 2006).	21
Πίνακας 2.2.1: Τα υπό εξέταση επιλέξιμα άρθρα που ισχυρίζονταν για αλληλεπιδράσεις φύλου-γονιδίου στον τίτλο τους.....	30
Πίνακας 2.2.2: Παραδείγματα ισχυρισμών με ανεπαρκή υποστήριξη.....	39
Πίνακας 2.2.3: Παραδείγματα εσφαλμένων ισχυρισμών.....	40
Πίνακας 2.2.4: Ιστορικό επιβεβαίωσης των ισχυρισμών αλληλεπιδράσεων γονιδίων-φύλου με τη φαινομενικά καλύτερη εσωτερική εγκυρότητα.....	47



~~SECRET~~

~~1. The purpose of this document is to provide information on the current status of the project and to identify the key areas of concern.~~
~~2. The information contained herein is classified as "Secret" and is intended for the eyes of authorized personnel only.~~
~~3. It is the policy of this organization that all information of this nature be kept confidential and not be disseminated to the public.~~
~~4. Any unauthorized disclosure of this information could result in the compromise of national security and the safety of our personnel.~~
~~5. Therefore, it is requested that you exercise the highest degree of discretion and care in handling this information.~~
~~6. If you have any questions or concerns regarding this document, please contact the appropriate authority.~~



ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ

ΛΑ: λόγος αναλογιών

ΔΕ: διάστημα εμπιστοσύνης

ΕΤΕ: ενδοτερτατημοριακό εύρος

ΤΚΕ: ταχύτητα καθίζησης ερυθροκυττάρων

LDL: low density lipoprotein

HDL: high density lipoprotein

SNP: single nucleotide polymorphism

PUFA: Polyunsaturated Fatty Acids

GWA: genome-wide association study

HuGE: Human genome epidemiology

HuGE Pub Lit: HuGE Published Literature

apoB: apolipoprotein B

apoA1 : apolipoprotein A1

VLDL: very low density lipoprotein



ΚΕΦΑΛΑΙΟ 1: ΓΕΝΙΚΟ ΜΕΡΟΣ

1.1 - Εισαγωγή στη Γενετική Επιδημιολογία

Η επιδημιολογία μπορεί να οριστεί ως “η μελέτη της κατανομής και της εξέλιξης διαφόρων νοσημάτων στον ανθρώπινο πληθυσμό και των παραγόντων που τις διαμορφώνουν ή μπορούν να τις επηρεάσουν”. Η γενετική επιδημιολογία είναι το κομμάτι εκείνο της κλασσικής επιδημιολογίας που εστιάζεται στη μελέτη των γενετικών παραγόντων και των αλληλεπιδράσεών τους και των συσχετίσεών τους με νοσήματα ή φαινοτύπους. [1]

Η ιστορία της γενετικής επιδημιολογίας ξεκινάει πολύ πριν την ανακάλυψη της δομής του DNA, από την παρατήρηση πως ο κίνδυνος εμφάνισης ενός νοσήματος ήταν μεγαλύτερος αν υπήρχε θετικό οικογενειακό ιστορικό για το νόσημα αυτό. Το γεγονός αυτό οδήγησε στη διατύπωση των μοντέλων κληρονομικότητας κατά Mendel και την επιτυχία της εξήγησης πολλών μονογονιδιακών νοσημάτων. [2] Ο τρόπος κληρονομικότητας (mode of inheritance), ή αλλιώς γενετικό μοντέλο (genetic model), καθορίζει τον τρόπο με τον οποίο ένα γνώρισμα (νόσος ή φαινότυπος) εκφράζεται σε σχέση με τον γονότυπο του ατόμου. Ένα γνώρισμα κληρονομείται ως επικρατές (dominant) εάν η ύπαρξη της υπεύθυνης γενετικής πληροφορίας σε ένα και μόνο γονέα είναι ικανή για την εκδήλωση του. Εάν είναι αναγκαία η ύπαρξη και στους δυο γονείς, τότε το γνώρισμα αυτό κληρονομείται ως υπολειπόμενο (recessive). Μερικές φορές τα άτομα που είναι ετερόζυγοι για την επίμαχη γενετική πληροφορία, δηλαδή φέρουν μόνο ένα αντίγραφο αυτής, έχουν ενδιάμεσο κίνδυνο εμφάνισης του γνωρίσματος σε σχέση με τους ομόζυγους για αυτή (δύο αντίγραφα) και το μοντέλο ονομάζεται συνεπικρατές (co-dominant). Η ανακάλυψη του DNA, η αποκωδικοποίησή του καθώς και άλλες ανακαλύψεις στο χώρο της γενετικής έδωσαν τεράστια ώθηση στην ανάπτυξη της γενετικής επιδημιολογίας, κυρίως σε νοσήματα άλλα πέραν των μονογονιδιακών.



1.2 Το ανθρώπινο γονιδίωμα και η ποικιλότητά του

Το ανθρώπινο γονιδίωμα αποτελείται από ολόκληρη την αλληλουχία του DNA. Το απλοϊδές (haploid) γονιδίωμα εκτείνεται σε περίπου 3.3 δισεκατομμύρια ζεύγη βάσεων (base pairs, bp). Περίπου 3% του γονιδιώματος αποτελεί την αλληλουχία που κωδικοποιεί κάποιο τελικό προϊόν, δηλαδή 30000-40000 γονίδια που κωδικοποιούν πρωτεΐνες. [3-5] Το 99.9% του γονιδιώματος δυατόμων χωρίς καμιά συγγένεια είναι πανομοιότυπο, αλλά το υπόλοιπο 0.1% μπορεί να διαφέρει με πολλούς τρόπους. Αυτές οι περιοχές του DNA ή loci που διαφέρουν μεταξύ των ατόμων ονομάζονται πολυμορφικές. Οι εναλλακτικές αλληλουχίες που βρίσκονται σε μια πολυμορφική περιοχή (locus) ονομάζονται αλλήλια. Σε κάποιες πολύ πολυμορφικές περιοχές οι περισσότεροι άνθρωποι είναι ετεροζυγότες (έχουν διαφορετικά αλλήλια στα δύο χρωμοσώματα). Η μέση ετεροζυγωτία στο ανθρώπινο γονιδίωμα είναι 0.1%-0.4%, δηλαδή μία σε κάθε 250 έως 1000 βάσεις είναι πολυμορφική.

Συνήθως ο όρος πολυμορφισμός χρησιμοποιείται για εκείνες τις παραλλαγές που είναι σχετικά συχνές (βρίσκονται σε τουλάχιστον 1% του πληθυσμού) και δεν έχουν ισχυρά δηλητηριώδεις συνέπειες. Ισχυρά δηλητηριώδεις σπάνιες παραλλαγές ονομάζονται συνήθως μεταλλάξεις. [6] Ο περισσότερο κοινός τύπος πολυμορφισμού στο ανθρώπινο γονιδίωμα είναι ο πολυμορφισμός μοναδικού νουκλεοτιδίου (single nucleotide polymorphism, SNP) ή σημειακή μετάλλαξη. Μεταβάσεις, η αντικατάσταση μια πουρίνης από μια άλλη (A↔G) ή μιας πυριμιδίνης από μια άλλη (C↔T), είναι ο πιο συχνός τύπος SNP.

Υπάρχουν αρκετά επιχειρήματα για τη χρησιμοποίηση των SNPs, παρά άλλων τύπων γενετικών πολυμορφισμών, για τη συσχέτιση με νοσήματα. [7-9] Πρώτον, τα SNPs είναι πάρα πολλά και σκορπισμένα σ' όλο το γονιδίωμα, σε εξόνια, ιντρόνια, προμότερες, βελτιωτές, και ενδογονιδιακές περιοχές, [10, 11] μερικοί δε από αυτούς τους πολυμορφισμούς μπορεί να είναι λειτουργικοί. Δεύτερον, ομάδες γειτονικών SNPs μπορεί να παρουσιάζουν πρότυπα συσχέτισης που θα μπορούσαν να βελτιώσουν τη γενετική χαρτογράφηση [11] και θέσεις ανασυνδυασμού. [12] Τρίτον, ενδοπληθυσμιακές διαφορές στις συχνότητες των SNPs μπορούν να χρησιμοποιηθούν σε πληθυσμιακές μελέτες. [13, 14] Τέταρτον, τα SNPs είναι λιγότερο μεταλλάξιμα σε σχέση με τα άλλα είδη των πολυμορφισμών. [15, 16]



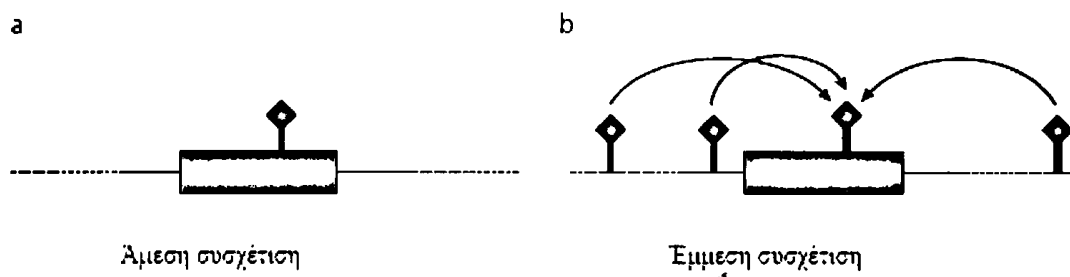
1.3 Η περίπτωση των πολυσύνθετων νοσημάτων

Σε αντίθεση με τα απλά γνωρίσματα (όπως η κυστική ίνωση) που μπορούν να εξηγηθούν με την παραλλαγή μια γενετικής παραλλαγής, τα πολυσύνθετα γνωρίσματα (χοληστερίνη, έμφραγμα του μυοκαρδίου, αυτοάνοσα νοσήματα κλπ) προκαλούνται από μια πληθώρα παραγόντων, τόσο γενετικών όσο και περιβαλλοντικών, καθώς και από την πιθανή αλληλεπίδραση αυτών μεταξύ τους ή με άλλα χαρακτηριστικά, όπως το φύλο. Τα περισσότερα σύνθετα νοσήματα φαίνεται να έχουν ετερογενή αιτιολογία. Διαφορετικοί γενετικοί παράγοντες μπορεί να προκαλούν γενετική προδιάθεση στο ίδιο νόσημα αλλά η επίδραση του καθενός να είναι πολύ μικρή. Αυτό σημαίνει πως για να βρεθούν οι προδιαθεσικοί γενετικοί παράγοντες ενός νόσηματος πρέπει να γίνουν μεγάλοι αριθμοί αναλύσεων, καθώς η γνώση μας για την αιτιολογία των περισσότερων νοσημάτων δεν είναι πλήρης. Αν κανείς προσπαθήσει να ελέγξει τα εκατομμύρια των γενετικών παραλλαγών, καθώς και τις αλληλεπιδράσεις αυτών μεταξύ τους ή με άλλα χαρακτηριστικά, ο αριθμός των αναλύσεων που προκύπτει είναι τουλάχιστον αστρονομικού μεγέθους. [17]



1.4 Μελέτες γενετικής συσχέτισης

Η κλασική επιδημιολογία αναζητεί αν μια μετρούμενη έκθεση σ' ένα πληθυσμό συσχετίζεται με την εμφάνιση ενός γνωρίσματος (νόσημα ή άλλο χαρακτηριστικό, πχ. ύψος). Οι μελέτες συσχέτισης στην γενετική επιδημιολογία κάνουν το ίδιο πράγμα για τις γενετικές “εκθέσεις” ή παράγοντες. Οι περισσότερο χρησιμοποιούμενοι σχεδιασμοί μελετών συσχέτισης είναι οι μελέτες δείκτου-ελέγχου, οι συγχρονικές και οι μελέτες κοόρτης. Ένας έλεγχος συσχέτισης μπορεί να είναι πληροφοριακός ακόμα και αν βασίζεται σε μη λειτουργικές γενετικές παραλλαγές. Μπορεί επίσης να είναι χρήσιμος στην ανίχνευση ανισορροπίας σύνδεσης μεταξύ ενός νοσήματος και ενός μη λειτουργικού δείκτη. [18] Μια ανάλυση συσχέτισης βασισμένη σε μια υποθετικά λειτουργική γενετική παραλλαγή ονομάζεται άμεση, και μια βασισμένη σε ανισορροπία σύνδεσης με ένα δείκτη έμμεση (εικόνα 1). [19-21]



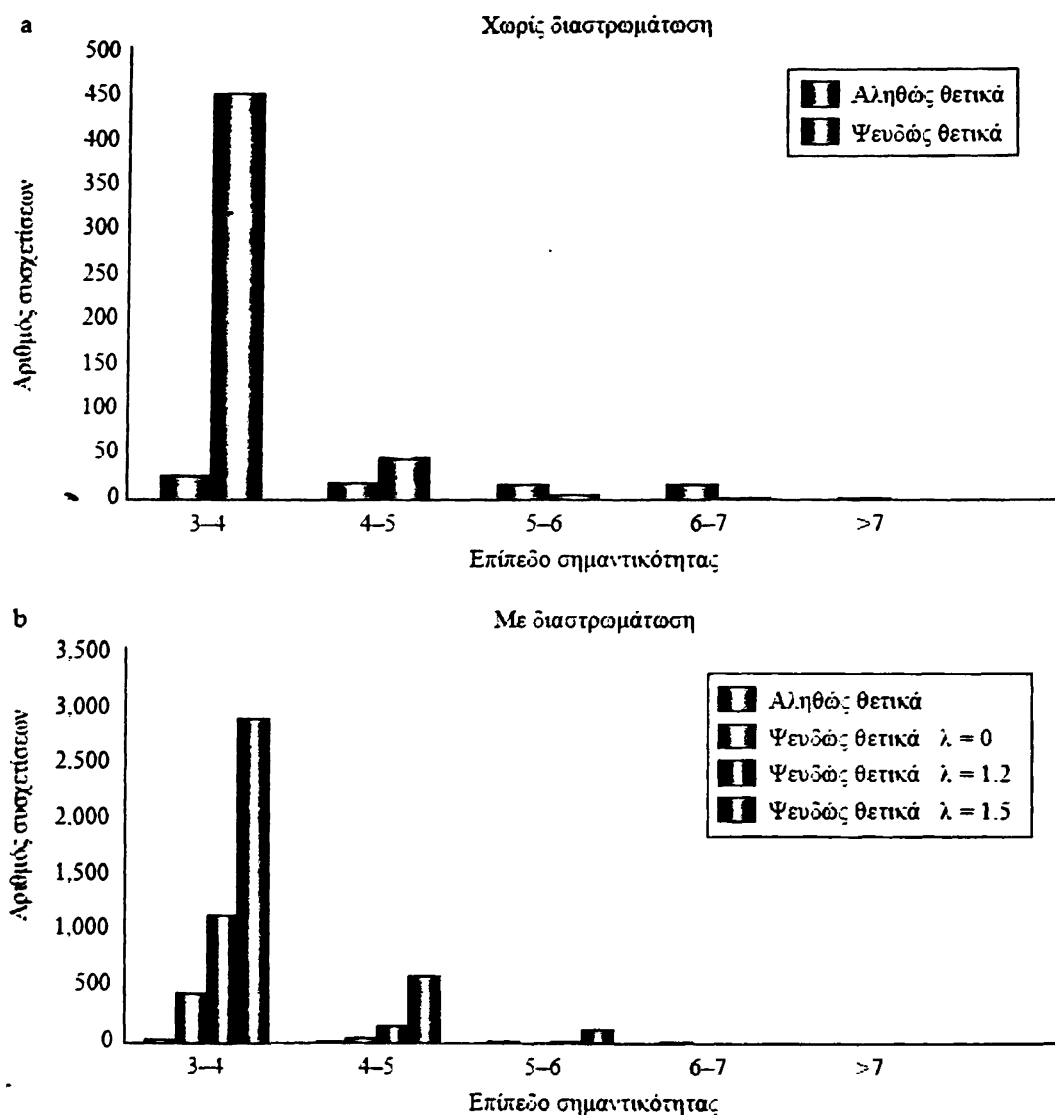
Εικόνα 1: Έλεγχος SNPs για άμεση και έμμεση συσχέτιση.

a. Περίπτωση στην οποία ένα ελεγχόμενο SNP (κόκκινο) εξετάζεται άμεσα για συσχέτιση με ένα νόσημα. b. Τα ελεγχόμενα SNPs (κόκκινο) επιλέχθηκαν να βάση την ανισορροπία σύνδεσης (ΑΣ, linkage disequilibrium) για να παρέχουν πληροφορίες για όσο το δυνατόν περισσότερα άλλα SNPs. Το SNP με μπλε χρώμα ελέγχεται έμμεσα για συσχέτιση, εφόσον είναι σε ΑΣ με τα άλλα τρία.

Υπάρχει όμως και η περίπτωση η παρατηρηθείσα συσχέτιση να οφείλεται στην ύπαρξη κάποιας διαστρωμάτωσης του πληθυσμού ή επιμιξία αυτού (εικόνα 2). Σε έναν τέτοιο πληθυσμό σε διαφορετικά στρώματα αυτού οι ιδρυτικοί (founder) πληθυσμοί έχουν διαφορετικές γενετικές επιδράσεις και οποιαδήποτε γενετική περιοχή με διαφορετικές συχνότητες αλληλίων μεταξύ των



στρωμάτων ή των ιδρυτικών πληθυσμών μπορεί να συσχετιστεί με ένα νόσημα, παρόλο που μπορεί να μην έχει απολύτως καμία σχέση με αυτό.



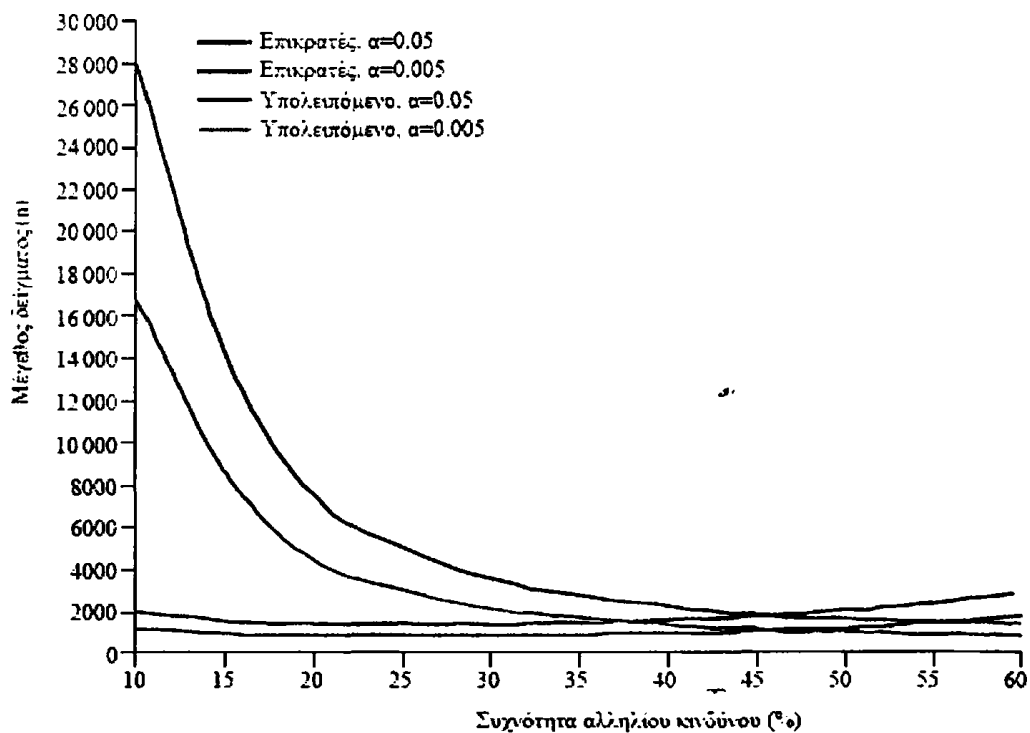
Εικόνα 2: Επίδραση της πληθυσμιακής διαστρωμάτωσης σε μελέτες γενετικής συσχέτισης σε όλη την έκταση του γονιδιώματος (genome-wide association studies, GWAS). Προσομοίωση της επίδρασης της πληθυσμιακής διαστρωμάτωσης στο λόγο αληθώς θετικών / ψευδώς θετικών αποτελεσμάτων σε μια GWA. α. Αναμενόμενος αριθμός αληθώς θετικών και ψευδώς θετικών συσχετίσεων σε προσομοιώσεις με διαφορετικά επίπεδα σημαντικότητας (εμφάνιση ως $-\log(p)$). Αποτελέσματα από προσομοιώσεις μελέτης 1,000 ασθενών και 1,000 μαρτύρων, όπου γονοτυπήθηκαν 500,000 SNPs αλλά μόνο 100 αντιπροσωπεύουν πραγματικές συσχετίσεις (συχνότητα αλληλίου 15% και λόγος αναλογιών από 1.2-1.5). β. Προσομοίωση του ίδιου σεναρίου κάτω από δύο διαφορετικά επίπεδα ήπιας διαστρωμάτωσης, $\lambda=1.2$ και $\lambda=1.5$.



1.5 Προβλήματα στη γενετική επιδημιολογία

Οι μελέτες συσχέτισης στη γενετική επιδημιολογία είναι επιρρεπείς σε μεγάλο αριθμό συστηματικών σφαλμάτων.

Στη γενετική των σύνθετων νοσημάτων τόσο το σφάλμα τύπου I όσο και το σφάλμα τύπου II πρέπει να είναι περιορισμένο. [7, 22, 23] Η ισχύς των μελετών συσχέτισης εξαρτάται κυρίως από το μέγεθος του δείγματος, το μέγεθος της επίδρασης του γενετικού παράγοντα, την ισχύ της ανισορροπίας σύνδεσης με ένα δείκτη και τις συχνότητες των προδιαθεσικών και δεικτών (το αλληλίο που γονοτυπήθηκε) αλληλίων. [18] Η εικόνα 3 δείχνει τα μεγέθη δείγματος που χρειάζονται για την ανίχνευση ενός λόγου αναλογιών (odds ratio) της τάξης του 1.5 με ισχύ 80% και σφάλμα τύπου I (α) είτε 0.05 είτε 0.005.



Εικόνα 3: Εκτιμήσεις μεγέθους δείγματος για αναλύσεις SNPs σε μελέτες δείκτη-ελέγχου. Το δείγμα είναι το άθροισμα των ασθενών και μαρτύρων, με λόγο ασθενών:μαρτύρων ίσο με ένα. Ανιχνεύσιμη διαφορά $\Lambda A > 1.5$. Ισχύς 80%. α = σφάλμα τύπου I.

Ένα άλλο πρόβλημα είναι οι πολλαπλοί έλεγχοι, δηλαδή όταν πολλαπλά SNPs σε ένα ή περισσότερα γονίδια ελέγχονται για συσχέτιση με το ίδιο γνώρισμα (trait). [24] Η διόρθωση του ουδού σημαντικότητας από $\alpha=0.05$ σε πχ 0.005, είναι περισσότερο ρεαλιστική και προστατεύει από υψηλά ποσοστά σφάλματος τύπου I. Η επίτευξη τόσο χαμηλών τιμών p, σε συνδυασμό με τις χαμηλές επιδράσεις των γενετικών παραγόντων, απαιτεί δείγματα μεγάλου μεγέθους. Παρόλα αυτά, οι μελέτες γενετικής συσχέτισης είναι συνήθως αρκετά υποδεέστερες σε θέμα ισχύος από ο,τι θα έπρεπε να είναι.[9, 25-28]

Οι μελέτες γενετικής συσχέτισης για σύνθετα νοσήματα συχνά είτε αποτυγχάνουν να εντοπίσουν προδιαθεσικούς γενετικούς παράγοντες είτε δεν μπορούν να επιβεβαιώσουν μελέτες που εντόπισαν κάποιους παράγοντες. [22, 29, 30] Παρόλη την ευρεία χρήση των μελετών δείκτου-ελέγχου στη γενετική επιδημιολογία, η ασυνέπεια των αποτελεσμάτων των είναι ένα αναγνωρισμένο μειονέκτημα.[31, 32] Αυτή η έλλειψη επαναληψιμότητας αποδίδεται στα μικρά μεγέθη δείγματος με ανεπαρκή ισχύ, στη βιολογική και φαινοτυπική πολυπλοκότητα, στην ανισορροπία σύνδεσης, και τη πληθυσμιακή διαστρωμάτωση. [7, 29, 32] Άλλες αιτίες αποτελούν η ετερογένεια στο σχεδιασμό των μελετών, τις αναλυτικές μεθόδους και τον ορισμό του φαινότυπου. Για τα περισσότερα σύνθετα νοσήματα η ύπαρξη πολλαπλών γενετικών παραγόντων μικρών ατομικών επιδράσεων, αλληλεπιδράσεων γονιδίων, αλληλεπιδράσεων γονιδίων-περιβαλλοντικών παραγόντων και η χαμηλή στατιστική ισχύς, σημαίνει πως τόσο η αρχική ανίχνευση όσο και η επιβεβαίωση θα είναι αρκετά δύσκολη.



1.6 Αλληλεπιδράσεις γονιδίων και αναλύσεις υποομάδων σε σύνθετα νοσήματα: το παράδειγμα της σύγκρισης φύλων.

Το φύλο είναι ένας παράγοντας που έχει εμπλακεί εκτεταμένα στο παρελθόν ως διαμορφωτής επιδράσεων στη κλινική έρευνα. Παρόλα αυτά, εμπειρικά δεδομένα από τυχαιοποιημένες κλινικές μελέτες έχουν δείξει ότι πολλές ισχυριζόμενες διαφορές υποομάδων βασισμένες στο φύλο είναι πλασματικές και έχουν οδηγήσει σε σοβαρές παρερμηνείες.[33] Για παράδειγμα, για περισσότερο από 10 χρόνια υπήρχε η πεποίθηση πως η ασπιρίνη ήταν αναποτελεσματική στη δευτεροπαθή πρόληψη του εγκεφαλικού στις γυναίκες βασισμένη σε μια ανάλυση υποομάδων μειωμένης ισχύος.[34]

Στην εποχή του ανθρώπινου γονιδιώματος, πολλές φορές έχουν εμφανιστεί μελέτες που υποστηρίζουν πως κάποια συχνή γενετική παραλλαγή (variant) μπορεί να έχει διαφορετική επίδραση στους άνδρες απ' ότι στις γυναίκες. Πολλά νοσήματα ή γνωρίσματα (traits) με ισχυρό γενετικό υπόβαθρο έχουν διαφορετικό επιπολασμό στα δυο φύλα. Για παράδειγμα, τα αυτοάνοσα νοσήματα, οι ενδοκρινοπάθειες και η μακροζωία είναι πιο συχνές στις γυναίκες, ενώ οι καρδιαγγειακές νόσοι, το ισχαιμικό εγκεφαλικό επεισόδιο και οι χοληστερινοπάθειες πιο συχνές στους άνδρες.[35] Αυτές οι παρατηρήσεις δε σημαίνουν απαραίτητα πως μια συγκεκριμένη γενετική παραλλαγή θα πρέπει να έχει επίσης διαφορετική επίδραση στα δυο φύλα. Για τους περισσότερους φαινότυπους, πολλές συχνές γενετικές παραλλαγές είναι πιθανόν να είναι υπεύθυνες για τον καθορισμό της προδιάθεσης στο νόσημα.[36] Ανάμεσα στις αυτοσωμικές παραλλαγές μόνο κάποιες από αυτές ίσως να αλληλεπιδρούν με το φύλο. Παρόλα αυτά, δεδομένου ότι πληροφορίες για το φύλο είναι σχεδόν πάντα διαθέσιμες σε κάθε μελέτη γενετικής επιδημιολογίας, είναι εύκολο για κάποιον να ελέγξει αν το φύλο επηρεάζει τις γενετικές επιδράσεις. Τελικά, ένας μεγάλος αριθμός ισχυρισμών γίνονται για διαφορές στα δυο φύλα.



1.7 Σκοπός

Σκοπός της παρούσας διατριβής ήταν η μελέτη και περιγραφή μελετών γενετικής επιδημιολογίας για πολυσύνθετα νοσήματα και η περιγραφή των χαρακτηριστικών αυτών. Ιδιαίτερη μνεία γίνεται για πολλαπλούς γενετικούς παράγοντες και τις αλληλεπιδράσεις αυτών. Ειδικότερα περιγράφεται και αξιολογείται η εγκυρότητα ισχυρισμών για διαφορετικές επιδράσεις γενετικών παραγόντων με βάση το φύλο (αλληλεπίδραση φύλου-γονιδίου) σε ένα μεγάλο δείγμα μελετών γενετικής επιδημιολογίας σύνθετων νοσημάτων.



ΚΕΦΑΛΑΙΟ 2: ΕΙΔΙΚΟ ΜΕΡΟΣ

2.1 - Εμπειρική αποτίμηση της περιγραφής των μελετών γενετικής επιδημιολογίας

2.1.1 Σκοπός

Η ανθρώπινη γενετική επιδημιολογία είναι ένα ραγδαία αναπτυσσόμενο επιστημονικό πεδίο που εξετάζει την επιρροή της γενετικής ποικιλότητας στην ανθρώπινη υγεία.[37-40] Παρόλο που ένας μεγάλος και αυξανόμενος αριθμός μελετών έχει ερευνήσει τις συσχετίσεις μεταξύ γενετικών παραλλαγών και της προδιάθεσης για συχνά νοσήματα, λίγες στατιστικά σημαντικές συσχετίσεις έχουν επιβεβαιωθεί σε πολλαπλές άλλες μελέτες. [41, 42] Διάφανη περιγραφή των πληθυσμών μιας μελέτης, των μεθόδων συλλογής δεδομένων, των αναλυτικών μεθόδων και των συμπερασμάτων αυτής μπορεί να βοηθήσει τους αναγνώστες να αναγνωρίσουν καλύτερα τα σημεία που επηρεάζουν την επαναληψιμότητα των μελετών γενετικής συσχέτισης. Στην παρούσα διατριβή επιχειρήθηκε λεπτομερής αξιολόγηση των πρακτικών περιγραφής των μελετών συσχέτισης της γενετικής επιδημιολογίας.

2.1.2 Μέθοδοι

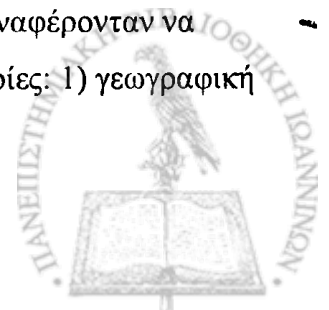
Το 2001 το Δίκτυο Ανθρώπινης Γενετικής Επιδημιολογίας (Human Genome Epidemiology Network, HuGENet) ίδρυσε τη βιβλιογραφική βάση HuGE Published Literature (HuGE Pub Lit), μια συνεχώς ανανεούμενη, αναζητήσιμη, διαδικτυακή βάση άρθρων γενετικής επιδημιολογίας βασιζόμενων σε πληθυσμούς. [43] Σχετικές μελέτες αναγνωρίζονται εβδομαδιαία από τη PubMed (www.pubmed.gov) από ένα γενετικό επιδημιολόγο που καταγράφει το σχεδιασμό της μελέτης, τα γονίδια και τα εξεταζόμενα νοσήματα. Έως τις 27 Μαΐου 2007 αυτή η βάση περιείχε ένα σύνολο 27,386 άρθρων που εξέταζαν συσχετίσεις γονοτύπου-φαινοτύπου (τόσο ποσοτικών όσο και ποιοτικών χαρακτηριστικών), δημοσιευμένων σε 2773 περιοδικά. Για την επιλογή των άρθρων αναζητήσαμε στη HuGE Pub Lit για πληθυσμιακές μελέτες που χρησιμοποίησαν περιγραφικούς σχεδιασμούς (πχ μελέτες δείκτου-ελέγχου, κοόρτης) για να ερευνήσουν συσχετίσεις γονιδίων-νοσημάτων και αλληλεπιδράσεις μεταξύ γενετικών παραγόντων. Οι οικογενειακές μελέτες σύνδεσης (linkage) δε συλλέγονται συστηματικά από τη HuGE Pub Lit, οπότε δε συμπεριελήφθησαν στην αξιολόγηση. Επιπρόσθετα, η ανάλυση περιορίστηκε σε μελέτες πλήρους κειμένου διότι οι μελέτες που δημοσιεύονται ως συνοπτικές περιγραφές (πχ περιλήψεις ή επιστολές) πιθανώς θα αυξάνανε την ετερογένεια του δείγματος.



Η αξιολόγηση σχεδιάστηκε το 2004 και η συλλογή των δεδομένων και οι αναλύσεις έγιναν μεταξύ 2004-2007. Για την κύρια ανάλυση επιλέχθηκε ένα τυχαίο δείγμα που αντιστοιχούσε στο 5% των άρθρων που δημοσιεύτηκαν από το 2001-2003 και συμπεριλήφθηκαν στη HuGE Pub Lit πριν τις 30 Μαΐου του 2004 (n=8,115). Το τυχαίο αυτό δείγμα ήταν 406 άρθρα. Επίσης, επιλέχθηκαν τυχαία 40 άρθρα δημοσιευμένα το 2006 που καταλογογραφήθηκαν στη HuGE Pub Lit πριν τις 18 Μαΐου 2007 (n=5,353), ώστε να ερευνηθούν τυχόν βελτιώσεις στην περιγραφή αυτών των μελετών με την πάροδο του χρόνου.

Μετά την ανάγνωση κάθε άρθρου, 91 από το 2001-2003 και 12 από το 2006 αποκλείστηκαν από την ανάλυση για τις κάτωθι αιτίες: μη δημοσιευμένα στα Αγγλικά (2001-2003: n=28, 2006: n=6), μελέτες πληθυσμιακής ανίχνευσης (screening) (2001-2003: n=23, 2006: n=0), κλινικές μελέτες ή μελέτες φαρμακογενωμικής (2001-2003: n=16, 2006: n=3), μη διαθέσιμο πλήρες κείμενο (2001-2003: n=11, 2006: n=0), αποτυχία πλήρωσης των κριτηρίων της HuGE Pub Lit μετά από λεπτομερέστερη εξέταση (2001-2003: n=6, 2006: n=1), οικογενειακές μελέτες (2001-2003: n=3, 2006: n=0), μετα-αναλύσεις (2001-2003: n=2, 2006: n=0).

Για τα άρθρα του 2001-2003, μια σταθμισμένη φόρμα εξαγωγής αναπτύχθηκε και δοκιμάστηκε σε 10 άρθρα. Η φόρμα διορθώθηκε ανάλογα με τα συμπεράσματα της πιλοτικής αυτής φάσης ώστε οι ορισμοί για τα εξαγόμενα δεδομένα να είναι ξεκάθαροι και σαφείς. Τα στοιχεία της τελικής φόρμας σχεδιάστηκαν για τη συλλογή πληροφοριών σχετικά με την περιγραφή του σχεδιασμού της μελέτης, της μεθόδου γονοτύπησης, την εξέταση για την ύπαρξη πληθυσμιακής διαστρωμάτωσης, των αναλυτικών μεθόδων (συμπεριλαμβανομένης της ταυτόχρονης ανάλυσης πολλαπλών γενετικών παραγόντων), και των συμπερασμάτων των μελετών. Επιπρόσθετα, η φόρμα επέτρεπε την ύπαρξη διαφορετικών περιγραφικών σχεδιασμών, πολλαπλών ομάδων συμμετεχόντων και την εξέταση πολλαπλών γενετικών παραγόντων. Τα άρθρα κωδικοποιήθηκαν να έχουν πιθανή δυσταξινόμηση για το εξεταζόμενο νόσημα όταν το άρθρο δεν ανέφερε ξεκάθαρα τα κριτήρια διάγνωσης. Όταν αναφέρονται πολλαπλές ομάδες συμμετεχόντων σε μια μελέτη επιλέχθηκε η μεγαλύτερη ομάδα για μελέτες κοόρτης ή συγχρονικές και η μεγαλύτερη ομάδα ασθενών και μαρτύρων για μελέτες δείκτου-ελέγχου. Τα δεδομένα συλλέχθηκαν ξεχωριστά για κάθε ομάδα ασθενών και μαρτύρων για μελέτες δείκτου-ελέγχου και για όλους τους συμμετέχοντες άσχετα με το αν νοσούσαν ή όχι για μελέτες κοόρτης ή συγχρονικές. Για τον σκοπό της ανάλυσης, τα συλλεγμένα δεδομένα για τις ομάδες των ασθενών και μαρτύρων συνδυάστηκαν ώστε τα στατιστικά να μπορούν να υπολογιστούν για όλους τους συμμετέχοντες. Επιπρόσθετα, για τις μελέτες δείκτου-ελέγχου, καταγράφηκε εάν οι ασθενείς και οι μάρτυρες αναφέρονταν να προέρχονται από τον ίδιο πληθυσμό, σύμφωνα με μια ή περισσότερα κατηγορίες: 1) γεωγραφική



περιοχή, 2) κλινικός πληθυσμός, 3) γενικός πληθυσμός (πχ εθνική ομάδα), ή εάν η πληροφορία για την επιλογή κατάλληλων μαρτύρων έλειπε ή ήταν ελλιπής.

-Δεκατέσσερα στοιχεία εξετάστηκαν στα άρθρα του 2006. Ανάμεσα σε αυτά ήταν το μέγεθος των συμμετεχόντων, τα γονίδια κι οι πολυμορφισμοί. Επιπρόσθετα, επιλέχθηκαν δέκα στοιχεία που υπήρχαν σε όλους τους σχεδιασμούς μελετών και είχαν περιγραφεί σε ποσοστό μικρότερο του 50% των άρθρων του 2001-2003.

Οι συγκρίσεις των άρθρων του 2006 έναντι αυτών του 2001-2003 έγιναν με Mann-Whitney U test για συνεχείς τιμές και με Fisher's exact test για δυαδικές μεταβλητές.

2.1.3 Αποτελέσματα

2.1.3.1 Άρθρα ανθρώπινης γενετικής επιδημιολογίας δημοσιευμένα από το 2001 έως το 2003

Τα 315 που επιλέχθηκαν για την ανάλυση δημοσιεύτηκαν σε 194 περιοδικά και περιέγραφαν ευρήματα από 227 μελέτες δείκτου-ελέγχου, 32 κοόρτες και 56 συγχρονικές μελέτες. Όπως φαίνεται στον πίνακα 2.1.1, τα περισσότερα (75.9%) περιέγραφαν μεγέθη δείγματος μικρότερα από 500 συμμετέχοντες· 9.2% μεγέθη δείγματος μεγαλύτερα ή ίσα με 1,000 συμμετέχοντες (διάμεσος: 265, ενδοτερτατημοριακό εύρος: 142-471). Η στατιστική ισχύς αναφέρονταν στο 12.7% των άρθρων. Πολλαπλοί εξεταζόμενοι πληθυσμοί (πχ περισσότερες από μια ομάδα ασθενών ή μαρτύρων) περιείχαν το 25.4% των άρθρων. Οι περισσότερες μελέτες δίναν πληροφορίες τουλάχιστον για την προέλευση (87.9%) και τα κριτήρια συμμετοχής (97.5%) των συμμετεχόντων. Πληροφορίες για το φύλο περιείχαν τα τρία τέταρτα των άρθρων, ενώ η διάμεση ή μέση ηλικία των συμμετεχόντων και ένα μέτρο διασποράς γύρω από αυτή την τιμή περιγράφονταν στο 65.4% και 54.6%, αντίστοιχα. Μία στις έξι μελέτες ξεκάθαρα δήλωναν πως οι συμμετέχοντες δεν είχαν καμιά συγγένεια μεταξύ τους. Το 11.8% των άρθρων μπορεί να είχαν δυσταξινομήσει την ενδιαφερόμενη έκβαση.



Πίνακας 2.1.1: Περιγραφόμενα χαρακτηριστικά των σχεδιασμών μελέτης για 315 τυχαία επιλεγμένα άρθρα γενετικής επιδημιολογίας (2001-2003).

Περιγραφόμενο χαρακτηριστικό	Αριθμός	Ποσοστό
Αριθμός συμμετεχόντων		
<100	49	15.6
100-499	190	60.3
500-999	47	14.9
>= 1000	29	9.2
Περιγραφή της ισχύς της μελέτης		
Όχι	275	87.3
Ναι	40	12.7
Περιγραφή χρησιμοποίησης πολλαπλών πληθυσμών ή ομάδων δείκτου-ελέγχου		
Όχι	235	74.6
Ναι	80	25.4
Παροχή πληροφοριών για την προέλευση των συμμετεχόντων		
Όχι	38	12.1
Ναι	277	87.9
Παροχή πληροφοριών για τα κριτήρια συμμετοχής των συμμετεχόντων		
Όχι	8	2.5
Ναι	307	97.5
Περιγραφή κατανομής φύλου για όλους τους συμμετέχοντες		
Όχι	84	26.7
Ναι	231	73.3
Περιγραφή μέσης ή διάμεσης ηλικίας για όλους τους συμμετέχοντες		
Όχι	109	34.6
Ναι	206	65.4
Περιγραφή τυπικής απόκλισης ή ενδοτερτατημοριακού εύρους ηλικίας για όλους τους συμμετέχοντες		
Όχι	143	45.4
Ναι	172	54.6
Δήλωση πως χρησιμοποιήθηκαν συμμετέχοντες χωρίς καμιά συγγένεια μεταξύ τους		
Όχι	259	82.2
Ναι	56	17.8
Πιθανότητα για δυσταξινόμηση έκβασης		
Όχι	272	86.3
Μη ξεκάθαρο	6	1.9
Ναι	37	11.8

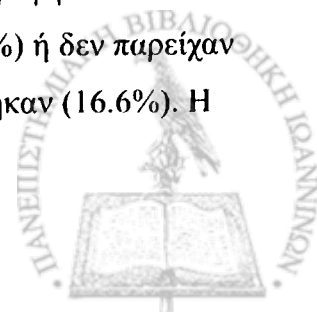


Επτά τοις εκατό των άρθρων περιέγραφαν πως τα αποτελέσματα της γονοτύπησης επικυρώθηκαν με τη χρήση επαναληπτικών δειγμάτων και ένα επιπρόσθετο 9.8% πως μια διαφορετική τεχνική επικύρωσης χρησιμοποιήθηκε (Πίνακας 2.1.2). Σπάνια περιγραφόταν τυφλή αξιολόγηση του γενετικού τεστ σε σχέση με την έκβαση (11.1%) ή της έκβασης σε σχέση με το γενετικό τεστ (3.8%). Λίγα άρθρα ανέφεραν πως κάποιιο συμμετέχοντες είχαν αποκλειστεί (11.8%) ή σχολίαζαν τον αριθμό των δειγμάτων που δε μπορούσαν να γονοτυπηθούν (15.6%).

Πίνακας 2.1.2: Περιγραφόμενα χαρακτηριστικά των μεθόδων γονοτύπησης για 315 τυχαία επιλεγμένα άρθρα γενετικής επιδημιολογίας (2001-2003).

Περιγραφόμενο χαρακτηριστικό	Αριθμός	Ποσοστό
Αναφερόμενη επικύρωση των αποτελεσμάτων γονοτύπησης με τη χρήση διπλών δειγμάτων		
Όχι	293	93.0
Ναι	22	7.0
Αναφερόμενη επικύρωση των αποτελεσμάτων γονοτύπησης με τη χρήση διαφορετικής μεθόδου		
Όχι	284	90.2
Ναι	31	9.8
Αναφορά πως η αξιολόγηση των γενετικών ελέγχων ήταν τυφλή στις εκβάσεις ή στους φαινότυπους		
Τυφλοποίηση	35	11.1
Ασαφής	280	88.9
Αναφορά πως η αξιολόγηση της έκβασης ή του φαινότυπου ήταν τυφλή στα αποτελέσματα του γενετικού ελέγχου		
Τυφλοποίηση	12	3.8
Ασαφής	303	96.2
Αναφορά για αποκλεισμό ατόμων από τις αρχικές ομάδες συμμετεχόντων		
Όχι	278	88.2
Ναι	37	11.8
Αναφορά αποτυχίας γονοτύπησης δειγμάτων		
Όχι	266	84.4
Ναι	49	15.6

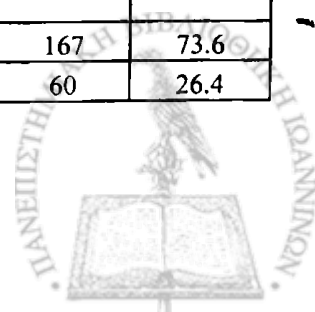
Όπως φαίνεται στον πίνακα 2.1.3, σχεδόν 60% των άρθρων ανέφεραν πως όλοι οι συμμετέχοντες προέρχονταν από τον ίδιο εθνικό πληθυσμό, ενώ το 9.5% ανέφεραν πως ο υπό μελέτη πληθυσμός περιελάμβανε περισσότερους του ενός εθνικούς πληθυσμούς. Τα περισσότερα άρθρα (76.7%) είτε έκαναν αναλύσεις διαστρωμάτωσης με βάση την εθνική καταγωγή είτε διορθώσανε γι'αυτή. Παρόλα αυτά, λίγα μελετήσανε τις εθνικές ομάδες μαζί (6.7%) ή δεν παρείχαν ξεκάθαρες πληροφορίες πως τα δεδομένα διαφορετικών εθνικών ομάδων αναλύθηκαν (16.6%). Η



χρήση ασύνδετων γενετικών δεικτών (unlinked genetic markers) για τον έλεγχο της πληθυσμιακής διαστρωμάτωσης ήταν εξαιρετικά σπάνια (0.6%). Ανάμεσα στις μελέτες δείκτου-ελέγχου (n=227), τα δυο τρίτα ανέφεραν πως οι ασθενείς και οι μάρτυρες προέρχονταν από την ίδια γεωγραφική περιοχή· ένα στα πέντε ανέφεραν πως οι ασθενείς και οι μάρτυρες προέρχονταν από τον ίδιο κλινικό πληθυσμό και ένα τρίτο ανέφεραν πως οι ασθενείς και οι μάρτυρες προέρχονταν από τον ίδιο γενικό πληθυσμό. Περισσότερα από το ένα τρίτο των άρθρων δεν ήταν ξεκάθαρα σχετικά με την προέλευση των πληθυσμών ή δεν παρείχαν καμία πληροφορία.

Πίνακας 2.1.3: Περιγραφόμενα χαρακτηριστικά περί πληθυσμιακής διαστρωμάτωσης για 315 τυχαία επιλεγμένα άρθρα γενετικής επιδημιολογίας (2001-2003).

Περιγραφόμενο χαρακτηριστικό	Αριθμός	Ποσοστό
Ξεκάθαρη δήλωση πως όλοι οι συμμετέχοντες προέρχονταν από τον ίδιο εθνικό πληθυσμό		
Ασαφές	130	41.3
Ναι	185	58.7
Ανάλυση με βάση διαφορετικές εθνικές ομάδες		
Όχι	285	90.5
Ναι	30	9.5
Αν περιλαμβάνονταν διαφορετικές εθνικές ομάδες, πως χειρίστηκαν τη διαφορετική εθνικότητα στην ανάλυση (n=30)		
Διαστρωμάτωση ή στάθμιση για εθνικές ομάδες	23	76.7
Σύμπτυξη των εθνικών ομάδων	2	6.7
Ασαφές	5	16.6
Περιγραφή πώς ασύνδετοι γενετικοί δείκτες χρησιμοποιήθηκαν για την ανίχνευση πληθυσμιακής διαστρωμάτωσης		
Όχι	313	99.4
Ναι	2	0.6
Αναφορά πώς οι ασθενείς και οι μάρτυρες προέρχονταν από τον ίδιο γεωγραφικό πληθυσμό (n=227)		
Όχι	79	34.8
Ναι	148	65.2
Αναφορά πως οι ασθενείς και οι μάρτυρες προέρχονταν από τον ίδιο κλινικό πληθυσμό (n=227)		
Όχι	180	79.3
Ναι	47	20.7
Αναφορά πως οι ασθενείς και οι μάρτυρες προέρχονταν από τον ίδιο γενικό πληθυσμό (n=227)		
Όχι	167	73.6
Ναι	60	26.4

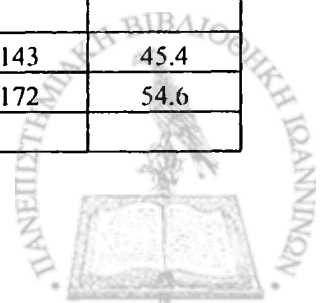


Αναφορά ασαφών ή ελλειπών στοιχείων για την προέλευση του πληθυσμού των ασθενών και μαρτύρων (n=227)		
Όχι	142	62.6
- Ναι	85	37.4

Περίπου τα μισά άρθρα δήλωναν πως εξέτασαν αν οι πληθυσμοί βρίσκονταν σε ισορροπία Hardy-Weinberg. Από αυτά το 6.6% δήλωνε πως οι συχνότητες γονοτύπων παρέκκλιναν από την αναμενόμενη ισορροπία Hardy-Weinberg (Πίνακας 2.1.4). Περιληπτικά δεδομένα (πχ συχνότητες γονοτύπων/αλληλίων παρουσιαζόμενες σε πίνακα) περιέχονταν για όλους τους γενετικούς παράγοντες για όλες τις ενδιαφερόμενες εκβάσεις στο 87% των άρθρων. Η ανάλυση βασισμένη στα γονίδια ήταν λιγότερο συχνή (54.6%) από ό,τι η ανάλυση γονοτύπων (85.7%). Κατά την ανάλυση γονοτύπων, το 20.7% των άρθρων χρησιμοποιούσε συγκεκριμένες γενετικές συγκρίσεις βασισμένες στο επικρατές ή το υπολειπόμενο μοντέλο. Από τα τελευταία, το 44.1% έδινε κάποια εξήγηση για την χρησιμοποίηση του συγκεκριμένου μοντέλου. Ένα στα δέκα άρθρα περιέγραφε διόρθωση για πολλαπλές συγκρίσεις, με πιο συχνή μέθοδο (70%) τη διόρθωση κατά Bonferroni, ενώ ένα μόνο άρθρο ανέφερε πως χρησιμοποίησε τα τεστ Tukey's και Scheffe για τον έλεγχο των πολλαπλών συγκρίσεων. [44]

Πίνακας 2.1.4: Περιγραφόμενα χαρακτηριστικά των αναλυτικών μεθόδων και των αποτελεσμάτων για 315 τυχαία επιλεγμένα άρθρα γενετικής επιδημιολογίας (2001-2003).

Περιγραφόμενο χαρακτηριστικό	Αριθμός	Ποσοστό
Αναφορά πως όλοι οι γενετικοί παράγοντες εξετάστηκαν για ισορροπία Hardy-Weinberg		
Όχι	164	52.1
Ναι	151	47.9
Αν αναφέρονταν η ισορροπία Hardy-Weinberg, απέτυχε να την εκπληρώσει κάποιος γενετικός παράγοντας;		
Όχι	141	93.4
Ναι	10	6.6
Αναφορά συνολικών δεδομένων για όλους τους γενετικούς παράγοντες και όλες τις εκβάσεις		
Όχι	41	13.0
Ναι	274	87.0
Χρήση αλληλίων για τις γενετικές συγκρίσεις		
Όχι	143	45.4
Ναι	172	54.6
Χρήση γονοτύπων για τις γενετικές συγκρίσεις		



Όχι	45	14.3
Ναι	270	85.7
Αν χρησιμοποιήθηκαν γονότυποι για τις αναλύσεις, έγιναν όλες οι συγκρίσεις ή μόνο επιλεκτικές (n=270)		
Όλες οι δυνατές	214	79.3
Επιλεκτικές	56	20.7
Δικαιολόγηση για την επιλογή συγκεκριμένων συγκρίσεων (n=56)		
Όχι	33	58.9
Ναι	23	41.1
Διόρθωση για πολλαπλές συγκρίσεις		
Όχι	276	87.6
Ναι	39	12.4
Αν έγινε διόρθωση για πολλαπλές συγκρίσεις, τι τύπος ήταν (n=40, ένα άρθρο χρησιμοποίησε 2 μεθόδους)		
Bonferroni	28	70.0
Fischer's post hoc	1	2.5
Monte Carlo simulations	1	2.5
Scheffe's test	2	5.0
Tukey's test	3	7.5
Ασαφές	5	12.5
Συζήτηση των επιπτώσεων των ευρημάτων στη δημόσια υγεία, την ιατρική ή την κλινική πράξη		
Όχι	193	61.3
Ναι (οποιαδήποτε αναφορά)	122	38.7
Αναφορά πως ήταν η πρώτη μελέτη στο θέμα		
Όχι	266	84.4
Ναι	49	15.6
Ξεκάθαρη αναφορά στην πρώτη μελέτη (n=266)		
Όχι	243	91.4
Ναι	23	8.6
Ξεκάθαρη αναφορά σε συστηματική ανασκόπηση		
Όχι	297	94.3
Ναι	18	5.7
Ξεκάθαρη αναφορά σε μη-συστηματική ανασκόπηση		
Όχι	309	98.1
Ναι	6	1.9

Συνολικά, λιγότερο από το 40% των άρθρων συζητούσε για τις επιπτώσεις στη δημόσια υγεία, την ιατρική ή κλινική πράξη των ευρημάτων. Λιγότερες από μια στις έξι μελέτες ισχυρίζονταν πως ήταν οι πρώτες που ανέλυναν τη συγκεκριμένη συσχέτιση. Από τα άρθρα που δεν έκαναν αυτόν τον ισχυρισμό το 8.6% έκανε αναφορά στην πρώτη μελέτη του συγκεκριμένου θέματος. Το 6% των άρθρων ανέφεραν μια συστηματική ανασκόπηση και το 1.9% μια μη-συστηματική ανασκόπηση.

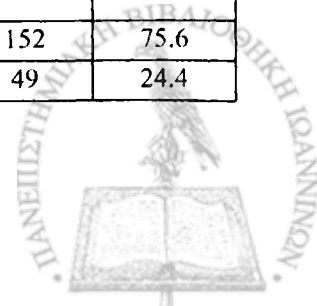


2.1.3.2 Άρθρα με πολλαπλούς γενετικούς παράγοντες

Σχεδόν τα δυο τρίτα των άρθρων (n=201) εξερευνούσαν πολλαπλούς γενετικούς παράγοντες, συχνά σε περισσότερα του ενός γονίδια (Πίνακας 2.1.5). Τέσσερις στις έξι μελέτες περιείχαν γενετικούς παράγοντες από τουλάχιστον 2 γονίδια. Ανάμεσα σ' αυτά που μελετούσαν 2 ή περισσότερους πολυμορφισμούς περιγραφή ανισορροπίας σύνδεσης υπήρχε μόνο στο 22.9%. Ανάλυση απλοτύπων αναφέρονταν σε ένα αντίστοιχο ποσοστό (21.4%). Οι αλληλεπιδράσεις γονιδίων δεν ήταν αρκετά συχνές επίσης. Μόνο το 24.4% των μελετών με 2 ή περισσότερους πολυμορφισμούς περιέγραφε τον έλεγχο για αλληλεπίδραση μεταξύ αυτών. Όταν άρθρα ανέφεραν αλληλεπιδράσεις γονιδίων σχεδόν τα μισά χρησιμοποιούσαν το λόγο αναλογιών ως μέτρο κινδύνου, μόνο το 4.1% την απόλυτη διαφορά και κανένα το αποδοτέο κλάσμα (attributable fraction). Τα υπόλοιπα δεν ανέφεραν κάποιο μέτρο κινδύνου. Τα μισά άρθρα που ανέφεραν αλληλεπιδράσεις γονιδίων περιέγραφαν τουλάχιστον μια στατιστικά σημαντική αλληλεπίδραση.

Πίνακας 2.1.5: Περιγραφόμενα χαρακτηριστικά της ανάλυσης πολλαπλών γενετικών παραγόντων για 315 τυχαία επιλεγμένα άρθρα γενετικής επιδημιολογίας (2001-2003).

Περιγραφόμενο χαρακτηριστικό	Αριθμός	Ποσοστό
Αριθμός των αναλυόμενων γονιδίων		
1	188	59.7
2	71	22.5
≥3	56	17.8
Αριθμός των αναλυόμενων γενετικών παραγόντων		
1	114	36.2
2	85	27.0
3	56	17.8
4	21	6.6
≥5	39	12.4
Αναφορά για ανισορροπία σύνδεσης (ανάμεσα σ' αυτά που μελετούσαν 2 ή περισσότερους πολυμορφισμούς) (n=201)		
Όχι	155	77.1
Ναι	46	22.9
Αναφορά για την ανάλυση απλοτύπων (ανάμεσα σ' αυτά που μελετούσαν 2 ή περισσότερους πολυμορφισμούς) (n=201)		
Όχι	158	78.6
Ναι	43	21.4
Αναφορά για αλληλεπίδραση γονιδίων (ανάμεσα σ' αυτά που μελετούσαν 2 ή περισσότερους πολυμορφισμούς) (n=201)		
Όχι	152	75.6
Ναι	49	24.4



Αν αναφέρονταν αλληλεπίδραση γονιδίων, χρησιμοποιήθηκε ο λόγος αναλογιών ή ο λόγος κινδύνων (n=49)		
Όχι	22	44.9
Ναι	27	55.1
Αν αναφέρονταν αλληλεπίδραση γονιδίων, χρησιμοποιήθηκε η απόλυτη διαφορά (n=49)		
Όχι	47	95.9
Ναι	2	4.1
Αν αναφέρονταν αλληλεπίδραση γονιδίων, χρησιμοποιήθηκε το αποδοτέο κλάσμα (n=49)		
Όχι	49	100.0
Ναι	0	0.0
Αν αναφέρονταν αλληλεπίδραση γονιδίων, ο ισχυρισμός έγινε για μια στατιστικά σημαντική αλληλεπίδραση (n=49)		
Όχι	24	49.0
Ναι	25	51.0

2.1.3.3 Σύγκριση των άρθρων του 2001-2003 με εκείνα του 2006

Ο αριθμός των συμμετεχόντων, τα γονίδια και οι πολυμορφισμοί υπό ανάλυση ήταν παρόμοια μεταξύ των δυο περιόδων (Πίνακας 2.1.6). Τα άρθρα από το δείγμα του 2006 συχνότερα χρησιμοποιούσαν μεγέθη δείγματος κάτω από 500 συμμετέχοντες (75%), περιέγραφαν ένα μόνο γονίδιο (75%) και πολλαπλούς γενετικούς παράγοντες (64.3%).

Τρία από τα δέκα στοιχεία που αναφέρονταν σε λιγότερο από το 50% των άρθρων του 2001-2003 περιγράφονταν στατιστικά σημαντικά συχνότερα σε εκείνα του 2006 (Πίνακας 2.1.6). Οι μελέτες που ήταν δημοσιευμένες το 2006 ήταν περισσότερο πιθανόν να αναφέρουν την ισχύ της μελέτης (2001-2003: 12.7%, 2006: 28.6%, $p=0.03$), τη χρήση συμμετεχόντων χωρίς συγγένεια (2001-2003: 17.8%, 2006: 35.7%, $p=0.03$) και την επικύρωση των αποτελεσμάτων γονοτύπησης χρησιμοποιώντας διπλά δείγματα (2001-2003: 7%, 2006: 21.4%, $p=0.02$). Παρόλα αυτά, κάθε στοιχείο εκτός από την ισορροπία Hardy-Weinberg αναφέρονταν σε λιγότερα από τα μισά άρθρα του δείγματος του 2006.



Πίνακας 2.1.6: Σύγκριση των περιγραφόμενων χαρακτηριστικών άρθρων γενετικής επιδημιολογίας από δυο διαφορετικές χρονικές περιόδους (2001-2003 vs. 2006).

Περιγραφόμενο χαρακτηριστικό	2001-2003		2006		p-value ^a
	Αριθμός	Ποσοστό	Αριθμός	Ποσοστό	
Βασικά περιγραφικά στοιχεία					
Αριθμός συμμετεχόντων					
<100	49	15.6	5	17.9	0.90
100-499	190	60.3	16	57.1	
500-999	47	14.9	4	14.3	
≥1000	29	9.2	3	10.7	
Αριθμός των αναλυόμενων γονιδίων					
1	188	59.7	21	75.0	0.12
2	71	22.5	4	14.3	
≥3	56	17.8	3	10.7	
Αριθμός των αναλυόμενων γενετικών παραγόντων					
1	114	36.2	10	35.7	0.29
2	85	27.0	4	14.3	
3	56	17.8	6	21.4	
4	21	6.6	1	3.6	
≥5	39	12.4	7	25.0	
Στοιχεία που αναφέρονταν σε λιγότερο από το 50% των άρθρων της περιόδου 2001-2003					
Αναφορά της ισχύος της μελέτης					
Όχι	275	87.3	20	71.4	0.03
Ναι	40	12.7	8	28.6	
Ξεκάθαρη δήλωση για τη μη-συγγένεια των συμμετεχόντων					
Όχι	259	82.2	18	64.3	0.03
Ναι	56	17.8	10	35.7	
Αναφερόμενη επικύρωση των αποτελεσμάτων γονοτύπησης με τη χρήση διπλών δειγμάτων					
Όχι	293	93.0	22	78.6	0.02
Ναι	22	7.0	6	21.4	
Αναφερόμενη επικύρωση των αποτελεσμάτων γονοτύπησης με τη χρήση διαφορετικής μεθόδου					
Όχι	284	90.2	24	85.7	0.32
Ναι	31	9.8	4	14.3	
Αναφορά πως η αξιολόγηση των γενετικών ελέγχων ήταν τυφλή στις εκβάσεις ή στους φαινότυπους					
Τυφλοποίηση	35	11.1	3	10.7	0.62
Ασαφές	280	88.9	25	89.3	



Αναφορά πως η αξιολόγηση της έκβασης ή του φαινότυπου ήταν τυφλή στα αποτελέσματα του γενετικού ελέγχου					
Τυφλοποίηση	12	3.8	2	7.1	0.32
Ασαφές	303	96.2	26	92.9	
Αναφορά για αποκλεισμό ατόμων από τις αρχικές ομάδες συμμετεχόντων					
Όχι	278	88.2	26	92.9	0.86
Ναι	37	11.8	2	7.1	
Αναφορά αποτυχίας γονοτύπησης δειγμάτων					
Όχι	266	84.4	23	82.1	0.46
Ναι	49	15.6	5	17.9	
Περιγραφή πως ασύνδετοι γενετικοί δείκτες χρησιμοποιήθηκαν για την ανίχνευση πληθυσμιακής διαστρωμάτωσης					
Όχι	313	99.4	28	100.0	1.00
Ναι	2	0.6	0	0.0	
Αναφορά πως όλοι οι γενετικοί παράγοντες εξετάστηκαν για ισορροπία Hardy-Weinberg					
Όχι	164	52.1	12	42.9	0.23
Ναι	151	47.9	16	57.1	

^a Οι τιμές p υπολογίστηκαν από δίπλευρα Mann-Whitney U tests για περιγραφικά στοιχεία και από μονόπλευρα Fisher's Exact Tests για τα στοιχεία που αναφέρονταν σε λιγότερο από το 50% των άρθρων του 2001-2003.

2.1.4 Συζήτηση

Πολλοί δημοσιευμένοι ισχυρισμοί για συσχετίσεις γονιδίων-νοσημάτων δεν έχουν επιβεβαιωθεί όταν μελετήθηκαν σε ανεξάρτητα δείγματα. [41, 42] Θεωρούμενες αιτίες αυτής της ασυνέπειας περιλαμβάνουν την αξιολόγηση της στατιστικής σημαντικότητας χωρίς τον συνυπολογισμό της χαμηλής εκ των προτέρων πιθανότητας για συσχέτιση, τη χαμηλή στατιστική ισχύ, την ακατάλληλη επιλογή του δείγματος, τα σφάλματα μέτρησης, τους συγχυτικούς παράγοντες και η επιλεκτική παρουσίαση των αποτελεσμάτων. [17, 30, 37, 38, 41, 42, 45-48] Προηγούμενες αναλύσεις έδειξαν πως πολλά δημοσιευμένα άρθρα γενετικής επιδημιολογίας δε παρέχουν επαρκείς πληροφορίες για την αξιολόγηση αυτών. [49-51] Παρόλα αυτά, τα αποτελέσματα αυτών των αναλύσεων περιορίζονταν σε ένα συγκεκριμένο φαινότυπο (πχ σήψη) [49, 50] ή είναι ξεπερασμένα. [51] Η παρούσα ανάλυση παρέχει μια ανανεωμένη ανασκόπηση της αναφοράς αυτών των σημαντικών στοιχείων σε δυο αντιπροσωπευτικά δείγματα της γενετικής επιδημιολογίας.



Οι περισσότερες μελέτες είχαν μικρό αριθμό δείγματος. Αρκετές μετα-αναλύσεις έχουν βρει στατιστικά σημαντικές διαφορές μεταξύ των αποτελεσμάτων μικρών και μεγάλων μελετών γενετικής συσχέτισης. [29, 42] Αυξανόμενα τεκμήρια συνηγορούν πως ατομικοί γενετικοί παράγοντες έχουν μέτρια επίδραση στον κίνδυνο ανάπτυξης σύνθετων, πολυπαραγοντικών νοσημάτων.[18, 52, 53] Συνεπώς, μπορεί να χρειαστεί η συμμετοχή πολλών χιλιάδων ατόμων για να επιτευχθεί η αναγκαία ισχύς για την αναγνώριση και επιβεβαίωση πραγματικών γενετικών συσχετίσεων.[30, 52, 54, 55]

• Η δυνατότητα αξιολόγησης πιθανών σφαλμάτων επιλογής ήταν σημαντικά περιορισμένη σε πολλές εξεταζόμενες μελέτες. Παρόλο που οι περισσότερες μελέτες παρείχαν κάποια ποιοτική πληροφορία για τους συμμετέχοντες (όπως η καταγωγή και τα κριτήρια εισαγωγής), η περιγραφή αυτής ήταν σποραδική ακόμα και για απλά στοιχεία, όπως η ηλικία και το φύλο. Δυνάμει σημαντικές πληροφορίες, όπως ο αριθμός των αποκλεισμένων συμμετεχόντων ή ο αριθμός των δειγμάτων που δεν μπορούσαν να γονοτυπηθούν, συχνά παραλείπονταν.

Η δυσταξινομήση μπορεί να περιορίσει σημαντικά την ισχύ της μελέτης και να επηρεάσει τα αποτελέσματα.[20, 47, 54, 56, 57] Ένα μικρό ποσοστό των μελετών περιέγραφαν μέτρα όπως η γονοτύπηση επαναληπτικών δειγμάτων και τυφλοποίηση του ερευνητικού προσωπικού,[57, 58] ώστε να διασφαλιστεί πως τα γενετικά δεδομένα δεν είχαν δυσταξινομηθεί. Παρόλο που η πρακτική του ελέγχου για γονοτυπικά σφάλματα με τη χρήση της παρέκκλισης από την ισορροπία Hardy-Weinberg είναι υπό συζήτηση, [57, 59-61] περίπου οι μισές μελέτες περιέγραφαν τέτοια αποτελέσματα.

Η πληθυσμιακή διαστρωμάτωση μπορεί να συμβεί όταν οι συμμετέχοντες σε μια μελέτη επιλέγονται από υποπληθυσμούς με διαφορετικό επιπολασμό των φαινοτύπων και γονοτύπων.[19, 27, 53, 60] Παρόλο που η έκταση κατά την οποία η πληθυσμιακή διαστρωμάτωση οδηγεί σε εσφαλμένα ευρήματα παραμένει άγνωστη, [62-64] τα περισσότερα άρθρα του δείγματος παρείχαν πληροφορίες για την εθνική προέλευση των συμμετεχόντων και σχεδόν όλες οι μελέτες ασθενών μαρτύρων ανέφεραν πως οι μάρτυρες και οι ασθενείς προέρχονταν από τον ίδιο πληθυσμό. Ελάχιστες μελέτες περιέγραφαν τη χρήση ασύνδετων δεικτών για τον έλεγχο της πληθυσμιακής διαστρωμάτωσης. Οι μελέτες συσχέτισης εκτεταμένου γονιδιώματος (genome wide association studies) παρέχουν εξαιρετικά γενετικά δεδομένα για την εξέταση και τη διόρθωση για πληθυσμιακή διαστρωμάτωση.[65-67]

Ως αποτέλεσμα της επιλεκτικής αναφοράς στατιστικά σημαντικών αποτελεσμάτων από πολλαπλές μετα-αναλύσεις και το συστηματικό σφάλμα αναφοράς, η έκταση του σφάλματος τύπου



Η στη δημοσιευμένη βιβλιογραφία μπορεί να είναι μεγάλη.[17, 47, 53, 54, 56, 58, 60, 68, 69] Μόνο μια μειοψηφία των άρθρων ανέφερε τη χρήση διόρθωσης για πολλαπλές συγκρίσεις. Αναφερόμενες δικαιολογίες για συγκεκριμένες γενετικές συγκρίσεις μπορεί να σημαίνει πως βασίζονταν σε εκ των προτέρων σχεδιασμό και δεν ήταν προϊόν επιλεκτικής περιγραφής. Παρόλα αυτά, λιγότερες από τις μισές μελέτες που χρησιμοποίησαν το επικρατές ή υπολειπόμενο μοντέλο παρείχαν αιτιολόγηση για τη χρήση αυτών των συγκρίσεων και όχι άλλων. Ανάμεσα στα άρθρα που περιέγραφαν αλληλεπιδράσεις γονιδίων ένα σημαντικό ποσοστό ανέφερε στατιστικά σημαντικά αποτελέσματα. Παρόλα αυτά, πολλά από αυτά πιθανώς είναι εσφαλμένα, δεδομένης της περιορισμένης ισχύος των περισσότερων μελετών να αναγνωρίσουν αληθινές συσχετίσεις, πόσο μάλλον αλληλεπιδράσεις.[19, 52] Η μεγάλη συχνότητα των “θετικών” αποτελεσμάτων στην παρούσα ανάλυση μπορεί να αντιπροσωπεύει ένα συνδυασμό πολλαπλών συγκρίσεων, επιλεκτική αναφοράς και συστηματικού σφάλματος δημοσίευσης.[47]

Η χρήση κοινών προτύπων αναφοράς θα μπορούσε να βελτιώσει τη διαφάνεια της ερευνητικής μεθοδολογίας, βοηθώντας έτσι να αναγνωριστεί η επιλεκτική αναφορά και οι πηγές συστηματικών σφαλμάτων και συγχυτικών παραγόντων, ενώ θα επέτρεπε μια πιο ολοκληρωμένη σύνθεση των δεδομένων σε συνασπισμούς ή μετα-αναλύσεις.[37, 45, 70]

Με τη δημοσίευση επιπρόσθετης πληροφορίας διαδικτυακά, τα περιοδικά μπορούν να παρέχουν στους συγγραφείς την ευκαιρία να παρουσιάσουν τη μεθοδολογία της μελέτης τους και τα αποτελέσματα σε μεγαλύτερη λεπτομέρεια από το έντυπο άρθρο.[71, 72]

Συνοψίζοντας, τα αποτελέσματα της παρούσας διατριβής παρέχουν τεκμήρια πως πολλές πληροφορίες, που είναι απαραίτητες για την αξιολόγηση της εγκυρότητας των επιστημονικών ευρημάτων, δεν αναφέρονται συνεπώς στη βιβλιογραφία της γενετικής επιδημιολογίας. Η χρήση στανταρισμένων οδηγιών αναφοράς θα βοηθούσε τον αναγνώστη να αξιολογήσει καλύτερα τα επιστημονικά ευρήματα.



2.2 Αναλύσεις υποομάδων σε πολυπαραγοντικά νοσήματα: το παράδειγμα της σύγκρισης φύλων.

2.2.1 Σκοπός

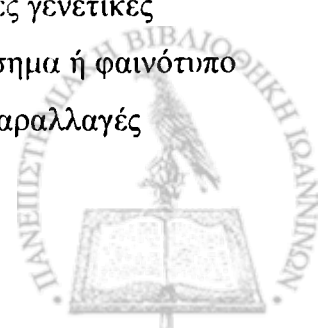
Η εμπειρική περιγραφή ενός μεγάλου δείγματος προεξεχόντων ισχυρισμών για διαφορετικές γενετικές επιδράσεις στα δυο φύλα. Η αξιολόγηση κατά πόσο αυτοί οι ισχυρισμοί είναι μεθοδολογικά ισχυροί ή βασιστήκανε σε επιλεκτικές και/ή υποδεέστερες αναλύσεις και ανεπαρκή ή εσφαλμένη τεκμηρίωση. Επίσης, ο έλεγχος αν οι ισχυρισμοί που είχαν την καλύτερη εσωτερική εγκυρότητα είχαν επιβεβαιωθεί από άλλες μελέτες.

2.2.2 Μέθοδοι

2.2.2.1 Επιλογή των μελετών

Σκοπός ήταν η συλλογή ενός δείγματος μελετών που ισχυρίζονταν διαφορετικότητα των δυο φύλων σε συσχετίσεις γονιδίων με νόσους και στα οποία ο ισχυρισμός ήταν τόσο προεξέχων που εμφανίζονταν ακόμα και στον τίτλο του άρθρου. Οι συγκρίσεις υποομάδων με βάση το φύλο είναι ένα πολύ κοινό τοπίο στην επιδημιολογική βιβλιογραφία. Η συλλογή όλων αυτών ή ενός συστηματικού αποσπάσματος θα ήταν απαγορευτική. Κάθε ηλεκτρονική αναζήτηση θα απέδιδε μόνο ένα μικρό δείγμα από τα εκατοντάδες άρθρα που αξιολογούν διαφορές υποομάδων με βάση το φύλο, διότι αυτές συνήθως εμφανίζονται ως δεύτερες ή τρίτες αναλύσεις. Αντιθέτως, με την εστίαση στον τίτλο ενός άρθρου, θα ευνοούνταν η επιλογή των περισσότερο προεξαρχόντων διαφορών υποομάδων ανάμεσα στα φύλα. Αυτές οι μελέτες είναι εκείνες στις οποίες οι συγγραφείς (και πιθανώς οι κριτές και οι εκδότες) είναι περισσότερο βέβαιοι για την ισχύ των παρατηρούμενων διαφορών υποομάδων με βάση το φύλο.

Η αναζήτηση έγινε στο Pubmed έως τις 6 Ιουλίου 2007. Το Pubmed θεωρείται πως καλύπτει περισσότερο από ικανοποιητικά τις γενετικές επιδημιολογικές μελέτες.[43] Χρησιμοποιήθηκε μια στρατηγική αναζήτησης που θα είχε υψηλή ειδικότητα στη συλλογή ενός δείγματος επιλέξιμων μελετών: `polymorphism* [ti] AND (gender [ti] OR sex [ti])`. Μελετήθηκαν οι τίτλοι και οι περίληψεις και, σε περίπτωση αμφιβολιών, το πλήρες κείμενο για πλήρωση των κριτηρίων εισαγωγής. Θεωρήθηκαν επιλέξιμες όλες οι μελέτες που ισχυρίζονταν διαφορετικές γενετικές επιδράσεις στα δυο φύλα για μία ή περισσότερες γενετικές παραλλαγές σε ένα νόσημα ή φαινόμενο στους ανθρώπους. Μελετήθηκαν τόσο οι διαλληλικές όσο και οι πολυαλληλικές παραλλαγές



(συμπεριλαμβανομένων και των απλοτύπων) καθώς και οι δυαδικοί και συνεχείς φαινότυποι και νοσήματα.

Αποκλείστηκαν μελέτες που δεν ήταν γραμμένες στα αγγλικά, μελέτες για ένα μόνο φύλο, μελέτες στις οποίες το φύλο χρησιμοποιούνταν ως ανεξάρτητος εκτιμητής της νόσου, μελέτες που δεν προσπαθούσαν να συσχετίσουν γενετικές παραλλαγές με ένα νόσημα ή φαινότυπο, μελέτες που δεν είχαν ανθρώπους ως δείγμα και μελέτες στις οποίες οι συγγραφείς δεν κάνανε κανένα ισχυρισμό για διαφορές μεταξύ των δυο φύλων. Οι μελέτες που πληρούσαν τα κριτήρια εισαγωγής συμπεριελήφθησαν ανεξάρτητα από την έκταση της ποσοτικής πληροφορίας που παρείχαν για να υποστηρίξουν τους ισχυρισμούς τους.

2.2.2.2 Εξαγωγή των δεδομένων

Εξήχθησαν τα ακόλουθα δεδομένα από κάθε επιλέξιμη μελέτη: πρώτος συγγραφέας, περιοδικό δημοσίευσης και έτος, συνολικό δείγμα, ποσοστό γυναικών, γονίδια και παραλλαγές για τα οποία υπήρχαν ισχυρισμοί, και νοσήματα/φαινότυπους.

Για κάθε ζεύγος γενετικής παραλλαγής και φαινότυπου για το οποίο υπήρχε ένας ισχυρισμός, καταγράφηκε η ακριβής διατύπωση του ισχυρισμού και κάθε νύξη πως η διαφορά ελέγχθηκε βασισμένη σε σχεδιασμό εκ των προτέρων (a priori) ή ως τμήμα μεταγενέστερων αναλύσεων (post hoc)· το είδος των γενετικών παραλλαγών (διαλληλικές, πολυαλληλικές και απλότυποι)· το είδος του φαινότυπου (δυαδικό ή συνεχές γνώρισμα). Ειδικά για τις διαλληλικές παραλλαγές καταγράψαμε τη χρήση κάποιου γενετικού μοντέλου [υπολειπόμενο (recessive), επικρατές (dominant), αθροιστικό (additive), αλληλικό (allele-based), ελεύθερο μοντέλο (model-free), ή άλλο/αδιευκρίνιστο]. Επίσης, καταγράφηκε για κάθε φύλο το μέγεθος των επιδράσεων (effect sizes) και τα μέτρα της αβεβαιότητας αυτών ή τα απόλυτα δεδομένα που θα μπορούσαν να χρησιμοποιηθούν για την επιβεβαίωση της παρουσίας των ισχυρισμών.

Για κάθε ισχυρισμό διαφοράς μεταξύ φύλων καταγράφηκε εάν ήταν βασισμένος σε ανάλυση ολόκληρου του δείγματος ή υποομάδας αυτού. Στην περίπτωση του τελευταίου, καταγράφηκε ο ορισμός της υποομάδας και αν βασιζόταν σε εκ των προτέρων σχεδιασμό. Επίσης, καταγράφηκε εάν οι υποομάδες ορίζονταν βασισμένες σε γενετική πληροφορία για την ίδια γενετική παραλλαγή με αυτή του ισχυρισμού (πχ σύγκριση των AA έναντι των αα ομόζυγων χωρίς εκτίμηση των Aa ετερόζυγων, σύγκριση του απλότυπου 1 έναντι του 4 χωρίς εκτίμηση των 2 και 3), γενετική πληροφορία για κάποια άλλη γενετική παραλλαγή, άλλα χαρακτηριστικά του δείγματος (πχ ηλικία), άλλους παράγοντες ή συνδυασμούς των παραπάνω.



Για κάθε ισχυρισμό διαφορών μεταξύ φύλων καταγράφηκε εάν είχε επαρκή υποστήριξη (documentation), ανεπαρκή υποστήριξη, ή εάν ήταν εσφαλμένος. Για να έχει ένας ισχυρισμός επαρκή υποστήριξη θα έπρεπε να πληρούνται 3 κριτήρια. Πρώτον, ο ισχυρισμός θα έπρεπε να βασίζεται στην ίδια γενετική σύγκριση μεταξύ των δυο φύλων. Δεύτερον, δε θα σύγκρινε διαφορετικές υποομάδες στα δυο φύλα (πχ ηλικιωμένοι άνδρες έναντι νεαρών γυναικών). Τρίτον, θα έπρεπε να περιγράφουν είτε ένα στατιστικά σημαντικό τεστ (τιμή $p < 0.05$) που θα έλεγχε την αλληλεπίδραση (interaction) φύλου-γονιδίου, είτε η αλληλεπίδραση θα έπρεπε να ήταν εμφανής διότι τα περιγραφόμενα όρια εμπιστοσύνης των επιδράσεων για κάθε φύλο δε θα ήταν αλληλοεπικαλυπτόμενα. Ανεπαρκώς υποστηριζόμενοι ισχυρισμοί ήταν εκείνοι που πληρούσαν μόνο τα δυο πρώτα κριτήρια, ενώ εσφαλμένοι αυτοί δε θα πληρούσαν είτε ένα είτε και τα δύο πρώτα κριτήρια.

Ανεπαρκής υποστήριξη δε σημαίνει πως ο ισχυρισμός είναι απαραίτητα ακατάλληλος και εσφαλμένος. Ένας ανεπαρκώς υποστηριζόμενος ισχυρισμός μπορεί να είναι σωστός (να υπάρχει πραγματικά στατιστικά σημαντική αλληλεπίδραση) ή εσφαλμένος (να μην υπάρχει στατιστικά σημαντική αλληλεπίδραση) αλλά αυτό δεν είναι ξεκάθαρο από τις περιγραφόμενες αναλύσεις που παρουσιάζονται στο αντίστοιχο άρθρο.

2.2.2.3 Επανεξέταση των ισχυρισμών

Για κάθε ισχυρισμό που δεν ήταν εσφαλμένος και υπήρχαν διαθέσιμες πληροφορίες, επιχειρήθηκε να ελεγχθεί εάν όντως υπήρχε στατιστικά σημαντική διαφορά επιδράσεων ανάμεσα στα δυο φύλα (αλληλεπίδραση φύλου-γονιδίου).

Οποτεδήποτε ένα επικρατές, υπολειπόμενο, ή αλληλικό μοντέλο υπονοούνταν, χρησιμοποιήθηκε για κάθε φύλο ο φυσικός λογάριθμος του λόγου αναλογιών (odds ratio) για δυαδικές εκβάσεις ή η απόλυτη διαφορά για συνεχείς εκβάσεις και οι αντίστοιχες διακυμάνσεις (variance). Οποτεδήποτε δεν περιγράφονταν μεγέθη επιδράσεων και για τα δυο φύλα, έγινε προσπάθεια να υπολογιστούν από τις παρουσιαζόμενες πληροφορίες για τα διαστήματα εμπιστοσύνης, την τυπική απόκλιση, ή το τυπικό σφάλμα του μέσου, και τον αριθμό των παρατηρήσεων. Έπειτα υπολογίστηκε η τιμή του z ως λόγος του οποίου αριθμητής είναι η διαφορά των κανονικοποιημένων (normalized) επιδράσεων και παρονομαστής η τετραγωνική ρίζα της διακύμανσης της διαφοράς. Σε περίπτωση αθροιστικού μοντέλου μοντελοποιήθηκαν μοντέλα τάσης (trend model) για κάθε φύλο και χρησιμοποιήθηκαν οι παραγόμενοι συντελεστές και τα αντίστοιχα τυπικά σφάλματα για να υπολογιστεί η τιμή του z . Τελικά, σε περίπτωση ελεύθερου μοντέλου χρησιμοποιήθηκε ανάλυση διακυμάνσεων με αλληλεπίδραση φύλου-γονιδίου.



Όλες αυτές οι αναλύσεις διεξήχθησαν χρησιμοποιώντας δεδομένα που αναφέρονταν στα ίδια άτομα για τα οποία είχε γίνει ο ισχυρισμός από τους συγγραφείς. Κατά συνέπεια, εφόσον ο ισχυρισμός για το φύλο είχε γίνει για ολόκληρο τον υπό μελέτη πληθυσμό ελέγχθηκε για αλληλεπίδραση φύλου-γονιδίου στον ίδιο ολόκληρο πληθυσμό. Εάν ο ισχυρισμός είχε γίνει για μια υποομάδα ελέγχθηκε για αλληλεπίδραση φύλου-γονιδίου στην ίδια υποομάδα. Όταν αστάθμιστες και σταθμισμένες εκτιμήσεις επιδράσεων ήταν διαθέσιμες χρησιμοποιήθηκαν οι αστάθμιστες, εκτός και αν περιγράφονταν ως ανεξάρτητοι ισχυρισμοί.

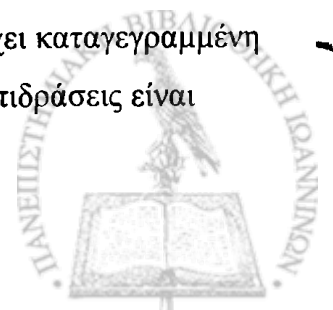
2.2.2.4 Επιβεβαίωση των στατιστικά σημαντικών αλληλεπιδράσεων φύλου-γονιδίων από άλλες μελέτες

Ακόμα κι αν μια αλληλεπίδραση φύλου-γονιδίου βασίζεται στη βέλτιστη στατιστική κι αναλυτική υποστήριξη, δε σημαίνει πως αυτή η υποστήριξη είναι όντως πραγματική. Οι γενετικές επιδράσεις είναι εκτεθειμένες σε μια εκτεταμένη πολλαπλότητα ελέγχων και σε διαφορετικά λάθη και συστηματικά σφάλματα.[17] Επομένως, η επιβεβαίωση από επιπρόσθετες ανεξάρτητες μελέτες θεωρείται απαραίτητη για την ενίσχυση της αξιοπιστίας των γενετικών επιδράσεων.[29] Για τους παραπάνω λόγους ελέγχθηκε κατά πόσο οι προτεινόμενες αλληλεπιδράσεις φύλου-γονιδίου στο δείγμα είχαν όντως αξιολογηθεί από επιπρόσθετες μελέτες και κατά πόσο τα αποτελέσματα αυτών συμφωνούσαν με τις προτεινόμενες αλληλεπιδράσεις.

Η εστίαση έγινε σε ισχυρισμούς διαφορών μεταξύ των δυο φύλων που πληρούσαν όλα τα ακόλουθα κριτήρια: οι αναλύσεις τους βασίζονταν σε εκ των προτέρων σχεδιασμό, πλήρη δεδομένα ήταν διαθέσιμα, οι αλληλεπιδράσεις φύλου-γονιδίου ήταν στατιστικά σημαντικές στο επίπεδο $p=0.05$ στην επανεξέταση των δεδομένων και είτε είχε αναλυθεί ολόκληρο το δείγμα είτε μια υποομάδα που είχε επιλεγεί με βάση εκ των προτέρων σχεδιασμό. Αυτοί οι ισχυρισμοί ήταν εκείνοι που προφανώς είχαν την βέλτιστη πιθανή εσωτερική εγκυρότητα.

Μελετήθηκε λεπτομερώς κάθε άρθρο που έκανε τέτοιους ισχυρισμούς και καταγράφηκε κατά πόσο είχαν αναφέρει προηγούμενες μελέτες που ερευνούσαν την ίδια αλληλεπίδραση φύλου-γονιδίου. Επιπρόσθετα, αναζητήθηκαν στη βάση ISI Web of Knowledge άρθρα που ανέφεραν άρθρα που πληρούσαν τα ανωτέρω κριτήρια. Κατόπιν αυτά μελετήθηκαν και ελέγχθηκε κατά πόσο είχαν αξιολογήσει την ίδια αλληλεπίδραση φύλου-γονιδίου, κι αν ναι, τι βρήκαν.

Η στρατηγική αναζήτησης δεν είχε 100% ευαισθησία να ανασύρει όλες τις μελέτες που ελέγξαν τέτοιες αλληλεπιδράσεις φύλου-γονιδίου. Παρόλα αυτά, δεν υπάρχει καταγεγραμμένη αξιόπιστη αναζήτηση να ανασύρει όλα τα σχετικά άρθρα. Επίσης, οι αλληλεπιδράσεις είναι



πιθανώς συχνά “θαμμένες” μέσα στο κείμενο ή δεν περιγράφονται καθόλου, ειδικά όταν δεν είναι στατιστικά σημαντικές. Όμως, προηγούμενες μελέτες που υποστηρίζουν τον ίδιο ισχυρισμό είναι πολύ πιθανό να έχουν αναφερθεί σε άρθρα που αλληλεπιδράσεις φύλου-γονιδίου είναι το κύριο θέμα ή αναφέρονται στον τίτλο του άρθρου, όπως το δικό μας δείγμα μελετών. Παρόμοια, επακόλουθες μελέτες που εντόπισαν τις ίδιες αλληλεπιδράσεις φύλου-γονιδίου, είναι πιθανό να αναφέρουν ένα άρθρο του οποίου το κεντρικό θέμα ήταν αυτές οι αλληλεπιδράσεις. Συνολικά, η στρατηγική ευνοεί περισσότερο την ανάσυρση μελετών που υποστηρίζουν παρά διαφωνούν με τις προτεινόμενες αλληλεπιδράσεις.

2.2.3 Αποτελέσματα

2.2.3.1 Επιλέξιμο δείγμα μελετών και ισχυρισμοί με βάση το φύλο

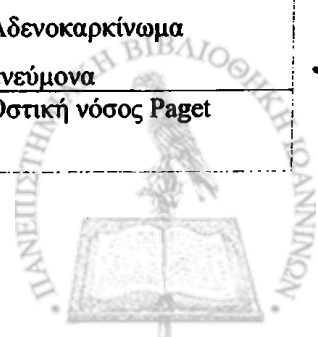
Η ηλεκτρονική αναζήτηση ανέδειξε 215 παραπομπές εκ των οποίων 138 αποκλείστηκαν μετά την αξιολόγησή τους: 5 μελέτες δεν ήταν γραμμένες στα αγγλικά, 34 άρθρα αξιολογούσαν μόνο το ένα φύλο, 11 χρησιμοποιούσαν το φύλο ως ανεξάρτητο εκτιμητή, 28 δεν ήταν μελέτες συσχέτισης, 54 δεν αναφέρονταν σε ανθρώπους, 4 δεν έκαναν κανένα ισχυρισμό και για 2 δεν ήταν δυνατό να εντοπιστεί το πλήρες κείμενό τους. Εβδομήντα επτά άρθρα [73-149] ήταν επιλέξιμα και περιείχαν συνολικά 432 διακριτούς ισχυρισμούς για διαφορές μεταξύ των δυο φύλων (διάμεσος=4, ενδοτερτατημοριακό εύρος [ETE], 2-7 ισχυρισμοί ανά άρθρο).

Αυτές οι μελέτες ήταν δημοσιευμένες μεταξύ 1994 και 2007 σε 63 διαφορετικά περιοδικά με διάμεσο δείκτη επιρροής (impact factor) 3.868 (ETE 2.826-5.699). Το διάμεσο μέγεθος δείγματος των μελετών ήταν 560 (ETE 274-921) και διάμεσο ποσοστό γυναικών 49% (ETE 44%-53%). Οι 432 ισχυρισμοί ήταν για 63 διαφορετικά γονίδια. Οι ισχυρισμοί αναφέρονταν σε ένα εύρος νοσημάτων και φαινοτύπων, με τα πιο κοινά να είναι η λοίμωξη από τον ιό της ηπατίτιδας C (n=32 ισχυρισμοί), η υψηλής πυκνότητας λιποπρωτεΐνη χοληστερίνης (HDL) (n=26), καρκίνος πνεύμονα (n=21), ο σακχαρώδης διαβήτης τύπου II (n=19), η σκλήρυνση κατά πλάκας (n=18), η υπέρταση (n=14), η χαμηλής πυκνότητας λιποπρωτεΐνη χοληστερίνης (LDL) (n=13), η οστική νόσος Paget (n=12) και η διαβητική νεφροπάθεια από διαβήτη τύπου I (n=10) (Πίνακας 2.2.1).

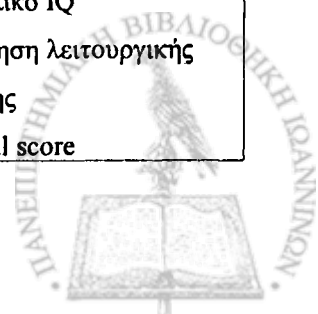


Πίνακας 2.2.1: Τα υπό εξέταση επιλέξιμα άρθρα που ισχυρίζονται για αλληλεπιδράσεις φύλου-γονιδίου στον τίτλο τους.

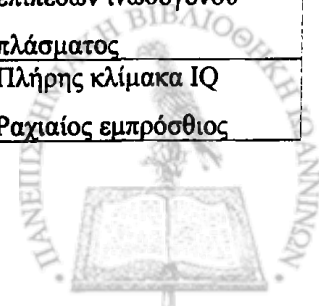
Πρώτος συγγραφέας	Έτος	Συνολικό δείγμα	Γυναίκες %	Αριθμός ισχυρισμών	Γονίδιο	Πολυμορφισμός/απλότυπος	Νόσημα/Γνώρισμα
Mepian C	2007	75	56	23	SePP	Ala234Thr	Selenoprotein P πλάσματος Glutathione peroxidase 3 Thioredoxin reductase 1 Erythrocyte thioredoxin reductase 1† Glutathione peroxidase 1 activity Glutathione peroxidase 1 protein Glutathione peroxidase 1 activity/protein ratio Lymphocyte glutathione peroxidase 4
Hayashi K	2007	200	58	10	ESR1	30T/C 401T/C	Συστολική αρτηριακή πίεση Μέση αρτηριακή πίεση HDL χοληστερόλη Ταχύτητα παλμού-κόματος βραχιονίου-αστραγάλου
Baud P	2007	612	58	4	COMT	Val158Met	Θυμός Έλεγχος θυμού
Yang KD	2007	898		2	CTLA-4	+49A/G	Επίπεδα IgE ομόλογο λώρου ≥0.5 kU/L
Glorioso N	2007	712	48	8	ATP1A1 Dear ATP1A1	SNP1 SNP10 SNP11 SNP12 CC halpotype SNP1-SNP2 TT halpotype TG halpotype	Υπέρταση Κανονική πίεση
Gallagher CJ	2007	1095	46	3	UGT2B17	ins/del	Καρκίνος πνεύμονα Αδενοκαρκίνωμα πνεύμονα
Beyens G	2007	302 690	44 53	12	TNFRSF1 IB	rs1485286 C/ T (SNP 11)	Οστική νόσος Paget



		992	50			rs2073617 rs6415470 rs11573869 GA rs2073618 TGGACGC TCGATGC haplotypes distribution	
Leung KH	2007	560	26	2	SLC11A1	SLC6a/b	Φυματίωση
Bolufer P	2007	897	47	9	GSTM1 NQO1 GSTT1	ins/del *2 variant ins/del	Οξεία μυελοβλαστική λευχαιμία Οξεία λεμφοβλαστική λευχαιμία
Yamaguchi S	2007	4854	45	19	THBS2 F3 ADIPOQ PON1	T>G(3' UTR) -603A>G G>T(intron 2) A>G (Arg160Gly)	Σακχαρώδης διαβήτης τύπου II
Russo P	2007	1604	50	4	CYP11B2	-344C/T	Συστολική αρτηριακή πίεση Διαστολική αρτηριακή πίεση Υπέρταση
Froehlich TE	2007	174	49	2	DRD4	7 allele	Μάθηση κανόνων και αντιστροφή, στάθμιση για συνολικές δοκιμές Μάθηση κανόνων και αντιστροφή, στάθμιση για ολοκληρωμένα επίπεδα
Dedoussis GV	2007	173	53	8	PPARg2	Pro12Ala	Ολικά τριγλυκερίδια (TC) Ολική χοληστερόλη HDL apoB λόγος TC/HDL λόγος apoB/apoA1
Barnett JH	2007	8707		6	COMT	Val108/158M et	Επιλεκτική προσοχή Λεκτικό IQ Συνολικό IQ Μέτρηση λειτουργικής μνήμης Global score



							Αντίθετες λέξεις
Schott E	2007	1274	52	11	CTLA-4	318C;49A A49G	Λοίμωξη από ιό της ηπατίτιδας C
Niemi M	2006	32	44	9	SLCO1B1	c.521T>C	Φαρμακοκινητική της πραβαστατίνης
Asselbergs FW	2006	2527	53	6	PAI AT1R Bradykini n B(2) receptor	4G/5G A1166C 58CT	ln(plasminogen activator inhibitor I levels) ln(tissue plasminogen activator)
Yiannakour is N	2003	118	53	2	Leptin	G-2548A	Επίπεδα διαλυτού υποδοχέα λεπτίνης Δείκτης ελεύθερης λεπτίνης
Seripa D	2006	1408	53	3	ApoE	e4 allele	Ηλικία ≥60 έτη (μακροζωία)
Paladino N	2006	495	43	21	IL-10	G-1082A	Λοίμωξη από ιό της ηπατίτιδας C
Korner A	2007	943	52	3	FAS	Val1483Ile	Δείκτης μάζας-σώματος HDL Λόγος LDL/HDL
Sundar PD	2006	1691	64	3	ABCA1	R219K R219K; G-17C	Πρόσφατης έναρξης νόσος Alzheimer
Ozawa T	2006	992	62	5	KCNQ1	G643S	Ρυθμός καρδιάς QTf (διόρθωση κατά Fridericia) Διάστημα κύματος T Λόγος Tpe/Qt
Ben Assayag E	2006	545	52	10	Fibrinoge n Bb	G455A	Ποσοστό ερυθροκυττάρων Ακτίνα κενού TKE Ινωδογόνο Συσχέτιση μεταξύ vacuum radius και επιπέδων ινωδογόνου πλάσματος Συσχέτιση μεταξύ ποσοστού ερυθροκυττάρων και επιπέδων ινωδογόνου πλάσματος
Kates WR	2006	58	45	3	COMT	Val158Met	Πλήρης κλίμακα IQ Ραχιαίος εμπρόσθιος

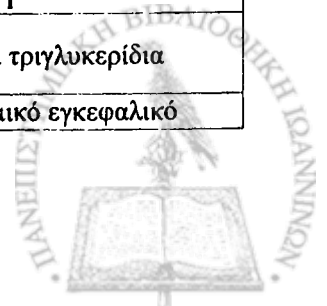


							όγκος εγκεφάλου Τροχιακός εμπρόσθιος όγκος εγκεφάλου
Mizuno T	2006	194	49	2	5-HTTLPR	long/short	Ευαισθησία στο στρες (κατάσταση) Ευαισθησία στο στρες (γνώρισμα)
Mlynarski WM	2005	794	48	9	CCRS	A59029G 32 bp ins/del A;Ins haplotype A;Del haplotype G;Ins	Διαβητική νεφροπάθεια σε διαβήτη τύπου I
Derzbach L	2005	308	50	4	ESR1	PvuII	Νεκρωτική εντεροκολίτιδα Ανοιχτός αρτηριακός πόρος Περίοδος παροχής οξυγόνου Ενδοκοιλιοτική αιμορραγία
Gloria- Bottini F	2005	337	46	1	ACPI		Λόγος βάρους γέννησης/βάρους πλακούντα
Sjoberg RL	2006	200	60	5	5-HTTLPR	I/s	Αυτοεκτιμώμενη κλίμακα κατάθλιψης
Tan EC	2005	616	36	5	Nogo	CAA insertion/dele tion TATC insertion/dele tion CAA+TATC- CAA+TATC+ CAA-TATC-	Σχιζοφρένεια
Shj Q	2005	2192	49	20	MTHFR	C677T A1298C	Καρκίνος πνεύμονα
Foltynie T	2005	291	40	1	BDNF	Val66Met	Ικανότητα σχεδιασμού στη νόσο Parkinson
Kajinami K	2005	338	40	5	ESR1 ApoA1 ESR1; APOA1	PvuII-XbaI+ +83 variant PvuII-XbaI+;	HDL χοληστερίνη

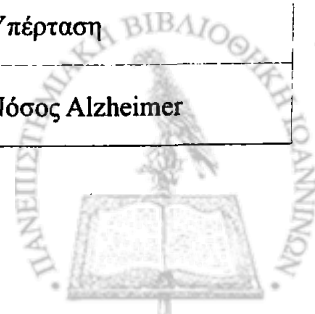


						+83 CA12/G haplotype CA13/A D12S313;3'(3 25)*G>A I1(761)*CAn ;D12S2510 D12S2510;D 12S2511 3'(325)*A I1(761)*CA1 2 I1(761)*CA1 3 3'(325)*G	
Kantarci OH	2005	861	66	13	IFNG		Σκλήρυνση κατά πλάκας
Chang JC	2004	644	47	2	CTLA-4	G49A	Επίπεδα IgE ομφάλιου λώρου
Cui J	2004	1000	48	1	TSP-4	A389P (29926G>C)	Έμφραγμα μυοκαρδίου
Schrijver HM	2004	477	52	5	TGFB1	T869C 869T;915G 869C;915C	Σκλήρυνση κατά πλάκας
Liu G	2004	1949	50	4	SOD2 SOD2;MP O	Ala16Val Ala16Val; MPO variant	Μη-μικροκυτταρικό καρκίνωμα
Szczeklik W	2004	857	60	1	COX-2	G-765C	Άσθμα
Corder EH	2004	5615	44	2	ApoE	Epsilon4	Γεροντική πλάκα
Chen W	2004	745	43	2	XBPI	197C/G	Σχιζοφρένεια
Kajinami K	2004	344	40	5	MDRI	C3435T G2677T/A; C3435T haplotype GC	HDL χοληστερίνη LDL χοληστερίνη
Yang KD	2004	1333	50	6	CTLA-4	G49A	Επίπεδα IgE Επίπεδα logIgE Επίπεδα IgE ≥100KU/L Αλλεργική ρινίτιδα
Gong MN	2004	189	33	2	SP-B	Exon 4 variant	Σύνδρομο οξείας αναπνευστικής δυσχέρειας Άμεσος πνευμονικός τραυματισμός
Ko YL	2004	716	46	4	HL	-514C/T	HDL χοληστερίνη

						-250G/A	
Nakanishi S	2004	249	51	5	FABP2	Ala54Thr	Ολικά τριγλυκερίδια Ολική χοληστερίνη LDL χοληστερίνη Δείκτης μάζας σώματος
Chen Y	2004	277	47	1	CCL2	-2518; -2076	Νόσος Behcet
Espino-Montoro A	2003	104	43	11	ApoC3	S1/S2	Ολικά τριγλυκερίδια VLDL τριγλυκερίδια LDL τριγλυκερίδια HDL τριγλυκερίδια Ολική χοληστερίνη VLDL χοληστερίνη LDL χοληστερίνη Λόγος TG-VLDL/HDL-C Ολική apoB Γλυκόζη εκκίνησης Ινσουλίνη εκκίνησης
Vandenbroeck K	2003	449	54	8	IFNG/IL2 6 region	D12S2510 IFNGCA*13; D12S2510*8; D12S2511*9	Ρευματοειδής αρθρίτιδα
Song J	2003	271	53	1	CYP11B2	C-344T	Επιβίωση νεφρών σε IgA νεφροπάθεια
Laule M	2003	1000	24	3	eNOS	CA repeat	Οξύ στεφανιαίο σύνδρομο
Reich H	2003	81	42	7	AGT1R	A1166C	Συστολική αρτηριακή πίση Διαστολική αρτηριακή πίση Μέση αρτηριακή πίση Αλλαγή στο ρυθμό σπειραματικής διήθησης Αύξηση αντίστασης νεφρικών αγγείων
Watson MA	2003	559	36	3	ApoE	Exon 4 variant	Ορθικός καρκίνος Προχωρημένοι όγκοι τάξης C&D κατά Dukes
Stankovic A	2002	385	45	3	ACE	I/D	Υπέρταση
Gyorffy B	2002	210	48	1	VDR	BsmI; ApaI; Tru9I	Σακχαρώδης διαβήτης τύπου I
Minihane AM	2002	135	44	1	APOC3	T2854G	Ολικά τριγλυκερίδια
Wu Y	2001	306	45	4	MTHFR	C677T	Ισχαιμικό εγκεφαλικό



Lio D	2002	450	57	1	IL-10 promoter	G-1082A	επεισόδιο Μακροζωΐα
Liu ZH	2002	417	62	10	ESR1	PvuII; XbaI	Νεφρίτιδα λύκου Αιματολογικές διαταραχές σε νεφρίτιδα λύκου Υπέρταση Σπειραματικοί θρόμβοι στη νεφρίτιδα του λύκου Σπειραματοσκλήρυνση στη νεφρίτιδα του λύκου Διάμεση αγγειίτιδα στην νεφρίτιδα του λύκου Διάμεσο τραύμα στην αγγειίτιδα του λύκου
Sciacca FL	2002	872	56	1	IL-1A	C889T	Ηλικία στη μυασθένια
Ordovas JM	2002	1577	52	7	APOA1	G75A	HDL χοληστερίνη
Oh JY	2001	1126	59	7	ApoE	exon 4 variant	Ολική χοληστερίνη LDL χοληστερίνη Περιφέρεια μέσης Κλάσμα Ολική/HDL χοληστερίνη
Ellsworth DL	2001	608	48	2	E-selectin	S128R	Απασβέσωση στεφανιαίας αρτηρίας
Du L	2000	186	59	5	5-HTTLPR 5-HTT	long/short VNTR	Ευσυνειδησία Νεύρωση
Bullido MJ	2000	371	62	3	LRP ApoE	exon 3 C>T A491T (-491-427) haplotype	Νόσος Alzheimer
Kark JD	2000	819	36	3	CETP	TaqI	HDL χοληστερίνη Apo A-I HDL-C και συσχετισμός με τριγλυκερίδια
Durlach A	1999	406	43	5	CETP	TaqI	HDL χοληστερίνη Πάθηση στεφανιαίων αγγείων σε διαβήτη τύπου II
Sagnella GA	1999	1366	54	2	ACE	I/D	Υπέρταση
Reynolds WF	1999	288	58	1	MPO	SpN	Νόσος Alzheimer



Freire MB	1998	242	43	1	AGT	M235T	Νεφροπάθεια σε διαβήτη τύπου I
Suarez F	1997	589		23	VDR	BsmI	Δείκτης μάζας σώματος Μήκος σώματος Βάρος σώματος Επιφάνεια σώματος
Lehtimaki T	1997	58	48	2	ApoE	Epsilon4	LDL χοληστερίνη Ολική χοληστερίνη
Carter AM	1997	502	54	4	Fibrinogen Bb	Arg448Lys	Οξύ εγκεφαλικό Επίπεδα ινωδογόνου στους 3 μήνες
Ferrieres J	1994	263	56	14	ApoE	Epsilon4	Ολική χοληστερίνη LDL χοληστερίνη Ολική apoB LDL apoB VLDL χοληστερίνη Ολικά τριγλυκερίδια
von Eckardstein A	1994	614	31	4	Apo A-IV	Gln360His	Apo A-I Apo A-IV Λεκιθίνη/χοληστερίνη Ακυλοτρανσφεράση
Hansen PS	1994	782	8	3	ApoB	all haplotypes(XbaI and Ins/Del)	Ισχαιμική καρδιακή νόσος

2.2.3.2 Αναφερόμενη εκ των προτέρων αξιολόγηση των ισχυρισμών

Από τους 432 ισχυρισμούς, 286 (66.2%) ήταν βασισμένοι σε εκ των προτέρων σχεδιασμένες συγκρίσεις των δυο φύλων, 68 (15.7%) σε μεταγενέστερη απόφαση και σε άλλους 78 (18.1%) το σχέδιο της ανάλυσης ήταν ασαφές.

2.2.3.3 Συγκρίσεις μεταξύ γενετικών παραλλαγών και φαινοτύπων

Οι γενετικές παραλλαγές ήταν διαλληλικές σε 328 ισχυρισμούς, πολυαλληλικές σε 44, απλότυποι σε 45 και πολλαπλοί πολυμορφισμοί εκτός απλοτύπων σε 15. Οι φαινότυποι ήταν δυαδικοί σε 212 ισχυρισμούς, συνεχείς σε 218 και διακριτοί σε άλλους 2.



Ανάμεσα στις διαλληλικές παραλλαγές ένα συγκεκριμένο γενετικό μοντέλο χρησιμοποιήθηκε σε 227 ισχυρισμούς (υπολειπόμενο σε 33, επικρατές σε 95, αθροιστικό σε 19, αλληλικό σε 16 και ελεύθερο μοντέλου σε 64). Οι υπόλοιποι 101 ήταν βασισμένοι σε άλλες, περισσότερο ιδιαίτερες συγκρίσεις (άγριου τύπου ομόζυγοι έναντι ομόζυγων παραλλαγής, $n=30$ · ετερόζυγοι έναντι ομόζυγων παραλλαγής, $n=8$ · ετερόζυγοι έναντι ομόζυγων άγριου τύπου, $n=13$ · και άλλες κατηγορίες, $n=50$).

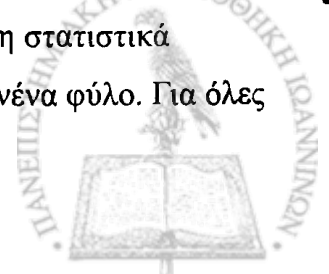
2.2.3.4 Πληθυσμοί υπό ανάλυση στους ισχυρισμούς

Το συνολικό μέγεθος δείγματος χρησιμοποιήθηκε σε 210 ισχυρισμούς (48.6%) και σε άλλους 222 χρησιμοποιήθηκαν υποομάδες αυτού. Η επιλογή της υποομάδας ήταν βασισμένη σε γενετική πληροφορία για την ενδιαφερόμενη γενετική παραλλαγή σε 60 ισχυρισμούς, σε γενετική πληροφορία για κάποια άλλη γενετική παραλλαγή σε δυο, σε άλλα χαρακτηριστικά του πληθυσμού σε 101 και σε συνδυασμούς των παραπάνω σε 59. Η επιλογή της υποομάδας αναφέρονταν ως σχεδιασμένη εκ των προτέρων σε 98 ισχυρισμούς (44.1%), σχεδιασμένη μετά την ανάλυση των δεδομένων σε 43 (19.4%) και αδιευκρίνιστη σε 81 ισχυρισμούς (36.5%).

2.2.3.5 Υποστήριξη των ισχυρισμών

Επαρκή υποστήριξη υπήρχε σε 55 ισχυρισμούς (12.7%). Αυτοί εμπεριείχαν 49 ισχυρισμούς που είχαν εκτελέσει στατιστικό έλεγχο για αλληλεπίδραση και άλλους έξι για τους οποίους δεν είχε γίνει στατιστικός έλεγχος, αλλά τα 95% διαστήματα εμπιστοσύνης των γενετικών επιδράσεων στα δυο φύλα δεν ήταν αλληλοεπικαλυπτόμενα. Σαράντα τέσσερις από τους 49 (89.8%) έδιναν τιμή p για την αλληλεπίδραση μεταξύ 0.01 και 0.05. Η μικρότερη τιμή p που περιγραφόταν ήταν 0.00008.

Ένα σύνολο 303 ισχυρισμών είχαν ανεπαρκή υποστήριξη. Σε 81 ισχυρισμούς, οι ερευνητές ανέφεραν πως στατιστικά σημαντικό αποτέλεσμα υπήρχε στο ένα φύλο αλλά όχι στο άλλο, ενώ και οι δυο επιδράσεις βρίσκονταν προς την ίδια κατεύθυνση, πχ προδιάθεση. Σε 46 ισχυρισμούς, οι ερευνητές ανέφεραν πως στατιστικά σημαντικό αποτέλεσμα υπήρχε στο ένα φύλο αλλά όχι στο άλλο, ενώ και οι επιδράσεις βρίσκονταν προς διαφορετικές κατευθύνσεις. Σε 16 ισχυρισμούς, οι ερευνητές ανέφεραν πως μια μεγαλύτερη ή περισσότερο στατιστικά σημαντική επίδραση υπήρχε στο ένα φύλο αλλά όχι στο άλλο, ενώ και οι δυο επιδράσεις ήταν στατιστικά σημαντικές προς την ίδια κατεύθυνση. Σε 107 ισχυρισμούς, μια στατιστικά σημαντική επίδραση υπήρχε στο ένα φύλο, αλλά δεν υπήρχε καμιά πληροφορία για στατιστική σημαντικότητα στο άλλο φύλο. Επιπρόσθετα, 53 ισχυρισμοί ανέφεραν στατιστικά σημαντική επίδραση στο ένα φύλο και μη στατιστικά σημαντική επίδραση στο άλλο, αλλά η κατεύθυνση δεν περιγραφόταν για κανένα φύλο. Για όλες



αυτές τις περιπτώσεις, ο τρόπος που διατυπωνόταν ο ισχυρισμός αποδείκνυε την ύπαρξη στατιστικά σημαντικής αλληλεπίδρασης φύλου-γονιδίου. Χαρακτηριστικά παραδείγματα παραθέτονται στον Πίνακα 2.2.2. Ανάμεσα στους ισχυρισμούς με ανεπαρκή υποστήριξη, εννέα περιγράφανε κάποιου είδους έλεγχο για αλληλεπίδραση φύλου-γονιδίου που ήταν στατιστικά μη σημαντικός στο επίπεδο του 0.05 για την τιμή του p (εύρος τιμής p: 0.088-0.983).

Πίνακας 2.2.2: Παραδείγματα ισχυρισμών με ανεπαρκή υποστήριξη.

Τύπος ανεπαρκούς υποστήριξης	Παράδειγμα*	Αναφορά
Στατιστικά σημαντική επίδραση σε ένα μόνο φύλο, επιδράσεις στην ίδια κατεύθυνση	"... after stratification by gender significantly increased odds of developing ARDS were found in women with a variant SP-B allele (OR, 4.5; 95% CI, 1.1 to 18.8; p= 0.03) [Table 4], but not in men" (OR, 1.2; 95% CI, 0.4 to 3.8).	[118]
Στατιστικά σημαντική επίδραση σε ένα μόνο φύλο, επιδράσεις σε αντίθετες κατευθύνσεις	"Analysis of male study subjects provided a statistically significant risk reduction for the ε3/ε3 genotype (OR=0.53; 95% CI 0.32–0.89)... whereas absolutely no effect was revealed in women (OR=1.04; 95%CI 0.52–2.09)."	[127]
Μεγαλύτερη ή περισσότερο στατιστικά σημαντική επίδραση για το ένα φύλο από ότι στο άλλο, στατιστικά σημαντικές επιδράσεις και στα δύο φύλα	"...in the male group...there was a slight and significant (p<0.05) increase in total, VLDL, LDL, and HDL-triglycerides and an increase of VLDL-cholesterol in carriers of the S1S2 genotype (p<0.05). ...the women carrying the S1S2 mutation had much more marked differences regarding plasma lipids and lipoprotein composition" Γυναίκες: p<0.001 για την ολικά, VLDL, και LDL-τριγλυκερίδια, και VLDL-χοληστερίνη; p<0.05 για HDL-τριγλυκερίδια (δεδομένα από πίνακα)	[122]
Στατιστικά σημαντική επίδραση στο ένα φύλο, καμιά πληροφορία για την κατεύθυνση στο άλλο φύλο	"Regression analysis showed no association between neonatal morbidity and genotype in girls. However, boys carrying "p" allele were at lower risk for patent ductus arteriosus (OR [95% CI]: 0.24 [0.05-0.971], P<0.05)."	[101]
Στατιστικά σημαντική επίδραση στο ένα φύλο, άγνωστη κατεύθυνση των επιδράσεων και στα δύο φύλα	"When significance was assessed separately for the two sexes, however, Taq1B polymorphism was found to affect HDL-C significantly in males (F =3.640, P=0.028) but not in females (F=0.947, P=0.390)"	[141]

* Παρατίθεται το πλήρες κείμενο στην αγγλική γλώσσα.



Συνολικά 74 ισχυρισμοί ήταν εσφαλμένοι. Οι κατηγορίες αυτών παρουσιάζονται στον Πίνακα 2.2.3. Οι πιο συχνές αιτίες ήταν η σύγκριση ανδρών ασθενών έναντι γυναικών με ένα συγκεκριμένο γονότυπο, διαφορώντας για τους άλλους γονότυπους (n= 28 ισχυρισμοί) και η σύγκριση ανδρών ασθενών απευθείας με γυναίκες ασθενείς διαφορώντας για την ομάδα ελέγχου (n= 8 ισχυρισμοί). Γενικά υπήρχε μια εκτεταμένη ποικιλία συγκρίσεων.

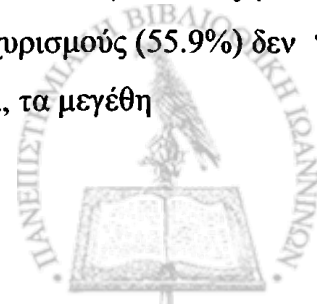
Πίνακας 2.2.3: Παραδείγματα εσφαλμένων ισχυρισμών.

Τύπος εσφαλμένου ισχυρισμού	Παράδειγμα*	Αναφορά
Σύγκριση των ανδρών ασθενών άμεσα με τις γυναίκες ασθενείς, αγνοώντας τους μάρτυρες	“Similarly, the frequency of 6/6 homozygotes was significantly higher (P= 0.023, EF= 0.24) and that of D12S2510*8 carriers lower (P= 0.025, PF= 0.26) in female patients compared with male patients.”	[123]
Σύγκριση ανδρών έναντι γυναικών για ένα δεδομένο γονότυπο αγνοώντας τους υπόλοιπους γονότυπους	“As a result, the difference in HDL-C between sexes, reconsidered on the basis of Taq1B genotype, was found to be significant at the P<0.0001 and P<0.0005 levels, respectively, in B1B1 and B1B2 subjects...”	[141]
Σύγκριση διαφορετικών γενετικών ομάδων στους άνδρες και στις γυναίκες	“... females of the AA genotype demonstrated the greatest fall in GFR compared with the other three groups (P= 0.01 versus males of the AC/CC genotype).”	[126]
Σύγκριση του ενός φύλου με μια υποομάδα του άλλου φύλου	“The distribution of haplotypes differed significantly among subgroups of patients, mainly because of a higher frequency of the Ins/X- and a lower frequency of the Del/X+ haplotypes in female and older male patients than in younger male patients and in reference men.”	[73]
Σύγκριση διαφορετικών υποομάδων κατηγοριών γονιδίων-έκθεσης	“For the MTHFR A1298C polymorphisms, the 1298CC genotype showed increased risk of lung cancer in those women who reported ever smoking (adjusted OR, 2.25; 95% CI, 1.19-4.23)... ” Men: adjusted OR, 0.44; 95% CI, 0.20-0.95 (data from table, not adjusted for smoking)	[105]

* Παρατίθεται το πλήρες κείμενο στην αγγλική γλώσσα.

2.2.3.6 Επανεξέταση των ισχυρισμών

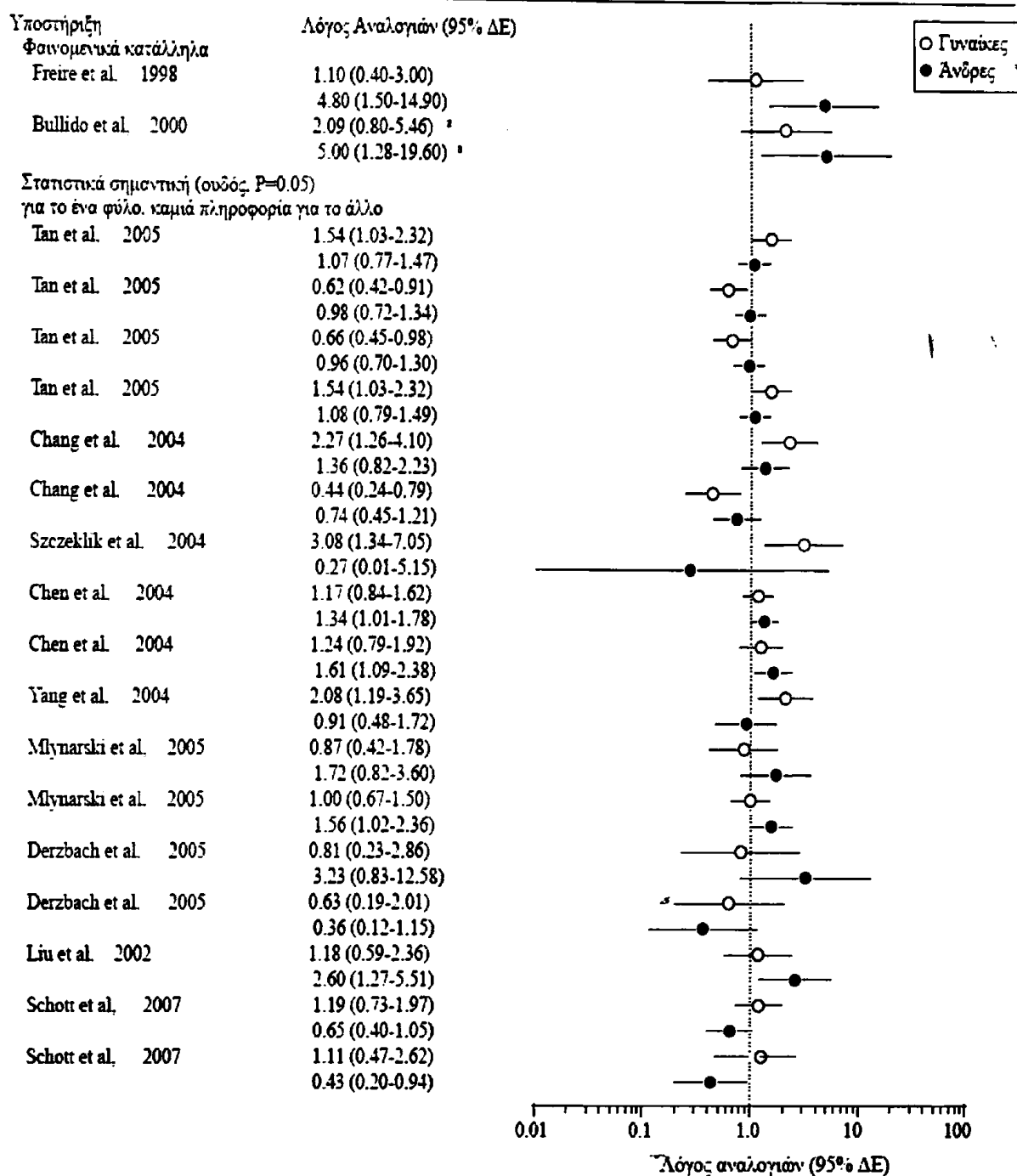
Διαθέσιμα δεδομένα για την επανεξέταση των ισχυρισμών υπήρχαν σε 188 περιπτώσεις (30 με επαρκή υποστήριξη και 58 με ανεπαρκή). Συνολικά, 105 από τους 188 ισχυρισμούς (55.9%) δεν ήταν στατιστικά σημαντικοί σύμφωνα με τις αναλύσεις μας. Χαρακτηριστικά, τα μεγέθη



επιδράσεων και τα αντίστοιχα διαστήματα εμπιστοσύνης παραθέτονται ανά ζεύγη ανδρών-γυναικών για τις περιπτώσεις που η αλληλεπίδραση δεν ήταν στατιστικά σημαντική και τα μεγέθη επιδράσεων ήταν λόγοι αναλογιών ($n=44$) (Εικόνες 4 και 5). Ογδόντα τρεις από τους 188 ισχυρισμούς ήταν στατιστικά σημαντικοί με βάση τις αναλύσεις μας, αλλά η πλειονότητα ($n=44$) είχε τιμές p μεταξύ 0.01 και 0.05, 25 μεταξύ 0.001 και 0.01, και μόνο 14 μικρότερες από 0.001.

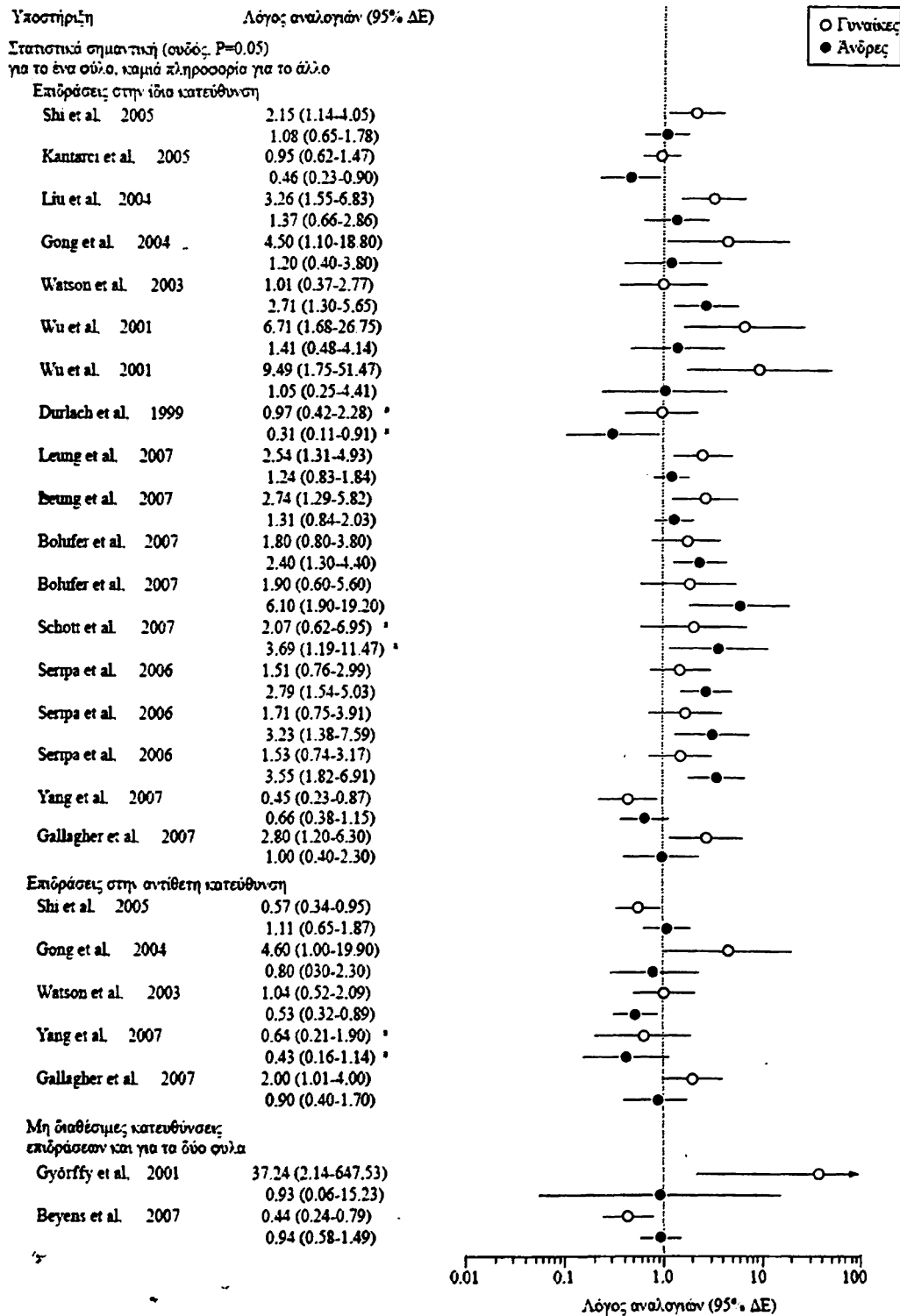
Σε 30 ισχυρισμούς οι ερευνητές παρείχαν κατάλληλη στατιστική υποστήριξη και τα πλήρη δεδομένα ώστε να μπορέσουμε να επανεξετάσουμε τις αλληλεπιδράσεις φύλου-γονιδίου. Από αυτούς, σε 23 ισχυρισμούς η διαφορά μεταξύ των φύλων ήταν στατιστικά σημαντική τόσο στο πρωτότυπο άρθρο όσο και στη δικιά μας επανεξέταση, ενώ σε άλλους 7 δε μπορούσαμε να επαληθεύσουμε τις αναφερόμενες στατιστικά σημαντικές διαφορές.





Εικόνα 4: Μεγέθη επιδράσεων για άνδρες και γυναίκες συμμετέχουσες σε μελέτες με φαινομενικά κατάλληλη υποστήριξη και εκείνες με στατιστικά σημαντικές επιδράσεις στο ένα φύλο και καμιά πληροφορία για το άλλο.

a. Βασισμένα στην επανεξέταση των δεδομένων b. Επανεξέταση όλων των εκτιμητών.



Εικόνα 5: Μεγέθη επιδράσεων για άνδρες και γυναίκες συμμετέχουσες σε μελέτες με στατιστικά σημαντικές επιδράσεις στο ένα φύλο αλλά όχι στο άλλο.

α. Βασισμένα στην επανεξέταση των δεδομένων β. Επανεξέταση όλων των εκτιμητών.



2.2.3.7 Επιβεβαίωση των ισχυρισμών με την καλύτερη εσωτερική εγκυρότητα

Από τους 432 ισχυρισμούς, μόνο οι 37 από 17 άρθρα βασίζονταν σε αναφερόμενο εκ των προτέρων σχεδιασμό, είχαν διαθέσιμα τα πλήρη δεδομένα η ανάλυση των οποίων ήταν στατιστικά σημαντική για την αλληλεπίδραση φύλου-γονιδίου και χρησιμοποιούσαν ολόκληρο το μέγεθος του δείγματος. Οποιαδήποτε είδους επιβεβαίωση εντοπίστηκε μόνο για τρεις από τους ισχυρισμούς αυτούς. Μια διαφορά μεταξύ των δυο φύλων είχε ήδη περιγραφεί σε δυο προηγούμενα άρθρα (ίδια κατεύθυνση επιδράσεων). Για έναν άλλον ισχυρισμό, μια προηγούμενη μελέτη δε βρήκε καμία επίδραση, ενώ μια επακόλουθη επιβεβαίωσε την αλληλεπίδραση. Ο τρίτος ισχυρισμός είχε ήδη περιγραφεί σε μια προηγούμενη μελέτη αλλά προς την άλλη κατεύθυνση, ενώ δυο προηγούμενες μελέτες ανέφεραν πως δεν υπήρχε διαφορετική επίδραση στα δυο φύλα. (Πίνακας 2.2.4).

Άλλοι 23 ισχυρισμοί (από 12 άρθρα) βασίζονταν σε αναφερόμενο εκ των προτέρων σχεδιασμό, είχαν διαθέσιμα τα πλήρη δεδομένα η ανάλυση των οποίων ήταν στατιστικά σημαντική για την αλληλεπίδραση φύλου-γονιδίου και η ανάλυση βασίζονταν σε μια υποομάδα του συνολικού πληθυσμού, που επίσης είχε οριστεί εκ των προτέρων. Από αυτούς, δυο ισχυρισμοί από το ίδιο άρθρο είχαν μια αναφερόμενη μελέτη που παρουσίαζε με στατιστικά σημαντικά αποτελέσματα για τους ίδιους ισχυρισμούς. Άλλος ένας είχε μια αναφορά με αποτελέσματα στην αντίθετη κατεύθυνση. Καμία μεταγενέστερη μελέτη που να επιβεβαιώνει αυτούς τους τρεις ισχυρισμούς δε βρέθηκε (Πίνακας 2.2.4, σελ 47).

2.2.4 Συζήτηση

Στην παρούσα διατριβή αξιολογήθηκαν εμπειρικά μελέτες γενετικής επιδημιολογίας που περιέγραφαν ισχυρισμούς για διαφορετική επίδραση γενετικών παραγόντων σε πολυσύνθετα νοσήματα ανάλογα με το φύλο. Οι ισχυρισμοί κάλυπταν ένα εύρος γονιδίων και εκβάσεων. Οι περισσότεροι συγγραφείς τους δήλωναν πως αυτές οι αναλύσεις είχαν σχεδιαστεί εκ των προτέρων. Παρόλα αυτά, η πλειονότητα των ισχυρισμών ήταν ανεπαρκώς υποστηριζόμενοι ή εσφαλμένοι, ενώ η περιγραφή ελέγχων για στατιστική αλληλεπίδραση φύλου-γονιδίου ήταν σπάνια. Όταν επανεξετάστηκαν τα διαθέσιμα δεδομένα, περισσότερες από τις μισές αλληλεπιδράσεις φύλου-γονιδίου δεν κατάφεραν να πετύχουν στατιστική σημαντικότητα στο επίπεδο $p=0.05$, ενώ οι περισσότερες από αυτές που το πέτυχαν είχαν μέτριας σημαντικότητας τιμές p . Ακόμα και οι ισχυρισμοί που είχαν τη φαινομενικά καλύτερη εσωτερική εγκυρότητα πολύ σπάνια είχαν επιβεβαίωση από άλλες μελέτες.

Οι συγκρίσεις υποομάδων έχουν αξιολογηθεί εκτεταμένα προηγουμένως, κυρίως στη



βιβλιογραφία των κλινικών μελετών. [150-152] Τη στιγμή της συγγραφής της παρούσας διατριβής καμία παρόμοια αξιολόγηση δεν υπήρχε στη γενετική επιδημιολογία, παρόλο που οι μελέτη γενετικών παραγόντων συχνών και πολυσύνθετων νοσημάτων και φαινοτύπων αντιπροσωπεύει μια ραγδαία αυξανόμενη βιβλιογραφία κλινικής έρευνας. [43] Στη βιβλιογραφία των κλινικών μελετών, οι αναλύσεις υποομάδων με βάση το φύλο ή άλλα χαρακτηριστικά είναι μια συχνή στρατηγική για την αλίευση και περιγραφή στατιστικά σημαντικών αποτελεσμάτων. Αρκετοί συγγραφείς έχουν συμφωνήσει στην ανάγκη της διάφανης διεξαγωγής και περιγραφής αναλύσεων υποομάδων, ειδικότερα αυτών που αφορούν το φύλο. [153] Παρόλα αυτά, η πλειονότητα των ισχυρισμών για αναλύσεις υποομάδων με βάση το φύλο είναι πιθανό να είναι τυχαία ευρήματα. [154, 155] Χαρακτηριστικά τυχαία ευρήματα περιγράφηκαν χρόνια πριν στην ανάλυση της κοόρτης ISIS-2 με βάση το αστρολογικό ζώδιο. [156] Η ίδια αρχή εξετάστηκε εκτενέστερα πρόσφατα σε προσομοιώσεις αναλύσεων υποομάδων που ανέδειξαν το μεγάλο κίνδυνο των ψευδώς θετικών αποτελεσμάτων σε παρόμοιες προσπάθειες. [157-159] Στην ανάλυση επιδράσεων ειδικών για το φύλο στη βιβλιογραφία της γενετικής επιδημιολογίας, οι ερευνητές πολύ συχνά έφευγαν σε κλασικές “παγίδες”. Το πιο τυπικό λάθος ήταν ο ισχυρισμός περί στατιστικής σημαντικότητας για ένα από τα δυο φύλα όταν η διαφορά των επιδράσεων μεταξύ των δυο φύλων δε διέφερε πέρα από την τύχη. Προσομοιώσεις έχουν δείξει πως μια σημαντική επίδραση σε μια υποομάδα μόνο μπορεί να εμφανίζεται συχνά ακόμα και σε μελέτες με επαρκές μέγεθος δείγματος. [157, 158] Η πλειονότητα αυτών των ισχυρισμών αναμένεται να είναι ψευδώς θετικοί.

Ένα άλλο σημαντικό πρόβλημα είναι η έλλειψη ισχύος να ανιχνευθούν αλληλεπιδράσεις επιδράσεων με το φύλο σ' αυτές τις μελέτες. Οι Brookes και συνεργάτες υπολόγισαν πως μια μελέτη με 80% ισχύ να αναδείξει τη συνολική επίδραση έχει μόνο 29% ισχύ να ανιχνεύσει την επίδραση μιας αλληλεπίδρασης του ίδιου μεγέθους. [158] Η καλοσχεδιασμένη αναζήτηση αλληλεπιδράσεων ίσως να απαιτεί ακόμα και 10 φορές μεγαλύτερο μέγεθος δείγματος συγκρινόμενο με αυτό που χρειάζεται για τις κύριες επιδράσεις. [157, 160, 161] Την ίδια στιγμή, οι μελέτες γενετικής συσχέτισης είναι ήδη 10 φορές μικρότερες από ό,τι θα επαρκούσε για να ανιχνευθούν ακόμα και κύριες επιδράσεις, βασιζόμενοι στην υπάρχουσα γνώση του μεγέθους των πιθανών επιδράσεων σε συχνές γενετικές παραλλαγές. [52] Οι περισσότερες κύριες επιδράσεις που έχουν προταθεί την τελευταία δεκαετία δεν έχουν επιβεβαιωθεί. [162] Κάτω από αυτές τις συνθήκες, μια στατιστικά σημαντική αλληλεπίδραση στο επίπεδο τιμής $p=0.05$, πιθανώς έχει πολύ μικρή θετική προγνωστική αξία για την παρουσία μια πραγματικής αλληλεπίδρασης.

Μερικοί περιορισμοί της παρούσας διατριβής θα πρέπει να συζητηθούν. Στην παρούσα



διατριβή χρησιμοποιήθηκε μια στρατηγική δειγματοληψίας που παρόλο που ήταν συστηματική ήταν επίσης οδηγημένη από την ευκολία. Η δειγματοληψία τείνει να επιλέγει τους πιο ισχυρούς ισχυρισμούς για διαφορές ανάμεσα στα δυο φύλα και κάποιος μπορεί να υποθέσει πως αυτοί είναι πιθανόν να έχουν καλή εσωτερική εγκυρότητα και καλύτερες πιθανότητες εξωτερικής επιβεβαίωσης. Παρόλα αυτά αυτό δεν μπορεί να αποδειχθεί. Το φύλο είναι ένα χαρακτηριστικό διαθέσιμο σχεδόν σε όλες τις μελέτες γενετικής συσχέτισης, άρα το παρόν δείγμα πιθανώς είναι μόνο η “κορυφή του παγόβουνου”. Δε θα ήταν πρακτικό να βρεθούν και αξιολογηθούν όλες οι συγκρίσεις φύλων, ακόμα και σε ένα περιγεγραμμένο δείγμα της βιβλιογραφίας. Η επιλεκτική παρουσίαση των υποομάδων και των δευτερογενών αναλύσεων είναι ολοένα και αυξανόμενα αναγνωρίσιμες πηγές συστηματικών σφαλμάτων, που μπορεί να οδηγήσουν σε μια προδιάθεση “θετικών” ευρημάτων στη επιστημονική βιβλιογραφία.[68, 163, 164] Τουλάχιστον, οι μελέτες που αξιολογήθηκαν είναι πιθανόν ανάμεσα σ'αυτές που οι συγγραφείς τους ήταν περισσότερο βέβαιοι για κάποιους, αν όχι όλους, τους ισχυρισμούς διαφορών ανάμεσα στα φύλα που παρουσίαζαν στα αποτελέσματά τους. Διαφορετικά δε θα τους ανέφεραν στον τίτλο των άρθρων τους.

Θα πρέπει επίσης να αναφερθεί πως για μερικούς ισχυρισμούς, ειδικά για αυτούς που διατυπώθηκαν πρόσφατα, η εξωτερική επιβεβαίωση μπορεί να μην έχει πραγματοποιηθεί προς το παρόν. Παρόλα αυτά, η γενετική επιδημιολογία είναι ένα γρήγορα εξελισσόμενο πεδίο και προσπάθειες επιβεβαίωσης πραγματοποιούνται με γρήγορους ρυθμούς.

Η παρούσα εμπειρική αποτίμηση εστιάστηκε στις αλληλεπιδράσεις φύλου-γονιδίων, αλλά οι επιπτώσεις εκτείνονται και σε άλλες αναλύσεις υποομάδων στη γενετική επιδημιολογία. Η γενετική επιδημιολογία είναι ένα πεδίο που περιέχει πολλές αναλύσεις υποομάδων, όχι μόνο με βάση το φύλο, αλλά και την ηλικία, την εθνική/φυλετική καταγωγή, άλλους πολυμορφισμούς, διαιτητικές συνήθειες, τρόπο ζωής και άλλες εκθέσεις.[165, 166] Παρόμοια προσοχή επιβάλλεται και στην ανάλυση, περιγραφή και ερμηνεία αυτών.

Η αναζήτηση αλληλεπιδράσεων φύλου-γονιδίων δε θα πρέπει να εγκαταλειφθεί. Ιδανικά, τέτοιες διαφορές θα πρέπει να βασίζονται σε εκ των προτέρων σχεδιασμένες, ξεκάθαρα ορισμένες και αρκετά ισχυρές υποομάδες. Μεταγενέστερες, βασιζόμενες σε ευρήματα αναλύσεις θα πρέπει να περιγράφονται ως τέτοιες εμφανώς. Τόσο εκ των προτέρων όσο και μεταγενέστερα σχεδιασμένοι ισχυρισμοί θα πρέπει να βασίζονται σε τεστ αλληλεπίδρασης και κατάλληλη εκτίμηση των πιθανών πολλαπλών συγκρίσεων. Ακόμα και τότε, τα αποτελέσματα θα πρέπει να εξηγούνται με προσοχή και θα πρέπει να επιβεβαιώνονται από αρκετές άλλες μελέτες πριν γίνουν αποδεκτά ως πραγματικοί παράγοντες κινδύνου.



Πίνακας 2.2.4: Ιστορικό-επιβεβαίωσης των ισχυρισμών αλληλεπιδράσεων γονιδίων-φύλου με τη φαινομενικά καλύτερη εσωτερική εγκυρότητα.

Πρώτος συγγραφέας, έτος	Γονίδιο	Φαινότυπος	Δείγμα	Σημειακός εκτιμητής ανδρών (95% ΔΕ)	Σημειακός εκτιμητής γυναικών (95% ΔΕ)	Μελέτες επιβεβαίωσης		Αναφορές
						Σημειακός εκτιμητής ανδρών (95% ΔΕ)	Σημειακός εκτιμητής γυναικών (95% ΔΕ)	
Stankovic, 2002	ACE I/D	Υπέρταση	Συνολικό	2.05 (1.07-3.91)	0.72 (0.33-1.6)	2.40 (1.30-4.50)	1.40 (0.72-2.69)	Προηγούμενη μελέτη (Higaki[167]): αλληλεπίδραση Προηγούμενη μελέτη (O'Donnell[168]): αλληλεπίδραση
Builido, 2000	LRP A491T	Alzheimer	Συνολικό	4.30 (1.80-10.20)*	1.50 (0.80-2.70)*	2.38 (0.81-7.00)	2.35 (0.52-9.41)	Προηγούμενη μελέτη (Ahmed[169]): καμιά αλληλεπίδραση Επακόλουθη μελέτη (Parra-Bonilla[170]): αλληλεπίδραση
Ordovas, 2002 †	APOA1 G75A	HDL χοληστερίνη	Συνολικό	0.00 (-0.20-0.03)	0.05 (0.02-0.07)	0.27 (0.08-0.46)	0.01 (-0.18-0.20)	Προηγούμενη μελέτη (Meng[171]): αντίστροφη αλληλεπίδραση Προηγούμενη μελέτη (Mata[172]): καμιά αλληλεπίδραση Προηγούμενη μελέτη (Xu[173]): καμιά αλληλεπίδραση.
Sagnella, 1999	ACE I/D	Υπέρταση	Αφρικανικής προέλευσης μόνο	0.79 (0.36-1.72)	2.54 (1.38-4.65)	1.36 (1.09-1.71)	1.07 (0.87-1.32)	Προηγούμενη μελέτη (O'Donnell[168]): αντίστροφη αλληλεπίδραση
Ferricres, 1994	apoE ε4	LDL χοληστερίνη	ε3/2 vs ε3/3	0.31 (-0.30-0.92)	1.73 (1.15-2.31)	NE	NE	Προηγούμενη μελέτη (Kotze[174]): καμιά αλληλεπίδραση
Ferricres, 1994	apoE ε4	Ολική χοληστερίνη	ε3/2 vs ε3/3	-0.05 (-0.72-0.62)	1.75 (1.30-2.20)	NE	NE	Προηγούμενη μελέτη (Kotze[174]): καμιά αλληλεπίδραση

* βασισμένα στην επανεξέταση των δεδομένων. Η αρχική μελέτη είχε ισχυριστεί ένα λόγο αναλογιών 3.2 (1.2-8.6) για τους άνδρες και δεν ανέφερε τον αντίστοιχο για τις γυναίκες.

NE: απουσία εκτιμητών † Οι Ordovas et al. ανέφεραν μια αλληλεπίδραση γονιδίου-φύλου-PUFA. Τα δεδομένα των Meng et al. και Mata et al. βασίζονται σε δίαιτα χαμηλών λιπαρών, ενώ τα δεδομένα των Xu et al σε δεδομένα βασικής γραμμής.



ΣΥΜΠΕΡΑΣΜΑΤΑ

- Από τη διατριβή αυτή εξάγονται τα παρακάτω συμπεράσματα:

1. Η γενετική επιδημιολογία των τελευταίων ετών περιλαμβάνει μελέτες που δεν αναφέρουν επαρκή στοιχεία απαραίτητα για την αξιολόγηση συστηματικών σφαλμάτων. Οι μελέτες κατά κόρον ήταν μικρές και εξέταζαν μόνο ένα ή λίγους γενετικούς παράγοντες σε ένα και μόνο γονίδιο. Η αλληλεπίδραση γονιδίων περιγράφονταν μόνο σε ένα μικρό ποσοστό μελετών που εξέταζαν πολλαπλούς γενετικούς παράγοντες. Τα στοιχεία αυτά φαίνεται να βελτιώνονται με την πάροδο του χρόνου.
2. Η περιγραφή αλληλεπιδράσεων γονιδίων-φύλου είναι αρκετά συχνή. Η αξιολόγηση ισχυρισμών για τέτοιες αλληλεπιδράσεις κατέδειξε πως το μεγαλύτερο ποσοστό των ισχυρισμών αυτών ήταν είτε ανεπαρκώς υποστηριζόμενοι είτε εσφαλμένοι. Η επανεξέταση 188 ισχυρισμών έδειξε ότι μόνο 83 (44.4%) από αυτούς είναι στατιστικά σημαντικοί στο επίπεδο $p=0.05$. Από τους 60 συνολικά ισχυρισμούς με την καλύτερη εσωτερική εγκυρότητα, μόνο ένας μπορούσε να επιβεβαιωθεί σε τουλάχιστον 2 άλλες μελέτες.

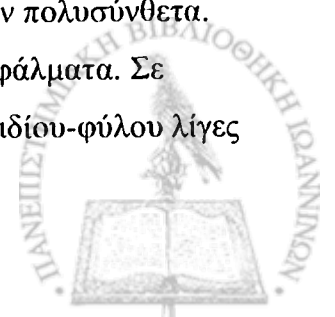


ΠΕΡΙΛΗΨΗ

-Η ραγδαία ανάπτυξη της γενετικής επιδημιολογίας τα τελευταία χρόνια οδήγησε στην προσπάθεια συσχέτισης μεγάλου αριθμού νοσημάτων ή άλλων γνωρισμάτων με ακόμα μεγαλύτερο αριθμό γενετικών παραγόντων. Η συντριπτική πλειονότητα αυτών των γνωρισμάτων είναι πολυσύνθετοι φαινότυποι ή νοσήματα για τα οποία υπάρχουν ισχυρές ενδείξεις για γενετική προδιάθεση, πχ αυτοάνοσα νοσήματα, καρκίνοι. Η αδυναμία εύρεσης ενός και μόνου γενετικού παράγοντα προδιάθεσης για τα γνωρίσματα έθεσε τις βάσεις για τη θεωρία των πολλών και συχνών γενετικών παραγόντων, δηλαδή πως τα γνωρίσματα αυτά προκαλούνται ή επηρεάζονται από πολλούς γενετικούς παράγοντες. Ως επακόλουθο υπάρχουν στη γενετική βιβλιογραφία σχεδόν όλοι οι πιθανοί συνδυασμοί συγκρίσεων γενετικών παραγόντων και γνωρισμάτων. Στην παρούσα διδακτορική διατριβή έγινε προσπάθεια συστηματικής καταγραφής στοιχείων της σχετικής βιβλιογραφίας (2001-2003 και 2006) και εξερεύνησης των πιθανών συστηματικών σφαλμάτων που μπορεί να την επηρεάσουν. Επίσης, έγινε ιδιαίτερη προσπάθεια καταγραφής της ταυτόχρονης περιγραφής πολλαπλών γενετικών παραγόντων στην ίδια μελέτη και αλληλεπιδράσεων αυτών. Ειδικότερα έγινε μια αξιολόγηση των ισχυρισμών για αλληλεπιδράσεις γονιδίων-φύλου, και επανεξέτασή τους.

Οι περισσότερες μελέτες που δημοσιεύτηκαν τη χρονική περίοδο 2001-2003 είχαν μικρό μέγεθος δείγματος και εξέταζαν μόνο ένα γονίδιο και λίγους πολυμορφισμούς. Λίγες μελέτες περιγράφανε μια αλληλεπίδραση γονιδίων (24.4%). Στο δείγμα του 2006 υπήρχε μια μικρή βελτίωση αλλά όχι στατιστικά σημαντική. Όσον αφορά τους ισχυρισμούς αλληλεπιδράσεων γονιδίων-φύλου, αυτοί ήταν εκ των προτέρων σχεδιασμένοι στην πλειονότητά τους (66.2%), ενώ περίπου οι μισοί αναφέρονταν σε όλο το δείγμα. Κατάλληλη στήριξη των ισχυρισμών υπήρχε μόλις στο 12.7%, ενώ οι περισσότεροι ήταν είτε ανεπαρκώς υποστηριζόμενοι είτε εσφαλμένοι. Από τους 60 συνολικά ισχυρισμούς με την καλύτερη εσωτερική εγκυρότητα, μόνο ένας μπορούσε να επιβεβαιωθεί σε τουλάχιστον 2 άλλες μελέτες.

Πολλά στοιχεία των μελετών γενετικής επιδημιολογίας δεν περιγράφονται επαρκώς στις μελέτες αυτές. Έως το χρονικό διάστημα πριν την έλευση των μελετών συσχέτισης σε όλη την έκταση του γονιδιώματος (genome wide association studies), οι μελέτες κατά κανόνα περιέγραφαν λίγα γονίδια ή γενετικούς παράγοντες και ελάχιστα εξέταζαν την αλληλεπίδραση μεταξύ γονιδίων, παρόλο που στην συντριπτική τους πλειονότητα τα υπό εξέταση γνωρίσματα ήταν πολυσύνθετα. Γενικά, σεβαστό ποσοστό αυτών ήταν εκτεθειμένες σε κλασικά συστηματικά σφάλματα. Σε εκείνες τις μελέτες όπου γίνονταν κάποιος ισχυρισμός πχ για αλληλεπίδραση γονιδίου-φύλου λίγες



ήταν οι επαρκώς τεκμηριωμένες περιπτώσεις και εξαιρετικά σπάνιες εκείνες με καλή εσωτερική και εξωτερική εγκυρότητα. Η μεθοδολογία που αναπτύχθηκε στην παρούσα διατριβή για την αξιολόγηση των ισχυρισμών για διαφορετικές επιδράσεις γενετικών παραγόντων στα δύο φύλα μπορεί να εφαρμοστεί και στη μελέτη αλληλεπιδράσεων γονιδίων μεταξύ τους ή με άλλα χαρακτηριστικά, όπως η ηλικία.



ABSTRACT

-The rapid growth of genetic epidemiology the last years led to the efforts for association of various traits with numerous genetic variants. The vast majority of these traits are complex diseases or phenotypes with known genetic predispose, e.g. cancer, autoimmune diseases. The failure to identify a single genetic variant or gene for these traits led to the theory of common genetic variants for complex diseases. Therefore in genetic literature there are all possible associations of genetic variants and traits. In this work we performed a systematic assessment of reporting practices of human epidemiology studies (2001-2003 and 2006). We were also interested in reporting of multiple genetic variants and their interactions. In a separate database of genetic epidemiology studies it was assessed whether claims for gene-gender interactions in complex diseases were properly documented.

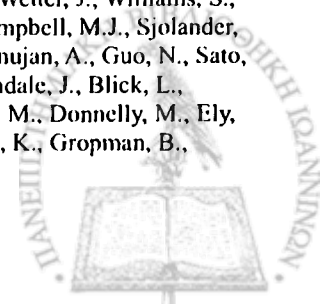
Most studies published in the period 2001-2003 were relatively small in sample size and examined one or more genetic variant within a single gene. Few studies reported a gene interaction (24.4%). In the 2006 sample there was an overall improvement in reporting most factors, although not statistically significant. The majority of gene-gender claims in the second database of studies were planned a priori, while almost half of these used the entire sample size. Appropriate documentation of gene-gender interaction was recorded in 12.7% of studies. Most of the claims were either insufficiently documented or spurious. Of 60 claims with seemingly the best internal validity, only 1 was consistently replicated in at least 2 other studies.

Studies of human genetic epidemiology, before the era of genome-wide association studies, examined few genetic variants or genes, although the vast majority of the traits of interest were of complex origin. In studies where a claim for gene-gender interaction was made, appropriately documented claims were few and ones with good internal and external validity rare.

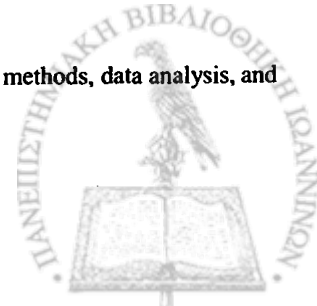


ΑΝΑΦΟΡΕΣ

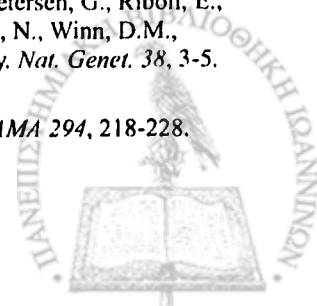
1. Burton, P.R., Tobin, M.D. and Hopper, J.L. (2005). Key concepts in genetic epidemiology. *Lancet* 366, 941-951.
2. Botstein, D. and Risch, N. (2003). Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat. Genet.* 33 Suppl, 228-237.
3. Carlson, C.S., Eberle, M.A., Rieder, M.J., Yi, Q., Kruglyak, L. and Nickerson, D.A. (2004). Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am. J. Hum. Genet.* 74, 106-120.
4. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczy, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J.P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Marra, M.A., Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chissoe, S.L., Wendl, M.C., Delchaunty, K.D., Miner, T.L., Delchaunty, A., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R.A., Muzny, D.M., Scherer, S.E., Bouck, J.B., Sodergren, E.J., Worley, K.C., Rives, C.M., Gorrell, J.H., Metzker, M.L., Naylor, S.L., Kucherlapati, R.S., Nelson, D.L., Weinstock, G.M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D.R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H.M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R.W., Federspiel, N.A., Abola, A.P., Proctor, M.J., Myers, R.M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D.R., Olson, M.V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G.A., Athanasiou, M., Schultz, R., Roe, B.A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W.R., de la Bastide, M., Dedhia, N., Blöcker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J.A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D.G., Burge, C.B., Cerutti, L., Chen, H.C., Church, D., Clamp, M., Copley, R.R., Doerks, T., Eddy, S.R., Eichler, E.E., Furey, T.S., Galagan, J., Gilbert, J.G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L.S., Jones, T.A., Kasif, S., Kasprzyk, A., Kennedy, S., Kent, W.J., Kitts, P., Koonin, E.V., Korf, I., Kulp, D., Lancet, D., Lowe, T.M., McLysaght, A., Mikkelsen, T., Moran, J.V., Mulder, N., Pollara, V.J., Ponting, C.P., Schuler, G., Schultz, J., Slater, G., Smit, A.F., Stupka, E., Szustakowski, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y.I., Wolfe, K.H., Yang, S.P., Yeh, R.F., Collins, F., Guyer, M.S., Peterson, J., Felsenfeld, A., Wetterstrand, K.A., Patrinos, A., Morgan, M.J., de Jong, P., Catanese, J.J., Osoegawa, K., Shizuya, H., Choi, S., Chen, Y.J. (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860-921.
5. Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., Gocayne, J.D., Amanatides, P., Ballew, R.M., Huson, D.H., Wortman, J.R., Zhang, Q., Kodira, C.D., Zheng, X.H., Chen, L., Skupski, M., Subramanian, G., Thomas, P.D., Zhang, J., Gabor Miklos, G.L., Nelson, C., Broder, S., Clark, A.G., Nadeau, J., McKusick, V.A., Zinder, N., Levine, A.J., Roberts, R.J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A.E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T.J., Higgins, M.E., Ji, R.R., Ke, Z., Ketchum, K.A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G.V., Milshina, N., Moore, H.M., Naik, A.K., Narayan, V.A., Neclam, B., Nusskern, D., Rusch, D.B., Salzberg, S., Shao, W., Shuc, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zhu, S., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A., Cheng, M.L., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S., Dodson, K., Doup, L., Ferreira, S., Garg, N., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hosten, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, I., Koduru, S., Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H., Reardon, M., Rodriguez, R., Rogers, Y.H., Romblad, D., Ruhfel, B., Scott, R., Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint, N.N., Tse, S., Vech, C., Wang, G., Wetter, J., Williams, S., Williams, M., Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J.F., Guigó, R., Campbell, M.J., Sjolander, K.V., Karlak, B., Kejariwal, A., Mi, H., Lazareva, B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S., Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yooseph, S., Allen, D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y.H., Coyne, M., Dahlke, C., Mays, A., Dombroski, M., Donnelly, M., Ely, D., Esparham, S., Fosler, C., Gire, H., Glanowski, S., Glasser, K., Glodek, A., Gorokhov, M., Graham, K., Gropman, B.,



- Harris, M., Heil, J., Henderson, S., Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C., Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D., Majoros, W., McDaniel, J., Murphy, S., Newman, M., Nguyen, T., Nguyen, N., Nodell, M., Pan, S., Peck, J., Peterson, M., Rowe, W., Sanders, R., Scott, J., Simpson, M., Smith, T., Sprague, A., Stockwell, T., Turner, R., Venter, E., Wang, M., Wen, M., Wu, D., Wu, M., Xia, A., Zandieh, A. and Zhu, X. (2001). The sequence of the human genome. *Science* 291, 1304-1351.
6. Cotton, R.G. and Scriver, C.R. (1998). Proof of "disease causing" mutation. *Hum. Mutat.* 12, 1-3.
 7. Risch, N.J. (2000). Searching for genetic determinants in the new millennium. *Nature* 405, 847-856.
 8. Collins, F.S., Patrinos, A., Jordan, E., Chakravarti, A., Gesteland, R. and Walters, L. (1998). New goals for the U.S. Human Genome Project: 1998-2003. *Science* 282, 682-689.
 9. Palmer, L.J. and Cookson, W.O. (2001). Using single nucleotide polymorphisms as a means to understanding the pathophysiology of asthma. *Respir. Res.* 2, 102-112.
 10. Collins, F.S., Guyer, M.S. and Chakravarti, A. (1997). Variations on a theme: cataloging human DNA sequence variation. *Science* 278, 1580-1581.
 11. Nickerson, D.A., Whitehurst, C., Boysen, C., Charmley, P., Kaiser, R. and Hood, L. (1992). Identification of clusters of biallelic polymorphic sequence-tagged sites (pSTs) that generate highly informative and automatable markers for genetic linkage mapping. *Genomics* 12, 377-387.
 12. Chakravarti, A. (1998). It's raining SNPs, hallelujah?. *Nat. Genet.* 19, 216-217.
 13. Kuhner, M.K., Beerli, P., Yamato, J. and Felsenstein, J. (2000). Usefulness of single nucleotide polymorphism data for estimating population parameters. *Genetics* 156, 439-447.
 14. McKeigue, P.M. (1998). Mapping genes that underlie ethnic differences in disease risk: methods for detecting linkage in admixed populations, by conditioning on parental admixture. *Am. J. Hum. Genet.* 63, 241-251.
 15. Brookes, A.J. (1999). The essence of SNPs. *Gene* 234, 177-186.
 16. Stallings, R.L., Ford, A.F., Nelson, D., Torney, D.C., Hildebrand, C.E. and Moyzis, R.K. (1991). Evolution and distribution of (GT)_n repetitive sequences in mammalian genomes. *Genomics* 10, 807-815.
 17. Ioannidis, J.P.A. (2003). Genetic associations: false or true?. *Trends Mol Med* 9, 135-138.
 18. Zondervan, K.T. and Cardon, L.R. (2004). The complex interplay among factors that influence allelic association. *Nat. Rev. Genet.* 5, 89-100.
 19. Cordell, H.J. and Clayton, D.G. (2005). Genetic association studies. *Lancet* 366, 1121-1131.
 20. Hattersley, A.T. and McCarthy, M.I. (2005). What makes a good genetic association study?. *Lancet* 366, 1315-1323.
 21. Palmer, L.J. and Cardon, L.R. (2005). Shaking the tree: mapping complex disease genes with linkage disequilibrium. *Lancet* 366, 1223-1234.
 22. Cardon, L.R. and Bell, J.I. (2001). Association study designs for complex diseases. *Nat. Rev. Genet.* 2, 91-99.
 23. Lander, E. and Kruglyak, L. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat. Genet.* 11, 241-247.
 24. Witte, J.S., Elston, R.C. and Cardon, L.R. (2000). On the relative sample size required for multiple comparisons. *Stat Med* 19, 369-372.
 25. Terwilliger, J.D. and Göring, H.H. (2000). Gene mapping in the 20th and 21st centuries: statistical methods, data analysis, and experimental design. *Hum. Biol.* 72, 63-132.



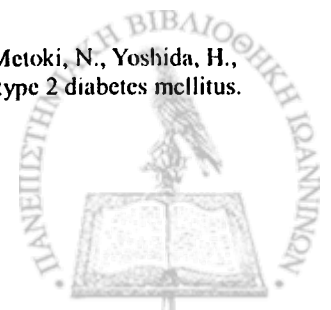
26. Risch, N. and Merikangas, K. (1996). The future of genetic studies of complex human diseases. *Science* 273, 1516-1517.
27. Cardon, L.R. and Palmer, L.J. (2003). Population stratification and spurious allelic association. *Lancet* 361, 598-604.
28. Goldstein, D.B., Tate, S.K. and Sisodiya, S.M. (2003). Pharmacogenetics goes genomic. *Nat. Rev. Genet.* 4, 937-947.
29. Ioannidis, J.P., Ntzani, E.E., Trikalinos, T.A. and Contopoulos-Ioannidis, D.G. (2001). Replication validity of genetic association studies. *Nat. Genet.* 29, 306-309.
30. Lohmueller, K.E., Pearce, C.L., Pike, M., Lander, E.S. and Hirschhorn, J.N. (2003). Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat. Genet.* 33, 177-182.
31. Kc, X., Durrant, C., Morris, A.P., Hunt, S., Bentley, D.R., Deloukas, P. and Cardon, L.R. (2004). Efficiency and consistency of haplotype tagging of dense SNP maps in multiple samples. *Hum. Mol. Genet.* 13, 2557-2565.
32. Weiss, K.M. and Terwilliger, J.D. (2000). How many diseases does it take to map a gene with SNPs?. *Nat. Genet.* 26, 151-157.
33. Rothwell, P.M. (2005). Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *Lancet* 365, 176-186.
34. Canadian Cooperative Study Group (1978). A randomized trial of aspirin and sulfinpyrazone in threatened stroke. The Canadian Cooperative Study Group. *N. Engl. J. Med.* 299, 53-59.
35. Harrison, T., Kasper, D., Braunwald, E. and et al (2004). Harrison's Principles of Internal Medicine. 16th ed, Volume , Edition (New York, NY: McGraw-Hill Education).
36. Yang, Q., Khoury, M.J., Friedman, J., Little, J. and Flanders, W.D. (2005). How many genes underlie the occurrence of common complex diseases in the population?. *Int J Epidemiol* 34, 1129-1137.
37. Little, J., Khoury, M.J., Bradley, L., Clyne, M., Gwinn, M., Lin, B., Lindegren, M. and Yoon, P. (2003). The human genome project is complete. How do we develop a handle for the pump?. *Am. J. Epidemiol.* 157, 667-673.
38. Khoury, M.J., Millikan, R., Little, J. and Gwinn, M. (2004). The emergence of epidemiology in the genomics age. *Int J Epidemiol* 33, 936-944.
39. Khoury, M.J. and Dorman, J.S. (1998). The Human Genome Epidemiology Network. *Am. J. Epidemiol.* 148, 1-3.
40. Khoury, M.J. (1999). Human genome epidemiology: translating advances in human genetics into population-based data for medicine and public health. *Genet. Med.* 1, 71-73.
41. Hirschhorn, J.N., Lohmueller, K., Byrne, E. and Hirschhorn, K. (2002). A comprehensive review of genetic association studies. *Genet. Med.* 4, 45-61.
42. Ioannidis, J.P.A., Trikalinos, T.A., Ntzani, E.E. and Contopoulos-Ioannidis, D.G. (2003). Genetic associations in large versus small studies: an empirical assessment. *Lancet* 361, 567-571.
43. Lin, B.K., Clyne, M., Walsh, M., Gomez, O., Yu, W., Gwinn, M. and Khoury, M.J. (2006). Tracking the epidemiology of human genes in the literature: the HuGE Published Literature database. *Am. J. Epidemiol.* 164, 1-4.
44. Xia, Y., Gueguen, R., Vincent-Viry, M., Siest, G. and Visvikis, S. (2003). Effect of six candidate genes on early aging in a French population. *Aging Clin Exp Res* 15, 111-116.
45. Ioannidis, J.P.A., Gwinn, M., Little, J., Higgins, J.P.T., Bernstein, J.L., Boffetta, P., Bondy, M., Bray, M.S., Brenchley, P.E., Buffler, P.A., Casas, J.P., Chokkalingam, A., Danesh, J., Smith, G.D., Dolan, S., Duncan, R., Gruis, N.A., Hartge, P., Hashibe, M., Hunter, D.J., Jarvelin, M., Malmer, B., Maraganore, D.M., Newton-Bishop, J.A., O'Brien, T.R., Petersen, G., Riboli, E., Salanti, G., Seminara, D., Smeeth, L., Taioli, E., Timpson, N., Uitterlinden, A.G., Vincis, P., Wareham, N., Winn, D.M., Zimmern, R., Khoury, M.J. (2006). A road map for efficient and reliable human genome epidemiology. *Nat. Genet.* 38, 3-5.
46. Ioannidis, J.P.A. (2005). Contradicted and initially stronger effects in highly cited clinical research. *JAMA* 294, 218-228.



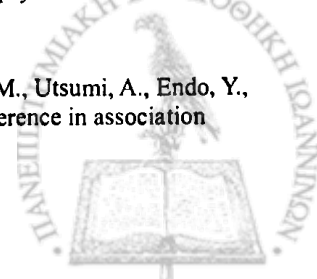
47. Colhoun, H.M., McKeigue, P.M. and Davey Smith, G. (2003). Problems of reporting genetic associations with complex outcomes. *Lancet* 361, 865-872.
48. Wacholder, S., Chanock, S., Garcia-Closas, M., El Ghormli, L. and Rothman, N. (2004). Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J. Natl. Cancer Inst.* 96, 434-442.
49. Peters, D.L., Barber, R.C., Flood, E.M., Garner, H.R. and O'Keefe, G.E. (2003). Methodologic quality and genotyping reproducibility in studies of tumor necrosis factor -308 G-->A single nucleotide polymorphism and bacterial sepsis: implications for studies of complex traits. *Crit. Care Med.* 31, 1691-1696.
50. Clark, M.F. and Baudouin, S.V. (2006). A systematic review of the quality of genetic association studies in human sepsis. *Intensive Care Med* 32, 1706-1712.
51. Bogardus, S.T.J., Concato, J. and Feinstein, A.R. (1999). Clinical epidemiological quality in molecular genetic research: the need for methodological standards. *JAMA* 281, 1919-1926.
52. Ioannidis, J.P.A., Trikalinos, T.A. and Khoury, M.J. (2006). Implications of small effect sizes of individual genetic variants on the design and interpretation of genetic association studies of complex diseases. *Am. J. Epidemiol.* 164, 609-614.
53. Hirschhorn, J.N. and Altshuler, D. (2002). Once and again-issues surrounding replication in genetic association studies. *J. Clin. Endocrinol. Metab.* 87, 4438-4441.
54. Newton-Cheh, C. and Hirschhorn, J.N. (2005). Genetic association studies of complex traits: design and analysis issues. *Mutat. Res.* 573, 54-69.
55. Davey Smith, G., Ebrahim, S., Lewis, S., Hansell, A.L., Palmer, L.J. and Burton, P.R. (2005). Genetic epidemiology and public health: hope, hype, and future prospects. *Lancet* 366, 1484-1498.
56. Page, G.P., George, V., Go, R.C., Page, P.Z. and Allison, D.B. (2003). "Are we there yet?": Deciding when one has demonstrated specific genetic causation in complex diseases and quantitative traits. *Am. J. Hum. Genet.* 73, 711-719.
57. Pompanon, F., Bonin, A., Bellemain, E. and Taberlet, P. (2005). Genotyping errors: causes, consequences and solutions. *Nat. Rev. Genet.* 6, 847-859.
58. Romero, R., Kuivaniemi, H., Tromp, G. and Olson, J. (2002). The design, execution, and interpretation of genetic association studies to decipher complex diseases. *Am. J. Obstet. Gynecol.* 187, 1299-1312.
59. Trikalinos, T.A., Salanti, G., Khoury, M.J. and Ioannidis, J.P.A. (2006). Impact of violations and deviations in Hardy-Weinberg equilibrium on postulated gene-disease associations. *Am. J. Epidemiol.* 163, 300-309.
60. Balding, D.J. (2006). A tutorial on statistical methods for population association studies. *Nat. Rev. Genet.* 7, 781-791.
61. Leal, S.M. (2005). Detection of genotyping errors and pseudo-SNPs via deviations from Hardy-Weinberg equilibrium. *Genet. Epidemiol.* 29, 204-214.
62. Freedman, M.L., Reich, D., Penney, K.L., McDonald, G.J., Mignault, A.A., Patterson, N., Gabriel, S.B., Topol, E.J., Smoller, J.W., Pato, C.N., Pato, M.T., Petryshen, T.L., Kolonel, L.N., Lander, E.S., Sklar, P., Henderson, B., Hirschhorn, J.N. and Altshuler, D. (2004). Assessing the impact of population stratification on genetic association studies. *Nat. Genet.* 36, 388-393.
63. Marchini, J., Cardon, L.R., Phillips, M.S. and Donnelly, P. (2004). The effects of human population structure on large genetic association studies. *Nat. Genet.* 36, 512-517.
64. Wacholder, S., Rothman, N. and Caporaso, N. (2000). Population stratification in epidemiologic studies of common genetic variants and cancer: quantification of bias. *J. Natl. Cancer Inst.* 92, 1151-1158.
65. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A. and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* 38, 904-909.
66. Gibbs, J.R. and Singleton, A. (2006). Application of genome-wide single nucleotide polymorphism typing: simple association and beyond. *PLoS Genet.* 2, e150.



67. WTCCC (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447, 661-678.
68. Chan, A., Hróbjartsson, A., Haahr, M.T., Gøtzsche, P.C. and Altman, D.G. (2004). Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 291, 2457-2465.
69. Khoury, M.J. (2004). The case for a global human genome epidemiology initiative. *Nat. Genet.* 36, 1027-1028.
70. Little, J., Bradley, L., Bray, M.S., Clync, M., Dorman, J., Ellsworth, D.L., Hanson, J., Khoury, M., Lau, J., O'Brien, T.R., Rothman, N., Stroup, D., Taioli, E., Thomas, D., Vainio, H., Wacholder, S. and Weinberg, C. (2002). Reporting, appraising, and integrating data on genotype prevalence and gene-disease associations. *Am. J. Epidemiol.* 156, 300-310.
71. Schriger, D.L., Ouk, S. and Altman, D.G. (2007). The use of the World Wide Web by medical journals in 2003 and 2005: an observational study. *Pediatrics* 119, e53-60.
72. Evangelou, E., Trikalinos, T.A. and Ioannidis, J.P.A. (2005). Unavailability of online supplementary scientific information from articles published in major journals. *FASEB J.* 19, 1943-1944.
73. Hansen, P.S., Klausen, I.C., Lemming, L., Gerdes, L.U., Gregersen, N. and Faergeman, O. (1994). Apolipoprotein B gene polymorphisms in ischemic heart disease and hypercholesterolemia: effects of age and sex. *Clin. Genet.* 45, 78-83.
74. Méplan, C., Crosley, L.K., Nicol, F., Beckett, G.J., Howie, A.F., Hill, K.E., Horgan, G., Mathers, J.C., Arthur, J.R. and Hesketh, J.E. (2007). Genetic polymorphisms in the human selenoprotein P gene determine the response of selenoprotein markers to selenium supplementation in a gender-specific manner (the SELGEN study). *FASEB J.* 21, 3063-3074.
75. Hayashi, K., Maeda, S., Iemitsu, M., Otsuki, T., Sugawara, J., Tanabe, T., Miyauchi, T., Kuno, S., Ajisaka, R. and Matsuda, M. (2007). Sex differences in the relationship between estrogen receptor alpha gene polymorphisms and arterial stiffness in older humans. *Am. J. Hypertens.* 20, 650-656.
76. Baud, P., Courtet, P., Perroud, N., Jollant, F., Buresi, C. and Malafosse, A. (2007). Catechol-O-methyltransferase polymorphism (COMT) in suicide attempters: a possible gender effect on anger traits. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 144B, 1042-1047.
77. Yang, K.D., Ou, C., Hsu, T., Chang, J., Chuang, H., Liu, C., Liang, H., Kuo, H., Chen, R. and Huang, E. (2007). Interaction of maternal atopy, CTLA-4 gene polymorphism and gender on antenatal immunoglobulin E production. *Clin. Exp. Allergy* 37, 680-687.
78. Glorioso, N., Herrera, V.L.M., Bagamasbad, P., Filigheddu, F., Troffa, C., Argiolas, G., Bulla, E., Decano, J.L. and Ruiz-Opazo, N. (2007). Association of ATP1A1 and dear single-nucleotide polymorphism haplotypes with essential hypertension: sex-specific and haplotype-specific effects. *Circ. Res.* 100, 1522-1529.
79. Gallagher, C.J., Muscat, J.E., Hicks, A.N., Zheng, Y., Dyer, A., Chase, G.A., Richie, J. and Lazarus, P. (2007). The UDP-glucuronosyltransferase 2B17 gene deletion polymorphism: sex-specific association with urinary 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol glucuronidation phenotype and risk for lung cancer. *Cancer Epidemiol. Biomarkers Prev.* 16, 823-828.
80. Beyens, G., Daroszewska, A., de Freitas, F., Fransen, E., Vanhoenacker, F., Verbruggen, L., Zmierzak, H., Westhovens, R., Van Offel, J., Ralston, S.H., Devogelaer, J. and Van Hul, W. (2007). Identification of sex-specific associations between polymorphisms of the osteoprotegerin gene, TNFRSF11B, and Paget's disease of bone. *J. Bone Miner. Res.* 22, 1062-1071.
81. Leung, K.H., Yip, S.P., Wong, W.S., Yiu, L.S., Chan, K.K., Lai, W.M., Chow, E.Y.D., Lin, C.K., Yam, W.C. and Chan, K.S. (2007). Sex- and age-dependent association of SLC11A1 polymorphisms with tuberculosis in Chinese: a case control study. *BMC Infect. Dis.* 7, 19.
82. Bolufer, P., Collado, M., Barragán, E., Cervera, J., Calasanz, M., Colomer, D., Roman-Gómez, J. and Sanz, M.A. (2007). The potential effect of gender in combination with common genetic polymorphisms of drug-metabolizing enzymes on the risk of developing acute leukemia. *Haematologica* 92, 308-314.
83. Yamaguchi, S., Yamada, Y., Matsuo, H., Segawa, T., Watanabe, S., Kato, K., Yokoi, K., Ichihara, S., Metoki, N., Yoshida, H., Satoh, K. and Nozawa, Y. (2007). Gender differences in the association of gene polymorphisms with type 2 diabetes mellitus. *Int. J. Mol. Med.* 19, 631-637.

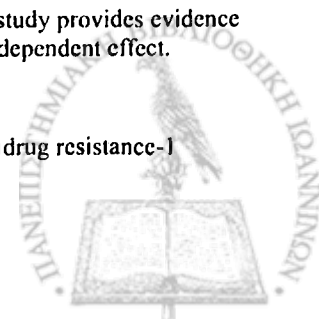


84. Russo, P., Loguercio, M., Lauria, F., Barba, G., Arnout, J., Cappuccio, F.P., Iacoviello, L., Siani, A. (2007). Age- and gender-dependent association of the -344C/T polymorphism of CYP11B2 with blood pressure in European populations. *J Hum Hypertens* 21, 333-336.
85. Froehlich, T.E., Lanphear, B.P., Dietrich, K.N., Cory-Slechta, D.A., Wang, N. and Kahn, R.S. (2007). Interactive effects of a DRD4 polymorphism, lead, and sex on executive functions in children. *Biol. Psychiatry* 62, 243-249.
86. Dedoussis, G.V.Z., Theodoraki, E.V., Manios, Y., Yiannakouris, N., Panagiotakos, D., Papoutsakis, C., Skenderi, K. and Zampelas, A. (2007). The Pro12Ala polymorphism in PPARgamma2 gene affects lipid parameters in Greek primary school children: A case of gene-to-gender interaction. *Am. J. Med. Sci.* 333, 10-15.
87. Barnett, J.H., Heron, J., Ring, S.M., Golding, J., Goldman, D., Xu, K. and Jones, P.B. (2007). Gender-specific effects of the catechol-O-methyltransferase Val108/158Met polymorphism on cognitive function in children. *Am J Psychiatry* 164, 142-149.
88. Schott, E., Witt, H., Hinrichsen, H., Neumann, K., Weich, V., Bergk, A., Halangk, J., Müller, T., Tinjala, S., Puhl, G., Neuhaus, P., Wiedenmann, B. and Berg, T. (2007). Gender-dependent association of CTLA4 polymorphisms with resolution of hepatitis C virus infection. *J. Hepatol.* 46, 372-380.
89. Niemi, M., Pasanen, M.K. and Neuvonen, P.J. (2006). SLCO1B1 polymorphism and sex affect the pharmacokinetics of pravastatin but not fluvastatin. *Clin. Pharmacol. Ther.* 80, 356-366.
90. Asselbergs, F.W., Williams, S.M., Hebert, P.R., Coffey, C.S., Hillege, H.L., Navis, G., Vaughan, D.E., van Gilst, W.H. and Moore, J.H. (2006). The gender-specific role of polymorphisms from the fibrinolytic, renin-angiotensin, and bradykinin systems in determining plasma t-PA and PAI-1 levels. *Thromb. Haemost.* 96, 471-477.
91. Yiannakouris, N., Melistas, L., Yannakoulia, M., Mungal, K. and Mantzoros, C.S. (2003). The-2548G/A polymorphism in the human leptin gene promoter region is associated with plasma free leptin levels; interaction with adiposity and gender in healthy subjects. *Hormones (Athens)* 2, 229-236.
92. Seripa, D., Franceschi, M., Matera, M.G., Panza, F., Kehoe, P.G., Gravina, C., Orsitto, G., Solfrizzi, V., Di Minno, G., Dallapiccola, B. and Pilotto, A. (2006). Sex differences in the association of apolipoprotein E and angiotensin-converting enzyme gene polymorphisms with healthy aging and longevity: a population-based study from Southern Italy. *J. Gerontol. A Biol. Sci. Med. Sci.* 61, 918-923.
93. Paladino, N., Fainboim, H., Theiler, G., Schroder, T., Muñoz, A.E., Flores, A.C., Galdame, O. and Fainboim, L. (2006). Gender susceptibility to chronic hepatitis C virus infection associated with interleukin 10 promoter polymorphism. *J. Virol.* 80, 9144-9150.
94. Körner, A., Ma, L., Franks, P.W., Kiess, W., Baier, L.J., Stumvoll, M. and Kovacs, P. (2007). Sex-specific effect of the Val1483Ile polymorphism in the fatty acid synthase gene (FAS) on body mass index and lipid profile in Caucasian children. *Int J Obes (Lond)* 31, 353-358.
95. Sundar, P.D., Feingold, E., Minster, R.L., DeKosky, S.T. and Kamboh, M.I. (2007). Gender-specific association of ATP-binding cassette transporter 1 (ABCA1) polymorphisms with the risk of late-onset Alzheimer's disease. *Neurobiol. Aging* 28, 856-862.
96. Ozawa, T., Ito, M., Tamaki, S., Yao, T., Ashihara, T., Kita, Y., Okamura, T., Ueshima, H. and Horie, M. (2006). Gender and age effects on ventricular repolarization abnormality in Japanese general carriers of a G643S common single nucleotide polymorphism for the KCNQ1 gene. *Circ. J.* 70, 645-650.
97. Ben Assayag, E., Bova, I., Berliner, S., Peretz, H., Usher, S., Shapira, I. and Bornstein, N.M. (2006). Gender differences in the expression of erythrocyte aggregation in relation to B beta-fibrinogen gene polymorphisms in apparently healthy individuals. *Thromb. Haemost.* 95, 428-433.
98. Kates, W.R., Antshel, K.M., Abdulsabur, N., Colgan, D., Funke, B., Fremont, W., Higgins, A.M., Kucherlapati, R. and Shprintzen, R.J. (2006). A gender-moderated effect of a functional COMT polymorphism on prefrontal brain morphology and function in velo-cardio-facial syndrome (22q11.2 deletion syndrome). *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 141B, 274-280.
99. Mizuno, T., Aoki, M., Shimada, Y., Inoue, M., Nakaya, K., Takahashi, T., Itoyama, Y., Kanazawa, M., Utsumi, A., Endo, Y., Nomura, T., Hiratsuka, M., Mizugaki, M., Goto, J., Hongo, M. and Fukudo, S. (2006). Gender difference in association

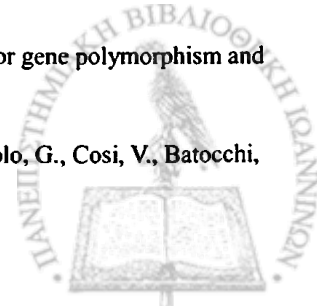


between polymorphism of serotonin transporter gene regulatory region and anxiety. *J Psychosom Res* 60, 91-97.

100. Mlynarski, W.M., Placha, G.P., Wolkow, P.P., Bochenski, J.P., Warram, J.H. and Krolewski, A.S. (2005). Risk of diabetic nephropathy in type 1 diabetes is associated with functional polymorphisms in RANTES receptor gene (CCR5): a sex-specific effect. *Diabetes* 54, 3331-3335.
101. Derzbach, L., Treszl, A., Balogh, A., Vásárhelyi, B., Tulassay, T. and Rigó J, J. (2005). Gender dependent association between perinatal morbidity and estrogen receptor-alpha PvuII polymorphism. *J Perinat Med* 33, 461-462.
102. Gloria-Bottini, F., Bottini, N., Cosmi, E., Cosmi, E.V. and Bottini, E. (2005). The effect of gender and ACP1 genetic polymorphism on the correlation between birth weight and placental weight. *Placenta* 26, 846-848.
103. Sjöberg, R.L., Nilsson, K.W., Nordquist, N., Ohrvik, J., Leppert, J., Lindström, L. and Orelund, L. (2006). Development of depression: sex and the interaction between environment and a promoter polymorphism of the serotonin transporter gene. *Int. J. Neuropsychopharmacol.* 9, 443-449.
104. Tan, E., Chong, S., Wang, H., Chew-Ping Lim, E. and Teo, Y. (2005). Gender-specific association of insertion/deletion polymorphisms in the nogo gene and chronic schizophrenia. *Brain Res. Mol. Brain Res.* 139, 212-216.
105. Shi, Q., Zhang, Z., Li, G., Pillow, P.C., Hernandez, L.M., Spitz, M.R. and Wei, Q. (2005). Sex differences in risk of lung cancer associated with methylene-tetrahydrofolate reductase polymorphisms. *Cancer Epidemiol. Biomarkers Prev.* 14, 1477-1484.
106. Foltynie, T., Lewis, S.G.J., Goldberg, T.E., Blackwell, A.D., Kolachana, B.S., Weinberger, D.R., Robbins, T.W. and Barker, R.A. (2005). The BDNF Val66Met polymorphism has a gender specific influence on planning ability in Parkinson's disease. *J. Neurol.* 252, 833-838.
107. Kajinami, K., Brousseau, M.E., Lamon-Fava, S., Ordovas, J.M. and Schaefer, E.J. (2005). Gender-specific effects of estrogen receptor alpha gene haplotype on high-density lipoprotein cholesterol response to atorvastatin: interaction with apolipoprotein AI gene polymorphism. *Atherosclerosis* 178, 331-338.
108. Kantarci, O.H., Goris, A., Hebrink, D.D., Heggarty, S., Cunningham, S., Alloza, I., Atkinson, E.J., de Andrade, M., McMurray, C.T., Graham, C.A., Hawkins, S.A., Billiau, A., Dubois, B., Weinshenker, B.G. and Vandenbroeck, K. (2005). IFNG polymorphisms are associated with gender differences in susceptibility to multiple sclerosis. *Genes Immun.* 6, 153-161.
109. Chang, J., Liu, C., Chuang, H., Ou, C., Hsu, T., Huang, E. and Yang, K.D. (2004). Gender-limited association of cytotoxic T-lymphocyte antigen-4 (CTLA-4) polymorphism with cord blood IgE levels. *Pediatr Allergy Immunol* 15, 506-512.
110. Cui, J., Randell, E., Renouf, J., Sun, G., Han, F., Younghusband, B. and Xie, Y. (2004). Gender dependent association of thrombospondin-4 A387P polymorphism with myocardial infarction. *Arterioscler. Thromb. Vasc. Biol.* 24, e183-4.
111. Schrijver, H.M., Crusius, J.B.A., García-González, M.A., Polman, C.H., Peña, A.S., Barkhof, F. and Uitdehaag, B.M.J. (2004). Gender-related association between the TGFB1+869 polymorphism and multiple sclerosis. *J. Interferon Cytokine Res.* 24, 536-542.
112. Liu, G., Zhou, W., Wang, L.I., Park, S., Miller, D.P., Xu, L.L., Wain, J.C., Lynch, T.J., Su, L. and Christiani, D.C. (2004). MPO and SOD2 polymorphisms, gender, and the risk of non-small cell lung carcinoma. *Cancer Lett.* 214, 69-79.
113. Szczeklik, W., Sanak, M. and Szczeklik, A. (2004). Functional effects and gender association of COX-2 gene polymorphism G-765C in bronchial asthma. *J. Allergy Clin. Immunol.* 114, 248-253.
114. Corder, E.H., Ghebremedhin, E., Taylor, M.G., Thal, D.R., Ohm, T.G. and Braak, H. (2004). The biphasic relationship between regional brain senile plaque and neurofibrillary tangle distributions: modification by age, sex, and APOE polymorphism. *Ann. N. Y. Acad. Sci.* 1019, 24-28.
115. Chen, W., Duan, S., Zhou, J., Sun, Y., Zheng, Y., Gu, N., Feng, G. and He, L. (2004). A case-control study provides evidence of association for a functional polymorphism -197C/G in XBP1 to schizophrenia and suggests a sex-dependent effect. *Biochem. Biophys. Res. Commun.* 319, 866-870.
116. Kajinami, K., Brousseau, M.E., Ordovas, J.M. and Schaefer, E.J. (2004). Polymorphisms in the multidrug resistance-1



- (MDR1) gene influence the response to atorvastatin treatment in a gender-specific manner. *Am. J. Cardiol.* 93, 1046-1050.
117. Yang, K.D., Liu, C., Chang, J., Chuang, H., Ou, C., Hsu, T. and Wang, C. (2004). Polymorphism of the immune-braking gene CTLA-4 (+49) involved in gender discrepancy of serum total IgE levels and allergic diseases. *Clin. Exp. Allergy* 34, 32-37.
 118. Gong, M.N., Wei, Z., Xu, L., Miller, D.P., Thompson, B.T. and Christiani, D.C. (2004). Polymorphism in the surfactant protein-B gene, gender, and the risk of direct pulmonary injury and ARDS. *Chest* 125, 203-211.
 119. Ko, Y., Hsu, L., Hsu, K., Ko, Y. and Lee, Y. (2004). The interactive effects of hepatic lipase gene promoter polymorphisms with sex and obesity on high-density-lipoprotein cholesterol levels in Taiwanese-Chinese. *Atherosclerosis* 172, 135-142.
 120. Nakanishi, S., Yamane, K., Kamei, N., Okubo, M. and Kohno, N. (2004). The effect of polymorphism in the intestinal fatty acid-binding protein 2 gene on fat metabolism is associated with gender and obesity amongst non-diabetic Japanese-Americans. *Diabetes Obes Metab* 6, 45-49.
 121. Chen, Y., Vaughan, R.W., Kondeatis, E., Fortune, F., Graham, E.M., Stanford, M.R. and Wallace, G.R. (2004). Chemokine gene polymorphisms associate with gender in patients with uveitis. *Tissue Antigens* 63, 41-45.
 122. Espino-Montoro, A., Barrios-Artillo, M., López-Chozas, J.M., Cayuela, A., Stiefel, P. and Villar, J. (2003). Influence of polymorphism (RFLP-sstI) at the apolipoprotein C-III gene locus on the lipoprotein metabolism and insulin resistance in essential hypertensive patients. Interaction between gender and genetic polymorphism. *Nutr Metab Cardiovasc Dis* 13, 194-201.
 123. Vandembroeck, K., Cunningham, S., Goris, A., Alloza, I., Heggarty, S., Graham, C., Bell, A. and Rooney, M. (2003). Polymorphisms in the interferon-gamma/interleukin-26 gene region contribute to sex bias in susceptibility to rheumatoid arthritis. *Arthritis Rheum.* 48, 2773-2778.
 124. Song, J., Narita, I., Goto, S., Saito, N., Omori, K., Sato, F., Ajiro, J., Saga, D., Kondo, D., Sakatsume, M. and Gejyo, F. (2003). Gender specific association of aldosterone synthase gene polymorphism with renal survival in patients with IgA nephropathy. *J. Med. Genet.* 40, 372-376.
 125. Laule, M., Meisel, C., Prauka, I., Cascorbi, I., Malzahn, U., Felix, S.B., Baumann, G., Roots, I., Stangl, K. and Stangl, V. (2003). Interaction of CA repeat polymorphism of the endothelial nitric oxide synthase and hyperhomocysteinemia in acute coronary syndromes: evidence of gender-specific differences. *J Mol Med* 81, 305-309.
 126. Reich, H., Duncan, J.A., Weinstein, J., Cattran, D.C., Scholey, J.W. and Miller, J.A. (2003). Interactions between gender and the angiotensin type 1 receptor gene polymorphism. *Kidney Int.* 63, 1443-1449.
 127. Watson, M.A., Gay, L., Stebbings, W.S.L., Speakman, C.T.M., Bingham, S.A. and Loktionov, A. (2003). Apolipoprotein E gene polymorphism and colorectal cancer: gender-specific modulation of risk and prognosis. *Clin. Sci.* 104, 537-545.
 128. Stanković, A., Zivković, M. and Alavantić, D. (2002). Angiotensin I-converting enzyme gene polymorphism in a Serbian population: a gender-specific association with hypertension. *Scand. J. Clin. Lab. Invest.* 62, 469-475.
 129. Györfy, B., Vászrhelyi, B., Krikovszky, D., Madácsy, L., Tordai, A., Tulassay, T. and Szabó, A. (2002). Gender-specific association of vitamin D receptor polymorphism combinations with type 1 diabetes mellitus. *Eur. J. Endocrinol.* 147, 803-808.
 130. Minihane, A.M., Finnegan, Y.E., Talmud, P., Leigh-Firbank, E.C. and Williams, C.M. (2002). Influence of the APOC3 -2854T>G polymorphism on plasma lipid levels: effect of age and gender. *Biochim. Biophys. Acta* 1583, 311-314.
 131. Wu, Y., Tomon, M. and Sumino, K. (2001). Methylene tetrahydrofolate reductase gene polymorphism and ischemic stroke: sex difference in Japanese. *Kobe J Med Sci* 47, 255-262.
 132. Lio, D., Scola, L., Crivello, A., Colonna-Romano, G., Candore, G., Bonafè, M., Cavallone, L., Franceschi, C. and Caruso, C. (2002). Gender-specific association between -1082 IL-10 promoter polymorphism and longevity. *Genes Immun.* 3, 30-33.
 133. Liu, Z., Cheng, Z., Gong, R., Liu, H., Liu, D. and Li, L. (2002). Sex differences in estrogen receptor gene polymorphism and its association with lupus nephritis in Chinese. *Nephron* 90, 174-180.
 134. Sciacca, F.L., Ferri, C., Veglia, F., Andreatta, F., Mantegazza, R., Cornelio, F., Franciotta, D., Piccolo, G., Cosi, V., Batocchi,



- A.P., Evoli, A. and Grimaldi, L.M.E. (2002). IL-1 genes in myasthenia gravis: IL-1A -889 polymorphism associated with sex and age of disease onset. *J. Neuroimmunol.* 122, 94-99.
135. Ordovas, J.M., Corella, D., Cupples, L.A., Demissie, S., Kelleher, A., Coltell, O., Wilson, P.W.F., Schaefer, E.J. and Tucker, K. (2002). Polyunsaturated fatty acids modulate the effects of the APOA1 G-A polymorphism on HDL-cholesterol concentrations in a sex-specific manner: the Framingham Study. *Am. J. Clin. Nutr.* 75, 38-46.
136. Oh, J.Y., Barrett-Connor, E. (2001). Apolipoprotein E polymorphism and lipid levels differ by gender and family history of diabetes: the Rancho Bernardo Study. *Clin. Genet.* 60, 132-137.
137. Ellsworth, D.L., Bielak, L.F., Turner, S.T., Sheedy, P.F.2., Boerwinkle, E. and Peyser, P.A. (2001). Gender- and age-dependent relationships between the E-selectin S128R polymorphism and coronary artery calcification. *J Mol Med* 79, 390-398.
138. Du, L., Bakish, D. and Hrdina, P.D. (2000). Gender differences in association between serotonin transporter gene polymorphism and personality traits. *Psychiatr. Genet.* 10, 159-164.
139. Bullido, M.J., Guallar-Castillón, P., Artiga, M.J., Ramos, M.C., Sastre, I., Aldudo, J., Frank, A., Coria, F., Rodríguez-Artalejo, F. and Valdivieso, F. (2000). Alzheimer's risk associated with human apolipoprotein E, alpha-2 macroglobulin and lipoprotein receptor related protein polymorphisms: absence of genetic interactions, and modulation by gender. *Neurosci. Lett.* 289, 213-216.
140. Kark, J.D., Sinnreich, R., Leitersdorf, E., Friedlander, Y., Shpitzen, S. and Luc, G. (2000). Taq1B CETP polymorphism, plasma CETP, lipoproteins, apolipoproteins and sex differences in a Jewish population sample characterized by low HDL-cholesterol. *Atherosclerosis* 151, 509-518.
141. Durlach, A., Clavel, C., Girard-Globa, A. and Durlach, V. (1999). Sex-dependent association of a genetic polymorphism of cholesteryl ester transfer protein with high-density lipoprotein cholesterol and macrovascular pathology in type II diabetic patients. *J. Clin. Endocrinol. Metab.* 84, 3656-3659.
142. Sagnella, G.A., Rothwell, M.J., Onipinla, A.K., Wicks, P.D., Cook, D.G. and Cappuccio, F.P. (1999). A population study of ethnic variations in the angiotensin-converting enzyme I/D polymorphism: relationships with gender, hypertension and impaired glucose metabolism. *J. Hypertens.* 17, 657-664.
143. Reynolds, W.F., Rhees, J., Maciejewski, D., Paladino, T., Sieburg, H., Maki, R.A. and Masliah, E. (1999). Myeloperoxidase polymorphism is associated with gender specific risk for Alzheimer's disease. *Exp. Neurol.* 155, 31-41.
144. Freire, M.B., Ji, L., Onuma, T., Orban, T., Warram, J.H. and Krolewski, A.S. (1998). Gender-specific association of M235T polymorphism in angiotensinogen gene and diabetic nephropathy in NIDDM. *Hypertension* 31, 896-899.
145. Suarez, F., Zeghoud, F., Rossignol, C., Walrant, O. and Garabédian, M. (1997). Association between vitamin D receptor gene polymorphism and sex-dependent growth during the first two years of life. *J. Clin. Endocrinol. Metab.* 82, 2966-2970.
146. Lehtimäki, T., Frankberg-Lakkala, H., Solakivi, T., Koivisto, A.M., Laippala, P., Ehnholm, C., Jokela, H., Koivula, T. and Nikkari, T. (1997). The effect of short-term fasting, apolipoprotein E gene polymorphism, and sex on plasma lipids. *Am. J. Clin. Nutr.* 66, 599-605.
147. Carter, A.M., Catto, A.J., Bamford, J.M. and Grant, P.J. (1997). Gender-specific associations of the fibrinogen B beta 448 polymorphism, fibrinogen levels, and acute cerebrovascular disease. *Arterioscler. Thromb. Vasc. Biol.* 17, 589-594.
148. Ferrières, J., Sing, C.F., Roy, M., Davignon, J. and Lussier-Cacan, S. (1994). Apolipoprotein E polymorphism and heterozygous familial hypercholesterolemia. Sex-specific effects. *Arterioscler. Thromb.* 14, 1553-1560.
149. von Eckardstein, A., Funke, H., Chirazi, A., Chen-Haudenschild, C., Schulte, H., Schönfeld, R., Köhler, E., Schwarz, S., Steinmetz, A. and Assmann, G. (1994). Sex-specific effects of the glutamine/histidine polymorphism in apo A-IV on HDL metabolism. *Arterioscler. Thromb.* 14, 1114-1120.
150. Assmann, S.F., Pocock, S.J., Enos, L.E. and Kasten, L.E. (2000). Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet* 355, 1064-1069.
151. Hernández, A.V., Steyerberg, E.W., Taylor, G.S., Marmarou, A., Habbema, J.D.F. and Maas, A.I.R. (2005). Subgroup analysis

